Sraffa and the Reconstruction of Economic Theory: Volume Two

Aggregate Demand, Policy Analysis and Growth

Edited by
Enrico Sergio Levrero,
Antonella Palumbo and
Antonella Stirati

Sraffa and the Reconstruction of Economic Theory: Volume Two

Aggregate Demand, Policy Analysis and Growth

Edited by

Enrico Sergio Levrero

Antonella Palumbo

and

Antonella Stirati

2013

palgrave
macmillan

Sraffa and the Reconstruction of Economic Theory: Volume Two

Aggregate Demand, Policy Analysis and Growth

Edited by

Enrico Sergio Levrero

Antonella Palumbo

and

Antonella Stirati

palgrave
macmillan

Contents

List of Contributors

Sergio Cesaratto, University of Siena

Roberto Ciccone, Roma Tre University

Óscar Dejuán, University of Castilla – La Mancha

Esther Dweck, IE, Federal University of Rio de Janeiro (UFRJ)

Fabio N. P. de Freitas, IE, Federal University of Rio de Janeiro (UFRJ)

Paolo Leon, Roma Tre University

Carlos Aguiar de Medeiros, IE, Federal University of Rio de Janeiro (UFRJ)

Antonella Palumbo, Roma Tre University

Fabio Petri, University of Siena

Franklin Serrano, IE, Federal University of Rio de Janeiro (UFRJ)

Matthew Smith, University of Sydney

Attilio Trezzini, Roma Tre University

Introduction

Antonella Palumbo

One of the fields of research to which the modern revival of the surplus approach has devoted much attention, in the fifty years following Sraffa's *Production of Commodities by Means of Commodities*, is the theory of output and the analysis of the processes of growth and development in capitalist economies. Although no explicit theorisation about the determinants of output levels is to be found in Sraffa's work, the revival of the classical approach to value and distribution and the critique of neoclassical theory prompted by his work, as shown in Garegnani's (1978–79) seminal contribution, have relevant implications for the analysis of output.

On the one hand, the critique of neoclassical mechanisms based on factor substitutability implies (i) the lack of the analytical basis on which traditional analysis affirms the automatic tendency of the system to the full utilisation of resources and (ii) a criticism of the reduction of Keynes's contribution to the analysis of the transitory incapability of the system of ensuring sufficient demand to absorb potential output. On the other hand, the fact that in the classical theory of value and distribution output levels are regarded as given when determining relative prices and the residual distributive variable (in the so-called 'core' of surplus analysis – Garegnani, 1984), implies a degree of freedom with respect to output determination and thus compatibility with different theories of output. In the words of Garegnani (1978, p. 340), 'Ricardo's theory of distribution is *open*, in the sense that it *neither* provides premises capable of justifying the tendency of investment to adjust to saving, *nor* depends on the existence of such a tendency. This "open" character sharply distinguishes the Ricardian theory from the subsequent marginalist theories, which saw distribution as the result of forces of demand for and supply of "factors of production".'

The eleven essays contained in this volume may be said to build on these original insights, by contributing new material to the general aim of explaining the determination of output levels and their evolution over time on the basis of two fundamental notions: the absence of any mechanisms whatsoever ensuring the tendency of the system to the optimal use of resources (and thus the necessity of studying the evolution of demand and the limits it may pose on the level and growth of output); and the fundamental independence between output determination and distribution, or, more accurately, the absence of any *necessary* relation between them.

This independence, which stems from the very structure of the classical theory of value and distribution, is a distinctive element of the long-period theory of output developed within the modern revival of the surplus approach, which sets it apart from other theories attempting the extension of the principle of effective demand to the long period on different bases.

One of the most influential of such attempts has been the 'Cambridge' theory of distribution (Kaldor, 1955–56; Robinson, 1956 and 1962) postulating a strict and necessary direct relation between the rate of accumulation and the rate of profit. Starting from the 'Keynesian hypothesis' (Kaldor, 1955–56, p. 195) according to which it is the level of investment, independently determined, which causes and generates the corresponding level of saving, the Cambridge approach interprets this causation as entailing that in the long period any autonomous increase in the rate of accumulation must necessarily be met by a corresponding increase in the ratio of saving to the stock of capital, which can only be obtained if the proportion of profits in national income grows, given the higher propensity to save out of profits. The Cambridge theory of distribution is thus based on the assumption that in the long period the overall level of output does not have sufficient elasticity in responding to autonomous changes in investment, and that for this reason the composition of output between consumption and saving must necessarily change to accommodate the higher rate of investment. This would happen through an increase of money prices relative to money wages (induced by the temporary excess demand) which will produce the necessary fall in real wages.

However, as shown by Garegnani (1992) and Garegnani and Palumbo (1998), such rigidity of long-run output to independent changes in demand – which in Kaldor (1955–56) takes the form of the full-employment hypothesis while in Robinson (1956, 1962) is represented by the assumption of the long-run rigidity of the output/capital ratio – is in reality unwarranted when taking seriously the conception of the growth process as driven by demand. An essential part of this conception is in fact the notion that productive resources, far from being exogenously determined, are mainly the result of the long-run level of activity itself. This happens through the processes of creation or destruction of productive capacity that a prolonged over-utilisation or under-utilisation of installed capacity entail, while other mechanisms (such as migration, changes in the participation rate, transfers between low-productivity and high-productivity sectors – see Kaldor, 1985) are at work allowing for a similar flexibility in the supply of labour force in response to changes in the demand for labour.

Such mechanisms typically operate over long periods, so that some discrepancies between available resources and their utilisation are always observable; at the same time, the fact that unused resources tend slowly to disappear from sight implies that there is no need to observe great such discrepancies in order to recognise the wide margins of elasticity with which output may accommodate changes in demand. The necessary link between accumulation and distribution that the Cambridge theory of distribution postulates thus disappears, and the necessary trade-off between investment and consumption which characterises that theory disappears as well. Not only does this allow for a separate determination of distribution on the basis of a complex of social and institutional forces, as entailed by the classical surplus analysis (see the essays in Volume I); it also implies that the fundamental conclusion of Keynes's analysis as regards the short period, namely that a *direct* relation of causation may be postulated between investment and consumption, is confirmed and extended to the long period. All the authors contributing to this volume share this common theoretical ground and interpret the long-run principle of effective demand as entailing that the variations in the level of output act as the adjusting mechanism between investment and saving, exactly as postulated in Keynes's short-run analysis.

Similarly to what happens in the determination of distribution, the ultimate determinants of growth are to be sought in a complex of socio-historical circumstances which defy any simple and mechanical analysis of causation. Growth cannot be considered in merely quantitative terms as the increase over time of an aggregate indifferentiated amount of product; it has rather to be regarded as 'a complex process, entailing structural change of the economic system, such that it can only be plausibly explained in concrete terms by reference to social, politico-institutional and technological factors' (as Smith maintains in his essay in this volume). This also entails that the clear-cut distinction traditionally to be found in economic literature between the theory of growth and the theory of development loses part of its meaning.

This theoretical attitude with respect to the analysis of growth is not only an expression of the above-mentioned separability and mutual independence between the determination of distribution and the determination of output, which allows a plurality of determining factors to be taken into account in each: it is also an expression of the particular method of analysis which, following the lead of classical economists, is adopted in the modern revival of the surplus approach when dealing with such questions. According to Garegnani's (1984) reconstruction (see the Introduction to Volume I), outside the purely deductive and analytical 'core' of their theories (devoted to the study of the general and necessary relations between relative prices and distribution), the classical economists used to address such questions as the determinants of distribution or of accumulation (the 'intermediate data' of the core) by means of an analysis made up of a mixture of deduction, observation and generalisation from experience; where no precise and mechanical relations are postulated between variables, but complex systems of influences and interrelations are considered (this also implies the possibility of considering mutual interrelations between accumulation and distribution, provided they are part of this complex system of influences and are not regarded as necessary and mechanical relations).

This very methodological characteristic implies that in the theory of output and in the study of the determinants of growth, similarly to what happens in the theory of distribution, the classical approach requires a method of analysis in which institutional and social factors play a relevant role, so that no abstract general theorisation is possible without reference to the historical conditions in which economies develop. In principle, this implies that no neat distinction can be traced, when analysing the determinants of distribution or those of output, between pure theory and applied analysis.

To different degrees, the essays contained in this volume (as well as in Part II of the first volume, 'The Revival and Development of the Classical Theory of Distribution') bear out this methodological attitude; and this is also the reason why no sharp line has been drawn, in conceiving this volume, between a theoretical and an applied part. Rather, the two parts into which the volume is divided deal respectively with 'Demand-Led Growth in the Classical Approach' (the first seven essays) and 'Understanding the International Economic Order' (the last four essays).

In the first essay of Part I Ciccone addresses a problem which currently has a relevant impact on collective choices and policy debates, i.e., the role and management of public debt. The aim of his chapter is to show how deeply the attitude towards public debt and its policy implications may change when regarded from the theoretical angle assumed in this volume.

According to the 'view of total output as governed by aggregate demand independently of the availability of resources', public debt cannot be seen as subtracting savings to private accumulation, so that no trade-off arises in principle between public and private expenditure. The chapter highlights the potential expansionary effect of public deficits, which act on output both directly and indirectly by inducing higher capacity-building investment, and critically reviews the conclusions of mainstream literature on fiscal multipliers. On these premises, Ciccone shows analytically the potential perverse effects of restrictive fiscal policies aimed at reducing the debt/GDP ratio by comparing the long-run effects of different policy regimes, and maintains that lower values of the debt ratio might be rather achieved, under certain conditions, through expansionary fiscal policies.

Petri's chapter is conceived as a critique of the neoclassical theory of investment. Apart from the decisive critiques that have been directed at it following the capital debates, which show the fallacious bases on which the decreasing demand for factors is obtained when heteregeneous capital is considered among them, Petri finds a second weakness in the neoclassical theory of investment, which would be present even conceding the working of the factor substitution mechanism, and has to do with the necessary dependence of investment on output.

He shows that factor substitution, once the durability of fixed capital is taken into account, necessarily operates only on the investment flows. This implies that factor substitution operates slowly and that during the very long adjustment process its effects interact with the effects of other forces (like the influence of investment on aggregate demand and the influence of demand on the level of production), producing in the end the possibility that the relationship between the rate of interest and investment is in some circumstances opposite to what theory prescribes. More accurately, the rate of interest proves insufficient to determine investment as long as the quantity of labour employed at each stage of the adjustment process is indeterminate. The analysis also implies that no mechanism can be immediately envisaged, even conceding factor substitution, that automatically increases the demand for labour in response to falling wages.

The following five chapters all deal, in a more or less direct way, with the controversial methodological issue of the analytical tool to be used in studying the long-run tendencies of the system, which has been the object of much debate within the approach.

Controversy has particularly centred around the notion of 'supermultiplier', originally developed by Hicks (1950), which has been proposed by Serrano (1995) as a useful representation of the long-period relationship between the level of output and autonomous demand that in his view can be postulated in the analysis of growth within the modern revival of the surplus approach (other supermultiplier models are to be found in Bortis, 1997 and Dejuán, 2005). The supermultiplier is determined on the hypothesis that capacity-building investment is to be regarded, in the long run, as an entirely induced variable, while other components of demand (especially current public expenditure and exports) may be regarded as completely autonomous. This kind of 'unproductive' (in the sense of being unable to modify productive capacity) or 'final' expenditure would be the driving force of growth, since its variations induce in the long period changes in output both through consumption and by inducing changes in investment aimed at adjusting capacity to demand.

The advocates of the supermultiplier model stress its capability of replicating the simple Keynesian formula of the multiplier in an equally simple relation valid for the long run, where the capacity effects of investment are considered; the possibility of restating the autonomous role of demand in driving the growth process, since variations in the growth rate of autonomous demand are accommodated by changes in the relative shares of autonomous expenditure and induced investment in total output; and the independence of distribution from accumulation entailed by the model. The critics (especially Trezzini, 1995, 1998; Palumbo and Trezzini, 2003) question the assumption of average normal utilisation of capacity, by showing that any change in the rate of growth of autonomous demand implies a process of adjustment whereby the degree of capacity utilisation proves different from normal, even on average; consequently they believe that the validity of the model is limited to the extremely artificial case of autonomous expenditure growing for an unlimited long period at a constant rate (Trezzini, 1998, p. 66), and regard such an assumption as in contrast with the assumed autonomy of demand. More generally, Palumbo and Trezzini (2003) show that, from a logical point of view, the variability (both in the short and the long run) of the degree of utilisation of productive capacity is the way in which the autonomy of demand manifests itself, by producing the very discrepancies between demand and capacity that imply the need for adjustment of the latter. Analysis of the actual way in which the process of adjustment operates shows its lengthy and uncertain nature, due to the fact that the process itself is bound to generate offsetting forces which may render the complete adjustment between production and capacity an unrealistic occurrence, even without considering the variability of capacity requirements induced by technical progress. While not challenging the tendency to a uniform rate of profit and the gravitation of market prices towards normal prices, which occur on the flow of newly produced capacity (on the issue of gravitation, see the Introduction to Volume III), the systematic incapability of the system of attaining 'fully adjusted positions' (the definition is due to Vianello, 1985) challenge their role as reference points in the analysis of the evolution of quantities.

The chapter by Trezzini contained in this volume addresses this methodological question from a general point of view, by investigating the double relationship, in the analysis of growth, between actual and theoretical magnitudes, on the one side, and cyclical and trend positions of the economy, on the other side. His main thesis is that the prevailing method of studying the trend of produced quantities 'through growth paths whose characteristics ... are defined regardless of the quantities actually produced' derives from an 'unjustified *transposition* of the relation between theoretical and actual magnitudes developed in the theory of prices into the theory of the evolution of quantities'. On the basis of the different nature of the forces at stake in the determination of produced quantities, he advocates a different method of analysis capable of taking into account the fact that the pattern of cyclical fluctuations may affect the economy's trend by inducing irreversible changes (such as the presence of new capacity or its destruction, or the acquisition of consumption standards) which may not be considered as merely accidental. Both the absence of any theoretical reason to consider the determinants of trend magnitudes as persistent and independent of cyclical phenomena, and the unrealism of the hypothesis that the components of autonomous demand should grow along a smooth regular path, imply that there is no solid foundation to the attempt at establishing a relationship between normal and actual quantities analogous to the relationship between normal and actual prices.

The main thrust of Palumbo's chapter is the critique of the notion of 'potential output' which is to be found in mainstream literature and the critical analysis of the methodologies used by international institutions to obtain empirical estimates of potential output. Estimation methods are shown to be deeply influenced by the dominant theoretical framework in which potential output is exogenously determined by supply factors; on the other hand, the actual content of empirical estimates is often heavily dependent on the time series of actual output, in the theoretical (unproven) belief that actual output cannot have deviated too markedly from potential, the latter being a strong attractor of the former. Some puzzling results of this literature, especially as regards the output–inflation relationship assumed by mainstream theory, are examined. Against the limitations of the supply-side approach, potential output is regarded within the demand-led growth framework as endogenous and dependent on the economy's actual realisations. It is in this context that Palumbo examines the debate about the analytical role of the fully adjusted positions within the demand-led growth approach and concludes that a meaningful definition and quantification of potential output can only be referred to the short period, the long-run evolution of potential being a path-dependent phenomenon.

In his contribution, Smith is also critical of the limitations of the assumption of normal utilisation in the demand-led growth approach. After having exposed the limitations of a rigid supermultiplier model, he proposes an alternative growth model with variable utilisation both in the short and the long run. Though in a simplified way, the model contains an attempt at historical periodisation of the relationship between autonomous demand and aggregate output. This is achieved by letting the value of the supermultplier change period by period (each historical time period characterised by specific relations between variables is called an 'epoch'; each epoch is assumed to be long enough for fixed productive capacity to fully adjust to expected demand). Unlike in 'steady-state' models, the growth rate depends in Smith's model not only on the growth rate of autonomous demand but also on the long-run change in the value of the supermultiplier which may be influenced by such phenomena as changes in income distribution and technical change. Average utilisation of capacity over the whole timespan in the end proves to be different from normal, although the model does not allow for changes in utilisation within each period. The simplicity of the model cannot, however, obscure, in Smith's opinion, the complex interactions that exist among the different variables influencing the growth process, but rather aims at representing the basic relation between autonomous demand and output as historically contingent.

Differently from the chapters just described, in the last two chapters of Part I the supermultiplier model, together with the assumption of normal utilisation of capacity, is considered as a correct and useful representation of the actual relationship between output and autonomous demand, and a fundamental element of the theory of output in the modern classical approach.

Dejuán's paper contains a restatement of the supermultiplier model with a number of qualifications and clarifications. He particularly discusses the issue of stability, by attributing the extreme instability of Harrod's (1939) model to the excessive influence of the actual rate of capacity utilisation on the expected rate of growth. In his view, 'models that incorporate the accelerator mechanism will be stable provided demand expectations depend on objective factors (like the autonomous trend) instead of the ups and downs of current income and capacity utilization'. In order for the autonomous trend to represent an objective factor to which expectations may be anchored, however, the same rate of growth of autonomous demand has to last long enough to be perceived as persistent by entrepreneurs: 'Certainly, in epochs without a well established autonomous trend, the economy will suffer from macroeconomic instability'. A second qualification is that the model must be formulated in terms of a flexible supermultiplier if the latter has to result in a stabilising mechanism. This is obtained by making investment depend not only on the expected growth of demand but also on the deviations from normal capacity.

In their paper Freitas and Dweck use the supermultiplier demand-led growth model to analyse the pattern of Brazilian economic growth from 1970 to 2005. The basic model is integrated with the idea that at least in some sub-periods the balance of payments was a binding constraint on Brazil's growth. However, the demand-led growth structure allows different possible growth regimes to be considered. In the case of Brazil, the authors identify both a balance of payments-constrained demand-led growth process, and a policy-constrained (or pure) demand-led growth process (in which the availability of foreign exchange is not the binding constraint on growth) as relevant for different sub-periods.

The authors also propose an exercise in 'demand-side' growth accounting. Their decomposition of the rate of growth aims both at assessing the role of domestic and external demand in the different phases of the growth process (with the conclusion that Brazilian growth was mainly led by the public sector contribution), and at showing the role in the different phases, respectively, of changes in autonomous demand and changes in the supermultiplier. The authors conclude that the latter has to be regarded as a changing magnitude, which implies that the identification of the historical factors capable of determining its changes must be an essential part of the analysis of growth determinants.

The four chapters in Part II of this volume address the changing features of the international economic order and use the theoretical insights provided by the approach as interpretative tools of observed phenomena.

Serrano focuses in particular on the behaviour of the commodities market and on the sustained increase of dollar commodity prices in the 2000s. He maintains that this trend cannot be explained by the great rise in Chinese demand, and that demand factors (whether final demand or demand for speculation) seem not to be strong enough to account for the intensity and the timing of the phenomenon. Nor do any explanations in terms of natural scarcity of the various commodities seem relevant. Rather, on the basis of the classical analysis of the cost of production as the main determinant of price, he draws attention to the various factors which have contributed to an increase in the relative cost of production of commodities *comprehensive of rents*: the deliberate restriction of the supply of oil, obtained through the convergent action of OPEC policies and 'the revival of natural resource nationalism', the fast growth of real wages in commodity-exporting countries, the real appreciation of their currencies with respect to the dollar, the stagnation in unit costs of manufactured exports due to slow growth of real wages in advanced capitalist countries and fast growth in industrial productivity in the emerging countries which are new exporters of industrial products. He also emphasises the historically contingent nature of the new international economic order of the 2000s, and stresses some elements of continuity with the preceding phase such as the continuing heavy dependence of developing countries on American monetary policy and American financial markets.

Medeiros's contribution aims at assessing the decline of the developmental state and the transformations that governments and public agencies are undergoing in emerging countries as actors in the process of development. The perspective adopted is 'based on classical political economy, on Latin American structuralism and on a Gramscian perspective on the state's formation'. On these bases, he emphasises the role of different economic and social structures in influencing institutions, the uneven structure of the international environment with the hegemonic state's policies playing a crucial role, the nature of the state as 'a central institution where the dominant class or some of its sectors lead a coalition of power, building a hegemonic project compatible with a particular accumulation strategy'.

He identifies three different phases in the post-war period, the first based on national capital coordinated by national states (state-led industrialisation); the second based on the internationalisation of productive and finance capital and the opening of developing countries' markets to foreign capital (the neo-liberal phase); the third characterised by the return to expansionary Keynesian policies in developing countries in a more internationalised production context.

Cesaratto's contribution is centred on the different views of international relations that it is possible to find in the history of economic thought. He classifies the various theories into two main categories, those believing 'in the harmonic virtues of international laissez-faire' and those implying a conflictual view of international relations. He especially deals with this second group of theories, in search of insights that may contribute to a deeper understanding of current international arrangements, and particularly the current relationship, within the euro area, between Germany and Southern European countries. He reviews the contributions of various strands of thought (Mercantilism, the German Historical School, International Political Economy), and maintains that the Sraffian approach (together with Kalecki's theory of exports) may provide the useful insights scattered in these approaches with the sound analytical base they seem to lack.

In analysing some aspects of the great crisis of 2007 and the following recession, Leon's chapter addresses the issue of recent trends in distribution which is also dealt with in other contributions (Serrano in this volume, and Stirati and Pivetti in Volume I). His main thesis is that the direction of causation that some authors (among whom Pivetti) envisage between deteriorating distribution and increase in household debt should be reversed, with the latter potentially being a relevant force in determining the former. This is based on the idea that the continuous increase in wealth market values has produced an easy transformation of wealth into income, thus relaxing the liquidity constraint, and weakening the bargaining power of trade unions by rendering wage income less necessary for workers. This would explain, in Leon's view, why the lowering of real wages happened in a context of very high employment, supposedly entailing a relatively strong wage-bargaining position for workers.

References

Bortis, H. (1997) *Institutions, Behaviour and Economic Theory* (Cambridge: Cambridge University Press).

Dejuàn, Ó. (2005) 'Paths of accumulation and growth: towards a Keynesian long-period theory of output', *Review of Political Economy*, 17: 231–52.

Garegnani, P. (1978) 'Notes on consumption, investment and effective demand: I', *Cambridge Journal of Economics*, 2: 335–53.

Garegnani, P. (1979) 'Notes on consumption, investment and effective demand: II', *Cambridge Journal of Economics*, 3: 63–82.
Garegnani, P. (1984) 'Value and Distribution in the Classical Economists and Marx', *Oxford Economic Papers*, 36: 291–325.
Garegnani, P. (1992) 'Some Notes for an Analysis of Accumulation', in J. L. Halevi, D. Laibman, and E.J. Nell (eds), *Beyond the Steady State. A Revival of Growth Theory* (London: Macmillan).
Garegnani, P. and Palumbo, A. (1998) 'Accumulation of Capital', in H. D. Kurz and N. Salvadori (eds), *The Elgar Companion to Classical Economics* (Cheltenham: Edward Elgar).
Harrod, R., (1939) 'An essay on dynamic theory', *Economic Journal*, 49: 14–33.
Hicks, J. R. (1950) *A Contribution to the Theory of the Trade Cycle* (Oxford: Clarendon).
Kaldor, N. (1955–56) 'Alternative theories of distribution', *Review of Economic Studies*, 23: 63–100.
Kaldor, N. (1985) *Economics without Equilibrium* (Armonk, New York: Sharpe).
Palumbo, A. and Trezzini, A. (2003) 'Growth without normal capacity utilization', *European Journal of the History of Economic Thought*, 10: 109–35.
Robinson, J. (1956) *The Accumulation of Capital* (London: Macmillan).
Robinson, J. (1962) *Essays in the Theory of Economic Growth* (London: Macmillan).
Serrano, F. (1995) 'Long period effective demand and the Sraffian supermultiplier', *Contributions to Political Economy*, 14: 67–90.
Trezzini, A. (1995) 'Capacity utilisation in the long run and the autonomous components of aggregate demand', *Contributions to Political Economy*, 17: 33–66.
Trezzini, A. (1998) 'Capacity utilisation in the long run: some further considerations', *Contributions to Political Economy*, 17: 53–67.
Vianello, F. (1985) 'The pace of accumulation', *Political Economy*, 1: 69–87.

Part I

Demand-Led Growth in the Classical Approach

1
Public Debt and Aggregate Demand: Some Unconventional Analytics

Roberto Ciccone

1.1 Introduction

This chapter explores some basic questions about the effects on the economic system of financing public expenditure by issuing debt. It develops within a theoretical framework differing from that which is currently predominant both in pure theory and in applications to specific problems such as those addressed here. In particular, the approach in this chapter rests on two basic and closely related premises.

The first is the 'classical' explanation of income distribution in terms of circumstances of a social and institutional character, and hence the rejection of the view of distribution as determined by the forces of demand and supply held by the dominant theory in its various formulations. The second is the application of the Keynesian principle of effective demand to the long-run or trend levels of total output, which are regarded here as dependent on the size of the aggregate demand ultimately determined independently of output in conditions of full employment (or 'natural unemployment'). Aggregate demand therefore sets an upper limit on the levels of activity obtainable in the long run (and, needless to say, in the short run too), which would generally be insufficient to allow for the full employment of available resources.[1]

The close relationship between the two premises stems from the natural compatibility of the classical theory of distribution with a determination of output levels into which the available quantity of labour and other resources do not enter as directly relevant circumstances. The analytical structure of classical theory, characterised by the absence of relations of a *necessary* character between the determination of distribution and the determination of output levels, makes it possible to study changes in variables such as the aggregates of output, consumption, investment, public expenditure and so on with no need to consider the effects that those changes may have on distribution and the price system. The analysis will therefore be carried out on the assumption of a given distribution of income and associated price configuration.

Within the theoretical framework outlined above, the work seeks to single out some basic relations and propositions concerning the effects of public deficits and public debt that might, if valid, serve as a foundation to address the issue under the conditions possibly set by specific contexts.

1.2 Public debt as additional private wealth

Let us start by defining some relations between the relevant magnitudes. In a closed economy the necessary equality between aggregate expenditure and aggregate output, i.e., between total savings and total investment, entails the identity

$$S_p = I + D$$

where S_p is private savings, I private investment and D public deficit.

In our theoretical framework, the level of income is determined by the level of aggregate demand, i.e., by the sum of private and public expenditure, which implies that that identity is satisfied, through changes in the level of income, by the adjustment of the flow of private savings to the sum of private investment and public deficit. An increase in public deficit ΔD caused by a given increase in public expenditure is therefore counterbalanced by an equal increase in the flow of private savings ΔS_p. In other words, through the influence of public expenditure on aggregate demand and hence on the level of income, an increase in public deficit generates *additional* private savings, savings that would have not been formed in the absence of the increase in the deficit.

The flow of private savings is in turn equivalent to the addition to the stock of private sector wealth in the period, gross of capital depreciation:

$$\Delta W_p = S_p = I + D$$

where ΔW_P stands for the gross variation in private wealth. Our point is that public deficits add to private investment by generating savings, and therefore wealth, in the hands of the private sector. As a result, the size of private wealth at any given time would be larger, the larger the sum of the public deficits built up in the past.

It should be evident how this conclusion relates to the view of total output as governed by aggregate demand independently of the availability of resources. In neoclassical theory, where the level of income is determined (apart from temporary deviations) by the supply of resources with consequent adjustment of aggregate demand, the size of private wealth is instead basically independent of the flows of public deficits. In that theoretical framework, the condition of full employment, or 'natural' unemployment, entails a corresponding level of income with the associated amount of savings. Private savings and the formation of private wealth are therefore set independently of the sum of public deficits and private investment, with a trade-off arising between the latter two 'uses' of private savings – unless it is maintained, as in the 'Ricardian equivalence' argument, that public deficits cause an increase in the share of savings out of full-employment income, thus displacing private consumption rather than private investment.[2]

We can say therefore that within the theoretical framework adopted here, no constraint is imposed on the size of public debt by the size of private wealth, since the latter increases to the same extent as the increase in public debt. Nor is any particular significance to be attached to a variation in the proportion of private wealth constituted by government debt, as an increase in that proportion, for example, would be no more than the arithmetical consequence of the fact that the stock of public bonds held by private agents has grown proportionally more than other types of assets constituting private wealth, such as real capital, and should not be mistaken for a negative influence on capital accumulation.

While public expenditure and deficits may play a role in raising the level of aggregate demand and income, no negative consequences need therefore arise for the economic system from the related accumulation of public debt as such. It is, however, precisely the capacity of public expenditure to produce relevant and lasting effects on aggregate demand that is questioned in much of the current literature, often by claiming that, at least in the long run, the values of 'fiscal multipliers' are quite low, if not indeed zero or negative, due to the depressing effect of fiscal expansion on some components of private expenditure.[3] Examination of the arguments asserting this allegedly negative influence shows, however, that they are based either on relations peculiar to neoclassical theory and unwarranted in the analytical framework adopted here or on quite arbitrary premises or procedures.

The most traditional case of displacement of private expenditure is obviously the 'crowding out' of private investment due to the higher interest rate caused by an increased public deficit. To start with, the dependence of the level of investment on the interest rate is a typical neoclassical concept and one that finds no analytical foundation, at least in the guise of a relation of general character (as implied by the functional form it is usually given), in our theoretical context. Moreover, the debates on capital theory have shown that the idea of an inverse relation between the interest rate and investment is questionable even within the boundaries of neoclassical theory. In turn, the supposed rise in the interest rate is itself part of the same conceptual picture, since it would be generated by the 'competition' between public deficit and private investment as alternative uses for scarce private savings, the amount of the latter being established by the full-employment level of income. This scarcity has instead no place in our framework, where private savings adjust to the sum of public deficit and private investment through changes in the level of income, which implies, as pointed out above, analogous adjustment between the corresponding magnitudes of stock. A different and theoretically more 'neutral' reason is sometimes put forward to explain the rise in the interest rate following an increase in public deficit (or even in tax-financed public expenditure), namely that the expansionary effect of public spending feeds the demand for transaction money and thus pushes the interest rate up if the supply of money remains unchanged. While the analytical premises for the result are not strictly neoclassical in this case, and can indeed be found in a purely Keynesian setting, it is the very notion of the supply of money remaining unchanged in the presence of an increasing demand for money that is wholly arbitrary. The particular stock of money in existence is in no sense 'natural' and rather reflects the targets of the monetary authority, among which the level of the interest rate is certainly paramount. The assumption that the central bank would not adjust the money supply and thus let the interest rate rise therefore amounts to the assumption of an unjustified switch to a more restrictive monetary policy. It follows that a reasonable *ceteris paribus* condition would rather be to take the central bank's targets as given and hence assume an adjustment of the supply of money designed to keep the interest rate at the same level.

Another argument in support of the depressive influence of public deficit on the expenditure of private agents concerns the latter's consumption. The grounds for maintaining that private consumption would be reduced by an increase in public deficit lie in the so-called 'intertemporal budget constraint' imposed on the government, whereby the present value of future debt of the public sector (as well as the future debt of any agent) must be zero. If this constraint is to be met, each present flow of public deficit must have as a counterpart an equal present value of future budget surpluses and hence of tax increases, if we refer for simplicity to unchanged future levels of expenditure. On the assumption that households are forward-looking, the expected future change in the fiscal regime would induce them to increase savings to the same extent as the present value of the additional future taxes, coinciding with the present increase in public deficit. It is thus claimed that due to this behaviour on the part of households, the increase in public deficit is offset by an equal fall in private consumption with no net effect on aggregate demand and the level of income.

Closer examination reveals, however, that the irrelevance of aggregate demand with respect to the determination of income is actually a *presupposition* rather than a consequence of the behaviour of households outlined above. In order to have a rational foundation, this behaviour must be based on the 'Ricardian equivalence',[4] the substance of which can be seen as consisting of two elements: a) households perceive that wealth in the form of public bonds is bound to be cancelled out by the additional future taxes imposed by the government's intertemporal budget constraint; b) they wish to leave a given amount of wealth to their heirs. As a result of a), households are prompted to increase their savings (reducing consumption) to the same extent as they subscribe new public debt, so that the latter is bequeathed *in addition to*, and not in the place of, the (real) wealth they wish to hand down to the next generation. As should already be evident, the need for households to reduce consumption in order to increase savings reveals that the level of income is conceived as independent of aggregate demand from the very outset. The role played by the reduction of private consumption is in fact to prevent the public deficit from crowding out private investment and therefore the real wealth to be bequeathed. If the level of income is instead seen as depending on aggregate demand, an increase in deficit spending would generate additional income and private savings, and the invariance of private investment would require no reduction in private consumption, which would, on the contrary, be free to increase together with the increase in income. Therefore, even with regard to the argument just discussed, the ineffectiveness of fiscal policy proves to be embedded in its very premises.

In conclusion, there appears to be no general reason within the theoretical framework adopted here to presume that the influence of public deficit on aggregate demand is offset by corresponding contractions of private expenditure, either for investment or consumption. On the contrary, recognition that deficit spending helps to keep aggregate demand, and therefore the level of income, higher than it would otherwise be makes it possible to attribute it with a positive influence not only on private consumption but also on investment levels, especially in a long-run perspective. Once demand is seen as the factor governing levels of activity also in the long run, it is reasonable to suppose that at least the 'dimensional' share of investment, namely the part connected with the size of productive capacity, is affected by the prospective levels of demand and therefore by the trajectories of fiscal policy (and economic policy in general) involved in determining them.

1.3 Debt–output ratio and public expenditure reductions: the case of a single period

If, as is usual, the ratio of public debt to domestic product is taken as a 'measure' (albeit not necessarily a meaningful measure) of debt, the influence of deficit spending on both terms of the ratio renders its relation with that measure more complex than is commonly acknowledged.

Let us consider alternative fiscal policies, characterised by different levels of public expenditure, while assuming unchanged tax rates. Let the index *A* denote magnitudes that would obtain under the current policy and indexes *R* and *E* magnitudes that would obtain respectively under a restrictive policy (with lower flows of public expenditure) and an expansionary policy (with higher flows). We assume a closed economy in which public expenditure and private investment do not depend in any predetermined way on the level of output and private consumption is a constant proportion of the private sector's current disposable income. Moreover, the share of the latter consisting of interest on public bonds is entirely saved, a simplifying assumption that does not affect our results qualitatively because it in fact cuts off one of the channels through which public deficits can feed aggregate demand and therefore acts *against* our argument.

Apart from its theoretical premises, a further difference of our argument with respect to conventional treatment of the subject is that the effects of a change in fiscal policy will be studied by comparing the alternative values that the ratio of debt to domestic product would take under one policy or the other in the same period or series of periods. This way of reasoning is at variance with the more usual procedure whereby the debt ratio resulting, or expected to result, over a period of time from a policy change is compared with the value of the ratio calculated at some date prior to the policy change. As we shall see later (Section 1.7), the reason for adopting this approach is that policy changes apply to some previously planned policy and the counterparts of the results obtainable by the switch are therefore the effects that the older policy would continue to produce if still in force rather than the past values of relevant variables.

If we assume for the moment that private investment flows are not affected by fiscal policy and can be therefore taken as given in the face of changes in the latter, the level of income at time t under the currently planned fiscal policy would be

$$Y_t^A = m(I_t + G_t^A)$$

where I and G stand respectively for private investment and public expenditure, and m is the income multiplier. Given the assumptions stated above,

$$m = \frac{1}{1 - c(1 - z)}$$

where c is the (marginal and average) share of consumption out of private disposable income and z the (marginal and average) tax rate.[5] The multiplier is therefore assumed to produce all its effects on income within the same period in which any given level of 'autonomous' expenditure, either private or public, takes place. The admitted arbitrariness of this assumption loses much of its relevance when the analysis extends to several periods, as it does here, and its global results do not therefore depend on how the direct and indirect effects of individual flows of expenditure are actually spread out over time.

With a restrictive policy entailing lower public expenditure $G_t^R < G_t^A$, the level of income at time t would instead be

$$Y_t^R = m(I_t + G_t^R)$$

with a difference in the level of income equal to

$$\Delta Y_t^R = Y_t^A - Y_t^R = m(G_t^A - G_t^R) = m\Delta G_t^R$$

On the other hand, the lower public expenditure reduces the deficit at time t with respect to that which would result with policy A:

$$\Delta D_t^R = D_t^A - D_t^R = (1 - mz)\Delta G_t^R$$

As expressed by the term $(1 - mz)$, the cut in expenditure reduces the deficit less than its amount, as each unit of lower spending also reduces public revenues by its multiplied effect m on the level of income times the tax rate z.[6]

If we assume for simplicity that interest on public debt starts to be paid one period after issue, the stock of debt at t with policy R is lower than with policy A by the same amount as the reduction in the deficit:

$$\Delta B_t^R = B_t^A - B_t^R = (1 - mz)\Delta G_t^R$$

where B stands for the stock of debt.

Using the previous relations, we can now express the ratio of public debt to domestic product that would obtain at t with policy R:

$$\frac{B_t^R}{Y_t^R} = \frac{B_t^A - (1 - mz)\Delta G_t^R}{Y_t^A - m\Delta G_t^R}$$

The purpose of the adoption of the restrictive policy is to lower the ratio of debt to domestic product below the level that would be obtained with the actual policy, i.e., to obtain

$$\frac{B_t^R}{Y_t^R} < \frac{B_t^A}{Y_t^A}$$

The latter condition evidently requires the reduction in the domestic product to be proportionally smaller than the reduction in the debt, i.e.,

$$\frac{m\Delta G_t^R}{Y_t^A} < \frac{(1 - mz)\Delta G_t^R}{B_t^A}$$

which, after substituting for m and simple algebraic manipulation, reduces to

$$\frac{B_t^A}{Y_t^A} < (1 - c)(1 - z)$$

The quantity $(1 - c)(1 - z)$ thus sets the upper limit of the ratio of debt to domestic product within which a reduction in public expenditure in a given period would be effective in reducing the debt ratio below the level that would obtain in the same period under the planned spending policy. For values of the ratio above that limit, a lower level of public expenditure would result in a higher ratio, whereas a higher level would result in a *lower* ratio. The idea behind this result is quite simple: the larger the size of debt with respect to domestic product, the smaller the effect that a given reduction of public expenditure would produce on the former in relation to the latter. Beyond a certain threshold, the negative effect of less public spending on domestic product would be proportionally larger than on the stock of debt and the ratio would consequently prove higher than it would have been with an unchanged public deficit (and vice versa). In other words, the magnitude $(1-c)(1-z)$, which represents the marginal (and average) 'propensity to save' of the private sector in relation to domestic product in our analysis, also provides the ratio between the difference in the stock of private wealth and the difference in domestic product, both differences resulting from the implementation of policy *R* instead of policy *A* in period *t*:

$$(1-c)(1-z) = \frac{\Delta W_{pt}^{R}}{\Delta Y_{t}^{R}}$$

where ΔW_{pt}^{R} represents the amount by which private wealth at time *t* would be less under policy *R* than it would under the current policy *A*.[7] The lesser formation of private wealth would, however, be precisely the counterpart of the lower accumulation of public debt engendered by policy *R*, i.e. $\Delta W_{pt}^{R} = \Delta B_{t}^{R}$, and the last equality entails

$$(1-c)(1-z) = \frac{\Delta B_{t}^{R}}{\Delta Y_{t}^{R}}$$

The marginal rate of private savings is therefore equal to the marginal difference in the ratio of debt to domestic product generated by policy *R* as against policy *A*. It is obvious that the debt ratio obtainable with policy *R* will prove lower than that obtainable with policy *A* only if the latter is lower than the marginal difference generated by policy *R*, i.e., only if

$$\frac{B_{t}^{A}}{Y_{t}^{A}} < \frac{\Delta B_{t}^{R}}{\Delta Y_{t}^{R}}$$

which explains why the magnitude $(1-c)(1-z)$ functions as the upper limit of the *A* debt ratio for the restrictive policy to be effective in reducing the debt ratio.

The quantity $(1-c)(1-z)$ can thus be taken as the threshold for the current debt ratio, above which the effects of changes in the level of public expenditure on the debt ratio would be opposite to those commonly expected. It should be noted that, making allowances for our simplifying assumptions, the values that the threshold can plausibly take are quite small. For c=0.70 and z=0.30, for example, its magnitude would be as small as 0.21, which is much lower than the ratios of public debt to domestic product generally observable in real economies.

1.4 Expenditure reductions over a number of periods

So far we have considered the question of the effects of a change in the level of public spending that takes place in a single period. As relations in the longer run are obviously what matter most, we shall now go on to study the effects of variations in public expenditure extending over a series of periods.

As regards the absolute amount of debt, a series of lower flows of deficit spending obviously reduces the stock of debt below the level it would otherwise have found. This effect, which increases with the length of the series of reductions and the scale of their amounts, is enhanced by the increasing reduction in interest outlays caused by the cumulative effect on the stock of debt. We shall now consider a policy R that cuts the levels of public expenditure from time 1 to time t with respect to the previously planned policy A and express the stock of public debt at the final time t in relation to the debt which would have obtained at t with policy A. If we assume a zero interest rate for the moment, the relation between the two stocks of debt at time t would be:

$$B_t^R = B_t^A - (1-mz)\sum_{s=1}^{s=t} \Delta G_s^R$$

where $\Delta G_s^R = G_s^A - G_s^R$ is the reduction of public spending in period s which policy R entails in comparison with policy A.

If we now introduce a positive interest rate i on public debt, each reduction of debt in comparison with policy A generates further reductions as time goes on through the lower flows of outlay on compound interest entailed.[8] In greater detail, one unit of lower debt at time s generates a lower debt at time t by an amount of $(1+i)^{t-s}$, and the relation between the debt stocks therefore becomes

$$B_t^R = B_t^A - (1-mz)\sum_{s=1}^{s=t} \Delta G_s^R (1+i)^{t-s}$$

In order to identify the condition that must be fulfilled in order for the restrictive policy *R* to lower the ratio of public debt to domestic product at time *t*, let us now express the ratio as

$$\frac{B_t^R}{Y_t^R} = \frac{B_t^A - (1-mz)\sum_{s=1}^{s=t} \Delta G_s^R (1+i)^{t-s}}{Y_t^A - m\Delta G_t^R}$$

and impose the condition that it must prove smaller than the debt ratio with policy *A*:

$$\frac{B_t^A - (1-mz)\sum_{s=1}^{s=t} \Delta G_s^R (1+i)^{t-s}}{Y_t^A - m\Delta G_t^R} < \frac{B_t^A}{Y_t^A}$$

This condition requires the reduction in the domestic product to be proportionally smaller than the reduction in the stock of debt:

$$\frac{m\Delta G_t^R}{Y_t^A} < \frac{(1-mz)\sum_{s=1}^{s=t} \Delta G_s^R (1+i)^{t-s}}{B_t^A}$$

from which, by rearranging and expressing the value of *m*, we obtain:

$$\frac{B_t^A}{Y_t^A} \frac{\Delta G_t^R}{\sum_{s=1}^{s=t} \Delta G_s^R (1+i)^{t-s}} < (1-c)(1-z)$$

We thus find once again a condition containing the ‘inversion threshold’ of the debt ratio in relation to changes in the level of public expenditure. In comparison with the result obtained by considering a single period, the magnitude that must remain below the threshold if policy *R* is to bring about a lower debt ratio is still the value that the debt ratio would assume at *t* with policy *A*, this time multiplied by a quantity evidently lower than 1. (It should be remembered that the reductions ΔG_s^R are regarded as *positive* values.) As might be expected, the condition to be fulfilled is therefore less restrictive than in the case of a single period as a result of the cumulative effects of the series of reductions in public spending.

The relation we have just obtained gains in transparency if the magnitude that expresses the effects of the succession of reductions in expenditure is transferred to the right-hand side, thus reinstating this side in the position of an ‘inversion threshold’ for the current – policy debt ratio as in the single-period case:

$$\frac{B_t^A}{Y_t^A} < (1-c)(1-z)\frac{\sum_{s=1}^{s=t} \Delta G_s^R (1+i)^{t-s}}{\Delta G_t^R} \qquad [1.1]$$

The quantity on the right-hand side of **[1.1]** can be seen as the value of the inversion threshold of the debt ratio at time t 'corrected' by the cumulative effects of the lower flows of public spending. The 'correction' operates in such a way as to make the threshold variable in relation to the series of reductions in expenditure with respect to policy A, increasing its value, from the second period of the sequence, above the value $(1-c)(1-z)$ initially assumed by the threshold. In actual fact, the fraction on the right-hand side generally increases with the number of periods and therefore the series of expenditure reductions, even though it might show some decreasing stretches after a sufficient number of periods and depending on the actual values of ΔG_s^R, but in any case remaining above the initial value $(1-c)(1-z)$.[9] Figure 1.1 shows the behaviour of the threshold, which we shall indicate as $T(t)$, leaving aside any possibility of a drop and therefore assuming its monotonicity.

Let us then suppose that over time policy A would produce the debt ratio $\frac{B_t^A}{Y_t^A}$ shown in Figure 1.2, starting from a value higher than the initial level of $T(t)$ and then proceeding along a rising trend, which is assumed to be linear for simplicity. We shall also assume that the spending reductions entailed by policy R are such as to make the threshold $T(t)$ increase more rapidly than the debt ratio resulting from policy A, so that the threshold cuts the trend of the debt ratio after a certain number of periods and the latter finds its way below the threshold. The intersection point in τ therefore marks a divide between the two regions in which the values of the R debt ratio would stand: the northwest region H, where the R debt ratio would exceed the A ratio, and the southeast region L, where the R ratio would be lower than the A ratio.

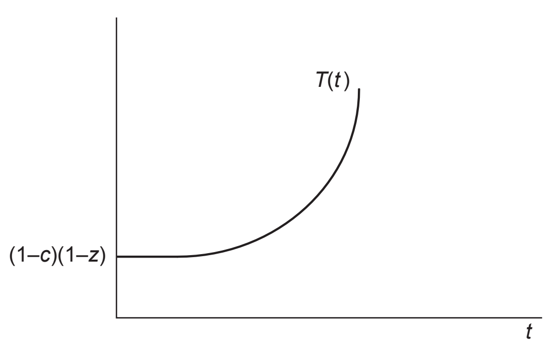


Figure 1.1 The threshold

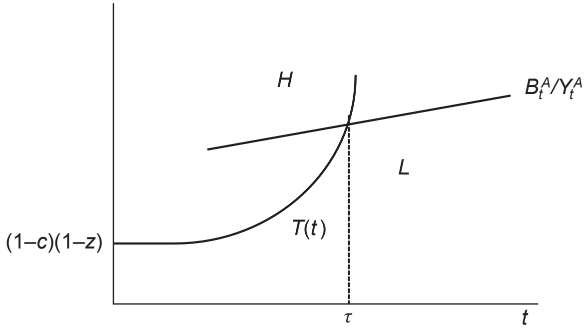


Figure 1.2 Threshold and present-policy debt ratio

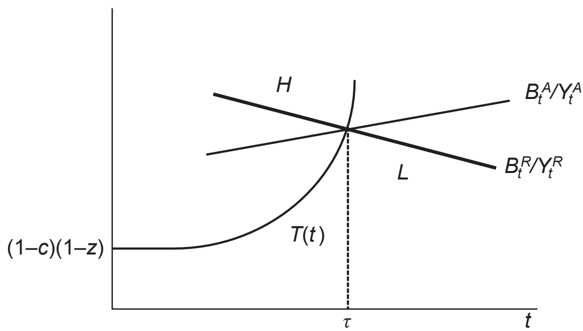


Figure 1.3 Effects of a restrictive policy (investment given)

The specific trend of the *R* debt ratio would of course depend both on the values of the spending flows entailed by policy *R*, which are those defined by the assumed reductions from the levels planned by policy *A*, and on the assumed trend of the *A* debt ratio. In order to simplify the graphic exposition, let us represent also the trend of the *R* debt ratio in linear form and assume that it is downward sloping, as shown by the thick line in Figure 1.3. Neither the decreasing course of the curve nor its monotonic behaviour is therefore a necessary characteristic of the trend of the *R* ratio, not least because of the influence of public spending on aggregate demand and output. The conclusions here are, however, not strictly dependent on the shape of the *R* ratio curve, and the assumption that the slope of the latter is negative is no more special than any other.[10]

We can thus see that as long as the debt ratio associated with policy *A* stands above the threshold, the debt ratio produced by policy *R* would in turn stay *above* the *A* ratio, and up to that point the effects of the restrictive policy could therefore be said to be 'perverse', as they are opposite to those expected.

On the other hand, the above exercise shows that after a number of periods the debt ratio produced by policy *R* would fall below what would be generated by policy *A*. This result might be taken to mean that the restrictive policy is ultimately effective in bringing about a debt ratio lower than otherwise, the delay – and indeed the opposite outcome – that the process goes through before reaching the target being conceivably not so relevant if the main concern is the long-run behaviour of magnitudes. Furthermore, the reductions in the levels of public spending seem to be capable, other things remaining equal, of making the $T(t)$ curve quite steep, and therefore rapidly increasing.[11]

Appealing to the temporary nature of the 'perverse' effects of a restrictive policy in order to diminish their importance, a reaction often observed when such consequences take place in actual economies, implicitly rests, however, on a belief in an inherent tendency of aggregate output to move towards its potential level and hence denies the autonomous role of demand in determining output. In the above analysis, the obstacle that prevents the lower levels of public spending from reducing the debt ratio, even to the point of raising it, is in fact the negative effect they have on levels of output. This depressing influence is permanent and the comparative diminution that the ratio eventually undergoes is obtained *despite* such influence, rather than because of its disappearance. As a matter of fact, the reduction of the debt ratio with respect to what it would be under policy *A* is obtained through a comparative reduction in *both* the stock of debt and domestic product, with the former decreasing proportionally more than the latter after a number of periods due to the cumulative action of spending cuts and lower interest outlays. The 'perverse' effects on the debt ratio are therefore the manifestation of a change in the trend levels of output caused by the restrictive change in fiscal policy and should be kept distinct from any sort of short-run consequences due to temporary circumstances or agents' reactions to that policy. This is reinforced by the consideration that, as will be argued shortly, it is precisely the persistence of the depressive effects of the restrictive policy that is liable in the long run to cause further changes enhancing those effects and consequently extending the number of periods for which the policy could generate a comparatively *higher* debt ratio.

1.5 Further on public expenditure reductions: negative feedbacks on private investment

The effects of given reductions in public spending have been examined so far on the assumption that the tax rate, the propensity to consume of the private sector and the interest rate on public bonds all remain constant. It has also been assumed that the other component of 'autonomous expenditure', namely private investment, does not alter the series of its magnitudes in relation to the adoption of a different fiscal policy and that therefore the effects on output levels of sequences of different levels of public spending are limited to those produced by the latter differences alone (the action of the income multiplier being of course taken into account).

The last of these assumptions appears to be the most restrictive. As already pointed out, there is no need in the theoretical framework adopted here for the interest rate to depend either on the size of the public debt, however it may be measured, or on measures of fiscal policy or more generally on changes in the levels of activity, and it is therefore legitimate to take it as given in the event of changes in those circumstances. The same holds all the more for the tax rate, which is completely under government control. As regards the private propensity to consume, while there could be reasons in support of its dependence on long-run levels of income, taking this dependence explicitly into account would only make the value of the multiplier variable without altering the principle of which the multiplier is the quantitative expression, namely that the level of aggregate output is governed by demand. The general character of the analysis undertaken here does not require us to go beyond the principle as such, and treating the propensity to consume as given constitutes a simplifying device that imposes no substantial restrictions on the results obtained.

Different considerations instead apply to the assumption that the levels of output generated by a given fiscal policy do not affect the investment plans of the private sector. Given our key principle that output is limited, in the long run no less than in cycle fluctuations, by the level of demand, it follows as a natural corollary that demand also influences the level of private investment. At least as regards the share of gross investment that could be described as 'dimensional', i.e., related to the desired scale of productive capacity, strict dependence is quite obvious on the levels of output that firms expect to be profitably attained or maintained in the periods ahead and hence on future levels of demand. It is quite reasonable, on the other hand, to presume that the trends of expected demand are influenced by the levels of demand that have prevailed in the past periods, so that lower levels of public spending will negatively affect private investment through their depressing impact on aggregate demand. Moreover, private investment may be influenced even by the *planned* levels of public spending if, as is conceivable, firms take government planning into account in forming their views about future market conditions and therefore output trends.

The conclusion that private investment is negatively affected by lower levels of public expenditure has important consequences for our analysis. Let ΔI_s^R indicate the reduction (taken as algebraically positive) in gross private investment that would be obtained in period s under the restrictive fiscal policy R compared with what it would be in the same period under policy A. Due to the influence of the lower levels of investment on levels of income and therefore on public revenues, the absolute amount of public debt that would result at time t from the application of policy R since time 1 is now larger. As above, this can be expressed in relation to what it would have been under policy A:

$$B_t^R = B_t^A - (1-mz)\sum_{s=1}^{s=t} \Delta G_s^R (1+i)^{t-s} + mz\sum_{s=1}^{s=t} \Delta I_s^R (1+i)^{t-s}$$

where the last term on the right-hand side represents the effect on the final stock of debt of the lower revenues (with consequent higher outlay on interest) caused by the lower flows of private investment expenditure. Correspondingly, the level of income at time t falls by comparison with what it would be under policy A because of the lower levels of public spending and private investment in the period:

$$Y_t^R = Y_t^A - m(\Delta G_t^R + \Delta I_t^R)$$

For the debt ratio at time t under policy R to be lower than it would be under policy A, the decrease in the level of income must be less than proportional to the decrease in the stock of debt, that is

$$\frac{Y_t^A - Y_t^R}{Y_t^A} < \frac{B_t^A - B_t^R}{B_t^A}$$

and, therefore, using the above relations:

$$\frac{m(\Delta G_t^R + \Delta I_t^R)}{Y_t^A} < \frac{(1-mz)\sum_{s=1}^{s=t} \Delta G_s^R (1+i)^{t-s} + mz\sum_{s=1}^{s=t} \Delta I_s^R (1+i)^{t-s}}{B_t^A}$$

Expressing the value of the multiplier m, the latter inequality may be written in the form:

$$\frac{B_t^A}{Y_t^A} < (1-c)(1-z)\frac{\sum_{s=1}^{s=t} \Delta G_s^R (1+i)^{t-s}}{\Delta G_t^R + \Delta I_t^R} - z\frac{\sum_{s=1}^{s=t} \Delta I_s^R (1+i)^{t-s}}{\Delta G_t^R + \Delta I_t^R} \qquad [1.2]$$

The expression on the right-hand side of relation **[1.2]** is the value now assumed by the 'inversion threshold' $T(t)$ of the debt ratio. Comparison with inequality **[1.1]**, where no changes in the levels of private investment are considered, reveals two differences. In addition to the reduction in public spending in period t, the denominator of the first fraction contains the reduction in private investment in the same period, which obviously reduces the value of the fraction (besides reinforcing the possibility of some decreasing stretches in its behaviour). The magnitude of the threshold is further reduced by the negative term in z, which could, if its values were sufficiently large, even give the threshold a decreasing course, in this case not necessarily restricted to limited stretches alone.[12] The size of this negative effect on the value of the threshold evidently depends on the size of the decreases in the flows of private investment that policy R generates by comparison with policy A. Even though this kind of influence does not lend itself to shaping in a general form, it is reasonable to expect that both the amounts and the duration of reductions in public spending would, through their impact on aggregate demand, directly affect the amounts by which private investment flows would be lower under policy R than policy A.

In conclusion, by augmenting the negative influence of the restrictive policy on output levels, reductions in private investment engender two effects. The first is a drop in the values of the threshold curve, which switches to the right and may become *decreasing* from a certain point on.[13] The second is an increase in the values of the debt ratio ensuing from the restrictive policy in any given period due to the simple fact that lower income levels generate lower public revenues. This effect would be seen graphically in an upward movement of the line representing the trend of the R debt ratio.

In Figure 1.4, $T_0(t)$ indicates the threshold curve ensuing with unchanged investment levels and T_1 the threshold associated with a series of investment reductions (by comparison with what would happen under policy A) starting from period t_1, while the two downward-sloping lines intersecting them represent the trends of the R debt ratio that would obtain respectively under the first and the second hypothesis about investment behaviour. As we can see, with no change in

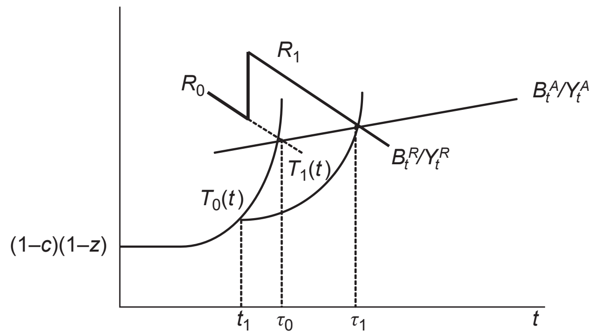


Figure 1.4 Effects of a restrictive policy (investment reacting once)

investment flows, the *R* policy would potentially bring the debt ratio below the level entailed by policy *A* in period τ_0 (line R_0). As a result of the curtailment of private investment flows at t_1, however, the threshold curve becomes $T_1(t)$ from then on with the corresponding upward movement in the trend of the *R* debt ratio (line R_1). The moment at which the latter would fall below the policy *A* ratio thus changes from period τ_0 to period τ_1 and the course of the *R* debt ratio would be represented by the bold broken line.

There is, however, nothing to rule out the possibility of the continuation of lower levels of demand and output, and their additional reduction entailed by the lower investment levels, subsequently generating even larger reductions in private investment flows. In Figure 1.5, the decreases in investment, again by comparison with the investment flows under policy *A*, are supposed to have increased still further since not only period t_2 but also period t_3 with consequent successive switches of both the threshold curve and the trend of the *R* debt ratio. The latter would thus remain above the trend of the *A* debt ratio for a number of periods, falling below it only if and when the reduction of private investment flows (by comparison with what would occur under policy *A*) stops for a sufficient length of time. Moreover, this result would eventually be obtained with levels of domestic output and private investment that could be lower by any degree than those resulting from policy *A*.

Even more drastically, policy *R* could prove totally incapable of bringing the trend of the debt ratio below that of policy *A* if the reduction in private investment flows were to become sufficiently large eventually to cause a decreasing course of the *T*(*t*) curve. This case is represented in Figure 1.6, where the decrease in private investment induced by policy

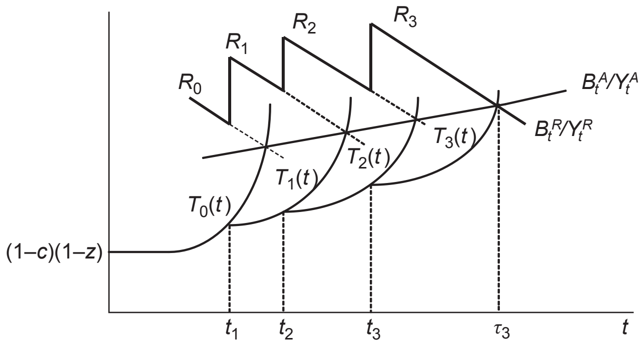


Figure 1.5 Effects of a restrictive policy (investment reacting over time)

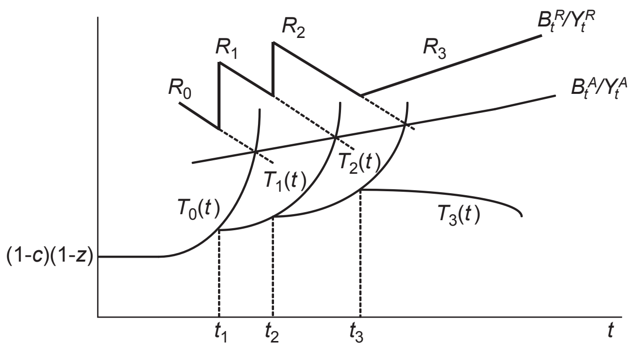


Figure 1.6 Effects of a restrictive policy (threshold eventually downward sloping)

R in period t_3, compared with the investment levels that would obtain under policy A, is supposed to be such as to generate a downward-sloping threshold curve. In this case, the latter would not intersect the trend of the A debt ratio and consequently, despite the continuation of a restrictive policy, the trend of the R debt ratio would continue to stand above that of the A ratio.[14]

1.6 The case of increases in public expenditure

We have thus seen that the attainment of a comparatively lower debt ratio is not guaranteed by the extension of the number of periods for which a policy of reduced public spending is implemented. In the absence of compensatory changes in other elements of autonomous demand, which there would be no reason to expect, precisely the prolongation of such a policy, with its negative influence on long-run demand trends, would possibly if not indeed inevitably lead to a drop in flows of private investment. In the case represented in Figure 1.5, this shifts the goal of reducing the debt ratio further into the future with no guarantee that the process will not fall into a tailspin, thus moving the position of the threshold curve $T(t)$ further and further to the right. The already uncertain reduction of the debt ratio becomes impossible in the case represented in Figure 1.6, where sufficiently strong reductions in flows of private investment cause the slope of the $T(t)$ curve to become negative and the restrictive policy gives rise to a permanent increase in the debt ratio. Furthermore, this uncertainty of results, even as regards their sign, is accompanied by a certain and permanent relinquishment of higher levels of income and employment by the community, something that could be described in the terms of the dominant theoretical framework as an instance of 'Pareto inferiority'.

These conclusions prompt examination of the situation that would instead arise with higher levels of public spending. Let E be an expansionary policy entailing increases in expenditure with respect to the current policy A. If we take into account the increases in private investment that would be induced by the higher levels of demand, the stock of public debt at time t under policy E, as against the debt that would result from policy A, is:

$$B_t^E = B_t^A + (1-mz)\sum_{s=1}^{s=t} \Delta G_s^E (1+i)^{t-s} - mz\sum_{s=1}^{s=t} \Delta I_s^E (1+i)^{t-s}$$

where ΔG_s^E and ΔI_s^E respectively represent the increases in public expenditure and private investment in the generic period s by comparison with what would obtain under policy A. The level of income in period t would correspondingly be higher by the amount

$$\Delta Y_t^E = m(\Delta G_t^E + \Delta I_t^E)$$

The condition required for policy E to produce a lower debt ratio at t than policy A is once again that the proportional increase in the level of income exceeds the proportional increase in the absolute stock of debt:

$$\frac{m(\Delta G_t^E + \Delta I_t^E)}{Y_t^A} > \frac{(1-mz)\sum_{s=1}^{s=t} \Delta G_s^E (1+i)^{t-s} - mz\sum_{s=1}^{s=t} \Delta I_s^E (1+i)^{t-s}}{B_t^A}$$

which, rearranging the terms and expressing the value of the multiplier m, provides the value that the debt ratio with policy A must be *above* in order for the said result to be obtained:

$$\frac{B_t^A}{Y_t^A} > (1-c)(1-z)\frac{\Delta G_s^E(1+i)^{t-s}}{\Delta G_t^E + \Delta I_t^E} - z\frac{\sum_{s=1}^{s=t} \Delta I_s^E(1+i)^{t-s}}{\Delta G_t^E + \Delta I_t^E}$$

The right-hand side of the above inequality represents the 'inversion threshold' of the $\frac{B_t^A}{Y_t^A}$ ratio, in this case a *lower* boundary, in order for the expansionary policy to generate a lower debt ratio at t. As can be seen, the expression of the threshold is strictly analogous to what was found in considering a restrictive policy, and it therefore behaves in the same way. On the assumption that private investment would respond positively to the higher aggregate demand triggered by the expansionary policy, in Figure 1.7 $T_0(t)$ represents the threshold curve assuming no change in investment flows as against what would obtain with policy A, and the curves $T_1(t)$, $T_2(t)$, and $T_3(t)$ instead exemplify the successive forms of the threshold curve resulting from increasingly large comparative increases in the flows of private investment. In the same figure, the line E_0 represents the course that the E debt ratio would take if investment flows were the same as under policy A, and the lines E_1, E_2, E_3 stand for the paths of the ratio generated by the increasing levels of investment entailed by the corresponding $T.(t)$ curve. As the figure shows, each further increase in investment flows moves both the threshold curve and the E debt-ratio line to the right, thus shifting the time

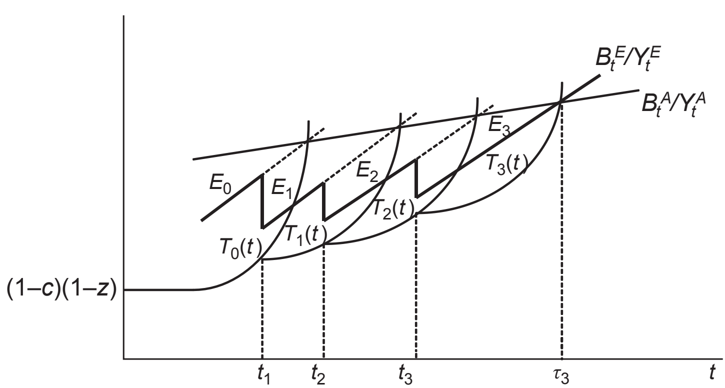


Figure 1.7 Effects of an expansionary policy (investment reacting over time)

at which the two curves jointly intersect the trend line of the A debt ratio and the E debt ratio overtakes the latter progressively from τ_0 to τ_1, then τ_2 and finally τ_3. As a result of the expansionary policy, the trend followed by the E debt ratio would therefore lie *below* the one generated by the current policy for a number of periods which would be larger, the larger and more persistent the effects produced directly and indirectly on aggregate demand.

In symmetry with the findings of our examination of a restrictive policy, one result of our analysis is that the trend of the expansionary policy debt ratio would lie permanently below that of the current policy if the resulting increases in private investment flows reached levels high enough to generate a downward-sloping $T(t)$ curve. This is the case represented in Figure 1.8, where the rise in private investment from period t_3 is supposed to be such as to make the $T_3(t)$ curve negatively sloped and prevent it from intersecting with the $\frac{B_t^A}{Y_t^A}$ line. In turn, the $\frac{B_t^E}{Y_t^E}$ line would never overtake the $\frac{B_t^A}{Y_t^A}$ line, and the expansionary fiscal policy would succeed in keeping the trend of the debt ratio below that of the current policy indefinitely.

1.7 Final remarks and conclusions

As stated at the outset, unlike the more usual procedure of comparing the results of a policy change with magnitudes calculated at some earlier

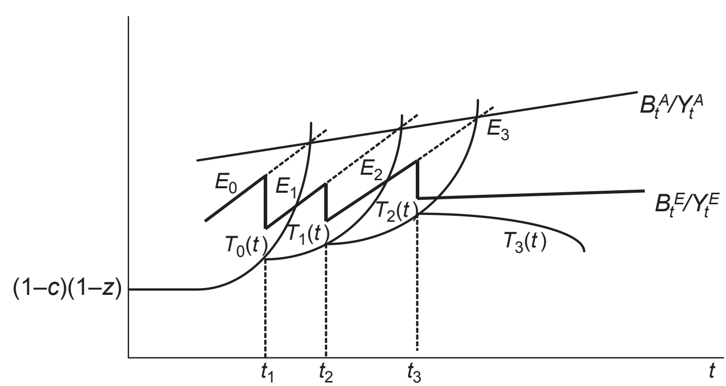


Figure 1.8 Effects of an expansionary policy (threshold eventually downward sloping)

time, our investigation compares a given trend of the debt ratio generated by the currently planned fiscal policy with the trend that would be generated by either a comparatively restrictive or a comparatively expansionary policy. The justification put forward for this procedure is the fact that a change in fiscal policy, and particularly in public spending, is generally formulated in terms of variations to be made on current policy. It is therefore more appropriate to evaluate the effects produced by the policy change over a given period through comparison with what would have been the result of current policy during the same period rather than with values prior to the change. This holds all the more if, as is evidently the case for the debt ratio and for objectives of fiscal policy in general, the level of overall output constitutes a relevant variable and is seen as demand determined by and hence dependent on the economic policies in force.

Nevertheless, the analysis can be easily supplemented if desired through the introduction of a specific value of the debt ratio that fiscal policy can be expected – according to some views of greater or lesser conceptual validity – to take as a benchmark for the actual debt ratio. On the obvious assumption that this target ratio is lower than the one obtainable with the current policy, it could be represented by a horizontal line drawn below the $\frac{B_t^A}{Y_t^A}$ line used in our graphs. The addition of such a line in Figures 1.5 and 1.6, which show how difficult if not indeed impossible it is for a restrictive fiscal policy to bring the trend of the debt ratio below the level resulting from the current policy, would then enable us to deduce analogous and actually stronger conclusions about the ability of this kind of policy to bring the debt ratio into line with an even lower target value. Adoption of the same procedure in Figures 1.7 and 1.8 would suggest that this adjustment would be more efficiently pursued by means of an expansionary policy.

Summing up, the chapter has endeavoured to show analytically that once a positive influence of public expenditure on total output is acknowledged, an *inverse* relation may hold, also in the long run, between levels of deficit spending and values of the ratio of public debt to domestic product. The results suggest that if a lower debt ratio constitutes a target of economic policy, as is currently the case in several countries, then regardless of the validity or otherwise of that objective, restrictive fiscal policies may prove ineffective and even self-defeating by generating higher rather than lower values of the debt ratio. On the contrary, expansionary policies may produce comparatively lower values of the ratio, and deficit spending would therefore be superior to fiscal retrenchment with regard not only to levels of activity and employment but also to what is often referred to as an index of solidity of the financial position of governments.[15]

The conclusions drawn here apply to policies implemented over series of periods and hold in this sense for the long run, not just the short. This is due to the fact that, contrary to what we find in much of the mainstream literature, the theoretical framework adopted here entails no need to counterbalance changes in fiscal policy with opposite changes in private expenditure, no matter how 'lagged'. The key idea of this chapter, namely that output is determined by demand, naturally implies instead that gross investment, the primary 'autonomous' component of private spending, directly depends in turn on the trend in levels of demand. Flows of private investment would thus respond to and hence assist changes in aggregate demand rather than counterbalancing them. By further enhancing the influence that fiscal policy can have in the long run on output levels, this behaviour of private investment reinforces the result that contrary to the 'orthodox' view, lower values of the debt ratio can be obtained more easily, and less painfully, through expansionary fiscal policies.

Notes

1. An application of the Keynesian principle of effective demand to the role of public spending and, more generally, to fiscal policy was put forward in a fairly large body of literature mainly dating from the 1940s and 1950s, which found perhaps its most 'radical' expression in Lerner's *functional finance* (Lerner, 1943). A major element distinguishing those contributions from this chapter is the fact that while the former often fail to specify the nature of the limit that aggregate demand imposes on total output, i.e., whether it is purely cyclical in character or extends to the trend of activity levels (cf. Ciccone, 2002, pp. 12–13, fn. 6), aggregate demand is explicitly regarded here as determining overall output in both the short and the long run.
2. For the premises implicit in the assumption that agents behave in this way, see Section 1.2 below.
3. A vast literature has developed focusing specifically on quantitative estimation of the change in the level of total output attributable to a change either in public expenditure or in tax rates or levels, a ratio generally labelled the 'multiplier of fiscal policy' or simply 'fiscal multiplier' (for a recent survey of this literature, see Ramey, 2011; see also Barba, 2001 for a critical discussion of the theoretical foundations of the claim that fiscal contractions can also be expansionary, which is still relevant to much of the work at issue). Many of these studies appear to be inspired by the assumption, implicitly or

explicitly derived from orthodox theory, that a change in fiscal policy can have no significant or in any case lasting influence on the level of demand, as it is bound to cause opposite changes in some components of the latter (private investment, private consumption or exports), so that the size of its effect on output is expected to be relatively small or even opposite in sign (but see Fazzari *et al.*, 2012 for an argument developed in a different direction). This is so by construction in the 'simulations' carried out through the use of models, the results of which obviously depend on the characteristics of the models adopted, which often reflect the set of relations peculiar to neoclassical theory. The size and sign of fiscal multipliers are alternatively investigated by means of pure empirical work relying on no theoretical structure and based exclusively on historical data. This is not without cost. Due to the dependence of fiscal magnitudes on several variables, including domestic output, a major problem encountered by this kind of research is the identification of changes in fiscal quantities (e.g. the primary balance) that can be attributed to deliberate policy actions and therefore treated as exogenous rather than resulting from the prevailing economic conditions (see the objections raised in this regard in IMF, 2010). The mass of work on 'fiscal multipliers' has in fact produced estimates that differ widely according to the assumptions and procedures adopted, e.g. ranging from −3.8 to +3.8 for public expenditure (cf. Van Brusselen, 2009, p. 18; estimates for tax-cut multipliers are reported to range between −4.8 and +3.0). The issue therefore still remains a matter for theory rather than something that can be decisively settled on empirical grounds. In the approach adopted here, which regards demand as determining the level of output, the influence of public spending on output can accordingly be expressed by the full action of the income multiplier with no need for counterbalancing variations in other constituents of demand.

4. As is well known, it was Barro (1974) who originally put forward the idea that households would reduce consumption by the same amount as the issuing of public debt, so that debt- and tax-financing of public expenditure would supposedly be equivalent in their effects on agents' behaviour. This article made no reference to Ricardo and the connection with an analogous point made by Ricardo was later noted by Buchanan (1976), since when the concept has often been referred to as 'Ricardian'. While claiming that the way in which public spending is financed should be a matter of indifference to the individual on logical grounds, Ricardo maintained, however, that this equivalence is not effectively perceived, as future taxes are in fact assigned less importance than present taxes. Unlike Barro, Ricardo thus maintains that debt financing would reduce savings by comparison with the situation in the case of tax financing (cf. Ricardo, 1951, pp. 186–7).
5. The rate z is here taken as representative of any kind of deduction from produced income accruing to the public sector as a whole as well as taxes and fees paid on purchases of goods and services. The constancy of z also rules out therefore any changes in public charges (e.g. for health services), which could furthermore come into conflict with the assumed invariance of income distribution.

6. Taking into account the value that the multiplier m assumes under our simple assumptions (see above in the text), the value of mz is lower than 1 for $z<1$ and $c<1$.
7. The difference in private wealth at t amounts to the difference in the flows of savings accruing to the private sector as a result of the difference in the level of domestic product under one policy and the other:

$$\Delta W_{pt}^{R} = (1-c)(1-z)\Delta Y_{t}^{R}$$

from which the equality stated in the text follows.
8. Interest outlays on public debt are considered net of taxes, any tax on them being no more than a clearing entry for the public sector.
9. If $T(h)$ and $T(h+1)$ are the values of the threshold at two subsequent times, their difference is

$$\begin{aligned} T(h+1) - T(h) &= (1-c)(1-z)\left[\frac{\Delta G_1^R(1+i)^h + \Delta G_2^R(1+i)^{h-1} + \cdots + \Delta G_{h+1}^R}{\Delta G_{h+1}^R}\right. \\ &\qquad \left. - \frac{\Delta G_1^R(1+i)^{h-1} + \Delta G_2^R(1+i)^{h-2} + \cdots + \Delta G_h^R}{\Delta G_h^R}\right] \\ &= (1-c)(1-z)\frac{1}{\Delta G_{h+1}^R}\left\{\left[(1+i) - \frac{\Delta G_{h+1}^R}{\Delta G_h^R}\right]\right. \\ &\qquad \left. \times \sum_{s=1}^{h} \Delta G_s^R(1+i)^{h-s} + \Delta G_{h+1}^R\right\} \end{aligned}$$

The algebraic sign of this quantity is determined by the quantity in braces, which can be negative if ΔG_{h+1}^R exceeds ΔG_h^R by a sufficient amount and the sum of $\Delta G_s^R(1+i)^{h-s}$ has attained a large enough value.
10. In point of fact, even with a linear form, the R debt ratio might be increasing with a slope lower than that of the A ratio (which is needed by the required intersection with the latter) or be flat. Being less favourable to our argument, the assumption of a decreasing trend is the one adopted here.
11. This can be easily seen under the simplifying assumption that the spending reductions are constant over time, i.e., $\Delta G_s^R = G^R$, in which case $T(t)$ reduces to

$$\begin{aligned} T(t) &= (1-c)(1-z)\frac{\sum_{s=1}^{s=t} \Delta G^R(1+i)^{t-s}}{\Delta G^R} = (1-c)(1-z)\sum_{s=1}^{s=t}(1+i)^{t-s} \\ &= (1-c)(1-z)\frac{(1+i)^t - 1}{i} \end{aligned}$$

the values of which, interestingly enough, are independent of the magnitude of the reduction in spending. This function can increase quite rapidly. For example, with $c=0.70$, $z=0.30$ and $i=0.03$, it takes the value 0.65 for $t=3$ and is already as high as 1.12 for $t=5$.

12. The difference between two subsequent values of the $T(t)$ is now

$$T(h+1)-T(h)=(1-c)(1-\mathrm{z})\left[\frac{\Delta G_1^R(1+i)^h+\Delta G_2^R(1+i)^{h-1}+\cdots+\Delta G_{h+1}^R}{\Delta G_{h+1}^R+\Delta I_{h+1}^R}\right.$$
$$\left.-\frac{\Delta G_1^R(1+i)^{h-1}+\Delta G_2^R(1+i)^{h-2}+\cdots+\Delta G_h^R}{\Delta G_h^R+\Delta I_h^R}\right]$$
$$-\mathrm{z}\left[\frac{\Delta I_1^R(1+i)^h+\Delta I_2^R(1+i)^{h-1}+\cdots+\Delta I_{h+1}^R}{\Delta G_{h+1}^R+\Delta I_{h+1}^R}\right.$$
$$\left.-\frac{\Delta G_1^R(1+i)^{h-1}+\Delta G_2^R(1+i)^{h-2}+\cdots+\Delta G_h^R}{\Delta G_h^R+\Delta I_h^R}\right]$$

By positing that this magnitude is negative and that $\dfrac{\Delta G_{h+1}^R+\Delta I_{h+1}^R}{\Delta G_h^R+\Delta I_h^R}=1+\sigma_{h+1}$,

we obtain,

$$(1-c)(1-\mathrm{z})\left\{\left[1-\frac{1+\sigma_{h+1}}{1+i}\right][\Delta G_1^R(1+i)^h+\Delta G_2^R(1+i)^{h-1}+\cdots+\Delta G_h^R(1+i)]+\Delta G_{h+1}^R\right\}$$
$$<\mathrm{z}\left\{\left[1-\frac{1+\sigma_{h+1}}{1+i}\right][\Delta I_1^R(1+i)^h+\Delta I_2^R(1+i)^{h-1}+\cdots+\Delta I_h^R(1+i)]+\Delta I_{h+1}^R\right\}$$

The above inequality, which implies the decreasing course of $T(t)$, can be satisfied by requiring no more than sufficient reductions in private investment (compared with the investment flows that would obtain under the currently planned fiscal policy). See also notes 13 and 14 below.

13. If the assumption that the reduction in public spending is constant over time (see n. 4 above) is combined for algebraic convenience with the assumption that the reduction in private investment is also constant, i.e. $\Delta I_s^R=\Delta I^R$, the expression of the threshold on the right-hand side of inequality [1.2] becomes

$$T(t)=(1-c)(1-z)\frac{\sum_1^t \Delta G^R(1+i)^{t-s}}{\Delta G^R+\Delta I^R}-z\frac{\sum_1^t \Delta I^R(1+i)^{t-s}}{\Delta G^R+\Delta I^R}$$
$$=\frac{(1+i)^t-1}{i}\left[\frac{(1-c)(1-z)}{1+\Delta I^R/\Delta G^R}-\frac{z}{1+\Delta G^R/\Delta I^R}\right]$$

where it is evident that the size of ΔI^R relative to ΔG^R affects the values of $T(t)$ negatively.

14. Under the specific assumptions of a constant decrease in public expenditure and a constant decrease in private investment, which allow us to use the formula of the threshold already indicated in n. 8, the table below exemplifies the progressive movement of the $T(t)$ curve and the eventual change of its slope into the negative as represented in Figure 1.6. The calculations are made assuming $c=0.70$, $z=0.30$ and $i=0.03$, and that the reduction in investment (expressed as a proportion of the public spending reduction) starts in period 3 and increases in steps consisting of intervals of three periods while remaining constant within each interval. The table shows that when the decrease in investment reaches

t \ $\Delta I^R/\Delta G^R$	$T_0(t)$	$T_1(t)$	$T_2(t)$	$T_3(t)$	$T_4(t)$	$T_5(t)$
	0.00	0.20	0.40	0.70	0.90	1.00
1	0.21					
2	0.43					
3	0.65	0.65				
4		0.78				
5		0.90				
6		1.04	1.04			
7			1.10			
8			1.17			
9			1.24	1.24		
10				1.24		
11				1.24		
12				1.24	1.24	
13					1.21	
14					1.18	
15					1.14	1.14
16						1.10
17						1.05
18						1.00
19						0.95

0.70 of the decrease in public spending, the threshold curve becomes flat and decreasing for further increments of the decrease in investment.

15. The analysis developed here is not immediately susceptible of transposition in terms of growth rates of variables, as is instead often the case in the literature concerning public debt and its ratio to domestic product. It is, however, easy to see that in such a reformulation of our basic relations (see Ciccone, 2002, ch. II) the rate of growth of public spending would be conceived as a policy variable positively affecting the rate of growth of overall output (see also Aspromourgos, 2007, section 2 for an analogous standpoint) and an *inverse* relation could therefore hold between the former rate and the ratio of public debt to domestic product.

References

Aspromourgos, T. (2007) 'Interest as an artefact of self-validating central bank beliefs', *Metroeconomica*, 58(4): 514–35.

Barba, A. (2001) 'Rethinking expansionary fiscal retrenchments', *Contributions to Political Economy*, 20: 57–86.

Barro, R. J. (1974) 'Are government bonds net wealth?', *Journal of Political Economy*, 82: 1095–117.

Buchanan, J. M. (1976) 'Barro on the Ricardian Equivalence Theorem', *Journal of Political Economy*, 84: 337–42.

Ciccone, R. (2002) *Debito pubblico, domanda aggregata e accumulazione* (Roma: Aracne).

Fazzari, S.M., Morley J. and Panovska I. (2012) 'State-Dependent Effects of Fiscal Policy', *Australian School of Business*, Research Paper 2012 ECON 27.

International Monetary Fund (2010) 'Will It Hurt? Macroeconomic Effects of Fiscal Consolidation', *World Economic Outlook*, ch. 3: 93–124.

Lerner, A. P. (1943) 'Functional finance and the federal debt', *Social Research*, 10: 38–51.

Ramey, V. (2011) 'Can government purchases stimulate the economy?', *Journal of Economic Literature*, 49(3): 673–8.

Ricardo, D. (1951) 'Funding System', in Piero Sraffa (ed.) with the collaboration of M. H. Dobb, *The Works and Correspondence of David Ricardo*, Vol. IV (Cambridge: Cambridge University Press): pp. 149–200.

Van Brusselen, P. (2009) 'Fiscal Stabilisation Plans and the Outlook for the World Economy', *NIME Policy Brief* 01-2009.

2

The Inevitable Dependence of Investment on Expected Demand: Implications for Neoclassical Macroeconomics

Fabio Petri

2.1 An internal critique of the neoclassical justification of Say's Law

The purpose of this chapter is to draw attention to a weakness, so far unnoticed, of the neoclassical argument in support of Say's Law – that is, of the thesis that investment *is determined by* savings,[1] and that therefore aggregate demand poses no obstacle to selling at cost-covering prices the aggregate supply of goods whatever the forces determining the latter. The neoclassical argument, relying upon an assumed negative interest elasticity of investment derived from the demand-for-capital function, neglects the problems with the marginalist or neoclassical conception of capital: as pointed out by the late Pierangelo Garegnani (1983, 1990), the discovery of reverse capital deepening undermines the foundations of Say's Law, because it undermines the belief in a negative interest elasticity of the demand for (value) capital, but then also the belief in a negative interest elasticity of aggregate investment; Garegnani concluded that the 'neoclassical synthesis' criticism of Keynes could not be accepted, and that aggregate demand had to be considered the determinant of employment and growth not only in the short period but also in the long run. In Petri (2004, ch. 7) I reinforced Garegnani's contention by showing that the attempts, after Keynes, to derive a negative interest elasticity of investment without relying on the traditional neoclassical conception of capital are all indefensible.[2] Here I intend to point out that the neoclassical argument meets grave difficulties even conceding the traditional neoclassical conception of capital–labour substitution[3] – and for a reason different from the ones adduced so far (such as 'malfunctioning' of financial intermediaries or irreducibly subjective expectations and animal spirits).

The reason is a fact hidden from sight in most current presentations of investment theory, the *inevitable* dependence of investment on desired capacity (and hence on expected demand and its variations) even in a neoclassical framework. I will point out that the marginalist, or neoclassical, approach to investment needs the continuous full employment of labour in order to arrive at a determinate influence of the interest rate upon investment; since the full employment of labour cannot be *assumed* as a starting point of the analysis (it can only be, if at all, a *result* of the analysis), labour employment must be considered variable, but then a given rate of interest leaves investment indeterminate even accepting the marginalist conception of capital–labour substitution, because the rate of interest can only determine the desired *ratio* of capital to labour, which leaves the capital to be invested in new plants indeterminate and to be determined by desired productive capacity. As I will show, some neoclassical economists (Jorgenson, and Dornbusch and Fischer) implicitly admitted it by having investment determined by expected demand, but did not see that then extremely serious problems arise concerning the capacity of wage flexibility to bring about the full employment of labour. What follows explains and expands on these points.

2.2 The neoclassical connection between demand for capital and investment

After Keynes it has become common to consider investment a function of the rate of interest *only*. Even when the influence of other variables is admitted, e.g. of income, it is generally seen as *additional* to the influence of the rate of interest, in the sense that the latter would suffice to determine investment if the other influences were very weak or absent. On the contrary a correct grasp of the neoclassical conception of capital–labour substitution implies that the rate of interest alone is unable univocally to determine investment, unless the full employment of labour is *assumed*. To see why, it is necessary to remember the traditional derivation of investment from the schedule of the demand for capital (conceived in the traditional marginalist way as a single factor, an amount of value).[4] The connection – often only implicit – between investment and demand for capital in J. B. Clark, Böhm-Bawerk, Wicksell, Marshall, Knight etc. has been described by Garegnani with a clarity that can hardly be surpassed (Garegnani, 1983: 34–7; 1990:

59–60). Investment was seen by these authors as the *flow* corresponding to the *stock* demand for capital, given that capital wears out and therefore needs a continuous flow of new capital goods for its stock to remain equal to the demand for it:

> The traditional theory implies that the delayed adjustments in the wages, rents, and prices of products do not fundamentally alter the terms of the question ... Hence the significance of the demand and supply functions for capital as a stock, which would exhibit the basic tendencies destined to emerge from the multiplicity of forces acting at any given moment in the savings investment market ... the traditional analyses of the demand and supply for capital were in effect intended to be an analysis of the demand and supply for savings, abstracting from the complications likely to operate at each particular moment of time in the savings–investment market. (Garegnani, 1990: 59–60)

Those authors had to admit of course that in a concrete economy any adjustment to a change in the data of equilibrium (e.g. labour immigration, or technical progress, or changes in the propensity to save) would also present the 'complications' Garegnani mentions, 'complications' due, for example, to differences in the age structure of fixed capital and connected irregularities of the need for replacement of scrapped plants; redistributions of purchasing power among social groups due to changes in the interest rate; possible interference of financial intermediaries; possible convenience of anticipated scrapping of fixed plants; mistaken expectations; slowness in adjustments of factor rentals; and so on. The effects of these 'complications' were to be studied if necessary at a second level of approximation; the demand-for-capital curve was believed to supply 'the basic tendencies', those emerging once the irregularities of the behaviour of prices owing to accidental or transitory disequilibrium phenomena had time to be sufficiently compensated or corrected, and therefore product and factor prices had become sufficiently close on average to their new normal levels, a process enforced by competition. For example, even without any change in optimal technologies a reduction in the rate of interest cannot but push freely competing firms to try and undercut their competitors by lowering product prices relative to money wages since average costs have decreased; if they don't, it will be new firms – whose birth will be stimulated by the persistence of prices higher than average costs – that will do it to gain market shares.

2.3 Capital–labour substitution requires capital to change 'form'

It is opportune to stress a number of important aspects of the determination of the marginalist *long-period investment function*, as I call the investment function generated by such an approach. The demand for capital is determined as the persistent demand for capital goods – aggregated in value terms – implied by the persistent demand for a given net product, this net product being the one produced when labour is fully employed, and production methods, output composition and prices being the normal ones associated with the income distribution determined by the full-employment marginal productivities of labour and capital (following general practice, I assume land is free, because land is not important for the issues to be discussed). Since at each given moment the endowment of 'capital' is crystallised in specific capital goods adapted to a specific productive method, a change (induced by a change in income distribution) in the desired (i.e., normal) capital–labour ratio in an industry can only be realised by replacement of the old durable capital goods with new ones of a different type, or for brevity, can only be realised in new plants (only in new plants can the marginal productivity of capital be determined, since only there can the normal *K/L* ratio be varied); if industry output is unchanged, the new plants will only be built to replace older plants that reach the end of their economic life and are scrapped, plants that are not so old continuing in operation as long as they earn non-negative residual quasi-rents. Changes in the output of an industry, whether due to changes in consumer choices or to changes in production methods in industries using that output as an input, will mostly be accommodated, in the short period, by changes in the degree of utilisation of existing plants, but if perceived as persistent will induce a desire to change productive capacity, and this will be the other main influence on gross investment (per unit of labour employed in new plants), affecting its composition through the desired composition of new capacity. The composition effect due to change of methods is part, in the traditional marginalist approach, of the overall operation of the so-called direct factor substitution mechanism, which changes the desired K/L ratio in the subsystem producing a given final good; the composition effect due to changes in consumer choice constitutes the indirect factor substitution mechanism. In either case, since in most cases it is impossible to utilise existing productive capacity for the production of goods different from those for which that capacity had been planned, the change in industry capacity can generally only be realised through the building of new plants where demand expands, and non-renewal of the scrapped older plants where demand contracts. Thus both the direct and the indirect substitution mechanisms between capital and labour can only operate by affecting the type and sector allocation of the *new* durable capital goods to be combined with the flow of labour gradually released by the scrapping of the durable capital goods that reach the end of their economic life. It is only through the replacement of the existing capital goods with capital goods adapted to produce different goods or adapted to a different technical method, that is, it is only through a change of the 'form' of capital, that the average economy-wide capital–labour ratio can change and a sufficiently elastic demand curve for factors can be obtained. For this reason Hicks (1932: 18–21) expressed strong doubts on the meaningfulness of a *short-period* demand curve for labour, and considered the notion of a demand curve for labour to be meaningful only if one allowed the 'form' of the given capital endowment of the economy to have the time to adapt to the changed real wage.

An implication of this view is that the process of change of the 'form' of capital brought about by a change in the rate of interest is necessarily *slow*, taking – in order to operate completely – the many years required for renewal of the entire stock of fixed plants of the economy: much longer, generally, than necessary for the new rate of interest to determine a gravitation to the new normal relative prices, a gravitation that only requires that the first plants built according to the new optimal factor proportions be capable of imposing a price equal to their lower average cost, obliging the older plants to accept that price and be content with residual quasi-rents. But economic conditions will seldom remain unchanged long enough for complete adaptation of *all* plants to an unchanging rate of interest; therefore it is implicit in this approach that in any concrete economy the rate of interest must be seen as determined, not so as to equalise the capital–labour ratio in factor demand to the ratio of the existing *total* endowment of capital to *total* labour supply, but rather so as to equalise the desired average ratio of capital to labour *in new plants* to the ratio between the flow of 'free' capital (savings) and the flow of labour 'freed' or 'released' by the gradual shutting down of old plants – a ratio that would coincide with the 'total' one only when and if the *entire* labour force were employed in plants embodying methods optimal for that income distribution, and generating productive capacities adapted to the composition of consumer demand. But since most gross investment would be generally motivated by unaltered replacement of used-up circulating and fixed capital, the implicit view of traditional marginalist economists was that the demand-for-capital schedule and its shifts would give a sufficient indication of the tendencies of the investment function.

Any non-negligible difference of actual investment from the long-period investment function would be part of those transitional or irregular 'complications' mentioned by Garegnani.[5] The most important aspect of the long-period investment function, its negative interest elasticity, could in any case be argued certainly to hold for the actual investment function too, since the *K/L* ratio employed on average in new plants would certainly tend to decrease with rises in the interest rate, while the flow of gradually 'freed' labour as well as the speed of completion of changes in industry dimensions would hardly increase.

2.4 The need for the full employment of labour

It should now be clear that the long-period investment function crucially depends on the assumption of full employment of labour. If for simplicity we assume 'investment' *I* to indicate only the part of total investment whose ratio to labour and to output will respond to changes in distribution, that is, to consist only of gross investment going to new plants,[6] and if $L^\wedge$ stands for the flow of labour employed in new plants, the optimal *K/L* ratio determined by the rate of interest determines $I/L^\wedge$, but *I* remains indeterminate unless $L^\wedge$ at the denominator is given. The long-period investment function assumes $L^\wedge$ to correspond to the flow of labour gradually 'freed' by the closure of the oldest plants in a situation of continuous full employment of labour. It is this given $L^\wedge$ that allows the *K/L* ratio corresponding to the given rate of interest to determine investment.

If the possibility of labour unemployment is admitted, then a given average *K/L* ratio in new plants does not suffice to determine investment, because new plants can employ less, or (if there already is some unemployment) more, labour than the flow 'released' by the closure of old plants, correspondingly gradually reducing or increasing employment. A given rate of interest, without some assumption fixing $L^\wedge$, leaves investment indeterminate.

2.5 A simple model with putty-clay capital

A very simple model, that stresses only the direct factor substitution mechanism, can give concreteness to the above considerations. Assume an economy where a single good is produced by labour and putty-clay capital; production within each period adapts to the demand forthcoming in that period (the analysis is in discrete time); the output can be consumed, or it can be invested, i.e., costlessly transformed into capital, but the newly produced capital becomes productive only at the beginning of the following period. The capital–labour ratio must be chosen at the moment of transformation of output into capital, from the possibilities determined by an *ex ante* production function $Y=F(K,L)$, and the resulting capital good allows only one output–labour ratio, which is constant as labour employment per unit of capital varies from zero up to a maximum corresponding to the capital–labour ratio originally chosen. (Thus there may be less than full-capacity utilisation of some or all capital goods.) Capital goods last 10 periods with constant efficiency, independently of the K/L ratio chosen at the time of their creation and of the level of utilisation of the capital good during its life. I assume no technical progress.

The economy is initially in stationary full-employment equilibrium with capital goods fully utilised: at the end of every period the oldest one-tenth of the capital goods are scrapped and replaced by new capital goods of the same type, produced during the period; the newly installed capital goods utilise in the following period the one-tenth of the labour force which is 'freed' by the scrapping of the oldest capital goods. The real wage equals the marginal product of labour in new plants; once the real wage is fixed, the real rate of interest (I neglect risk) is univocally determined (by – owing to the presence of fixed capital – rather complicated equations into which we need not enter).

Then, let us assume, at the beginning of one period the real wage unexpectedly rises (trade unions or political decisions impose this rise, without a change in labour supply) and it is expected to remain at the new level for many periods, and the real interest rate adapts rapidly, so the optimal K/L ratio in new plants rises; the quantity of output destined to investment, let us assume, does not change (this allows us to consider the quantity of capital as not changing); from the subsequent period onwards, part of the one-tenth of the labour supply 'freed' by the scrapping of the oldest plants remains unemployed; the other nine-tenths of the labour force remain employed by the already existing plants, which I assume still yield positive quasi-rents because the wage increase is small. Assume (i) that savings keep translating without difficulty into investment; (ii) that the amount of output destined to gross investment does not decrease in subsequent periods in spite of the decrease in labour employment, so the stock of capital (in the physical sense of total amount of output from which it was created) does not change and (iii) that the real wage does not change. Then after 10 periods the total physical capital K_{Tot} of the economy, measured in physical terms as the sum of the given-up consumption that allowed its creation, has not changed, and labour employment (which is less than labour supply) corresponds to the new lower L/K ratio multiplied by the aggregate capital measured as indicated. All employed labour now produces output at the new Y/L

ratio. The final labour employment as a function of the real wage is indicated by a labour demand curve that traces the marginal productivity of labour when the given physical supply K_{Tot} of capital is introduced into the economy's production function $F(\cdot)$. This is the labour demand function that, as Hicks requested, allows the 'form' of the given quantity of capital to become adapted to the real wage.

(It would not be unrealistic to interpret the length of the period of this analysis as at least a year – fixed plants often last much longer than 10 years –, so the wage change would take at least 10 years to exert its full effect on employment. The slowness of the adjustment implicit in this theory is seldom fully perceived, so its important consequences escape general recognition. One consequence is that, in order to avoid implausible enormous falls of wages whenever unemployment were to arise, the theory must admit the presence of social forces that render wage decreases very slow (Petri 1991: 272–3). But then it is unclear why those same social forces – custom, solidarity, feelings of fairness, aversion to accepting reductions of wage relativities as stressed by Keynes, bargaining power of trade unions, threat of violence, etc. – might not be capable of totally preventing wage falls even in the presence of unemployment, thus constituting the basis for a determination of wages alternative to the neoclassical tendency toward a supply-and-demand equilibrium, and very much in line with the views of the first attentive observer of capitalism, Adam Smith. Another consequence is that even the neoclassical economist has little reason to presume that the *negative* effect, to be discussed later, of a decrease of real wages on employment through its negative effect on aggregate demand will be slower and weaker in its action than the positive effect on the demand for labour coming from capital–labour substitution.)

The assumption that production takes one period (with all productions started at the beginning of a period and ending at the end of the period) means that in each period t the output $Y_t = C_t + I_t$ cannot include the output of plants created by I_t. So Y_t is the result of the full utilisation of the plants that the economy has *at the beginning* of the period, each vintage producing and employing labour depending on the amount of capital good of that vintage and on the K/L ratio chosen for that vintage. Thus in order to determine the demand for labour the reasoning takes Y_t in each period as given, determined by the full utilisation of beginning-of-period capacity. (Changes of the real wage at the beginning of the period have no effect on labour employment in that period, at least as long as the wage change does not cause anticipated scrapping of plants.)

Let us now remember the considerable elasticity of the output of the several industries in real economies, in response to variations in demand (the elasticity that makes the working of the Keynesian multiplier possible). Variations of demand will be met at first by variations of inventories and then by variations of output levels tending to bring inventories to normal – and, in manufacturing industry, generally with little or no change in prices. The premises of this elasticity are not represented in the above model, which lacks inventories, but this elasticity should nonetheless be admitted. And it is well known that firms plan productive capacity for a level of utilisation which is considerably less than the technical maximum level (and is nonetheless esteemed to be optimal for the reasons pointed out in the literature on optimal capacity utilisation: Marris, Betancourt and Clague, Winston, Heinz Kurz etc.), so that not only under-utilisation of plant, but also above-normal utilisation is a possibility. Therefore what we described above (see p. 50) as the *maximum* output/capital ratio corresponding to the chosen *K/L* ratio must more realistically be reinterpreted as the *normal* output/capital ratio, which can be exceeded if demand is particularly high. And 'full-capacity output' must be interpreted to mean *normal*-utilisation output, not an upper limit to actual output.

Once this elasticity of output in response to demand is admitted, then there is no obstacle to admitting an autonomous influence of investment upon output, in either direction. An investment less than normal-capacity savings will encounter no obstacle in causing *Y* to be less than normal-capacity output even if initially there was full employment of a rigid labour supply. An investment greater than normal-capacity savings will cause *Y* to be greater than normal-capacity output as long as an increase in labour employment is possible.[7] Savings will adjust to investment via the variation of *Y* induced by the multiplier.

2.6 Implications of the possibility of unemployment

Once the basic intuition is grasped, it is convenient to abandon the picture of production as consisting of rigidly separate cycles and to admit, more realistically, continuous production and continuous scrapping. The scrapping of old plants causes a flow of 'freed' labour; new plants absorb a flow of new employment. The moment the possibility of unemployment is admitted, even with a constant employment in the already existing plants that continue to be utilised, the second flow can be smaller than the first, causing a gradual rise in unemployment, or greater than the first, with a gradual reduction in unemployment. The indeterminacy of labour employment in new plants implies that a given *K/L* ratio in new plants leaves investment indeterminate, as we saw above (see p. 49). This confirms the conclusion that even conceding the neoclassical conception of capital–labour substitution, income distribution is insufficient to determine investment, and therefore employment, too, is not determined. A given real wage (and corresponding real interest rate and normal relative prices) determines only the ratio *K/L* in new plants, it does not suffice to determine investment (and labour absorption) in new plants. As for employment in already existing plants, the rigid output–labour ratio implies that an assumption of decreasing marginal product of labour cannot be accepted, hence employment cannot be considered as determined by the real wage;[8] more realistically, employment will depend on output which will be determined by sales, and therefore, through the multiplier, by investment. And since the desired *K/L* ratio in new plants leaves investment indeterminate, there seems to be little alternative to considering investment as determined by the desire to reach normal-capacity utilisation, i.e., by the expected level and variations of demand.[9]

But before examining some implications of this view of investment, let us note how the above considerations destroy the neoclassical demand curve for labour. What emerges is that no incompatibility exists between a rise of real wages and a constant or increasing labour employment, *even accepting the neoclassical conception of capital–labour substitution*. Capital–labour substitution can operate only in new plants, and a greater *K/L* ratio in new plants implies a lower absorption of labour in new plants and no increase in employment elsewhere only if investment and the other autonomous components of aggregate demand remain constant or decrease. There is no reason why they should: the elasticity of output makes an increase of employment in existing plants perfectly possible if, for example, public expenditure, or investment, increases (in fact, as I argue later and have written elsewhere (cf. Petri, 2004: 320), investment *will* increase):

> the flexibility of production in response to changes in demand implies that there is no necessary influence, in the short as well as in the long period, of changes in real wages on the demand for labour. In existing plants, where capital already has a given 'form', higher real wages will bring about little or no change in output per unit of labour: employment will depend on capacity utilization which will depend on aggregate demand. In new plants, the flexibility of production of

> capital goods industries will generally pose no problem with obtaining the inputs required by the adoption of the new most profitable methods of production on the scale suggested by the expected level of aggregate demand, even if the latter is increasing considerably. Thus (apart from political reactions) there generally is no incompatibility between more employment and higher wages, all that is required is that the higher wages be accompanied by a stimulus to aggregate demand. This will be so even when it were the case that a higher wage implied a shift to more value-capital-intensive techniques and therefore required more savings: the increase in savings will be brought about by the increase in aggregate output. (Thus one might say, in neoclassical language, that owing to the adaptability of production to demand, relative factor proportions adapt to income distribution rather than the other way round.) (Petri 2011a: 411; 416, footnote 36)

Because of the above, empirical enquiries confirming that in most industries wages equal the marginal revenue product of labour would be no confirmation that the marginal product of labour *determines* real wages, because the causality must be understood to go the other way: owing to the adaptability of production to demand it will be prices and methods of production (i.e., the capital goods utilised by firms) that will adapt to a given real wage, so as to render the marginal revenue product of labour equal to the wage.

2.7 A decrease of real wages reduces investment

But – the neoclassical economist will object – all the above is based on *not* assuming the full employment of labour, and this can be at most a transitory state if the labour market is competitive: the decrease of real wages will increase the demand for labour. But will it really? I have just argued that the neoclassical decreasing demand curve for labour is destroyed by the analysis developed so far. So the effect on employment of a tendency of real wages to decrease in the presence of non-frictional unemployment must be examined anew;[10] and a readiness of workers to accept wage decreases as the normal answer to the emergence of unemployment will not be credible if, as I will argue, such decreases do not generally bring about an increase in employment even accepting neoclassical capital–labour substitution.

In existing plants, I have argued that labour employment depends on sales, not on the real wage; an increase of employment requires an increase of sales i.e., leaving government intervention and changes in the propensity to consume aside, an increase of investment. In new plants, conceding the neoclassical conception of capital, the decrease of real wages reduces the *K/L* ratio. Assume that investment is motivated by desired productive capacity and that the economy has been stationary for some time so initially entrepreneurs have little reason to expect anything but the same demand also for the next few years. Assume initially that already existing plants continue to be utilised normally. Then the new plants can only aim at satisfying the same demand that was satisfied by the scrapped plants they are replacing. Let us initially consider only the direct-substitution mechanism. The decrease of the *K/L* ratio in new plants planned for a given output corresponds to a shift on a given isoquant toward using more labour and less capital, hence *it reduces investment*. If aggregate demand did not decrease, this would not prevent an increase in the demand for labour in new plants (although a smaller increase than if *I* remained constant, see below) and a constant employment of labour in existing plants, and hence some increase in the total demand for labour. But *the decrease of investment reduces aggregate demand,* and then the assumption that the already existing plants keep being normally utilised turns out to be illegitimate, because the reduction of sales has a negative effect on employment in existing plants; thus even though the *flow* of employment in new plants increases, the overall *stock* (the level) of employment *decreases*.[11] Furthermore sooner or later the planned investment in new plants will be further revised downwards as expectations of unchanged sales turn out to be too optimistic; this further reduction of investment may well be small or even absent initially, but since *Y* remains lower than initially (its rise would need a rise in investment, while there is no incentive to such a rise) this will gradually persuade firms that they do indeed need a smaller productive capacity. Thus the decrease in wages starts a reduction in investment *and employment* that may continue for a long time.[12]

Now let us consider the indirect factor substitution mechanism. It is well known that this mechanism may not work in the direction needed by neoclassical theory, but neglecting for the sake of argument the possibility of 'perverse' income effects, the decrease in real wages changes the composition of consumption demand in favour of labour-intensive goods. The traditional derivation, from this change, of an increased demand for labour rests on an assumed unchanged total employment of capital, which in our framework where capital is putty-clay must mean an unchanged total investment. As in the direct-substitution mechanism, this assumption has no justification in view of the freedom with which investment can be decided. As in the other case, there is no reason for firms to expect future aggregate demand to be the greater one connected with *more* labour employment and an *unchanged* capital stock, since current aggregate demand is forthcoming from the income of the given capital and the not yet increased labour employment; only its composition is changing. The more plausible assumption is that the total value of expected demand for consumption goods is equal to the total current expenditure on them, and its changed composition corresponds therefore to a greater demand for labour and less demand for capital, that is, as in the direct-substitution case, less investment. Then the effect is the same as in the other case, a reduction of aggregate demand that causes a reduction of labour employment, with a likely subsequent further discouragement of investment.

2.8 Investment according to Dornbusch and Fischer

I am not the first to argue that even neoclassical theorists should admit an influence of expected sales on aggregate investment (in other words, a role for the accelerator broadly intended). This influence was indeed admitted in the first (1963) version of Jorgenson's 'neoclassical' approach to investment,[13] and it became the basis of the theory of investment in the popular macroeconomics textbook by Dornbusch and Fischer.

The basic idea of the approach of these economists was precisely, as I have argued, to take as given (expected) aggregate demand instead of labour employment in order to determine the desired capital stock and hence investment. Output is treated as if homogeneous and homogeneous with capital; then only the direct-substitution mechanism can be assumed to operate. The rate of interest selects the average capital–labour ratio on the aggregate isoquant corresponding to the planned level and composition of aggregate output; the desired capital stock changes if either the rate of interest, or planned output (i.e., expected demand), or both, change. Thus the desired capital stock is determined by the neoclassically determined capital/output average ratio, *and by the level of aggregate output*. A lower interest rate raises the desired K/Y ratio; with expected Y initially unchanged, the desired capital stock increases, although by less than if L, rather than Y, were kept fixed; the increase of the desired capital stock causes an increase of investment. Thus in the third edition (1984: 206–8) of their macro textbook Dornbusch and Fischer argue that, assuming a Cobb-Douglas aggregate production function $Y=L^{1-\gamma}K^{\gamma}$, the rental of capital (indicated as *rc*) causes a demand for capital K^* that depends on expected sales Y:

$$K^* = g(rc, Y) = \frac{\gamma Y}{rc} \qquad [2.1]$$

The role both of income distribution, and of *Y*, explains Dornbusch and Fischer's use of the term 'flexible accelerator' as an alternative denomination for what they also call the 'neoclassical approach' to investment. The approach of course requires the traditional and unacceptable marginalist conception of capital–labour substitution, and furthermore it is left with the problem of the speed with which the desired capital stock is reached when it changes discontinuously owing to a jump in the rate of interest;[14] but at least it avoids the frequent serious error, found in many current textbooks, of a derivation of the negative interest elasticity of the investment function from a given downward-sloping marginal-productivity-of-capital curve, forgetting that the marginal-productivity curve of capital needs a *given* labour employment, while the investment function is needed for the IS-LM model where labour employment is *variable*.[15] However, there is a price to be paid for avoiding this error: the consequences were alluded to in Section 2.7 and will now be explored further.

2.9 Criticism of the 'Keynes effect' mechanism

The view of Dornbusch and Fischer appears to have been that, since (if expected *Y* is given) the negative interest elasticity of desired capital and hence of investment obtains in their approach too, the 'neoclassical synthesis' criticism of Keynes is valid, a flexibility of money wages would ensure a tendency toward full employment. The well-known 'Keynes effect' mechanism at the heart of the 'neoclassical synthesis' relies on decreases of money wages in the presence of unemployment; this, according to Keynes, brings about some increase of employment in firms that initially expect to be able to sell more at a negligibly lower product price; this causes an excess of aggregate supply over aggregate demand since investment for the moment has not increased; the consequent decrease of the price level causes a decrease in the demand for money, hence a decrease in the rate of interest, hence an increase in investment. The same picture of how the tendency toward full labour employment operates if money wages are flexible downwards is obtained from Dornbusch and Fischer's textbook.

But their different approach to investment opens the way to a number of objections even if the neoclassical conception of capital–labour substitution is accepted without question.

First, the presence of an accelerator influence upon investment makes consideration of what has been happening to *Y* important. If, starting from a situation of desired capital–output ratio equal to the actual one, *Y* decreases for any reason (e.g. because of a decrease in exports, or in state expenditure) and remains low, then desired *K* is lower than actual *K*, and investment is discouraged; and this, through the multiplier, causes *Y* to decrease further, stimulating further decreases in desired *K*. The decrease in the rate of interest brought about by the 'Keynes effect' must then supply a *very* strong stimulus to investment to reverse this downward process. Such a strong stimulus cannot be expected, for two reasons. The first is that the increase in desired *K* is smaller than the one derived from the standard demand-for-capital curve, because the latter determines desired capital on the basis of a given employment of labour, while here firms move along a given (*K,L*) isoquant: this is shown in Figure 2.1 where the isoquant corresponding to a given *Y* is drawn, and a change in distribution that changes the optimal *K/L* ratio from α to β causes an increase of desired capital from K_1 to K_3 if labour employment is fixed at L_1, but only from K_1 to K_2 if output is fixed. The second reason is that the increase in the *K/L* ratio can be realised only in new plants, so it concerns only a very limited portion of productive capacity in every year. (The slowness of the change in the *K/L* ratio pointed out in Section 2.5 should not be forgotten: it is generally underestimated, owing to a mistaken

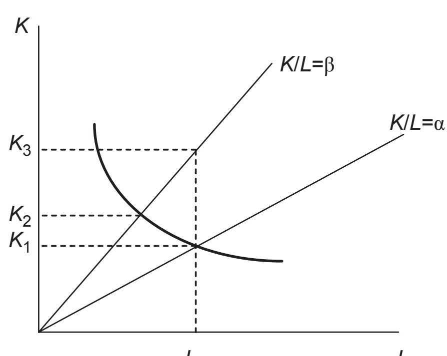


Figure 2.1 The change in the demand for *K* if *K/L* rises from α to β

tendency to conceive capital as putty-putty. Therefore the influence of Y on desired productive capacity and hence on investment has sufficient time to manifest itself.) Therefore even a neoclassical economist has little reason to expect the 'Keynes effect' to be more powerful than destabilising multiplier–accelerator interactions.

Second, the Dornbusch and Fischer approach implicitly recognises – in accordance with standard microeconomics – that the marginal products of the two factors labour and capital are tied together in such a way that if one marginal product increases, the other decreases, and that factor prices adjust to marginal products so that normal competitive extra-profits net of risk must be assumed to be (close to) zero when one studies investment.[16] This means that an increase in the desired *K/L* ratio will be associated with a change of relative factor prices consisting of a decrease in the real interest rate and an *increase* in the real wage. In order for the marginalist factor substitution mechanism to stimulate investment by raising the *K/L* ratio in new plants, the real rate of interest must decrease, i.e., *the real wage must increase*. On the contrary, the first stage of the 'Keynes effect' mechanism supposed to raise employment if money wages decrease consists of a *decrease* of real wages: firms raise employment and production because money wages decrease relative to prices that have not decreased yet; once prices start decreasing, since plausibly they decrease with some lag relative to the decrease of money wages, the real wage perhaps stops decreasing but remains lower than initially for all the deflationary period. As pointed out in Section 2.7, then investors have an incentive to adopt a *lower K/L* ratio in new plants, and this causes investment to *decrease*. To avoid this result, it would seem necessary that the decrease in real wages be strictly temporary, soon reversed by an *even greater* decrease in the price level (caused by prices rapidly adjusting to average costs including not only lower money wages but now also a lower rate of interest); then because of the rise in real wages the desired *K/L* ratio in new plants increases; this will hopefully stimulate investment, and *Y*, to the point of raising the demand for labour *in spite of* the rise in real wages. But note how one will then be admitting the possibility and indeed *necessity* of, at the same time, raising employment and real wages! Then it becomes difficult to deny that it must be the task of public intervention to secure such a result without the slowness and uncertainties of leaving it to the spontaneous workings of the market, which would in any case not be at all guaranteed to work in the required direction, because there is little reason to expect the necessary greater decrease in the price level to be sufficiently fast – firms are notoriously hesitant to decrease prices – and furthermore it is well known that price decreases raise the debt burden with possible negative effects on production and investment. (To all this one can add the well-known negative effect on the propensity to consume, and hence on the multiplier, associated with a decrease in real wages.) It is in any case striking that the rise in employment will have to be associated with a rise, not a decrease, in real wages. (Is this perhaps the reason why the Dornbusch and Fischer approach was not more widely adopted and was subsequently totally forgotten?)

2.10 A conclusion

These considerations should suffice to show how little one can trust that downward flexibility of money wages will reduce unemployment, even neglecting the Cambridge capital-theoretic criticisms, the moment one more consistently develops (full employment of labour not being initially assumed) the implications i) of the importance of durable capital; ii) of the inevitable influence of expected demand on investment; and iii) of the multiplier, and of possible multiplier–accelerator interactions broadly conceived.

To the above one must then add (i) the empirical evidence that consistently contradicts the presumption of a significant interest elasticity of investment; and (ii) the Cambridge results in capital theory, in particular the possibility of reverse capital deepening, that undermine the neoclassical conception of capital–labour substitution and show that the *theoretical* presumption of a negative interest elasticity of the demand for value capital per unit of labour has no solid foundation, so the lack of *empirical* support for such a presumption is not surprising. The conclusion must be that there is no reason at all to believe in a spontaneous tendency of market economies toward the full employment of labour.

2.11 Implications for wage flexibility and for growth theory

Two important implications of this conclusion can be pointed out.

The first is that the assumption that in the presence of unemployment money wages will decrease becomes implausible, and the Friedmanite thesis, that if in the presence of unemployment wages do not decrease then unemployment is voluntary, loses its analytical foundations. If reductions of wages have little or no effect on labour demand and can even have a negative effect, cumulative historical experience will have taught this fact to the labouring classes, ways will have been found to teach this knowledge and the consequent appropriate rules of conduct to the young, and it is then perfectly understandable that an unemployed worker will not, apart from exceptional circumstances, try to obtain a job by undercutting others. The generalised reduction in wages that wage undercutting would bring about would not reduce unemployment; it would only lower the incomes of employed workers – who often are the relatives of unemployed workers, and on whose income the living of the latter may depend. In such a situation it would be mistaken to define *unemployment* as voluntary: the absence of wage reductions is voluntary, but not unemployment. The unemployed worker, in refusing to accept a lower wage is *not* choosing the alternative 'no wage reduction, no job' over the alternative 'wage reduction, job'.

The second implication is the need to reconsider the theory of growth. The elasticity of output with respect to demand pointed out in Section 2.6 (see p. 53) strongly suggests a view of economic growth and capital accumulation as dependent on the evolution of the autonomous components of aggregate demand, because it implies that *aggregate* production can quickly adjust not only to decreases of aggregate demand, but also – within limits rarely approached – to *increases* in aggregate demand, so that it is generally possible, even in economies very close to full employment, to raise consumption *and* investment at the same time, if aggregate demand increases.[17] Hence investment is hardly ever constrained by savings; capital accumulation will result from the demand for additions to capital stocks due to increases in desired capacity, in turn due to increases of aggregate demand. A growing literature is developing these insights.

2.12 Implications for DSGE macro models

I conclude by briefly pointing out the relevance of the above analysis to the currently fashionable foundation of macro theory upon Dynamic Stochastic General Equilibrium (DSGE) models, where the problems for Say's Law pointed out in this chapter are pushed out of sight by an *assumption* of continuous full employment of the labour supply and of investment determined by savings. This assumption is generally justified by reference to the 'rigorous' microfoundations supplied by general equilibrium (GE) theory: the models are argued to be simplified renditions of the results one would derive from completely disaggregated intertemporal GE models, possibly made more realistic by the admission of adjustment costs, imperfect competition, and so on. The claimed premise of these models is therefore that intertemporal GE theory is a robust descriptive theory.

The curious thing is that the claimed consistency of this type of macro models with infinite-horizon GE theory is announced with pride, as supporting the trustworthiness of these models, while on the contrary more and more often general equilibrium specialists advance strong reservations on the descriptive validity of GE theory. One can mention Michio Morishima, Stephen Marglin, Duncan Foley and Alan Kirman as at one time convinced neoclassical theorists who have decidedly rejected GE theory. An implicit rejection or at least an agnostic attitude also emerge in the fact that the problems with uniqueness and stability have

> led many microeconomists to forsake the general equilibrium conceptualization altogether. As a result, microeconomic theory has, by and large, been reduced to a collection of techniques and tricks for resolving narrow, isolated microeconomic problems and the study of, also narrow and isolated, strategic behaviors. (Katzner, 2006, p. ix)

One can also mention the frequent denunciations by, for example, Frank Hahn or Franklin Fisher, of the sterility of stability studies based on adjustments that do not allow the implementation of disequilibrium decisions; but if time-consuming adjustments are allowed, the equilibrium becomes indeterminate because the data relative to the endowments of the several capital goods are no longer data, being altered by production. Also, many theorists are very uneasy about the utterly unrealistic assumption of complete futures markets or correct foresight; but the alternative of temporary equilibria without correct foresight, explored in the 1970s and early 1980s, is nowadays totally discredited (as evidenced by its complete disappearance from advanced micro textbooks), owing to the problems it encountered, which explains why Lucas, real business cycle theories, or DSGE models only refer to intertemporal equilibria as their 'rigorous' microfoundation.

One might then reject the DSGE approach in macroeconomics simply as a consequence of the rejection of intertemporal general equilibrium theory as a positive theory, a rejection motivated by this theory's need for the untenable assumption of complete markets or perfect foresight, by its lack of uniqueness or stability even granting the auctioneer, and by its inability to say anything on the distance between equilibrium paths and the behaviour of economies not continually perfectly in equilibrium (Petri, 1999: 50).[18]

But, as I have argued elsewhere (Petri, 1999: 53–4), it is difficult to understand the acceptance of intertemporal equilibria as descriptively valid without a more or less conscious belief that the undeniable occurrence, in actual economies, of disequilibrium and time-consuming adjustments does not destroy the neoclassical theses as to the *trend* the economy follows, which is reasonably approximated by the intertemporal equilibrium path. Only an idiot would deny that in actual economies there is no auctioneer and no complete futures markets, but rather time-consuming trial-and-error adjustments, mistakes, disequilibria, imperfect foresight; so DSGE theorists must believe that there are persistent forces that cause these disequilibria to be sufficiently corrected or compensated so that the trend the economy actually follows is not too far from the path described by their models.[19] But then the reference to disaggregated intertemporal equilibrium with perfect foresight as the 'rigorous' microfoundation of the models is only a smokescreen; behind it there is in fact a belief in the time-consuming adjustment mechanisms on whose basis the marginal approach was born and accepted, and that after Keynes were rehabilitated by monetarism carrying forward from the criticism of Keynes initiated by the neoclassical synthesis. Without some such belief the reference to intertemporal equilibria would be devoid of any justification, given that by themselves neo-Walrasian equilibria and their sequences tell us *nothing at all* about the actual path a market economy not continuously in equilibrium will follow.

For this reason, the arguments of the present chapter are relevant criticisms of DSGE models too, as well as of the whole development of neoclassical macroeconomics after Keynes. The characterisation of contemporary neoclassical macro models as simplified intertemporal general equilibrium models would, if taken seriously, deprive these models of any pretension to descriptive validity; such a pretension can only rely on traditional neoclassical macroeconomic tendencies, that is, on the same time-consuming adjustment mechanisms on which J. B. Clark or Wicksell or Pigou or Hayek, or the 'neoclassical synthesis' and monetarism, based their analyses. The relevance of the argument presented here, then, lies in its pointing out that, the moment the continuous full employment of labour is not assumed to start with, those adjustment mechanisms will not work as normally presumed, and Say's Law loses credibility even before one questions the neoclassical conception of capital–labour substitution: this was not realised because the correct implications of that conception for investment theory when labour employment is not given were not grasped. The recuperation of pre-Keynesian views initiated by the neoclassical synthesis and carried forward by monetarism, which is what lies behind the current faith in DSGE models, was made possible by a theory of investment which was mistaken not only in its foundation on an untenable conception of capital, but even if that conception is not questioned.

Notes

1. Of course I am not denying the possibility of a mutual influence (e.g. the propensity to save might depend on the rate of interest); but the basic idea of the theories I intend to criticise is that the aggregate amount of saving may well be given independently of investment, and then investment will adapt to it, while the converse (given investment, and savings adapting) does not happen; to insist on mutual determination misses this fundamental asymmetry.
2. A brief summary of my arguments is available in Petri, 2003, section 5.
3. I will not consider the attempts to defend a tendency of aggregate demand to adjust to aggregate supply on the basis of the Pigou (or 'real balance') effect: even Patinkin conceded that this effect is uncertain and anyway too weak. Cf. Petri, 2004, Appendix 7A2, pp. 292–5 for a confirmation of Patinkin's scepticism, based on recent estimates of the wealth effect.
4. I briefly recall and criticise here the two main attempts *before* Keynes to derive a negative interest elasticity of investment without having recourse to capital–labour substitution. In Walras future rentals of capital goods are treated as independent of the rate of interest, so the demand price of new capital goods (the discounted value of the given future rentals) rises when the rate of interest decreases, and this stimulates their production; but these given future rentals are an obviously indefensible assumption since the interest rate is one of the distributive variables and its changes alter the rentals of all other factors; as standard microeconomic theory teaches, product prices tend to minimum average costs, but then if the rate of interest decreases the rate of return on investments will tend to decrease too: the rentals of capital goods will decrease relative to their supply prices. The same objection applies to Irving Fisher, who assumes for each saver/investor a given series of alternative income streams among which the investor chooses – for each rate of interest – the one with the highest present value; prices are treated as given independently of the rate of interest, as in Walras. Thus Alchian (1955: 942) writes that Fisher's 'exposition ... is based on the supposition that one merely changes the rate of interest and holds other prices fixed', and correctly accuses such a procedure of logical inconsistency. Actually Fisher admits that changes in the rate of interest alter relative prices, but he dismisses the need for further discussion of the issue by writing that this influence is 'a factor which, after all, is more intricate than important' (*The rate of interest*, p. 168), a statement for which no support is supplied. On Keynes on investment cf. Petri (2004: 260–2).
5. Investment must anyway be determined over not too short a period, to avoid its being relevantly influenced by transitional phenomena such as, for example, a *decrease* of the rate of interest inducing an expectation of a further decrease in a few months' time, and inducing therefore a postponement, i.e., a *decrease*, of investment; or anticipated scrapping of new plants induced by the change in prices due to a *rise* in the rate of interest, that may induce a temporary *increase* in investment.

6. The rest of gross investment – partial replacement, without alteration, of durable capital components of existing plants that are not scrapped, plus purchase of raw materials to be used in already existing plants – will be generally rigidly determined by intended outputs and by the technology embodied in the plants, and will therefore be independent of changes in income distribution except in so far as these affect the composition of demand; accordingly, it can be taken as given (and for this reason it is permissible to neglect it) as long as normal utilisation of existing plants is assumed.
7. That is, as long as normal-capacity output, the one associated with the normal utilisation of existing productive capacity, is less than is necessary to employ the entire labour supply.
8. On this issue it is worth noticing the agreement between the empirically based criticism by Dunlop and Tarshis of the decreasing marginal productivity of labour in Keynes, and the earlier theoretically based rejection of a short-period decreasing demand curve for labour by Hicks (1932).
9. Of course innovation will be another fundamental determinant of investment, but its effects do not seem relevant for a discussion of the validity of Say's Law.
10. Keynes objected that, unless investment increases and absorbs the increased saving associated with the increased output brought about by the greater employment, the decrease in real wages and increase in employment will not happen, because workers can only reduce *money* wages, and the insufficient aggregate demand will cause prices to decrease in step with money wages. This argument rests on an assumed decreasing marginal product of labour in the several plants, so if real wages were capable of permanently decreasing, the demand for labour *would* rise; I wish to question the robustness of the neoclassical argument *even conceding* a decrease of real wages. Below (Section 2.9) I discuss money wages.
11. Consider the following rough example. Suppose I/Y is constant, the average life of plants is 10 years, and the reduced K/L ratio causes L/Y in new plants to rise by, say, 5 per cent; the first year the increased hiring of labour in new plants as a percentage of previous labour employment is 5 per cent of 10 per cent, i.e., one half of one percentage point. If investment decreases by, say, 4 per cent, this causes a reduction of Y (and plausibly of employment in existing plants) by 4 per cent, i.e., about a 3.5 per cent reduction of labour employment.
12. This conclusion is reached without considering the negative effect on Y due to the generally admitted rise in the average propensity to save caused by the decrease of the share of wages in national income, an effect which, if admitted, would strengthen the argument.
13. On the difference between Jorgenson's 1963 and 1967 approaches cf. Petri (2004: 287–90).
14. This speed is determined by Dornbusch and Fischer through a 'gradual adjustment hypothesis' that states that the larger the gap between the existing capital stock and the desired capital stock, the more rapid a firm's rate of investment. Empirical evidence is then referred to in order to estimate the speed of adjustment.

15. This is just one of the many grave mistakes to be found in the arguments for a negative interest elasticity of aggregate investment after Keynes (cf. Petri, 2004: 271–81).
16. Without this reciprocal adjustment of factor prices, the desired *K/L* ratio would not be given by equation [2.1]. Obviously the extra-profits to be considered are the ones on new plants, existing plants earn quasi-rents. (I use ‘extra-profits’ to mean what standard microeconomics calls ‘profits’, in order to avoid confusion with the classical meaning of ‘profits’.) On the need for such assumptions for the study of investment cf. Garegnani, (1983: 36; 46, fn. 25), also Petri (2011b: 67).
17. Labour constraints are usually not binding in the short run because of visible or hidden unemployment and underemployment, and over the longer run there are migrations, and structural social adaptations, e.g. changes in the participation of women.
18. A survey of these and other criticisms is in Petri (2011b).
19. Some such view is, for example, implicit in the numerous admissions by Lucas, Sargent and others that rational expectations make sense only for situations sufficiently persistent for agents to have had the time to learn how correctly to form their expectations – with the implication that during the learning mistakes are inevitable; but some learning is going on all the time, because of the continuous emergence of novelties (in each industry there may be technical progress, changes in tastes, etc.), hence those admissions imply that most markets are most of the time in disequilibrium, which can only be neglected if one looks at trends of the averages and one assumes that time-consuming adjustments operate which cause the trends to be sufficiently close to the equilibrium path.

References

Alchian, A. A. (1955) ‘The Rate of Interest, Fisher’s Rate of Return over Costs and Keynes’s Internal Rate of Return’, *American Economic Review*, 45 (December): 938–43.

Dornbusch, R. and Fischer, S. (1984) *Macroeconomics*, 3rd edn (New York: McGraw-Hill).

Garegnani, P. (1983) ‘Notes on Consumption, Investment and Effective Demand’, in J. Eatwell and M. Milgate (eds), *Keynes’s Economics and the Theory of Value and Distribution* (New York: Oxford University Press); repr. with modifications of a paper with the same title published in two parts in *Cambridge Journal of Economics*, 1978, 2(4): 335–53; and 1979, 3(1): 63–82, in turn a translation of ‘Note su consumi, investimenti e domanda effettiva’, *Economia Internazionale*, 1964–5.

Garegnani, P. (1990) ‘Quantity of Capital’, in J. Eatwell, M. Milgate and P. Newman (eds), *The New Palgrave: Capital Theory* (London: Macmillan).

Hahn, F. H. (1981) ‘Review of M. Beenstock “A neoclassical analysis of macroeconomic policy” (Cambridge University Press, 1980)’, *Economic Journal*, (December): 1036–9.

Hayek, F. A. (1932) ‘Money and capital: a reply’, *Economic Journal*, 42: 237–49.

Hicks, J. R. (1932) *The Theory of Wages* (London: Macmillan).

Jorgenson, D. V. (1963), 'Capital theory and investment behavior', *American Economic Review*, 53(2), (May): 247–59.

Katzner, D. W. (2006) *An Introduction to the Economic Theory of Market Behavior. Microeconomics from a Walrasian Perspective* (Cheltenham, UK: Edward Elgar).

Petri, F. (1991) 'Hicks's recantation of the temporary equilibrium method', *Review of Political Economy*, 3(3): 268–88.

Petri, F. (1999) 'Professor Hahn on the neo-Ricardian criticism of neoclassical economics', in G. Mongiovi and F. Petri (eds), *Value, Distribution and Capital: Essays in Honour of Pierangelo Garegnani* (London: Routledge): pp. 19–68.

Petri, F. (2003) 'Should the Theory of Endogenous Growth be based on Say's Law and the Full Employment of Resources?', in N. Salvadori (ed.), *The Theory of Economic Growth: A 'Classical' Perspective* (Cheltenham, UK and Northampton, MA, USA: Edward Elgar): pp. 139–60.

Petri, F. (2004) *General Equilibrium, Capital and Macroeconomics* (Cheltenham: Edward Elgar).

Petri, F. (2011a) 'On the Likelihood and Relevance of Reswitching and Reverse Capital Deepening', in N. Salvadori and C. Gehrke (eds), *Keynes, Sraffa and the Criticism of Neoclassical Theory. Essays in Honour of Heinz Kurz* (Abingdon and New York: Routledge): pp. 380–418.

Petri, F. (2011b) 'On the Recent Debate on Capital Theory and General Equilibrium', in V. Caspari (ed.), *The Evolution of Economic Theory. Essays in Honour of Bertram Schefold* (Abingdon and New York: Routledge), ch. 4: pp. 55–99.

3
The Meaning of Output Trends in the Analysis of Growth

*Attilio Trezzini**

3.1 Introduction

This chapter discusses the meaning of the output trends examined in the analysis of growth and their relation to actual levels of output and income. These issues are seldom explicitly examined in the literature, where the relation between actual and theoretical magnitudes assumed in the theory of prices appears to be automatically extended to the magnitudes addressed by theories of growth. In field of economic growth, however, the distinction between actual and theoretical magnitudes overlaps with the distinction between cyclical and trend positions of the economy. This gives rise to crucial misunderstandings, at least from the standpoint adopted here.

Any particular notion of a *trend level of output* assumed in the theory of growth implies a corresponding notion for all the components of output, including the trend level of aggregate consumption, investment and saving. Specific meanings are then attached to the ratios between aggregates: trend values for the ratio of saving to income, of consumption to income or investment to income.

The chapter opens with a discussion concerning some features of the relation between theoretical and actual magnitudes assumed in the theories of prices and distribution. This introductory section prepares the ground for the critical discussion then developed of the method widely used in the literature, whereby the trend of produced quantities is studied through growth paths whose characteristics – level and rate of growth – are defined regardless of the quantities actually produced. As

*I am grateful to R. Ciccone and P. Trabucchi for useful and interesting discussions and to an anonymous referee for the comments on a first version of the chapter.

argued in the second section, this method derives from the simple and unjustified *transposition* of the relation between theoretical and actual magnitudes developed in the theory of prices into the theory of the evolution of quantities.

The viewpoint of the classical and Keynesian approach to the analysis of growth is then assumed in the third part of the paper and solid grounds are found for rejection both of this transposition and of the assumption that the trends are independent of actual fluctuations in output.

3.2 Theoretical and actual magnitudes in the analysis of prices and distribution

Our starting point is the relation between theoretical and actual magnitudes in the theory of prices and distribution, where this relation has been developed and discussed and where its general terms are therefore sufficiently clear. The relation between theoretical and actual magnitudes is addressed here with a distinction drawn between short- and long-period magnitudes or between actual and natural or normal magnitudes.[1]

The magnitudes examined by theoretical analysis, i.e., normal prices, are studied with reference to ideal situations in which it is assumed that no transitory and purely accidental phenomena affect prices and that the process of competition has been completed. Normal prices imply a uniform rate of profit on capital invested in various industries.

The determining forces behind these theoretical magnitudes are more persistent than those determining temporary deviations of actual from theoretical magnitudes. They are thus sufficiently lasting to allow, in principle, the adjustment of actual prices to normal prices. The latter are therefore seen as centres of gravitation of actual prices and hence able to account almost completely for the determination of relative prices.

Actual prices differ in fact from normal ones due to numerous circumstances affecting the composition of output, which proves inadequate with respect to the composition of effectual demand.

The circumstances are transitory in nature and more or less accidental. They can in fact be determined by an inadequate composition of productive capacity inherited from the past, by the persistence in some firms of old techniques and other phenomena that competition tends to eliminate. Further divergence of the quantities brought to the market from the corresponding effectual demand can be determined by more accidental phenomena such as the effect of a stormy night on the day's fishing. All these phenomena are transitory, however, being automatically cancelled out either by the adjustment of production levels through the process of competition or by an accidental phenomenon with the opposite effect.

Relative persistence is an essential property of the causes that determine natural prices, allowing the necessary adjustments to take place in the case of phenomena such as the composition of productive capacity or the presence of obsolete techniques, and allowing purely accidental phenomena – a stormy night – to be cancelled out by accidental phenomena with the opposite effects. This property is thus essential if the natural prices determined by these persistent causes are to be seen as centres of gravitation for actual prices and these theoretical magnitudes are therefore to be considered significant in the interpretation of actual economies.

3.3 Relevant features of theoretical prices

Some *features* of the relation between theoretical and actual magnitudes in this theoretical field should be stressed.

The conditions determining normal prices are largely *independent* of the possible levels of actual prices. They can thus be studied separately from the causes determining the tendency of actual prices toward natural prices and separately from analysis of the ways in which this tendency takes place.

Prices and distribution are determined in classical political economy on the basis of the technical conditions of production, the level and composition of normal real wages, the level and composition of the quantities produced (the conditions of reproduction), and the tendency towards a uniform rate of return on capital. This set of conditions is independent of the possible levels of actual prices and rates of return. Normal prices remain the same regardless of whether some actual prices are temporarily higher or lower than normal and regardless of whether a divergence lasts for a longer or shorter period.

This property also implies that during any possible process of adjustment or gravitation of actual to normal prices, the latter do not change. A process of adjustment is thus conceivable and normal prices are independent of it.

Normal prices have an *independent existence* with respect both to the phenomena determining divergences of actual prices and to the paths followed by actual prices in adjustment or gravitation.

From another viewpoint, this independence of the determinants of natural prices from actual prices can be understood as the fact that the effects of the causes of divergence of actual prices from natural prices are *reversible*. When the quantity of a commodity brought to the market is greater than the corresponding effective demand, the actual price proves lower than the natural price. This has no effect on the determination of the natural price, however, and the effects of the causes determining the divergence (an error in the quantity produced) are therefore reversible. Once these causes are cancelled out, they no longer affect prices.

Similar considerations can be developed with respect to marginalist theories, according to which long-run prices are determined by consumer tastes, technical conditions of production and the quantities of factors of production available. These circumstances determine equilibrium prices (and quantities) independently of actual prices.

The marginalist determination of prices and distribution determines equilibrium levels and composition of *output* simultaneously. The study of *gravitation* is, in this context, more complex and less independent of actual quantities with respect to the corresponding theoretical process of classical political economy. The independence of normal prices from the corresponding actual magnitudes and the paths of adjustment therefore requires additional assumptions, such as the existence of an auctioneer and that actual exchanges take place only at equilibrium prices.[2]

This similar conception of the relation between theoretical and actual magnitudes in classical and marginalist analyses of prices is combined, however, with radical differences.

As is known, a major difference between the marginalist and classical approaches is the simultaneity of the former as opposed to the method based on separate logical stages that characterises the latter. In addition to the fundamental implication of divergent positions with respect to the tendency towards full employment, this difference also has relevant implications of a methodological nature.

The simultaneity of marginalist analyses inevitably prompts the adoption of a uniform method in different fields of economic analysis. All the relations are regarded as sufficiently simple and general to be represented – at least in their fundamental aspects – by means of simple and general mathematical relations.

Conversely, a meaningful distinction between the method used in analysis of the relations determining prices and distribution – the *core* – and the one used in the analyses of output and accumulation – like all the relations *outside the core* – was envisaged by Garegnani (1984) in classical political economy. Being automatically enforced by the working of competition, the relations *inside the core* are comparatively simple and sufficiently general to be represented by mathematical relations – price equations – that are by nature necessarily simple and general.

The relations between magnitudes outside the core appear to be dramatically affected by circumstances of a historical, institutional and political nature that are at the same time *fundamental* and susceptible of being different in different periods or different economies. This makes the use of quantitative relations to represent their influence highly implausible if not indeed misleading. The use of such relations is limited to the construction of *models* or *examples* for strictly *heuristic* purposes.

3.4 Theoretical and actual quantities vs. cyclical fluctuations and trend magnitudes

The lack of an explicit definition of the meaning of the theoretical magnitudes addressed – generally output trends – in the theory of growth is particularly problematic because, in the field of the evolution of produced quantities over time, the distinction between long-run (persistent and fundamental) and short-run (transitory and/or accidental) phenomena affecting the determination of prices overlaps with another relevant and distinct phenomenon, namely the fact that economies generally develop over time through fluctuations in the levels of produced output that describe the trend of the economy.

Unlike the situation with prices, a large amount of data is available for aggregate output. Figure 3.1 shows the actual data for Italy from 1960 to 2012. Real aggregate output clearly displays irregular evolution over time with an increasing trend.

Also evident is the existence of recessions of a more or less lasting and intense nature that interrupt the uneven growth. Each 'cycle' thus has specific different characteristics: average rates of growth and different 'patterns of fluctuations'.

Let us first consider this picture independently of the literature. Treating the entire span as one long, uniform period seems only to prevent us from explaining why the economy grows more intensely in one period than another. This appears to be precisely the object of theoretical analysis of the determinants of economic growth.[3] It therefore seems reasonable to divide the period of observation into sub-periods marked by recessions, which can thus be defined as cycles. A trend can then be extracted from each of these cycles, as presented in Figure 3.2 (by means of simple linear regression or any other more sophisticated statistical procedure).[4]

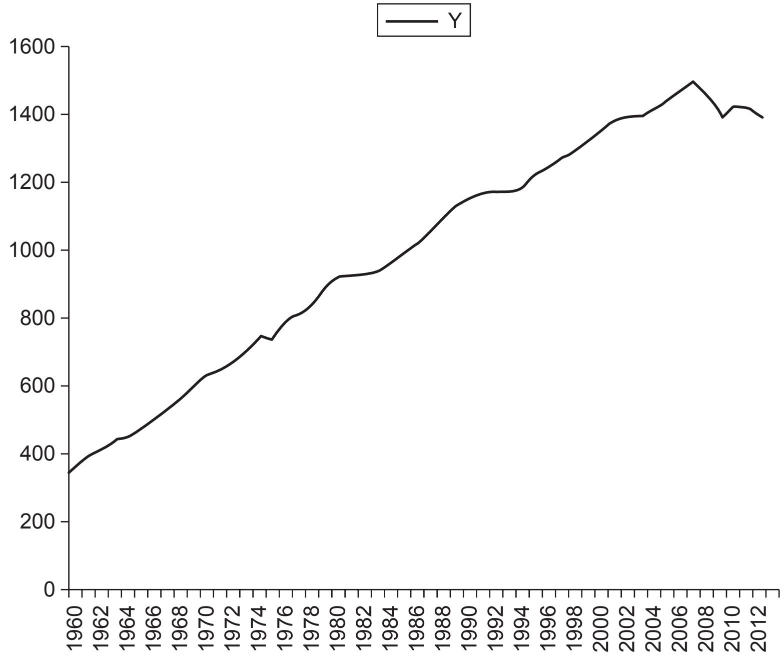


Figure 3.1 Real GDP (2000 prices), Italy, 1960–2012
Source: Ameco dataset.

These trends seem to be the phenomena requiring explanation. Why is economic growth more or less intense in one period than another? Why did the economy not grow more?

Figure 3.3 shows the actual data and their (linear) trend for real output of the first period into which it is plausible to divide the period of observation.

Any theory seeking to explain economic growth should be able to detect the major determinants of the phenomenon described in Figure 3.3. This necessarily entails a theoretical position as regards the determinants of the actual magnitudes represented by points like A in the figure, the determinants of the trend magnitudes represented by points like A*, and the relation between the two kinds of magnitude.

There is no explicit discussion of this methodological issue to be found in the literature on growth. What we are about to develop are therefore considerations about what appears to be the rationalisation of a method widely used in the literature. Apart from some exceptions

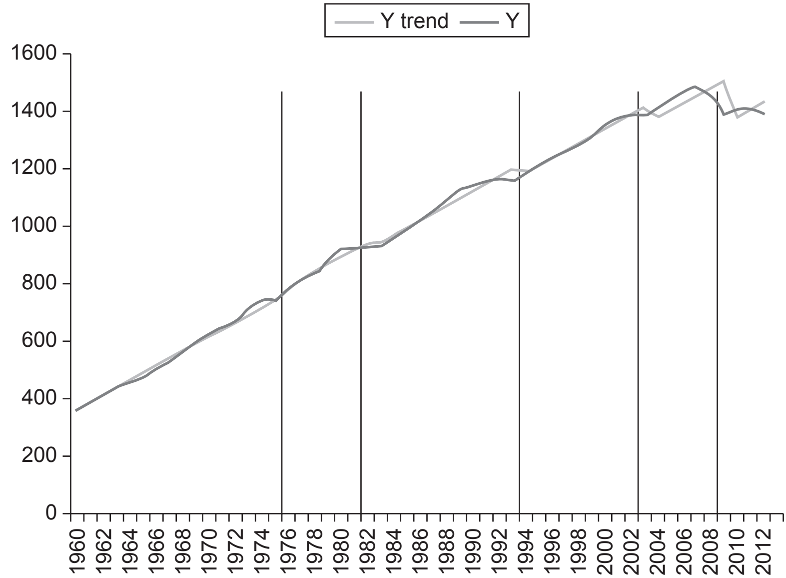


Figure 3.2 Actual data and linear trends for sub-periods: real GDP (2000 prices), Italy, 1960–2012
Source: Ameco dataset.

in both the mainstream (the Real Business Cycle) and Keynesian approaches,[5] what can be seen in most of the literature on growth is a tendency to study the determinants of the trend *independently* of the determinants of cyclical fluctuations. The causes of fluctuations are, at the same time, implicitly assumed as either *accidental* or *structural* phenomena that can in any case be studied separately and independently of the trend around which fluctuations occur.

This position seems to be based on automatic transposition of the method used in the theory of prices and distribution. The relations between the fluctuations of produced quantities (points A) and their trend levels (points A*) may in fact appear to be similar to the relations between short-period prices and the corresponding normal levels.

The assumption of the existence of trend levels of output independent of fluctuations in output is identical to the assumption of the existence of long-run or trend values for some *crucial* relations between quantities such as the capital–output ratio, K/Y, average propensity to save, S/Y, or investment–output ratio, I/Y, which can again be studied, and are actually determined in the models, independently of their actual values at different stages of the cyclical fluctuation.[6]

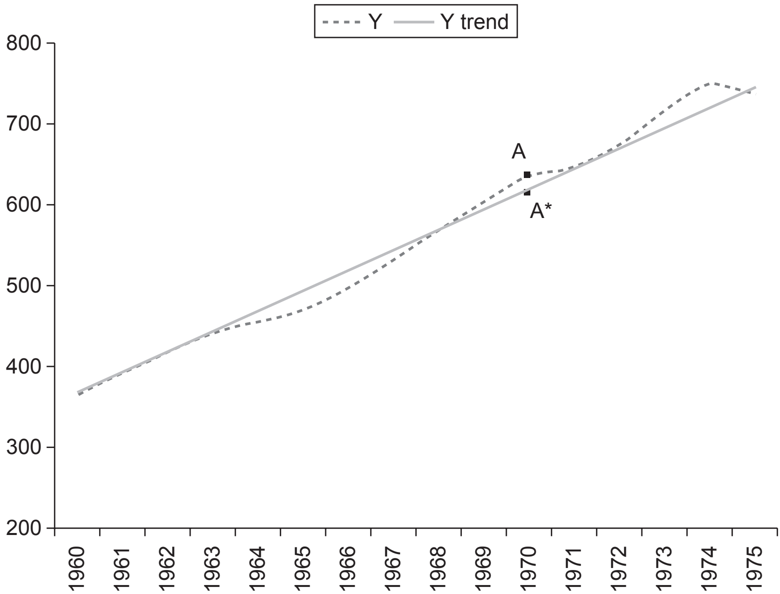


Figure 3.3 Actual data and linear trend: real GDP (2000 prices), Italy, 1960–1975
Source: Ameco dataset.

Our aim here is to consider this position, which is certainly plausible and widely accepted in the literature but may reveal some weaknesses on in-depth analysis.

3.5 The classical and Keynesian approach: the background

It is essential to state specifically that the argument presented in this chapter is developed within the framework of an approach to the analysis of growth in which the determining role of the evolution of aggregate demand is properly acknowledged. In particular, the approach adopted originates both from criticism of the marginalist idea of capital as a factor of production and from the classical surplus approach, defined as the classical and Keynesian approach to the analysis of growth.[7]

The critique of the marginalist concept of capital has been shown to be sufficient to deprive the mainstream theories of the most solid foundations for the alleged tendency towards full-employment equilibria. As such, this criticism is thus regarded as a 'second route' to the Keynesian principle of effective demand (Garegnani, 1978–79, 1983), the second in chronological terms but more solid from the theoretical viewpoint than the first opened up by Keynes himself. It is in fact completed by reference to the surplus theory of value and distribution, which is compatible with the principle of effective demand and thus makes it possible to extend the Keynesian premises to analysis of the long-run tendencies of economies.

These premises thus become the basic element for a classical and Keynesian approach to the analysis of growth and accumulation. The level of aggregate demand is conceived as the primary determinant of the level of output in any one period. It also determines the degree of utilisation of existing capacity and through it the evolution of capacity itself. Changes in aggregate demand generate changes in output through immediate changes in capacity utilisation, which lead in turn to the creation or destruction of capacity through higher or lower (nil or even negative) flows of investment. The latter in turn have effects on both demand and capacity. One central concept formulated is the *high elasticity* with which output responds to changes in aggregate demand in both the short and the long run. While this elasticity is determined in the short run by changes in the utilisation of capacity that allow different levels of production, the elasticity of output is made possible in the long run by changes in the degree of utilisation of capacity and by the consequent creation of new capacity.

3.6 The rigidity of output to demand and definition of the methodological issue

The concept of the elasticity of output with respect to changes in aggregate demand has a crucial bearing on the issue discussed here. The independence of the trend of output from cyclical fluctuations is in fact generally asserted on the basis of the assumption of the rigidity of output with respect to changes in aggregate demand. It is possible to identify two kinds of rigidity assumed in the literature.

The first is the rigidity asserted in the mainstream theories of growth due to the principle of the tendency towards full-employment equilibria. This principle leads to denial of any elasticity of output with respect to changes in demand and with it any role of aggregate demand in the growth process. In the context of mainstream analyses of growth, the tendency towards full employment is also sufficient to select a single path of evolution of output independently of any phenomenon determining fluctuations (whether accidental or systematic) superimposed on the tendency towards full employment. The full-employment trend of produced quantities therefore has an existence independent of fluctuations. This trend is also unaffected – at least in principle – by the actual positions of the economy during the cycle.[8]

The principles leading to the tendency towards full employment then lead to the *transposition* of the relation between theoretical and actual magnitudes developed in theories of prices into the analyses of growth. This transposition is consistent with the characteristic of marginalist theories that they are theories of prices, distribution and equilibrium quantities all at the same time. This simultaneity naturally and necessarily leads to extension of the same conception of the relation between long- and short-period prices to the analysis of relations between trend levels of output and their cyclical fluctuations.

The second is the rigidity assumed in most of the Keynesian growth models together with the assumption of steady-state conditions. In the analyses based on the Cambridge equation, the constancy of the ratio between capital and output is assumed as a condition defining the analysis of growth, and this assumption leads to steady-state paths being adopted as the object of analysis. This assumption is closely connected with the need to address the Harrod's instability, as discussed below, but means at the same time that the trend of the economy is conceived as independent of cyclical fluctuations in output.

The principle of the long-run elasticity of production with respect to changes in the levels of demand on which the classical and Keynesian analysis is based lies at the very root of the rejection of post-Keynesian analyses of growth. According to the authors of the classical and Keynesian approach, the denial of the elasticity of output to demand is incorrect and the role attributed on the basis of it to changes in distribution in the adjustment of savings to investment is to be rejected as misleading.[9]

3.7 Output trends and cyclical fluctuations in the classical and Keynesian approach to the analysis of growth

In the literature based on the classical and Keynesian approach to the analysis of growth, the definition of the theoretical magnitudes involved is discussed no more explicitly than in the other Keynesian approaches. This definition and the relation between the output trend and cyclical fluctuations thus appears controversial.[10]

Here we attempt to consider critically the possibility of assuming that the trend of output is independent of its cyclical fluctuations, as part of this literature does. Some elements of an alternative way of addressing the issue are then put forward. These are of course very much preliminary considerations, which may be regarded as elements for debate and discussion.

Our first step is to put forward what may be considered the most plausible rationalisation of a method based on the independence of the trend of output from its cyclical fluctuations.[11]

With respect to Figure 3.3, we first assume that the aim of the theory of growth is to explain the determination of points like A*. It is therefore necessary to see whether the relation between the actual magnitudes (point A) and the corresponding trend magnitudes (point A*) is such as to permit separate study of the determination of the trend and of divergences from it. As argued, this is the case for the actual/short-period levels of a price $\boldsymbol{p}$ and its theoretical value $\boldsymbol{p^*}$ when the viewpoint of classical political economy (and the original formulations of marginalist analyses) is adopted.

In order to separate the determination of trend, it is necessary to assume the existence of forces causing the economy to grow smoothly. The existence of long-run growth rates of the determining variables must be assumed. Theoretical mechanisms of adjustment permitting corresponding – and almost necessarily uniform – growth of the other variables should then be envisaged. This set of circumstances determines the trend of output.

Actual fluctuations around this trend should then be regarded as due in part to accidental phenomena (either temporary obstacles such as the composition of productive capacity or purely accidental phenomena represented by random variables with zero-mean values). These fluctuations appear to be of the same kind as the divergences of the actual levels of a price $\boldsymbol{p}$ from its theoretical value $\boldsymbol{p^*}$ and can therefore be overlooked completely in analysis of the long-run tendencies of the economy.

Fluctuations in output are generally seen, however, as the joint result of random phenomena and *structural* phenomena inherently endowed with cyclical behaviour.

The determinants of fluctuation patterns have in fact been examined by the theory of business cycles, in which many different theoretical explanations have been put forward. Monetary phenomena, inventory cycles, fixed investment cycles and political cycles are just some of the numerous principles suggested in this broad field of analysis to explain the fact that economies develop through fluctuations.

In order to separate the analysis of cyclical fluctuations from that of the trend of output, it is necessary to argue that the phenomena determining the irregularity of the evolution of output are in any case independent of those determining the trend. This appears to be the conception implicit in most of the large number of multiplier and accelerator models of the 1950s and 1960s and, more generally, in all the Keynesian analyses that adopt a steady-state path to represent the output trend.

The relation between the trend (A*) and cyclical/actual magnitudes (A) is assumed to possess the property identified as essential in the relation between normal and actual prices. The phenomena that determine those normal values on the one hand and trend magnitudes on the other are assumed to be *independent* of the phenomena determining the divergences from their theoretical or long-run counterparts of *actual* magnitudes on the one hand and *cyclical* magnitudes on the other. Though not absolute, the independence of the two kinds of phenomena must be sufficient to justify separate determination of the two sets of causes.

The trend is thus regarded as having independent existence and can be described as an autonomous entity with respect to actual fluctuations.

While this view can probably be challenged in any theoretical context, it does not appear to be sufficiently solid in the one assumed here, which correctly acknowledges the determining role of aggregate demand in the growth process.

3.8 An initial argument in favour of path dependence

When the role of aggregate demand in the growth process is acknowledged, the trend level of produced quantities does not appear to exist independently of the way in which it actually occurs.

As regards normal prices, it may be argued that they are not affected by the actual levels of market prices insofar as they are independent of the causes determining the divergence of actual prices. Longer periods of adjustment and greater differences (5 per cent rather than 1 per cent) between actual and normal prices are not regarded as capable of affecting normal prices, which continue to be determined by the same persistent and structural causes. A storm of two or three nights leading to a longer and more marked divergence of the actual price of fish does not alter the actual conditions of reproduction of fish.

The same reasoning does not seem to hold for the trend level of aggregate produced quantities.

Longer and more intense expansion appears to involve a number of phenomena determining the average rate of growth. Actual levels of aggregate demand that are more or less high for longer or shorter periods determine incomes that give rise to expenditure affecting the average level of aggregate demand. The same phenomena imply investments that affect not only demand but also capacity. These phenomena cannot be considered either transitory or irrelevant for the determination of output, which is the phenomenon addressed here. It therefore appears that longer and more intense expansions and shorter and less intense recessions are the way in which a steeper trend of growth is determined.

Moreover, the pattern of fluctuations affects the trend of the economy by generating phenomena that more or less *persistently* affect the determination of the levels of demand in the following periods. During a more intense or longer expansion, larger amounts of private wealth are accumulated and the incomes of households remain higher for a longer period, during which higher standards of consumption may be irreversibly acquired. A shorter period of expansion would prove insufficient to consolidate such standards. Symmetrically, longer and more intense investment booms imply that new additional productive capacity comes into existence, and this certainly affects both the possibility of producing and the decisions to invest in the following periods. This appears to be radically different from what happens when actual prices are higher to some degree over a more or less long period than their corresponding normal level.

Many phenomena generating *irreversible* effects become crucial when analysis focuses on the determination of output and its evolution. This makes it difficult to assume that a longer or more intense expansion has no effect on the trend of the economy.[12] A more detailed analysis of reality could result in a very long list of these irreversible effects.

Simple transposition of the relation between actual and normal prices into the theory of growth and accumulation does not appear to provide a solid basis for the reconstruction of an analysis correctly recognising the role of aggregate demand in growth.

In this context, the trend of output should probably be regarded simply as concisely representing a phenomenon that occurs through irregular periods of expansion and contraction. In this sense, it is similar to a simple average providing a concise indicator of how a phenomenon manifests itself in a population with different features. Neither the trend of produced quantities nor the statistical average exists independently of the variability through which the phenomena actually occur.

Conversely, this obviously implies that the features of the cycle are not independent of the growth process. They cannot be understood simply as random phenomena of greater or lesser intensity or a more or less intense manifestation of structural phenomena such as the inventory cycle or fixed investment cycle and so on. On unprejudiced observation of reality, they appear as the only way in which higher or lower *trend* rates of growth can occur, i.e., as the manifestation of the very causes of growth.

It should be stressed that the reference to classical political economy makes the classical and Keynesian approach more flexible by comparison with approaches that assume the tendency towards full employment in the long run. In the reappraisal of classical political economy, as stated above, particular importance is attached to the distinction between the method used in the *core* of the surplus theories and the one used in the analysis of relations *outside the core*. This distinction entails no necessary symmetry in the conception of the relation between theoretical and actual magnitudes in contexts of analysis that are actually different and separate.

3.9 The question of capacity adjustment

The transposition of the relation between actual and natural prices into the relation between the trend and cyclical fluctuations of output has also been fostered by another and closely connected question. The recognition of the tendency of productive capacity to adjust to aggregate demand found in most of the Keynesian approaches leads to the adoption of *fully adjusted positions* as the only ones of any relevance for long-run analyses.[13]

Another apparent symmetry is that the tendency of capacity to adjust to demand appears similar to the tendency towards a uniform rate of profit. Just as natural prices can only be associated with conditions in which the forces leading to uniformity of the rate of profit are 'at rest', the theoretical trend of the economy must be characterised by a normal degree of capacity utilisation. Any different degree of utilisation would in fact result in changes determined by the tendency to adjust capacity to demand, which would make a trend characterised by a non-normal degree of utilisation implausible as a representation of the primary forces of the growth process taking place through cyclical fluctuations.

This consideration lies ultimately at the origin of the crucial role played by Harrod's instability in the theory of growth.[14]

In the context of the Keynesian approaches to the analysis of growth, the principle of the tendency of capacity to adjust to demand has played a role similar to the one played by the tendency towards full employment in mainstream theories.

At the cost of drastic simplification, it is possible to state that in mainstream theories, due to the principles entailing the full employment of resources, given the initial amount of productive resources, the propensity to save and the production function (with all the possible changes determined by exogenous or endogenous technological changes), the path of growth around which fluctuations can occur is almost completely determined.

In the Keynesian approaches, the principle of the tendency of capacity to adjust to demand, insofar as it is interpreted as prompting us to associate normal utilisation with the positions representing the trend of output, leads to the conclusion that given the initial endowment of capital stock (productive capacity) and the ratio of saving to income, the path of growth is almost entirely determined. Different theories assert different roles of *aggregate demand* or its components in affecting growth by determining different ratios of investment to income and the corresponding ratio of actual savings to income.[15] The common assumption of normal utilisation of capacity (either constantly or on average over the business cycles) justifies the study of the trend independently of cyclical fluctuations in these approaches.

As is known, however, it has been suggested that the tendency of capacity to adjust to demand does not necessarily mean that normal utilisation is a condition to be presupposed in analysis of the long-run output of the economy.

An initial result in this direction reached by Ciccone (1986) is that the gravitation of market prices towards their normal levels does not necessarily entail full adjustment of capacity to the *level* and composition of effective demand. Normal prices may prevail when produced *output* has the right composition with respect to effective demand but aggregate productive capacity may be not normally utilised, even on the average over the fluctuations.[16] This principle was developed by challenging the validity of the Cambridge theory of distribution. Since then, however, other more general arguments have been developed in favour of the view that the economy does not necessarily tend to attain normal utilisation of capacity even on average over business fluctuations.[17] It would thus follow that it is misleading to study growth by means of paths characterised by normal utilisation even when it is implicitly assumed that this is not a constant condition, as in the steady-state models, but a condition arising on average over cyclical fluctuations, as in the supermultiplier analyses. This method would in fact lead to assuming as general relations between variables that may hold solely in very particular conditions, e.g. when demand just happens to grow constantly and homogeneously at a rate compatible with the normal utilisation of existing capacity.

The arguments against the idea that economies tend towards growth in conditions of normal utilisation include the fact that the process of adjustment of capacity to demand has effects on aggregate demand and its evolution. This adjustment occurs through investments that create both productive capacity and additional aggregate demand. It is thus very hard to argue that the demand effects of investment during this adjustment do not affect the position towards which the economy should converge. As discussed above, it is hard to maintain the necessary property of *independence* of theoretical positions with respect to the fluctuations involved in adjustment.

The process of adjustment of capacity to the evolution of aggregate demand therefore appears to be longer[18] and more uncertain than that of the adjustment of prices to their normal level.

Particular attention should be focused on the analyses based on the classical and Keynesian approach that share with the other Keynesian analyses this view of utilisation and the independence of trends from cyclical fluctuations. The analyses based on the supermultiplier assume that given the initial level of productive capacity, the capital/output coefficient, the propensity to save and the – regular and constant – growth rate of 'autonomous demand', it is possible to select a path of growth that will prove to be the trend around which cyclical fluctuations occur. These analyses regard the rate of growth (which is assumed to be constant) of autonomous demand $\mathbf{g_a}$ as the determining force and the whole process of growth is pegged to it. Capacity, output and all the components of aggregate demand are assumed to grow in accordance with this rate of growth.

The debate on the supermultiplier showed that this rate of growth of autonomous demand must be implicitly regarded as warranted (compatible with the existing productive capacity) and must remain the same forever. Any value of the rate determined independently of the level of existing capacity and/or any change in this (warranted) rate of growth would imply under/over-utilisation and a process of the adjustment of capacity to demand whereby the warranted rate adjusts to the autonomously determined (or altered) rate $\mathbf{g_a}$.[19]

It has also been pointed out just how essential the notion of 'relative persistence' is in this context.[20] In the process of gravitation of prices, it is possible to assume that the conditions determining normal prices are comparatively persistent with respect to the speed of the adjustment process. This is another feature of the relation between theoretical and actual magnitudes in the theory of prices that cannot be attributed to the relation between the trend and actual fluctuations of produced quantities. As noted above, the adjustment of capacity to aggregate demand appears to be so long, uncertain and complex that it becomes necessary, in order to argue in favour of a complete adjustment, to assume that the rate of growth $\mathbf{g_a}$ remains constant for an implausibly long period.

3.10 The intrinsic irregularity of the evolution of final demand components

A further point to be taken into consideration is that the view of the relation between trend and cyclical fluctuations under discussion here hinges on the possibility of assuming the existence of rates of growth of autonomous demand components that are themselves independent of the actual/cyclical levels of output. This is another assumption that appears far from plausible.

Examination of the evolution of the components of this final demand reveals that, by their very nature, they display an irregular pattern over time. For most of them, it is difficult to imagine a regular long-run rate of growth around which mainly random fluctuations occur. The asymmetric behaviour of this magnitude has been noted in studies on the expansion of aggregate consumption and taken as a crucial feature determining the evolution of the part of consumption that cannot be regarded as directly dependent on current income.[21] An analysis of government expenditure would certainly lead to the identification of cyclical and anti-cyclical components together with components for which a regular trend could be postulated. Different considerations may plausibly lead to similar conclusions as regards the components of aggregate demand related to international trade. The components of autonomous demand connected with the introduction of technical innovation would require yet another set of considerations, which could lead in any direction. Even though the analysis in this direction is far from being sufficiently developed, when all these components are considered without presupposing a methodology based on the assumption of a long-run constant rate of growth (necessarily uniform with the rates of growth of all the other magnitudes), it seems hard to argue that they tend to grow at a regular rate around which random fluctuations occur and thus generate a trend that can be studied independently of cyclical fluctuations. A growth rate of each of these magnitudes that proves to be more or less high on average over a period of time seems to be determined by longer and more intense expansions and shorter and less intense contractions of the corresponding magnitude.

3.11 The need to reconsider Harrod's instability

The critical considerations about the obstacles to full adjustment seem to make the assumption of an economic trend independent of the actual positions unnecessary and even misleading. They also clash, however, with the pressing need to focus on the theoretical 'adjusted' positions deriving from the idea of Harrod's instability. It is therefore necessary to examine the logical origins of this concept in order to evaluate its actual solidity. While this issue cannot be discussed here in detail for reasons of space, some relevant points can be recalled.

The concept rests on the assumption that any divergence between actual and desired utilisation generates an adjustment of capacity. It has been pointed out, however, that there may be different ways in which over- or under-utilisation can occur without creating any need to adjust capacity.[22] Moreover, it is possible to argue that the desired degree of utilisation can be affected by the same circumstances that generate increases in the actual utilisation;[23] developing this argument it is possible to argue that any divergence between the two tends to be much less relevant in determining instability.

This notion of instability is based on a conception of capital as qualitatively homogeneous and continuously divisible. This makes it possible to conceive the tendency of capacity to adjust to demand through a process of growth or decline of this homogeneous material, i.e., by increases or decreases in the rate of capital accumulation. Unprejudiced observation of reality would show that an adjustment of capacity to demand through an increase or decrease in the size of existing plants is an extremely rare phenomenon. Capacity generally adjusts through the radical destruction of entire plants and/or firms or the creation of new plants and/or firms. An individual plant may be under- or over-utilised all through its economic life and then disappear without its under- or over-utilisation ever being corrected by decreasing or increasing its size.

For these reasons it is possible to imagine under- or over-utilisation that lasts for many years (and even increases) and is corrected during a recession without making any difference – or only a limited difference – to the trend rate of accumulation.

Our scepticism about this idea of the adjustment of capacity to demand is strengthened when technical change is taken into account. The inadequacy of given productive capacity with respect to the demand for which it was installed may manifest itself when the same capacity is economically obsolete. In this case too, the inadequacy of capacity would not be corrected by an increase or decrease in size but by the creation of new plants.

The assumption of a general tendency of productive capacity to be determined by aggregate demand evolution does not ultimately appear to entail actual adjustment realisation (even on average over fluctuations). Theoretical output trends characterised by degrees of utilisation differing from normal on average and implying increasing or decreasing capacity utilisation over a period of time thus appear to be much less inconsistent than the concept of Harrod's instability makes them appear.

3.12 Conclusions: the outline of a possible alternative method of analysis

When the classical and Keynesian approach to the analysis of growth is assumed, the validity of grounding the analysis on paths of growth determined independently of actual positions of the economy seems to be challenged by the considerations developed here. This seems to be true despite the broad consensus on the use of this method. In the classical and Keynesian approach, this consensus may be based essentially on two elements: a) the transposition of the relation between theoretical and actual magnitudes assumed in the theory of prices and distribution into the analysis of the evolution of the quantities over time; and b) the symmetry between the tendency towards a uniform rate of profit in the theory of price and the tendency of productive capacity to adjust to the evolution of aggregate demand.

It has been argued that these elements are not sufficiently solid. In particular, in the context of demand-led growth, it does not seem to be possible to attribute the theoretical trend magnitudes with the property of independence from the actual magnitudes occurring in fluctuations, a property of independence similar to that attached to normal prices with respect to actual ones in classical theories of prices and distribution. The adjustment of capacity to demand has been shown to be not only longer and more complex and uncertain than the tendency towards uniformity of profits but also insufficient to guarantee the independence of the theoretical magnitude from the path of adjustment.

Let us now outline an alternative position emerging from these criticisms. Economic growth is a phenomenon consisting in the evolution of produced quantities that is likely to occur in an irregular manner, i.e., through fluctuations with different patterns. The study of the trend or of any average path of growth may not be separated from analysis of the way in which the relevant variables fluctuate in generating the trend, which is generally the object of theories. In this context, the trend has no *existence* independent of the actual irregular evolution of the variables. It is therefore a concise representation of a more complex occurrence. Growth theory should therefore investigate the change in variables and explain why, in a growing economy, output levels tend during expansion to reach levels higher than those achieved in the previous expansion and/or why recessions are interrupted at levels of output higher that those achieved in previous troughs. Symmetrically, growth theory should be able to explain why this does not occur and the opposite can even take place when the economy is not growing or is in decline.[24] Among other things, the actual evolution of produced quantities determines the actual rate of utilisation of productive capacity.[25]

Notes

1. Classical political economy and marginalist theories adopted different terms for the same concepts, partially as a reflection of differences in theoretical principles. We choose not to repeat or clarify the different definitions of the terminology used here – natural vs. normal or short-run vs. actual magnitudes – as this would only make the exposition more cumbersome without improving the completeness or solidity of the arguments presented.
2. This assumption is needed because the occurrence of exchange at other than equilibrium prices would entail a change in actual endowments and hence in the data determining equilibrium positions. See Bilancini and Petri (2008).
3. Though paradoxical, the approach of regarding a very long and heterogeneous series of time intervals as a single long period has actually been adopted in the literature more than once. It can be detected in some influential works of the 1950s, when Friedman (1957) and Modigliani and Brumberg (1954) studied the long-run propensity to save on the basis of Kuznets's data. It can also be seen in Kaldor (1957), who attempted to extract some *stylised facts* from studies of economic history to serve as driving features for the theory of growth.
4. A linear trend has been extracted through the linear interpolation of actual data. It is worth stressing that exponential interpolation should be used in order to study paths characterised by constant rates of growth.
5. References for these exceptions are given in footnotes 8 and 24.
6. This is another automatic transposition of a feature of the analyses of prices, where the relations between the variables determined by the theory – between normal prices and rate of profits or between quantities and prices – are in fact assumed as those prevailing in the long run. These relations are independent of the relationships between the same variables that occur transitorily during fluctuations of prices and rates of return.

7. This approach was originally sketched out in Garegnani (1962) (a work for Svimez, an Italian research centre for the development of southern Italy) and then developed in Garegnani (1978–9 and 1992). Various authors have since continued along the path opened up by Garegnani.
8. It is worth stressing that the assumption of the independence of the trend from actual positions of the economy has been also questioned in the *marginalist* debate on the business cycle. Particular significance attaches in this connection to Haberler (1938), esp. pp. 31–53. An important exception to the identification of full employment with the assumption of a trend determined independently of cyclical fluctuations can be found in the more recent versions of full-employment theories known as Real Business Cycle theories, where cycle and trend are studied as related phenomena both determined by stochastic shocks that affect full-employment equilibria. Even though the rigidity of output with respect to changes in demand is then asserted, this does not lead to the assumption of a trend independent of cyclical fluctuations. See Bronzi (2012).
9. See Ciccone (1986), Garegnani (1992), Vianello (1985). For the particular way in which the issue of steady state and the elasticity of output with respect to changes in aggregate demand is present in the neo-Kaleckian literature, see Trezzini (2011b).
10. The debate on the supermultiplier can be interpreted as the result of implicitly different conceptions of this issue. No explicit statement of the point is to be found, however, on either side.
11. As pointed out above, in the classical and Keynesian literature as well as most of the analyses of growth, it is virtually impossible to find an explicit treatment of the meaning of the trend of output. It is therefore unfair to attach this rationalisation of the method used directly to supermultiplier analyses. This chapter may also help to make explicit any other implicit rationalisation of the method by the authors involved.
12. The existence of irreversibility certainly matters when the viewpoint of demand-led growth is assumed and is not demonstrably irrelevant in supply-side analyses of growth. See n. 8 for traces of this problem in marginalist theories of the cycle.
13. *Fully adjusted positions* are to be understood as those in which the tendency of capacity to adjust to demand has been fully realised and the degree of capacity utilisation proves normal.
14. Once any divergence of the actual rate of growth from the warranted rate (of actual utilisation from normal) is seen as generating cumulative expansion or deflation, a path characterised by normal capacity utilisation appears the only one that can be taken as the theoretical trend of an economy. All the theories of growth have in fact attempted since Harrod's contribution to develop mechanisms capable of making the warranted rate of growth equal to the rate determined by the specific theory. In any case, however, the coinciding of the actual and warranted rates of growth is seen as a necessary condition. The exception is the case of the neo-Kaleckian models briefly discussed above.
15. At the cost of the simplification involved in any synthesis of different theories with important elements in common, we can say in post-Keynesian theories, a different rate of accumulation directly affects the ratio of actual

saving to income by altering the distribution of income. In neo-Kaleckian theories, different rates of accumulation directly affect the long-run degree of utilisation and thus determine differences in profits and savings. In the analysis based on the supermultiplier, finally, different rates of growth of autonomous demand indirectly affect the share of the amount of saving corresponding to normal capacity utilisation (capacity saving) absorbed by autonomous demand and directly affect the share of it capable of financing investment and capacity growth.

16. The composition of output may be attained with productive capacity adjusted in composition but not in level or in conditions in which new capacity has the right composition but works together with old capacity that produces output by means of obsolete techniques and under- or over-utilisation, thus generating 'quasi-rents'. Competition stops working when the rate of profit is uniform over the flows of new investment and not necessarily over the whole stock of capital.
17. Garegnani (1992, p. 58) argued in this sense as follows: 'What becomes clear in this connection is something which was perhaps not immediately evident, namely that the possibility that investment should generate a corresponding amount of savings through changes in aggregate productive capacity (the Second Keynesian Position) is inconsistent with the assumption of an economy working over time at the desired level of capacity utilisation, even only on an average taken over booms and slumps.' While this implication is evident in the extremely simplified case studied by Garegnani in which there is no autonomous component of aggregate demand, it is *only apparently* overcome – but does not completely disappear – when autonomous demand is assumed to exist (see Trezzini, 1995).
18. The length referred to is a theoretical one. What matters is the persistence of the determining forces (the data) with respect to the speed of the adjustment of the actual magnitudes to the theoretical ones. In the case of prices, the relevant adjustment of capacity to demand is limited to the new capacity created by investment flows. In the case of quantities produced, it is the whole stock of capital that has to change in composition and level. In this sense it not only appears longer in absolute terms but also is longer in relation to the persistence attributable to the determining forces: the rate of growth of accumulation or of autonomous demand. These determining forces do not appear to be persistent with respect to the length of the former adjustment, not least because they are not independent of the process of adjustment itself.
19. Trezzini (1995).
20. Trezzini (1998).
21. See Garegnani and Trezzini (2010) and Trezzini (2011a).
22. Trezzini (2011b).
23. Garegnani (1992) p. 56 and White (2006) pp. 156–58.
24. This 'path dependence' is essentially the same as that theorised in the 1950s and 1960s by authors working in a Keynesian framework (see Kaldor, 1951; Kalecki, 1968; Smithies, 1957; and Goodwin, 1967). The approach was essentially abandoned by the Keynesian tradition, unfortunately, and its revival is proving hard to achieve. It has been used more recently in some contributions based on the classical and Keynesian approach (see Garegnani and Trezzini, 2010).

25. These considerations do not necessarily mean that it is completely useless to construct models based on the assumption both of the independence of the trend of output with respect to its actual evolution and of the attainment on average of normal utilisation of capacity. These models must, however, be seen for what they are: abstractions that may help clarify some potential properties of a growing economy but involve numerous relations between variables that can neither represent nor fully explain those that occur in reality. On the other hand, this limited role is precisely what must be attributed to any type of *model* in the classical and Keynesian approach to the analysis of growth.

References

Bilancini, E. and Petri, F. (2008) 'The dynamic of general equilibrium: a comment on Professor Gintis', *Quaderno del Dipartimento di Economia Politica*, no. 538, Università di Siena.

Bortis, H. (1997) *Institutions, Behaviour and Economic Theory – A Contribution to Classical-Keynesian Political Economy* (Cambridge: Cambridge University Press).

Bronzi, M. (2012) 'Il nesso di causalità tra ciclo economico e trend: le teorie del ciclo reale e le teorie della domanda', Doctoral thesis in political economy, Università di Roma Tre.

Ciccone, R. (1986) 'Accumulation and capacity utilization: some critical considerations on Joan Robinson's Theory of Distribution', *Political Economy*, 1: 17–36.

Friedman, M. (1957) *A Theory of the Consumption Function* (Princeton: Princeton University Press).

Garegnani, P. (1962) *Il problema della domanda effettiva nello sviluppo economico italiano* Roma: Svimez.

Garegnani, P. (1978–79) 'Notes on consumption, investment and effective demand: I and II', *Cambridge Journal of Economics*, 2(4) and 3(1): 335–53, 62–82.

Garegnani, P. (1983) 'Two Routes to Effective Demand', in J. A. Kregel (ed.), *Distribution, Effective Demand and International Economic Relations* (London: Macmillan): pp. 69–80.

Garegnani, P. (1984) 'Value and Distribution in the Classical Economists and Marx', in *Oxford Economic Papers*, 36(2): 291–325.

Garegnani, P. (1992) 'Some Notes for an Analysis of Accumulation', in J. Halevi, D. Laibman and E. Nell (eds), *Beyond the Steady State: A Revival of Growth Theory* (Basingstoke and London: MacMillan).

Garegnani, P. and Palumbo, A. (1998) 'Accumulation of Capital', in N. Salvadori and H. D. Kurz (eds), *The Elgar Companion to Classical Economics* (Aldershot: Edward Elgar).

Garegnani, P. and Trezzini, A. (2010) 'Cycles and growth: a source of demand-driven endogenous growth', *Review of Political Economy*, 22(1): 119–25.

Goodwin, R. M. (1967) 'A Growth Cycle', in C. H. Feinstein (ed.), *Socialism, Capitalism and Economic Growth* (Cambridge: Cambridge University Press).

Haberler, G. (1938) *Prosperity and Depression: A Theoretical Analysis*, League of Nations, Geneva.

Harrod, R. F. (1939) 'An Essay in Dynamic Theory', *Economic Journal*, 49(193): 14–33.

Kaldor, N. (1951) 'Mr. Hicks on the Trade Cycle', *Economic Journal*, LXI, December: 833–47.

Kaldor, N. (1954) 'The relation of economic growth and cyclical fluctuations', *Economic Journal*, 64: 53–71.

Kaldor, N. (1957) 'A model of economic growth', *Economic Journal*, 67(268): 591–624.

Kalecki, M. (1968) 'Trend and the business cycle', *Economic Journal*, 78: 263–76.

Keynes, J. M. (1936) *The General Theory of Employment, Interest and Money* (London: MacMillan).

Modigliani, F. and Brumberg, R. (1954) 'Utility Analysis and the Consumption Function: An Interpretation of Cross-Section Data', in K.K. Kurihara (ed.), *Post-Keynesian Economics* (London: Allen & Unwin).

Palumbo, A. and Trezzini, A. (2003) 'Growth without normal capacity utilization', *European Journal of the History of Economic Thought*, 10: 109–35.

Serrano, F. (1995) 'Long-period effective demand and the Sraffian supermultiplier', *Contributions to Political Economy*, 14: 67–90.

Smithies, A. (1957) 'Economic fluctuations and growth', *Econometrica*, 25(1): 1–52.

Trezzini, A. (1995) 'Capacity utilization in the long run and the autonomous components of aggregate demand', *Contributions to Political Economy*, 14: 33–66.

Trezzini, A. (1998) 'Capacity utilization in the long run: some further considerations', *Contributions to Political Economy*, 17: 53–67.

Trezzini, A. (2011a) 'The irreversibility of consumption as a source of endogenous demand-driven economic growth', *Review of Political Economy*, 23(4), 537–56.

Trezzini, A. (2011b) 'Steady State and the Analysis of Long-Run Tendencies: The Case of Neo-Kaleckian Models', in R. Ciccone, C. Gehrke and G. Mongiovi (eds), *Sraffa and Modern Economics*, Vol. II (London and New York: Routledge): pp. 129–51.

Vianello, F. (1985) 'The Pace of Accumulation', *Political Economy*, 1(1): 69–87.

White, G. (2006) 'Demand-led growth and the Classical Approach to Value and Distribution: Are they Compatible?', in N. Salvadori (ed.), *Economic Growth and Distribution: On the Nature and Cause of the Wealth of Nations* (Aldershot: Edward Elgar): pp. 148–78.

4
Potential Output and Demand-Led Growth

*Antonella Palumbo**

4.1 Introduction

Potential output is generally understood as the desirable level of output, i.e., the highest which may be attained in any given situation without putting inflationary pressures upon the economy.

The political and practical relevance of the notion can hardly be overestimated. Potential output is frequently and increasingly used in the empirical analyses and policy prescriptions of such institutions as the International Monetary Fund (IMF) and the Organisation for Economic Cooperation and Development (OECD). Time series of the variable are reconstructed, both to evaluate the economy's performance and to project a potential growth path over a number of future periods,[1] against which the possible inflationary effects of policies are evaluated. It is used in the construction of derived indicators, such as the 'output gap' (the percentage deviation of actual from potential output, frequently used both in estimating inflation and for determination of the proper monetary policy);[2] and the 'cyclically adjusted budget balance' (a measure of the fiscal policy stance which is used in policy contexts such as the EU fiscal surveillance framework).[3]

Despite the frequent mention and use of potential output and related concepts, however, empirical estimation of the notions seems beset by difficulties and uncertainties ('a tricky business', according to Van Ark – see Gros *et al.* 2010, p.1).

---

*I wish to thank Anna Carabelli, Andrea Ginzburg and an anonymous referee for helpful comments, while retaining all responsibility for any remaining mistakes.

This chapter aims to examine the theoretical foundations of empirical research on potential output, especially that currently produced by international economic institutions and central banks. The purpose is twofold: first, to show that the notion of potential output changes significantly according to the theoretical framework within which it is defined, and how deeply empirical estimation methods are influenced by this. Second, it is maintained here that the current mainstream empirical literature, which starts from the theoretical presumption that the long-run path of potential output is supply-determined, provides empirical estimates which have generally little content apart from the representation of the theoretical belief that potential output must be a strong attractor of actual output, so that the unobservable variable is assumed (but not proved) to be some sort of moving average of actual output.

The related theoretical belief that there is a well-defined relationship between output gaps and changes in inflation, and the difficulty of detecting any systematic output–inflation relationship in actual data, is at the root of the variability and uncertainty of the estimates and of their poor performance as predictors of inflation.

Against these fundamental limitations, the notion of potential output will be analysed within the demand-led growth approach. The latter will here be defined in its broadest terms as that conception of growth according to which no automatic mechanisms ensure the tendency to full utilisation of resources, in the long as well as in the short run, so that the path of actual output is demand-constrained and determines in turn the evolution of potential output over time. This has relevant implications for empirical research. It will be maintained that not only is measurable potential output to be considered as a short-period notion, but no regular connection should be established or searched for between the output gap and inflation.

We start from the mainstream analysis of potential output, based on the idea of a permanent tendency of actual production to gravitate towards potential production, in accordance with the theories that describe growth as supply-constrained. Particular attention will be paid, respectively, to the notion of potential output in natural-rate and NAIRU models (Section 4.2) and in real business cycle and dynamic stochastic general equilibrium models (Section 4.3). Section 4.4 provides an overview of the results of current empirical estimates, almost exclusively based on supply-constrained growth models, and focuses on their limitations. In search of an alternative notion of potential output, Section 4.5 will be devoted to a brief analysis of the estimation method originally proposed by A. Okun in the 1960s, in which potential output was defined and measured on the basis of the Keynesian theory of effective demand, while Section 4.6 addresses the conception of potential output in the theories of demand-led growth and its implications for empirical research. Section 4.7 concludes.

4.2 Potential output in natural-rate and NAIRU models

The theoretical and empirical analyses on potential output which are currently conducted by central banks and international institutions[4] are largely based on the idea that potential output is determined by supply factors – the evolution of resources and productivity over time – while actual production, which can be affected by the state of demand and short-term shocks, tends to gravitate permanently towards it. This derives from the prevailing theoretical conception of the working of the economic system, which, based on the equilibrating mechanisms of neoclassical theory, regards growth as a supply-side phenomenon with no permanent limits posed by demand to the production of the economy's full potential.

Within this general view, however, different theories and models imply different specific notions of potential output and give rise to different methods for its estimation.

As a first fundamental qualification, the potential utilisation of labour is usually defined as implying natural unemployment, that is, according to Friedman's definition, the level of unemployment which is compatible with constant inflation. The behaviour of inflation on the two sides of the non-inflationary unemployment rate is regarded as symmetrical, with accelerating or decelerating inflation (and the related changes in expectations) playing the role of adjusting the system towards the natural (equilibrium) rate. Thus, monetarist and new classical models imply a strict relation between output and inflation. Demand shocks (especially monetary shocks) are regarded as causing deviations of actual from potential output, which may be more or less durable according to the rapidity of error correction in expectations, but are, however, of a temporary nature. Actual output fluctuates around the potential path, the latter depending exclusively on long-run (real) supply determinants.

The New Keynesian NAIRU models (Carlin and Soskice, 1990; Layard *et al.*, 1991) have partially different theoretical bases, relying on the hypothesis of imperfect competition or nominal and real rigidities which prevent the system from attaining efficient results. The NAIRU (Non-Accelerating Inflation Rate of Unemployment), though implying involuntary unemployment, represents equilibrium because it is the only rate of unemployment at which inflation is constant (the competing claims of workers and firms being such as to exactly exhaust the product). Spontaneous mechanisms based on accelerating or decelerating inflation[5] ensure the tendency of actual unemployment to the NAIRU. Involuntary unemployment at the NAIRU is not caused by insufficiencies of aggregate demand, but by the imperfections and market power characterising the labour (and product) market. It cannot therefore be corrected by means of demand policies, but rather through labour reforms or other competition-enhancing measures. Demand may cause deviations of the actual unemployment rate from the NAIRU.

In what follows, we will refer to the general type of NAIRU models, in which the NAIRU is an attractor of the actual unemployment rate, neglecting the possible hysteresis effects. In much of this literature hysteresis is in fact regarded at best as a medium-term phenomenon with the system converging in the end to the unique long-term NAIRU (see Layard *et al.*, 1991); see however below (Section 4.4.1) for discussion of the case of full hysteresis.

Notwithstanding their different theoretical bases, natural-rate models and NAIRU models have very similar empirical implications. First, in both types of models the equilibrium unemployment rate – whether natural rate or NAIRU – is regarded as an attractor of the actual rate, and potential output as an attractor of actual output.

Second, both approaches postulate a clear distinction between the long-run determinants of trend (long-run supply factors and labour market institutions) and the determinants of short-run fluctuations, which depend on demand shocks (policies included) and other factors, including temporary supply shocks (ECB, 2011, p. 80). Adjusting mechanisms ensure that the deviations induced by demand are only temporary[6] (see below for qualifications of this statement in the case of NAIRU models with full hysteresis).

Third, in both approaches a definite relation is postulated between the sign of output gaps and changes in inflation. Actual output being above potential necessarily implies accelerating inflation, while inflation must decelerate for negative output gaps.

On these theoretical bases, three distinct classes of methods can be identified for the empirical estimation of potential output (for a classification of methods, see Cotis *et al.*, 2005; Laxton and Tetlow, 1992):

a) On the basis of the assumed gravitation of actual output towards potential, it is possible to consider *trend output*, i.e., the growth path of actual output over time neglecting short-period fluctuations, as an empirical approximation of potential output. This defines the class of 'statistical methods' for potential output estimation, which are based on the extraction of the trend from the time series of actual output by means of various statistical techniques (see Section 4.3).

b) The theoretical presumption that the trend of potential output is exclusively determined by long-run supply factors and is totally independent of demand variables gives rise to an attempt to directly estimate the growth of such factors over time. This defines the class of 'economic' methods for potential output estimation, the most widely used of which is the *production function approach* (Giorno *et al.*, 1995; De Masi, 1997; Denis *et al.*, 2002; Billmeier, 2004; CBO, 2004). It consists in estimating, through specific hypotheses, the growth in time of the potential capital and labour inputs (the latter depending in turn on the calculation of the equilibrium rate of unemployment) and of a 'potential total factor productivity' supposedly representing the pace of technological innovation; the growth in time of potential output is then estimated by applying a production function to these hypothetical potential factors (see Section 4.4.3).

c) The supposed theoretical relationship between output gaps and inflation variations gives rise to an attempt to derive the estimates of potential output indirectly from the series of the inflation rate, by specifying a definite quantitative relationship between changes in the inflation rate and output gaps, i.e., a Phillips curve.[7] This produces the so-called 'multivariate' statistical methods for the estimation of potential output (to be distinguished from the univariate statistical methods seen above), which derive estimates of potential output from the time series of actual output but with corrections deriving from the behaviour of the inflation rate. The multivariate methods thus represent in fact a mixture between the statistical methods and the economic methods.

The output–inflation relationship is also essential for estimation of the NAIRU in the production function approach.

The essential characteristics of the notion of potential output in this literature can be summarised as follows:

- Potential output tends to be identified with the average level of actual output, so that output gaps should prove tendentially symmetrical.
- The behaviour of inflation is symmetrical on the two sides of potential output.
- The long-run path of potential output is exogenous and can be defined when knowing the laws of motion of resources and productivity.

As regards this latter point, it is worth noting that the possibility of projecting the growth path of potential output over time also depends on the fact that the majority of empirical analyses of this kind represent the economy by means of simple aggregative models with an extremely simplified definition of technical progress.[8]

4.3 Potential output in RBC and DSGE models

A different definition of potential output is to be found in real business cycle (RBC) models and gives rise to partially different estimation methods, or at least to different interpretations of the estimated values.

According to the RBC approach (Kydland and Prescott, 1982; Prescott, 1986), the economy's cyclical fluctuations are produced by the agents' optimal reaction to random real shocks. Trend and cycle have the same economic determinants and depend exclusively on supply factors, while money disturbances have negligible effects. Actual production thus always coincides with potential production: observed fluctuations in output should not be interpreted as *output gaps,* but as the fluctuating path of equilibrium itself. As for policy implications, 'costly efforts at stabilization are likely to be counterproductive. Economic fluctuations are optimal responses to uncertainty in the rate of technological change.' (Prescott, 1986, p. 21)

While according to this approach potential output should be regarded as the same thing as actual output, it is possible to obtain an estimate of *trend output* by applying a statistical filter to the time series of actual output. Actually, the technique of statistical filtering for estimating trend output was originally proposed by RBC authors (see especially Hodrick and Prescott, 1981), and has been largely utilised since.

Statistical filtering requires a (necessarily arbitrary) hypothesis of what has to be defined as *cycle* and what as *trend* in a time series. Different hypotheses give rise to different results: the time path of trend output could even be represented by a single number, expressing an average rate of growth over the whole estimation period. However, the Hodrick-Prescott filter (like other filters)[9] does generally assume that the trend itself varies in time and has a stochastic component.

The Hodrick-Prescott filter consists in the minimization of both deviations of actual output from potential and deviations of the rate of growth of potential output from a regular trend:

$$Min\left\{\sum_{t=1}^{T}(\ln Y_t - \ln Y_t^*)^2 + \lambda[(\ln Y_{t+1}^* - \ln Y_t^*) - (\ln Y_t^* - \ln Y_{t-1}^*)]^2\right\}$$

The estimation results are strictly dependent on the arbitrary value assigned to the parameter λ, determining the variability in time of the estimated series and its greater or lesser adherence to the series of actual output (Billmeier, 2004, pp. 18–19).

While in RBC models statistical filtering is regarded simply as defining the *trend* component of the actual time path of output, in natural-rate and NAIRU models the same method of estimation is regarded as approximately identifying the potential output path.

One relevant policy implication of RBC models is the virtual absence of any effects whatsoever and any possible scope for monetary policy. This extreme conclusion has been partly modified within a different class of models, the dynamic stochastic general equilibrium (DSGE) models, which stem from RBC models in assuming an intertemporal general equilibrium framework as the basis for modelling the economy. As in RBC models, fluctuations are described in terms of the optimal responses of agents to random shocks and a structure of lags implies a mechanism of propagation whereby stochastic disturbances have durable effects. However, in the so-called New Keynesian dynamic stochastic general equilibrium (NK-DSGE) models, it is assumed that nominal or real rigidities and market inefficiencies result in transitory deviations of actual output from potential output, implying scope for the active role of monetary policy in correcting output gaps (Woodford, 2003).[10]

In empirical evaluation and for policy purposes, policymakers need in the first place 'to assess the degree to which fluctuations in observed output reflect the optimal response to shocks that hit the economy, versus undesirable fluctuations' (Basu and Fernald, 2009, p. 2). Optimal fluctuations are defined as the fluctuations that would occur in case of full flexibility of prices and wages and zero mark-ups. In NK-DSGE analyses, a model is usually built representing the whole economy and then used to estimate, through simulation exercises, the following three notions of potential output (Vetlov *et al.*, 2011):

a) *Trend output*, which represents the time path of the economy net of business cycle fluctuations, the latter being partly induced by deviations of potential output itself with respect to a regular steady growth path. In fact, the identification of long-period trend does not amount, in these models, to estimating potential output but only its long-period component.
b) *Efficient output*, which is the path that output would follow if goods and labour markets were perfectly competitive (that is to say, if prices and wages were fully flexible and mark-ups were zero). Deviations of

actual from efficient output measure the losses of welfare associated with imperfect competition and nominal rigidities.

c) *Natural output*, which is the path that output would follow under flexible prices and wages but imperfectly competitive markets (so, differently from the efficient output, with different-from-zero mark-ups and mark-up shocks). Deviations of actual from natural output measure the relevance of nominal rigidities.

Efficient and natural output can deviate from actual output because the latter, unlike the former two, is determined by assuming sticky prices and/or wages. Policy should aim at identifying and correcting deviations from natural output, for the purpose of controlling inflation, and deviations from efficient output, which imply a loss of welfare, while optimal fluctuations should not be corrected but rather accommodated (Basu and Fernald, 2009; Galì *et al.*, 2011). The results of some empirical analysis carried on along these lines will be described in Section 4.4.4 below.

What is worth remarking is that the absolute majority of empirical estimates of potential output currently performed by central banks and international economic institutions are based on the kind of models which have been just described in this and the previous section and are thus based on the theoretical presumption that only supply factors matter in the determination of potential output.[11] As will be shown in the next section, this presumption deeply influences the results of estimation.

4.4 Problems and puzzles in the empirical estimates of potential output

In reviewing the results of empirical literature it is worth starting from the attempts at estimating the NAIRU, given the role that it plays also in estimation of potential output.

4.4.1 Chasing the NAIRU

There are numerous attempts in the literature to give empirical content to the notions of natural rate and NAIRU. It is worth noting that in applied research the two concepts are frequently identified with each other, notwithstanding the theoretical differences: empirically, they are both defined as the rate of unemployment that is compatible with constant inflation and to which actual unemployment tends (Gordon, 1989, p. 220; Cross, 1995).

It does not seem too far-fetched to maintain that empirical literature has generally failed to identify such equilibrium rate of unemployment

(Jenkinson, 1988). The first attempts assumed that the unobservable natural rate or NAIRU was a constant; the impossibility of detecting any such constant when analysing real economies has given rise to the notion of *time-varying NAIRU* (Gordon, 1997).[12]

Such difficulties became particularly evident in the course of attempts to explain the European experience of the 1980s and 1990s, when unemployment persistently remained much higher than in the preceding decades. This has created a dilemma: either the NAIRU tends to be quite stable, but in this case one should explain why wide and persistent deviations of the actual unemployment rate from the supposed equilibrium value do not tend to be corrected; or the NAIRU itself is very variable, but one should observe in reality some of the phenomena which according to theory justify such variability and especially its permanent changes of level (Pesaran and Smith, 1995, p. 219). The increase in actual average unemployment in the 1980s and 1990s in most European countries seems rather to have been accompanied by an increase and not a decrease in labour flexibility due to policy measures (see Blanchard and Katz, 1996, trying to justify the variability of the NAIRU, and the critical remarks in Solow, 2000, p. 5; Stirati, 2001, p. 428). Equally difficult to explain are the relevant differences in average unemployment among the different European economies (Gordon, 1989, p. 220).

Instead of identifying a single equilibrium rate of unemployment, data seem rather to show that in many cases different unemployment rates are compatible with stable inflation (Schreiber and Wolters, 2007; Karanassou *et al.*, 2003; Franz, 2005, p. 23; IMF, 2006, ch. 3).

Within the NAIRU models, one theoretical response to these empirical difficulties has been the notion of hysteresis, i.e. the supposed influence of actual unemployment rates on the equilibrium one (Blanchard and Summers, 1987; Jenkinson, 1988). This is justified in various ways, for example by assuming that high unemployment raises the share of long-term unemployment, determining loss of efficiency and competitiveness in part of the labour force and thus their inability to influence the bargained real wage (Layard *et al.*, 1991, pp. 74–5).

Hysteresis introduces a series of problems and contradictions in NAIRU models and opens up the possibility of completely overturning the very basic idea of the existence of an equilibrium unemployment rate. An equilibrium variable which is heavily influenced by the actual one is in fact a rather weak attractor of actual values, and thus an almost irrelevant theoretical concept. As for policy, full hysteresis, by subverting the relationship between actual and potential output, implies that demand policies may make the potential growth path shift permanently (Ball, 1996).

In the majority of NAIRU models, however, these radical conclusions are avoided by assuming that hysteresis is a short/medium-period phenomenon and that in the long period the system converges to a unique rate (Layard *et al.*,1991; Carlin and Soskice, 1990).

4.4.2 Estimates of the time-varying NAIRU

NAIRU is currently estimated – like potential output – by means of two different types of method, the statistical and the economic, or by some mixture of the two (Richardson *et al.*, 2000; Boone, 2000; De Masi, 1997; Billmeier, 2004; Denis *et al.*, 2002; Fabiani and Mestre, 2001). Statistical methods amount to defining the NAIRU as the trend component of the actual unemployment rate, while economic methods try to infer the NAIRU from equations representing economic relations, and particularly the Phillips curve.

The most widely used statistical methods are the Hodrick-Prescott filter, based on the double minimization both of unemployment gaps and of deviations of NAIRU from its regular growth path, and the Kalman filter, in which the actual series of unemployment rates is decomposed in a trend component, a cyclical component and error. Univariate filters produce estimates which automatically imply the tendency of actual unemployment rates to the NAIRU, but as a result the estimated NAIRU has no connection with the actual behaviour of inflation. Moreover, they have very scarce predictive powers as regards the value of the unemployment rate.

Multivariate statistical filters are based on the idea that univariate filters' estimates must be corrected by means of theory-based relationships (they are thus a combination of statistical and economic methods). The estimated NAIRU must be such that the estimated unemployment gaps are correlated with inflation changes. Such filters, however, are based on the assumption that deviations from estimated NAIRU are symmetrical over a sufficiently long period, thus amounting to calculation of a moving average of the series of actual unemployment rates, in which measures are taken to prevent estimated unemployment gaps from moving in a direction opposite to that predicted by theory.

Economic methods may use either structural-form equations or reduced-form equations to estimate the NAIRU. The former, which are usually based on a wage equation and a price equation (according to the specification proposed by Layard *et al.*, 1991) and try to infer the NAIRU from the equilibrium solution, should in principle provide a better (in the sense of more theory-based) estimate (Richardson *et al.*, 2000, p. 34). In practice they are scarcely used, because of both theoretical uncertainties surrounding the correct specification of the equations, and the instability of the resulting estimates and their lack of robustness due to the high number of unobservable parameters (Franz, 2005). A further problem has to do with the difficulty of estimating such variables as the desired wage, so that mainly estimation models based on reduced-form equations are used in applied research.

The latter usually consist of an expectation-augmented Phillips curve, where current inflation is explained by past inflation (approximating expected inflation on the assumption that inflation is a random walk – Layard *et al.*, 1991, p. 15), the unemployment gap ($u_t - u_t^*$) and other variables representing supply shocks (Gordon, 1997):

$$\pi_t = \pi_{t-1} + \alpha_1(u_t - u_t^*) + \alpha_2 z_t + e_t$$

where u_t^* is the unknown NAIRU, z_t summarises the supply shocks (z_t is normalised so as to have zero mean, since supply shocks are symmetrical by construction), while e_t is a white noise. The equation, which is typically more complex due to various lags (z_t may also be a vector of variables), is estimated by making use of two different filters, the multivariate Hodrick-Prescott filter or the multivariate Kalman filter.

The multivariate HP filter (Laxton and Tetlow, 1992) implies imposing the condition that the NAIRU is the trend component of the actual unemployment rate, according to:

$$u_t = u_t^* + \varepsilon_t$$

where ε_t is a white noise process; while the NAIRU itself follows a random walk:

$$u_t^* = u_{t-1}^* + v_t$$

where v_t is a white noise. With these restrictions, this 'economic' method of estimating the NAIRU is perfectly coincident with the application of a statistical multivariate filter to the series of actual unemployment. The quantity to minimize is:

$$Min\sum[(u_t - u_t^*)^2 + \lambda_1(u_t^* - u_{t-1}^*)^2 + \lambda_2 e_t^2]$$

(where e_t is the residual of the Phillips curve). With this 'economic' restriction the resulting NAIRU, though being by construction the moving average of the actual unemployment rate, is much more variable than the one obtained through univariate filters (Richardson *et al.*, 2000). Resulting estimates are influenced by the two weighting parameters λ_1 e λ_2: in particular, the higher the second, the higher the weight given to inflation data and the more variable the resulting NAIRU.

The multivariate Kalman filter estimates are obtained in the hypothesis that the trend (the NAIRU) is a random walk process with or without drift:

$$u_t^* = \mu_t + u_{t-1}^* + \varepsilon_t$$

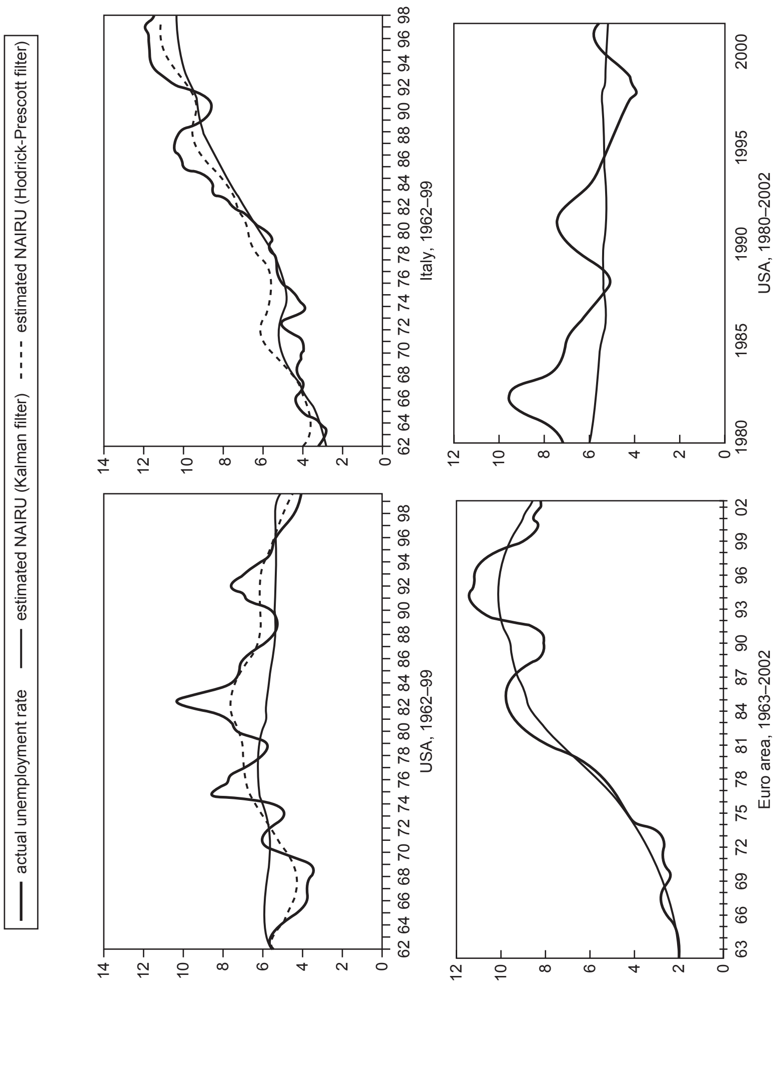
(the μ_t drift may either be absent, or deterministic, or a random walk process itself), with no economic restrictions on the trend; while the cyclical component is constrained by a Phillips curve-type relation. The Kalman filter is an iterative process in which the estimated NAIRU is progressively corrected on the basis on new information on the observed variable (see Richardson *et al.*, 2000, p. 42–3; Boone, 2000, p. 5; Denis *et al.*, 2002, p. 9ff. for details).

The whole procedure thus consists in attributing all the variability in actual unemployment rates which cannot be explained on the basis of inflation data to supposed changes in the unobservable NAIRU (apart from casual errors). Since the resulting estimates may show excessive variability, they are further corrected with smoothing procedures (Gordon, 1997; Richardson *et al.*, 2000, p. 44; Fabiani and Mestre, 2001, p. 11). As noted by Franz (2005, p. 17), 'If there is no limit on the ability of the NAIRU to fluctuate each time period, the time-varying NAIRU may jump up and down and soak up all the residual variation in the Phillips curve'.

In some of the estimation models, but not in all, further restrictions are imposed on the cyclical component to explicitly constrain it to have zero mean (Denis *et al.*, 2002, p. 10; Fabiani and Mestre, 2001, p. 10). Even if there is no such explicit hypothesis, however, the estimated NAIRU is built so as to absorb all the changes in the average level of the actual unemployment rate, automatically attributing them to supposed changes in the supply factors determining the NAIRU (Boone, 2000, p. 11; Franz, 2005, p. 18).

Figure 4.1, which reproduces some results of the estimates from different sources, confirms that the series of the estimated NAIRU always proves to be the trend component (though differently defined in different estimation models) of the actual unemployment rate series.[13]

A series of problematic aspects of NAIRU estimates are stressed in the literature: estimates are uncertain and not robust, i.e., change drastically


actual unemployment rate
estimated NAIRU (Kalman filter)
estimated NAIRU (Hodrick-Prescott filter)
USA, 1962–99
Italy, 1962–99
Euro area, 1963–2002
USA, 1980–2002

for a small change in specification (Staiger *et al.*, 1997, p. 34; CBO, 2004; Gordon, 1997, p. 21 and 24); most importantly from a theoretical point of view, they seem to be very inefficient in predicting inflation (Staiger *et al.*, 1997; Jenkinson, 1988). As seen above, inflation actually plays a relatively unimportant role in building NAIRU estimates, especially after smoothing procedures are applied. The resulting correlation between unemployment gaps and changes in inflation is indeed very low (Billmeier, 2004).

The dominant role is played instead by the actual unemployment rate: as noted by Gordon (1997, p. 28), 'fluctuations in the NAIRU seem too large to be plausible and seem mainly to mimic movements in the actual unemployment rate'. According to Galbraith (1997, p. 101), 'In general, the estimated NAIRU in a variety of studies has tracked the actual unemployment rate sluggishly. When unemployment rises, analysts tend to discover that the demographic characteristics of workers are deteriorating, or that the job-wage and wage-price dynamic has become unstable. And then the unemployment rate drifts down again, those flaws mysteriously begin to disappear, and a lower NAIRU is estimated.'[14]

4.4.3 Estimates of potential output: the production function approach

While the estimates of potential output which are obtained through statistical filters, both univariate and multivariate, are by definition the trend component of actual output, the 'production function approach' (see Section 4.2) should be based on economic relations and should identify the supply factors allegedly determining potential output. Thus in principle actual output should not automatically gravitate around the estimated potential, unless the theoretical presumption of gravitation were empirically confirmed.

In this approach, potential output is estimated through the following Cobb–Douglas production function with constant returns, applied to the time series of potential inputs:

$$Y_t^* = (N_t^*)^{\alpha} (K_t)^{1-\alpha} TFP_t^*,$$

Potential capital is identified with actual; potential labour input N_t^* is estimated by applying a statistical filter to the actual time series of labour force, so as to obtain the 'trend labour force' LF^*, and by correcting such trend labour force for the estimated NAIRU

$$N_t^* = LF_t^* (1 - NAIRU_t);$$

while 'potential' total factor productivity is obtained by applying a statistical filter to the series of the total factor productivity (TFP) which is estimated through application of a Cobb-Douglas of the same specification to actual data over the same period. Elasticity α is the observed average value of the wage share.

For what has just been seen as regards the estimates of the NAIRU, potential labour is in fact the trend component of actual employment, while potential TFP is the trend component of TFP (the residual) estimated on actual data.

Thus, despite the complication of the procedure, what the production function approach indirectly calculates is in fact a sort of moving average of the time series of actual output. Like estimates obtained through statistical filtering, these estimates are, disappointingly, very slightly correlated to inflation changes (Billmeier, 2009; ECB, 2011, p. 82–85). The attempt to define output gaps as more correlated to inflation gives rise to volatility of the estimated series of potential output, which then has to be smoothed.

It can be concluded that the tendential symmetry of output gaps which can be observed in these estimates is in reality built into the hypotheses,[15] which makes the estimation of potential output through the production function approach not significantly different from statistical filtering. Once again, the empirical content of the estimates is merely the theoretical presumption (but not the empirical proof) that actual output must necessarily gravitate, on average, around potential.

4.4.4 Estimates of potential output within NK-DSGE models

Attempts at estimating potential output within the NK-DSGE framework are still preliminary and do not offer a body of consolidated results (Vetlov *et al.*, 2011, p. 22).

As seen above, estimation in these models proceeds through simulation exercises once the 'correct' models for describing the economy have been identified. This implies 'calibrating' a general stochastic equilibrium model on the economy (thus finding, through econometric techniques, the value of the parameters which make the model mimic the true economy's fluctuations as closely as possible) and defining on this basis the nature of shocks which have hit the economy in the sample period (Basu and Fernald, 2009, p. 23). One way of obtaining the various measures of potential is to simulate the path of the economy in the counterfactual hypotheses of full flexibility of prices and wages and: a) presence of the low-frequency persistent shocks only (trend output); b) presence also of the high-frequency efficient shocks (efficient output); c) presence also of shocks on wage and prices mark-ups (natural output). Actual output may differ from all three measures due to price and wage stickiness, which, however, is regarded as a short-period phenomenon. Thus, with this estimation procedure, since the path that the economy has actually followed is theoretically defined as optimal (except for short-term fluctuations), the three definitions of potential output are in fact three different ways of smoothing the fluctuating path of actual output, and all three kinds of output gaps tend to have zero mean if the whole period is taken into consideration. Efficient and natural output are much more volatile than trend output for the fact of embodying a greater number of shocks in the definition of potential output, and thus account for a greater part of observed output variability; while trend output estimates prove to be very similar to those obtained through statistical filters (Vetlov *et al.*, 2011, p. 14).

This result may partly change depending on the hypotheses of the model and its specification. Galí and Gertler (2007, p. 30), for example, maintain that actual output gravitates around a natural inefficient level, due to the presence of mark-ups (caused by imperfect competition) systematically causing a loss of output and welfare if compared with the theoretical optimal path. This would imply an estimated efficient output systematically higher than natural and non-symmetrical output gaps: demand policy, however, would prove useless in closing them, while only supply-side policies aimed at enhancing competition would be effective.

In NK-DSGE models monetary policy is required to close the gap between actual and natural output, which should empirically prove to be highly correlated with inflation variations. However, the results of estimation in this respect are at best uncertain and mixed: besides being very volatile, natural output gives rise to output gaps whose correlation with inflation is often quite low (Vetlov *et al.*, 2011, p. 5–6).

Results of simulations are not robust and may vary widely for small changes in specification (Sims *et al.*, 2010, p. 211). The nature of shocks cannot be unequivocally defined empirically, making some arbitrariness in interpretation inevitable. Moreover, the basic model has been complicated with the introduction of many frictions, which makes the interpretation of the relationships among variables rather complex. The only reason to pursue these kind of empirical exercises rather than the traditional ones would be, according to Vetlov *et al.* (2011, p. 7), their allegedly superior theoretical consistency.

4.4.5 Estimates of potential output and the crisis

It may be concluded that the estimation of potential output according to its supply-side definition in general produces disappointing results.

The main theoretical hypothesis, i.e., that the sign and size of the output gap should be related to inflation, does not show in the data. If the estimated output gaps are constrained to be compatible with observed inflation changes, the resulting series is extremely volatile (the estimated potential oscillates wildly), which seems to contrast with the idea of potential output as a comparatively smooth series. On the other hand, estimated potential output is often nothing more than the statistical trend of actual output, or some elaborated variant of it. Its correlation with inflation changes is disappointingly low, which leads some authors to conclude that other factors, apart from the output gap, have become more relevant in influencing inflation (oil prices, prices of imports in general, expectations of inflation – see ECB, 2011, p. 85; Billmeier, 2009, p. 402).

One consequence of the way potential output is usually estimated shows in the effect on it of the deep contraction that has afflicted advanced economies in recent years. Although the recession was initially regarded mainly as a pronounced deviation from the trend of potential output, after some time it produced downward revisions of the potential itself (Gros *et al.*, 2010; ECB, 2009, p. 45; 2011, p. 73). This result, which is obviously implied in the very way estimates are built, is somehow puzzling for the mainstream theoretical conception of potential output. If the crisis is interpreted as a deep adverse demand shock, according to the theories of supply-side growth it should not be considered as capable of affecting the potential output. Explanations thus vary from regarding the whole change which has been detected in the potential level as something which will probably not affect the 'longer-term growth rate of potential output' if the economy reacts flexibly enough (ECB, 2011, p. 73), to looking for changes in supply-side determinants. These can be due either to low levels of investment in research and development, or to reduction of the labour force due to discouragement or similar effects (European Commission, 2009; ECB, 2011, p. 77).

In the end, the very content of the empirical estimates of potential output has forced applied research to admit the influence of actual on potential output, which was probably unexpected in principle.

We will now turn our attention to the theories of demand-led growth, in which the causation between actual and potential output is reverted. Although these theories have not given rise to a systematic body of applied research on the estimation of potential output, the insights they offer on the working of the system may serve to interpret some of the results and puzzles of mainstream empirical literature on the matter. As a premise, we briefly deal with the analysis of potential output proposed by A. Okun in the 1960s, which represents the first systematic attempt, along Keynesian lines, to estimate potential output for policy purposes.

4.5 Okun's definition of potential output

On the basis of macroeconomic Keynesian theory, Okun (1962; see also Council of Economic Advisors, 1962) regards potential output as the maximum attainable level of output in the short period, with given equipment and capacity. In giving empirical content to the notion, he assumes, however, that unemployment cannot be compressed below a certain threshold without causing unsustainable inflationary pressures, and arbitrarily (but following what he regarded as the consensus view of the US economy) puts such threshold at 4 per cent (Okun, 1962, p. 1). The 'unemployment gap' is then calculated as the difference between the 4 per cent non-inflationary unemployment rate, regarded by Okun as a lower bound, and the actual one. It must be noted that no specific hypothesis is introduced on the behaviour of inflation at unemployment rates above 4 per cent. In other words, 4 per cent is not regarded as the turning point which separates inflation from deflation (or accelerating inflation from decelerating inflation).[16]

A second element which is used in Okun's estimates is the so-called 'Okun's Law', establishing an empirical proportion (3:1, by Okun's own calculations) between changes in actual real output (real Gross National Product) and changes in unemployment.[17] The actual real GNP of a period in which unemployment has been 4 per cent may thus be defined as the initialising level for the series of potential output, which is then constructed by assuming that each percentage point unemployment gap implies a 3 per cent gap between potential and actual output. This allows both the level of potential output and its rates of growth over the estimation period to be defined. Some smoothing procedure is then used on the resulting estimated series in order to reduce its variability.

The results of Okun's definitions and hypotheses are shown in Figure 4.2, which is derived from his original data.

It is worth stressing the differences between Okun's estimates and the estimation methods which subsequently prevailed and are now used. First, while in natural-rate models potential output is a long-run concept, Okun (1962, p. 2) maintains that it can only be defined in the short period, on the basis of given technical conditions and given plants: 'to the extent', he maintains, 'that low utilization rates and accompanying low profits and personal incomes hold down investment in plant,

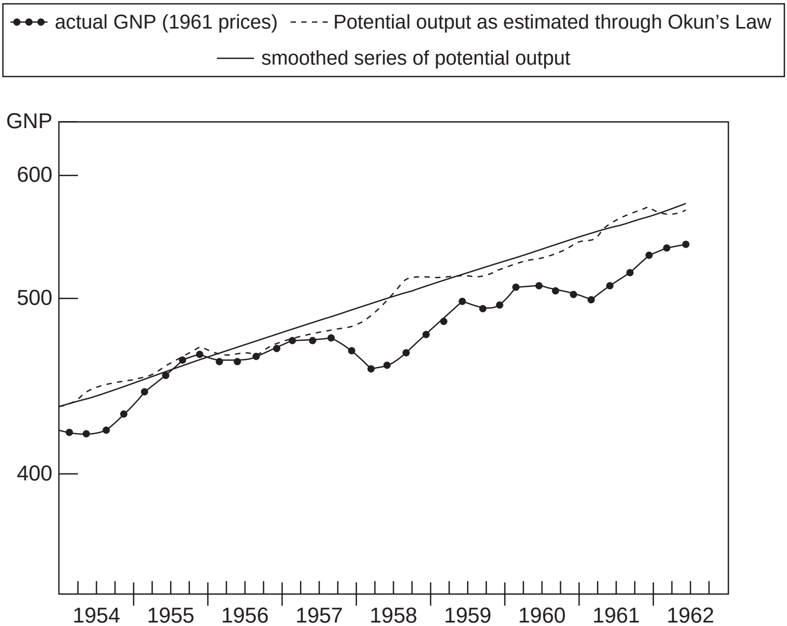


Figure 4.2 Estimates of potential output in Okun's definition, USA 1954–62
Source: Data from Okun (1962).

equipment, research, housing, and education, the growth of potential GNP will be retarded'. According to Okun, 'today's actual output influences tomorrow's productive capacity'.

Second, apart from exceptional periods of boom, actual output is systematically lower than the estimated potential; output gaps tend consequently to be mainly negative and non-symmetrical. Unlike in natural-rate and NAIRU models, potential output is theoretically characterised as the ceiling and not the average level of actual production. This implies that in Okun's view the non-inflationary rate of unemployment is not the *equilibrium* rate of unemployment.

Third, apart from the definition of the non-inflationary rate, Okun's procedure requires no use of inflation data in potential output estimation, thus implying no necessarily strict correlation between output gaps and inflation.

Okun's Law has also been used for estimation of potential output in later works. Even if no additional hypothesis is introduced on the level of potential output for initialising the series, it is possible to estimate a series of rates of change of potential output exclusively on the basis of Okun's Law (an exogenous definition of the potential rate of unemployment is, however, necessary). Such exercises are proposed for example by DeLong (2002) and Solow (2000).

4.6 Potential output in demand-led growth theories

Once set in the long period and taken as the basis for a theory of growth, the Keynesian principle of effective demand can be characterised as implying independence of investment from saving and the absence of any spontaneous tendency of the system to grow along a path with full utilisation of resources.[18]

However, some sort of mechanism must guarantee that the growth paths of actual output and potential output do not totally diverge over time. Although adjustment need be neither instantaneous nor complete, even in long periods, an adjustment mechanism must exist and this is represented by the endogeneity in the formation of resources.

Persistently under-utilised resources tend slowly to disappear – fixed capital shrinks if net investment is negative, labour may emigrate or participation rates may fall – while the pressure of high demand may, on the contrary, accelerate their formation (with the same phenomena of opposite sign; see Kaldor, 1985, pp. 35–6). Thus it may be maintained in the first place that for the principle of demand-led growth to be true it is not necessary to observe in reality wide and persistent gaps between actual production and full-capacity production (the latter being a variable quantity which can be slowly influenced by the former). At the same time, this does not imply that there are small margins of elasticity for production and capacity to adjust to demand: on the contrary, such margins are even wider because not only can the utilisation of existing resources vary, but they can be created or destroyed altogether.

Following Garegnani (1992), the loss of potential over a number of periods, due to under-utilisation of capacity in one period, is represented not only by the current under-utilisation, but also by the capacity which failed to materialise and to produce yet higher capacity. This implies that such loss is mostly invisible: borrowing an expression from Landes (1990, p. 6), we may say that 'the gap between what is and what can be is enormous'.

Equally wide are the margins for *increase* in the size of capacity, given the definition of normal utilisation as lower than technical full utilisation, and thus the possibility of over-utilisation leading to the installation of new capacity.

It may thus be maintained that it is the endogeneity of potential output and its dependence on actual that characterise demand-led growth theories. While output gaps would measure in each period the distance between actual output and capacity output, over a succession of periods they would not be sufficient to measure the distance between what has been actually realised and what could have been realised: they may at most represent the *visible* waste of capacity.

As regards inflation, no necessary and strict relation is postulated between output gaps and price changes. In the first place, the definition of normal utilisation of resources contains no reference to inflation, so that, while it can be maintained that in general prices rise more quickly on approaching full utilisation, no general and mechanic rule can be established. Inflation will be influenced by the distributive conflict (Stirati, 2001), the intensity of which need not bear any definite relation to the level of activity. As regards the long run, once admitting that high demand may produce pressures on capacity only for a limited period, before inducing the endogenous creation of new capacity, the conclusion likely follows that no permanent pressure can be exerted on costs and prices by demand (Serrano, 2006). It would thus be illegitimate to infer the measure of potential output from the actual behaviour of inflation.

4.6.1 Studying the long-run path of potential output in the demand-led growth approach

A matter of debate, within the demand-led growth approach, is whether the long-run trend of the system can be studied independently of its short-run behaviour. This has relevant implications for the possibility of defining potential output as a long-run or an exclusively short-run concept, or, more accurately, the possibility of projecting a long-run potential output path starting from an arbitrary point in time.

Some theories and models within the demand-led growth approach assume that the mechanism of capacity adjustment described above operates in such a way as to allow full adjustment between demand and capacity to be considered representative of the actual growth path of the economy (at least on average), so that the long-period relations between variables may be studied with exclusive reference to such situations of full adjustment. This is the case not only for steady-state growth models,[19] but also for the so-called 'supermultiplier' models (Serrano, 1995) in which the system's long-period growth is determined by the growth in time of autonomous demand, while investment is regarded as completely induced by the capacity adjustment mechanism. In principle, these models allow a long-period path of potential output to be defined, at each point in time, on the basis of the exogenous growth of autonomous demand, in case a very simple form of technical progress is assumed.

However, the very definition of growth as demand-determined implies that in any specific period demand may change autonomously and prove different from the previously installed capacity. Though capacity is continually built or destroyed on the basis of the strength of expected demand, normal utilisation of installed capacity is not guaranteed. The absence of automatic mechanisms for adjusting demand to capacity reproduces in each period the possibility of different-from-normal utilisation, while at the same time duration of fixed capital implies that the adjustment of capacity to demand may act slowly (Ciccone, 1986; Palumbo and Trezzini, 2003). In each period in which capacity is given, differences between actual output and capacity output may well occur and need not be symmetrical over any relevant observation period.

This implies the impossibility of defining exogenously a growth path of potential output. At each point in time, given installed capacity, it is always possible to define a hypothetical growth path that the economy would follow if, *assuming no technical change*, investment decisions allowed normal utilisation of existing capacity and if future demand were such as to allow normal utilisation of all capacity installed in each subsequent period. Such a 'normal utilisation path', however, would by no means represent the path towards which actual output gravitates, even neglecting innovation. Any under-utilisation would in fact determine a displacement of the whole path. Nor would the normal utilisation path even represent the highest possible growth path, given the upward flexibility of output and capacity.

In other words, at each point in time potential growth is in fact represented by a plurality of possible growth paths, even neglecting the role of technical progress. The path which the economy actually follows identifies the range of future growth possibilities (as noted by Setterfield, 2002, p. 5, this approach necessarily requires the notion of 'path dependence' in order to analyse long-run tendencies).

We may conclude that in the demand-led growth approach potential output may be a meaningful, quantifiable notion only with reference to the short period, in which it can be defined on the basis of installed capacity, while the long-period path of potential growth is an undetermined notion from an *ex ante* point of view. No exogenous path can be identified representing the long-run tendencies of the economy: trend and business cycle must be analysed together, with the latter influencing and determining the former.

4.7 Conclusions: policy implications

Although their uncertain and problematic nature is often recognised, empirical estimates of potential output play an extremely relevant role in shaping and binding policies. As has been shown, in the supply-side framework the actual content of such empirical estimates is essentially the *ex post* record of the economy's average actual growth path, which has led analysts to somehow recognise the adverse effect of the recent great crisis on potential output.

What matters are policy implications. Since in the supply-side framework demand policies may at best be aimed at stabilising the cycle, while playing no role whatsoever in the determination of potential, it follows that a lower estimated level of potential output, and thus a smaller negative output gap, implies less need and less scope for expansionary demand policies (ECB, 2011, p. 82). As maintained in the Economic Synopses of the Federal Reserve Bank of St Louis (2012, no 11, p. 2), 'the gap may be closing faster than we thought because potential GDP is lower than we thought. And, if potential GDP is lower than expected, then interest rates may have to rise sooner than expected to prevent an acceleration of inflation.'

As maintained by Solow (2000, p. 10), this *ex post* definition of potential output, whereby what the economy has actually realised is assumed to be not too far from potential, seems ultimately to respond only to the theoretical preconception according to which expansionary demand policies are normally to be avoided for their inflationary consequences, while all that is needed are a flexible labour market and structural reforms (ECB, 2011, p. 85).

The endogeneity of potential output and potential growth which are proper to the demand-led growth approach have entirely different policy implications. In the first place, if output is liable to be demand-constrained both in the short and the long run and the economy is devoid of self-regulating mechanisms, demand policies may prove useful not only for short-run stabilisation but also for growth. In the second place, the path dependence of the growth trajectories and the plurality of potential output paths at each moment in time imply wide scope for all kinds of industrial and capacity-building policies aimed at shaping and guiding the economy's growth along a socially chosen path.

In principle it is possible, in the demand-led growth context, to identify an estimation method of potential output which may provide a measure of the economy's performance and a guide for policy action.

Like Okun's method, it should aim at quantifying the 'visible' difference between what can be produced and what is actually produced in each short period with given capacity, while both theory and experience show the role of actual realisations in shaping the economy's future possibilities and so the impossibility of quantifying even approximately the growth of potential in time as an *ex ante* notion.

Unlike Okun's method, it should avoid any reference to inflation data in the estimation of potential output, if only in the definition of the single non-inflationary rate of unemployment as the potential one. In addition to the failure of mainstream models to establish a clear empirical relation between output and inflation, historical experience shows a variety of output–inflation relationships, such as phases of increasing unemployment at constant or even increasing inflation, decreasing inflation at persistently high unemployment rates, seemingly horizontal Phillips curves in some specific sub-periods (see IMF, 2006), and so on. It seems we must conclude that no general and systematic relationship is observable in reality,[20] and take this into account when attempting to quantify potential output.

Notes

1. See for example the Congressional Budget Office's annual *Update to the Budget and Economic Outlook* which contains 10-year projections for the US economy.
2. Reference is to the so-called 'Taylor rule', used extensively both in New Keynesian models and as a policy procedure by a number of central banks.
3. See the European Commission's bi-annual publication 'Cyclical Adjustment of Budget Balances'.
4. See, for some instances, the OECD's *Economic Outlook*, the IMF's *World Economic Outlook*, the ECB's *Monthly Bulletin*.
5. Increasing (decreasing) inflation is assumed to have contractionary (expansionary) effects on aggregate demand through the Keynes effect and the Pigou effect, provided it is not accommodated by monetary policy.
6. 'Business cycle effects and autonomous demand shocks of various kinds should wash out if we take a long enough period' (Nickell, 1997, p. 71–2).
7. In coherence with the hypotheses of the natural rate models, the Phillips curve used for estimation is assumed to be vertical in the long run while admitting only a short period trade-off between unemployment and inflation.
8. Technical progress is often exogenous and usually defined merely in terms of the growing ability of the economy to produce increasing quantities of the same single commodity with no attention paid to qualitative changes.
9. Some examples are the Baxter-King (1995) filter, which identifies high-frequency, medium-frequency and low-frequency changes in the time series;

the Beveridge-Nelson (1981) filter, which assumes a particular correlation between cycle and trend determinants; the Kalman filter based on decomposition of the series in trend, cycle and erratic component.

10. Woodford states that this class of models responds to the need for central banks to have a theoretical basis in justification and support of their action.
11. As already noted, this statement must be qualified for NAIRU models with full hysteresis. See below, Section 4.4.1.
12. 'When Milton Friedman first proposed the natural rate hypothesis ..., it sounded like royal edict had established the natural rate as another one of the universe's invariant constants. Today, there is general recognition that if a NAIRU exists, it must be changing over time' (Stiglitz, 1997, pp. 5–6).
13. The change in the level of Italian unemployment in the 1980s and 1990s compared to the previous period, for example, is entirely attributed to an exogenous change in the NAIRU.
14. Still more drastically, Franz (2005, p. 30) maintains: 'To put it differently, yet taking the risk of an oversimplification: What has been estimated with a highly sophisticated machinery is simply the trend unemployment rate. If so, this can be carried out much easier and within a few minutes, just by using an HP filter.'
15. 'Descriptively, the mean of the gap measure should be close to zero over longer time horizons' (Billmeier, 2009, p. 396). It must be further noted that such institutions as the IMF and the OECD usually revise the estimates according to specific knowledge of the economies considered: 'In all cases, estimates of potential output incorporate a substantial amount of judgement and country-specific expertise of desk officers' (De Masi, 1997, p. 10). 'Such measures may be qualified, sometimes heavily, by the judgement of country specialists' (Giorno *et al.*, 1995, p. 168).
16. It is thus a very different notion from the NAIRU. According to Tobin (1995, p. 39), 'in Keynes open-end inflation results from an "inflationary gap" in aggregate demand, while a "deflationary gap" leads to comparative stability of prices or price trends'.
17. The original formulation of Okun's Law is based on observation on US data of the period 1947–60. Okun (1962) explains the empirical relationship on the idea that increases in working hours, productivity and participation accompany the expansionary phases, while the reverse occurs in slumps.
18. For a survey of different theories within the demand-led growth approach see Setterfield (2002).
19. The absolute majority of demand-led growth models, among which those of the Cambridge tradition, are cast in terms of steady-state growth paths. In the Kaleckian models (see Blecker in Setterfield, 2002), however, the steady state is deemed compatible with a constant but non-normal degree of utilisation of capacity, thus implicitly contradicting the tendency to capacity adjustment.
20. Contrary to the widespread interpretation of the Phillips curve, it can be maintained that also in Phillips's (1958) original article the relationship between money wage variations and unemployment is described not as a definite function, but as a broad relationship which can change due to various different influences. See Palumbo (2010).

References

Ball, L. (1996) 'Disinflation and the NAIRU', *NBER Working Papers* 5520.

Basu, S. and Fernald, J. G. (2009) 'What Do We Know and Not Know about Potential Output?', *Federal Reserve Bank of San Francisco Working Paper* 2009-05, March.

Baxter, M. and King, R.G. (1995) 'Measuring Business Cycles Approximate Band-pass Filters for Economic Time Series', *NBER Working Papers* 5022.

Beveridge, S. and Nelson, C. R. (1981) 'A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the "business cycle"', *Journal of Monetary Economics*, 7: 151–74.

Billmeier, A. (2004) 'Ghostbusting: Which Output Gap Measure Really Matters?', IMF Working Papers 146, updated version in (2009), *International Economics and Economic Policy*, 6: 391–419.

Blanchard, O. and Katz, L. F. (1996) 'What We Know and do Not Know about the Natural Rate of Unemployment', *NBER Working Papers* 5822.

Blanchard, O. and Summers, L. (1987) 'Hysteresis and the European Unemployment Problem', *NBER Working Papers* W1950.

Boone, L. (2000) 'Comparing Semi-structural Methods to estimate Unobserved Variables', *OECD Economics Department Working Papers* 240.

Carlin, W. and Soskice, D. (1990) *Macroeconomics and the Wage Bargain* (Oxford: Oxford University Press).

Ciccone, R. (1986) 'Accumulation and capacity utilisation: some critical considerations on Joan Robinson's Theory of Distribution', *Political Economy*, 2: 17–36.

Congressional Budget Office (2004) 'A Summary of Alternative Methods for Estimating Potential GDP', *Background Paper*, March, http://www.cbo.gov/.

Cotis, J. P., Elmeskov, J. and Mourougane, A. (2005) 'Estimates of Potential Output: Benefits and Pitfalls from a Policy Prospect', in L. Reichlin (ed.), *The Euro Zone's Business Cycle: Stylised Facts and Measurement Issues*, CEPR.

Council of Economic Advisors (1962) *Economic Report of the President* (Washington: US Government Printing Office).

Cross, R. (ed.) (1995) *The Natural Rate of Unemployment* (Cambridge: Cambridge University Press).

De Masi, P. R. (1997) 'IMF Estimates of Potential Output: Theory and Practice', *IMF Working Papers* 177.

DeLong, B. (2002) 'Forecasting growth and unemployment', July, http://www.j-bradford-delong.net/

Denis, C., McMorrow, K. and Röger, W. (2002) 'Production Function Approach to Calculating Potential Growth and Output Gaps – Estimates for the EU Member States and the US', *European Commission, Directorate-General for Economic and Financial Affairs, Economic Papers* 176.

European Central Bank (2009) *Monthly Bulletin*, July.

European Central Bank (2011) *Monthly Bulletin*, November.

European Commission (2009) 'Impact of the Current Economic and Financial Crisis on Potential Output', *European Economy Occasional Papers* 49, June.

Fabiani, S. and Mestre, R. (2001) 'A System Approach for Measuring the Euro Area NAIRU', *ECB Working Papers* 65.

Franz, W. (2005) 'Will the (German) NAIRU Please Stand Up?', *Zentrum für Europäische Wirtschaftsforschung* (*Centre for European Economic Research*), Discussion Paper n. 03-35.
Galbraith, J. K. (1997) 'Time to Ditch the NAIRU', *Journal of Economic Perspectives*, 11: 93–108.
Galí, J. and Gertler, M. (2007) 'Macroeconomic Modeling for Monetary Policy' *Journal of Economic Perspectives*, 21: 25–46
Galí, J., Smets, F. and Wouters, R. (2011) 'Unemployment in an Estimated New Keynesian Model', *NBER Working Papers* 17084, National Bureau of Economic Research.
Garegnani, P. (1992) 'Some Notes for an Analysis of Accumulation', in J. Halevi, D. Laibman and E. Nell (eds), *Beyond the Steady-state* (Basingstoke and London: Macmillan).
Giorno, C., Richardson, P., Roseveare, D. and van den Noord, P. (1995) 'Potential Output, Output Gaps and Structural Budget Balances', *OECD Economic Studies* 24.
Gordon, R. J. (1989) 'Hysteresis in history: was there ever a Phillips curve?', *American Economic Review*, 79: 220–5.
Gordon, R .J. (1997) 'The time-varying NAIRU and its implications for economic policy', *Journal of Economic Perspectives*, 11: 11–32.
Gros, D., Alcidi, C., Biggs, M., Mayer, T., Van Ark, B. (2010) 'Forum: The Impact of the Financial Crisis on the Real Economy'; *Intereconomics*, 45(1), (January/February): 4–20.
Hodrick, R. J. and Prescott, E. C. (1981) 'Post-war U.S. Business Cycles: An Empirical Investigation', *Northwestern University Discussion Paper* 451, re-published in *Journal of Money, Credit and Banking*, 29 (1997): 1–16.
International Monetary Fund (2006) *World Economic Outlook*, April.
Jenkinson, T. J. (1988) 'The NAIRU: Statistical Fact or Theoretical Straitjacket?', in R. Cross (ed.), *Unemployment, Hysteresis and the Natural Rate Hypothesis* (Oxford: Blackwell).
Kaldor, N. (1985) *Economics without Equilibrium* (Armonk, NY: Sharpe).
Karanassou, M., Sala, H. and Snower, D. (2003) 'The European Phillips curve: Does the NAIRU exist?', *Applied Economics Quarterly*, 49(2): 93–121.
Kydland, F.. and Prescott, E. (1982) 'Time to build and aggregate fluctuations', E*conometrica*, 50: 1345–1370.
Landes, D. (1990) 'Why are we so rich and they so poor?', *American Economic Review*, 80: 1–13.
Laxton, D. and Tetlow, R. (1992) 'A simple multivariate filter for the measurement of potential output', *Technical Report* 59, Bank of Canada.
Layard, R., Nickell, S. and Jackman, R. (1991) *Unemployment, Macroeconomic Performance and the Labour Market* (Oxford: Oxford University Press).
Nickell, S. (1997) 'Unemployment and labor market rigidities: Europe versus North America', *Journal of Economic Perspectives*, 11: 55–74.
Okun, A. (1962) 'Potential GNP: its measurement and significance', *Cowles Foundation Paper* 190.
Palumbo, A. and Trezzini, A. (2003) 'Growth without normal capacity utilisation', *European Journal of the History of Economic Thought*, 10: 109–35.
Palumbo, A. (2010) 'La teoria del salario monetario di Phillips e la curva di Phillips', in G. Bonifati and A. Simonazzi, *Il ritorno dell'economia politica. Saggi in ricordo di Fernando Vianello* (Roma: Donzelli): pp. 79–102.

Pesaran, H. and Smith, R. (1995) 'The Natural Rate Hypothesis and its Testable Implications', in R. Cross (ed.), *The Natural Rate of Unemployment: Reflections on 25 years of the Hypothesis* (Cambridge: Cambridge University Press): pp. 203–30.

Prescott, E. (1986) 'Theory Ahead of Business Cycle Measurement', *Federal Reserve Bank of Minneapolis Quarterly Review*, Fall.

Richardson, P., Boone, L., Giorno, C., Meacci, M., Rae, D. and Turner, D. (2000) 'The Concept, Policy Use and Measurement of Structural Unemployment: Estimating a Time-Varying NAIRU across 21 OECD Countries', *OECD Economics Department Working Papers* 250.

Schreiber, S. and Wolters, J. (2007) 'The long-run Phillips curve revisited: is the NAIRU framework data-consistent?', *Journal of Macroeconomics*, 29: 355–67.

Serrano, F. (1995) 'Long-period effective demand and the Sraffian supermultiplier', *Contributions to Political Economy*, 14: 67–90.

Serrano, F. (2006) 'Mind the Gap: Hysteresis, Inflation Dynamics and the Sraffian Supermultiplier', Universidade Federal de Rio de Janeiro, mimeo.

Setterfield, M. (ed.) (2002) *The Economics of Demand-led Growth* (Cheltenham: Edward Elgar).

Sims, C. A., Edge, R.M., Gürkaynak, R. S. and Reis, R. (2010) 'How Useful Are Estimated DSGE Model Forecasts for Central Bankers?' *Brookings Papers on Economic Activity*, Fall, 209–59.

Solow, R. M. (1986) 'Unemployment: getting the questions right', *Economica*, 53, suppl., S23–S34.

Solow, R. M. (2000) 'Unemployment in the United States and in Europe. A contrast and the reasons', *CESifo Working Paper Series* 231.

Staiger, D., Stock, J. H. and Watson, M. W. (1997) 'The NAIRU, unemployment and monetary policy', *Journal of Economic Perspectives*, 11: 33–59.

Stiglitz, J. (1997) 'Reflections on the natural rate hypothesis', *Journal of Economic Perspectives*, 11: 3–10.

Stirati, A. (2001) 'Inflation, unemployment and hysteresis: an alternative view', *Review of Political Economy*, 13: 427–51.

Tobin, J. (1995) 'The Natural Rate as New Classical Macroeconomics', in R. Cross (ed.), *The Natural Rate of Unemployment* (Cambridge: Cambridge University Press).

Vetlov, I., Hlédik, T., Jonsson, M., Kucsera, H. and Pisani, M. (2011) 'Potential Output in DSGE Models', *ECB Working Paper* 1351.

Woodford, M. (2003) *Interest and Prices: Foundations of a Theory of Monetary Policy* (Princeton, NJ: Princeton University Press).

5
A Historical Approach to Demand-Led Growth Theory

*Matthew Smith**

5.1 Introduction

The central purpose of this chapter is to construct an analytical framework for explaining growth in concrete terms by reference to history consistent with the view that economic growth is fundamentally determined by the growth in aggregate demand.[1] As based on the Keynesian principle of effective demand, a demand-led theory of growth supposes that the level of aggregate output is determined in the long run by aggregate demand in which saving endogenously adjusts to autonomous demand through changes in income and output associated with the adjustment of productive capacity to aggregate demand. In this approach it is the growth in demand which determines the growth in output and the rate of capital accumulation in which it is supposed there is no technological constraint on output adjusting to demand growth. Key factors in explaining growth, notably, technical progress, are therefore conceived to contribute to economic growth through their effect on the growth in demand. From this standpoint, growth is a complex process, entailing structural change of the economic system, such that it can only be plausibly explained in concrete terms by reference to social, politico-institutional and technological factors. All these factors are seen to have an historical dimension in explaining growth.

The demand-led theory constructed here proceeds by building upon the Keynesian theory of demand-led growth consistent with the

*I am grateful to Tony Aspromourgos, John King, Heinz Kurz, Rod O'Donnell, Attilio Trezzini and an anonymous referee for some valuable comments on draft versions of this paper, without implicating them in the present product.

surplus approach to value and distribution of classical economics as reconstructed by Sraffa (1960). A critical feature of this theory, which clearly distinguishes it from supply-side growth theory developed on the basis of marginalist principles,[2] is the absence of any functional relationship between the quantity of inputs to be employed productively and the (relative) prices of those inputs which could provide a tendency for the economic system to gravitate toward full employment output. This feature stems from the fundamental analytical separability that exists in classical economics between, on the one hand, the determination of long-period normal prices and distribution and, on the other hand, the determination of outputs and the aggregate level of output as well as employment (see Garegnani 1984; 1990, pp. 122–32). In the framework presented this means that at any long-period positions along a growth path the determination of normal distribution and prices is conceived to correspond with the determination of long-run equilibrium levels of aggregate output at which demand is not necessarily sufficient to bring about the full employment of productive inputs.[3]

To begin with, in Section 5.2 we build a 'supermultiplier' growth model in which utilisation of an economy's productive capacity is assumed to always correspond to a *given* normal utilisation of capacity. On the basis of this assumption we derive a familiar 'steady-state' growth model in which output and the capital stock are conceived to grow constantly at the same rate for a given technique. We shall see, however, that such a growth model is not truly consistent with the fundamental Keynesian conception that in the long run demand is autonomous of saving in the determination of output along a growth path. In Section 5.3 we construct an alternative growth model in which the utilisation of capacity is conceived to vary both in the short and in the long run. Based on this conception the long-run average utilisation of capacity is endogenously determined and is systematically different from normal capacity utilisation. A novel feature of our analytical framework is that it is based on historical periodization in which trend economic growth from one period to the next is conceived to be determined not only by the growth rate of autonomous demand but also by long-run changes in the value of the supermultiplier. A short conclusion is provided in Section 5.4.

5.2 The growth model with normal utilisation

The Keynesian demand-led growth model employed incorporates a 'supermultiplier' of induced expenditure originally developed by Hicks (1950) which links quantitatively autonomous demand to equilibrium output and income. This particular supermultiplier model has been articulated in the literature, notably by Serrano (1995) and, from a critical viewpoint, by Trezzini (1995, 1998). A feature of the model is that productive capacity is determined by long-run aggregate demand. In the model there are three basic components of aggregate demand (AD_t), consisting of autonomous demand (A_t), induced consumption expenditure (c_tY_t) and induced investment (I^I_t) which also contributes to productive capacity:

$$AD_t = A_t + c_tY_t + I^I_t \qquad [5.1]$$

where c_t is society's marginal propensity to consume with values $0 < c_t < 1$. The first component, autonomous demand, consists of those expenditures that are explained independent of changes in income and output occurring over the same time period. It is essentially that part of aggregate demand which, along with induced investment, is accommodated by saving (including taxation) that is endogenously generated by income. In a closed economy these expenditures consist of government expenditure, autonomous investment and autonomous consumption, while in an open economy they include exports which, though different from other components, are accommodated by foreign income and when this is less than imports the margin of difference is accommodated by foreign saving. For simplicity, we shall assume a closed economy. Again for simplicity, we shall also assume that autonomous demand does not create additional productive capacity – that is, it is non-capacity-generating expenditure. The second component, induced consumption, is that consumption which is a positive function of the current level of income and output. As is well known, its relationship to income is defined by the marginal propensity to consume whose value is conceived to depend on socio-institutional factors, most notably, the distribution of income and, connectedly, the taxation and welfare system.

The third component is induced investment, through which productive capacity is conceived to adjust to aggregate demand. Based on the accelerator principle, induced investment will depend on the amount of productive capacity that needs to be installed for a given technique of production to ensure the level of output accommodates expected demand. For simplicity, we will initially assume there is zero depreciation on fixed capital and employ a rigid accelerator to express induced investment in a familiar way as:

$$I^I_t = (K_{t+1} - K_t) = a_t(Y^e_{t+1} - Y_t) \qquad [5.2]$$

where K_{t+1} is the capital stock required in the future period to accommodate the expected level of demand, Y^e_{t+1}, and a is the capital–output ratio. In order to account explicitly for the role of capacity utilisation the capital–output ratio can be expressed as a/u, determined as follows:

$$K_t/Y_t = (K_t/Y^*_t) \cdot (Y^*_t/Y_t) = a_t/u_t \qquad [5.3]$$

where a_t is the capital–output ratio, K_t/Y^*_t, when capacity is fully utilised (i.e. $u_t=1$) and u_t is the degree of capacity utilisation, defined *as the ratio of actual output to full-capacity output,* Y_t/Y^*_t, for a given capital stock. Rearranging equation [5.3] we obtain an expression for capacity utilisation:

$$u_t = a_t Y_t/K_t \qquad [5.4]$$

On the plausible assumption that, typically, firms install productive capacity to produce with spare capacity to meet peak demand as well as to enable an expansion in output to capture greater sales revenue in the event of persistent higher demand, the degree of normal (or desired) utilisation, u^n_t, will have a value between zero and full capacity (i.e. $0<u^n_t<1$).[4] Given that u^n_t is the ratio of the desired level of output produced to the full capacity of installed capital, Y^d_t/Y^*_t, then the desired capital–output ratio, a_t/u^n_t, is determined as follows:

$$K_t/Y^d_t = (K_t/Y^*_t) \cdot (Y^*_t/Y^d_t) = a_t/u^n_t \qquad [5.5]$$

And, therefore:

$$u^n_t = a_t Y^d_t/K_t \qquad [5.6]$$

In short, normal utilisation for an economic system is that which reflects the utilisation of capacity that firms determine will maximise their profit rates for a given technique, taking account of the possible fluctuations in actual demand and its impact on average costs over a period of time relevant to the installation of their existing capacity (see Ciccone, 1986, pp. 23–32). On this basis equation [5.2] for induced investment is re-written as:

$$I^I_t = (K^d_{t+1} - K_t) = (a_t/u^n_t)(Y^e_{t+1} - Y_t) \qquad [5.7]$$

where K^d_{t+1} is the capital stock *desired* in the future period based on the expected level of demand, Y^e_{t+1}, and the *desired* capital–output ratio

a_t/u_t^n as based on the dominant techniques of production and the normal utilisation of productive capacity. This accelerator relationship supposes that through time net investment ensures the capital stock adjusts to produce output levels according to the desired capital–output ratio.

By substituting equation [5.7] for I_t^I in equation [5.1] we obtain the following aggregate demand function:

$$AD_t = A_t + c_t Y_t + (a_t/u_t^n)\,(Y_{t+1}^e - Y_t) \qquad [5.8]$$

Solving for equilibrium income:

$$Y_t = A_t + c_t Y_t + (a_t/u_t^n)\,(Y_{t+1}^e - Y_t) \qquad [5.9]$$

and, with expected growth in output, $g_t^e = (Y_{t+1}^e - Y_t)/Y_t$ and, re-arranging, we obtain:

$$Y_t = A_t / [1 - c_t - (a_t/u_t^n) g_t^e] \qquad [5.10]$$

If it is then assumed firms have perfect foresight, expected growth, g_t^e, will be equal to the growth rate of output (and income), g_t^y. And if we substitute the propensity to save, s_t, for $1 - c_t$, the following expression is obtained:

$$Y_t = A_t / [s_t - (a_t/u_t^n) g_t^y] \qquad [5.11]$$

A positive value of Y_t requires that given $0 < s_t \leq 1$, $s_t > (a_t/u_t^n) g_t^y$. The equilibrium income so determined may be called 'capacity income' because it corresponds to a level of output produced at the normal utilisation of capacity. By re-arranging equation [5.11] a familiar growth equation for a supermultiplier model is obtained:

$$g_t^y = [s_t - (A_t/Y_t)] / (a_t/u_t^n) \qquad [5.12]$$

The equilibrium growth rate, g_t^y, is determined by the ratio A_t/Y_t for a range of possible values up to a maximum value of $g_t^y = s_t/(a_t/u_t^n)$, when $A_t = 0$. This equilibrium growth is that necessary to ensure capacity saving, being that level of saving which is generated from income when output is produced at the normal utilisation of capacity is equal to autonomous expenditure plus induced investment.[5] It is also the growth rate at which the degree of utilisation conforms continuously to normal utilisation along a steady-state growth path in which the capital stock and output continuously grow at the same rate for a given production technique.

A major problem with this steady-state growth model is that it is not really compatible with the fundamental Keynesian notion that demand is autonomous in the determination of the trend growth of output (Trezzini 1995, pp. 48–56). This can be explained by reference to the growth equation [5.12]. According to this equation, for the given ratio A_t/Y_t which determines g_t^y to remain constant along the trend growth path, the growth rate of autonomous demand, g_t^A, must be equal to the steady-state growth rate: i.e. $g_t^A = g_t^y$. However, this means that the growth rate of autonomous demand is limited in the sense that $g_t^A < s_t/(a_t/u_t^n)$ consistent with $A_t > 0$, where $s_t/(a_t/u_t^n)$ is capacity saving as a ratio of the capital stock.[6] The reason for this limitation in the model is that growth in autonomous demand, which is equal to or greater than $s_t/(a_t/u_t^n)$, cannot be accommodated by the growth in capacity saving necessary for equilibrium along the steady-state growth path. If demand is truly autonomous there appears no logical reason why its growth should be so bound by capacity saving. Connected to this is the peculiarity in equation [5.12] of the inverse relationship between the ratio A_t/Y_t and g_t^A on the basis of $g_t^A = g_t^y$ and for given values of s_t, a_t and u_t^n. Again, if demand is truly autonomous there is no plausible basis for supposing that its growth should systematically increase as the ratio of autonomous demand to capacity income (i.e. A_t/Y_t) decreases and, *vice versa* (Trezzini 1995, pp. 52–3). In the steady-state growth model the logic for this inverse relationship is that as the magnitude of the latter ratio decreases (increases) an increasing (decreasing) proportion of capacity saving can be devoted toward induced investment and, thereby, toward augmenting (diminishing) the growth in capacity, its output and, causally, in the demand necessary to realise equilibrium growth. Hence, in this steady-state model the growth in autonomous demand ultimately depends on saving which is generated by the equilibrium growth in capacity income. This underlies the lack of autonomy of demand in the growth process when the trend rate of output growth and the saving which is generated by it is based on a *given* normal utilisation of capacity.

Our argument then is that under steady-state conditions in which it is supposed that there is a given normal utilisation of capacity along the trend growth path, aggregate demand is denied an autonomous role in the determination of economic growth. Nevertheless, as shown by Garegnani (1992), this theoretical problem can be surmounted by allowing the degree of capacity utilisation to vary both in the short and long run so that any level of autonomous demand (investment) can be accommodated by the generation of saving induced through changes in income and output facilitated by changes in capacity utilisation as well as in productive capacity. By allowing for persistent as well as temporary variations in the utilisation of capacity, long-run output has the elasticity to accommodate changes in aggregate demand beyond the steady-state for a given propensity to save (*ibid.*).[7] Importantly, this variability in capacity utilisation ensures that aggregate demand has an autonomous role in the growth process, which is crucial to the Keynesian approach (see Trezzini 1995, pp. 48–57; Palumbo and Trezzini 2003, pp. 110–14). It is clear, though, that this conception of the growth process cannot be reconciled with steady-state growth since the capital stock and output will be systematically growing at different rates.

5.3 Alternative growth model with endogenous utilisation

In an attempt to incorporate long-run elasticity of output into a super-multiplier growth model, Serrano (1995) proposed that consistent with variability in capacity utilisation the *average* utilisation of capacity which emerged over time would be the same as the normal utilisation of capacity. Unlike the steady-state model, the utilisation of capacity is not assumed to be constant but rather the *given* normal utilisation of capacity is proposed to correspond to an average of its fluctuations over time. This conception therefore brings in historical time with all the variables expressed as averages, including the expected growth in demand of firms, in the determination of an average rate of growth. To express this conception equation [5.12] can be re-written as:

$$g_t^y = [s_t - (A_t / Y_t)] / (a_t / u_t^a) \qquad [5.13]$$

where u_t^a is the average utilisation of capacity and $u_t^a = u_t^n$.

The problem with this conception is that there appears to be no compelling reason why the average utilisation of capacity which emerges over time should be equal to the normal utilisation of capacity. Indeed, as shown by Trezzini (1998, pp. 59–66), even on the assumption of perfect foresight, any deviation of actual growth from the equilibrium growth rate associated with a normal utilisation of productive capacity will require average utilisation to vary significantly from normal over a considerable period of time to restore the equilibrium rate of growth. Furthermore, the process of adjustment itself can cause the growth in capacity to change in relation to output growth since changes in utilisation which, in the long run, affect capacity, simultaneously affect demand. Moreover, in the long-run adjustment of capacity to demand, investment induced by deviations of average from normal utilisation will simultaneously affect demand as well as productive capacity. Hence, as Palumbo and Trezzini (2003, pp. 115–20) have argued, once it is acknowledged that the utilisation of capacity may vary such as to ensure long-run elasticity in output that can accommodate any feasible level of aggregate demand, then the adjustment process of capacity to demand is a path-dependent one that means average utilisation is, except by rare coincidence, unlikely to be equal to normal utilisation, notwithstanding investment decisions by firms to achieve it.

The question is: where do these analytical issues leave the demand-led growth theory presented in this chapter? The answer is a more modest theory which does not pretend *to account fully* for the growth process but nevertheless provides a framework for analysing the central causes of *trend growth and economic development* in a more concrete way, consistent with the notion that the growth in aggregate output is fundamentally determined by the growth in aggregate demand. A historical approach is proposed, suggested in Serrano (1995),[8] in which the growth model provides a demand-led framework for a concrete explanation of the average growth rate over historical time periods, termed here as 'epochs'. While these epochs must at a minimum be long enough for the adjustment of fixed productive capacity by firms to longstanding demand conditions, they can otherwise be defined arbitrarily according to their significance in explaining growth trends by reference to key historical events as well as to the character of the demand-led forces which are ascertained to determine economic development and growth performance. Thus, for example, an epoch could be defined by reference to the event of war, a change in the international economic regime, a fundamental change in policymaking, an unprecedented structural change in the economic system or a combination of these or other such historically related events. The theoretical counterpart to epoch in our model is 'period', in which the long-run average growth rate is conceived to be determined by the persistent forces of demand specified in our demand-led theory. In this approach the average growth rate in each period is conceived to be linked to that of the previous period so that growth in period t can only be properly explained by reference to the history of the growth process in period $t-1$ and, prior to this, period $t-2$ and so on back to period $t-n$. Hence, the long-run average growth rate in any period is determined by demand-led forces which have a historical context and, in concrete terms, are to be explained by reference to history. As further clarified below, in correspondence with epochs, the long run of the periods used are at a minimum long enough for fixed productive capacity to adjust to expected demand conditions consistent with long-run growth.

Firstly, in accordance with the foregoing concept of long-run elasticity in which changes in the degree of capacity utilisation play an active role in the adjustment of saving to autonomous demand (and induced investment) along with output adjusting to aggregate demand, the model shall suppose that the utilisation of capacity is endogenously determined in the growth process. This means that in any period the average utilisation of capacity is conceived to be endogenously determined.[9] Based on the reasoning given above, the average utilisation so determined is not conceived, except by coincidence, to equal the normal utilisation of capacity upon which firms base their investment decisions in adjusting their capacity to demand. This conception requires us to suppose that expectations of future demand by firms are not normally realised. Secondly, therefore, in the model the unrealistic assumption of perfect foresight is dispensed with, such that the expected growth of demand for a firm is not necessarily equal to its actual growth (i.e. $g_t^e \neq g_t^y$). While it is acknowledged that firms will continuously adjust their expectations of growth in demand to historical growth rates, unless the growth rate is stable for a very long period of time it is not plausible to assume that their expectations will be systematically correct. Once steady-state growth is abandoned and the growth rate is conceived to be determined by demand in a path-dependent way perfect foresight has little plausibility. Thirdly, as already anticipated the model will take a more general form and suppose the existence of fixed capital. This means we must account for the effect of the rate of depreciation of the capital stock on induced investment by rewriting equation [5.7] above to:

$$I_t^I = (a_t / u_t^n)(Y_{t+1}^e - Y_t) + (a_t u_t^n d_t) Y_t \qquad [5.14]$$

where I_t^I is induced investment and d_t is the average rate of depreciation of *utilised* capital in period *t*. The second term on the right-hand side of equation [5.14] clearly expresses the notion that the rate of depreciation of the capital stock increases with its utilisation. With respect to the effect of depreciation on induced investment, the equation shows that induced investment in our model is conceived to be based on the rate of depreciation expected by firms to occur *at the normal utilisation* of the capital stock. However, because average utilisation will be systematically different from normal, the depreciation of the capital stock which occurs will be systematically different from that expected. As is elaborated below, this unexpected depreciation of the capital stock can influence future induced investment by, in turn, contributing to the deviation of average from normal utilisation.

Fourthly, since the degree of utilisation of capacity that is realised will, except by chance, be different from normal, we need to account for the effect of this systematic deviation on induced investment. The model is able to account, however mechanically, for the manner in which capacity is conceived to adjust to aggregate demand in a demand-led growth theory. It is proposed that deviations of average utilisation realised in the previous historical period $t-1$ from normal utilisation in period t will tend to induce a change in investment in the current period, t, by firms endeavouring to adjust their capacity to demand so as to establish normal utilisation. This means that if normal utilisation in period t remains unchanged from that in period $t-1$, firms overall adjust their capacity to the historical deviation between average and normal utilisation occurring in period $t-1$. However, it is envisaged that in each period normal utilisation will be revised by firms when installing new capacity according to a complex of factors so that u_t^n can be different from u_{t-1}^n and so on. The degree of divergence of average from normal utilisation in period $t-1$ will clearly be a major factor in revising the normal degree of utilisation in period t.[10] On this basis it is supposed that the deviation of average from normal utilisation in one period will tend to induce a change in capacity-adjusting investment in the next period. This capacity adjustment mechanism supposes that period $t-1$ is sufficiently long that any deviation between average and normal utilisation can be considered systematic and firms can feasibly adjust their capacity in period t to expected demand conditions in the future period $t+1$. Nevertheless, this mechanism represents a simplification of the capacity-adjusting process, since the deviation between average and normal utilisation, especially in a period of stagnant economic growth, may only reflect a disparity between the actual and expected frequency of fluctuations in demand with peak demand well below full capacity. On the other hand, equality between average and normal utilisation may merely mask a significant increase in the amplitude of fluctuations in peak demand, requiring firms to make additional investment in capacity. In this respect, an underlying assumption of our model is that firms tend to adjust capacity at discrete intervals in each period when they install planned spare capacity which will on average be utilised over a long period of time according to the expected future growth in demand. Incorporating this conception with the stated qualifications into the determination of induced investment, equation [5.14] above is re-written as:

$$I_t^I = (a_t / u_t^n)\ (Y_{t+1}^e - Y_t) + a_t u_t^n\ d_t Y_t + (a_t / u_t^n - a_t / u_{t-1}^a) Y_t \qquad [5.15]$$

where u^a_{t-1} is the average degree of utilisation realised in period $t-1$ and the term $(a_t/u^n_t - a_t/u^a_{t-1})Y_t$ reflects the adjustment of capacity to demand to restore normal utilisation.[11] By re-expressing equation [5.15] it can be easily shown that *net* of the expected depreciation of capital, induced investment is the difference between the capital stock desired by firms to accommodate expected demand in period $t+1$ at normal capacity utilisation, K^d_{t+1}, and what the capital stock will otherwise be in period t at the existing average utilisation of capacity determined in period t–1, denoted as K^r_t:

$$I^I_t - (a_t u^n_t d_t)Y_t = (K^d_{t+1} - K^r_t) = (a_t/u^n_t)Y^e_{t+1} - (a_t/u^a_{t-1})Y_t \qquad [5.16]$$

where $I^I_t - (a_t u^n d_t)Y_t$ is induced investment *net* of expected depreciation of the capital stock. It will be convenient here to employ equation [5.15] rather than equation [5.16].[12] However, what this latter equation shows is that whereas in steady-state and other models discussed above induced investment changes at the same constant rate as output (and income), in our model it changes at a different rate from one period to the next according to changes in average utilisation brought about by unexpected changes in the growth rate of demand.[13] Accordingly, for this reason alone, the capital stock tends to grow at a different rate from one period to the next in our model.

These elements can be represented in our model by rewriting equation [5.10] above as follows:

$$Y_t = A_t/[1 - c_t - (a_t/u^n_t)g^e_t - a_t u^n_t d_t - (a_t/u^n_t - a_t/u^a_{t-1})] \qquad [5.17]$$

where all variables are expressed as 'averages' so that g^e_t refers to the expected average growth in demand in period t and the condition $1 > [c_t + (a_t/u^n_t)g^e_t + a_t u^n_t d_t + (a_t/u^n_t - a_t/u^a_{t-1})]$ is met. In absence of perfect foresight, expected average growth in demand (and hence, in output) will not be equal to the average growth in output in period t, g^y_t, such that $g^e_t \neq g^y_t$. Given the values of c_t (or s_t), a_t, d_t, u^n_t, u^a_{t-1} and g^e_t, which together determine the supermultiplier, and the level of autonomous demand, A_t, long-run average income and output is determined.[14] On the basis of this datum and the historically *given* capital stock (i.e. K_t) employed to produce output (i.e. Y_t) in period t, the average utilisation of capacity, u^a_t, will be endogenously determined as follows:

$$u^a_t = a_t Y_t / K_t \qquad [5.18]$$

and with $g^e_t \neq g^y_t$, then $u^a_t \neq u^n_t$. Hence, except when $g^e_t = g^y_t$, the average utilisation of capacity will be systematically different from normal and average utilisation will vary from one period to the next, such that $u^a_t \neq u^a_{t-1}$. The capital stock to determine average utilisation of capacity in period t in equation [5.18] is itself determined historically in the following way:[15]

$$K_t = K_{t-1} + I^I_{t-1} + (u^n_{t-1} - u^a_{t-1})a_{t-1}d_{t-1}Y_{t-1} \qquad [5.19]$$

The term $(u^n_{t-1} - u^a_{t-1})\ a_{t-1}d_{t-1}Y_{t-1}$ in equation [5.19] is the average depreciation of the capital stock in period $t-1$ which was *not expected* by firms when, through induced investment (i.e. I^I_{t-1}), they installed capacity to accommodate expected demand in period t (i.e. Y^e_t). Unexpected depreciation so affecting the capital stock is conceived to be systematic on account of the systematic difference between average and normal utilisation. Hence, for example, if $u^a_{t-1} > u^n_{t-1}$ because $g^e_{t-1} < g^y_{t-1}$, the depreciation of the capital stock will be greater than anticipated and, thereby, not compensated by induced investment, will tend to reduce the stock of capital available in period t. By affecting capacity in this way, unexpected depreciation will contribute to a higher average rate of utilisation determined in period t (i.e. u^a_t)[16] and, thereby, tend to contribute to its deviation from normal utilisation which, in the manner explained above, firms will endeavour to correct through induced investment in period $t+1$.[17]

On the basis of the analysis above the average growth rate in period t will be equal to:

$$g^y_t = Y_t - Y_{t-1} / Y_{t-1} \qquad [5.20]$$

where current average output, Y_t, is determined in equation [5.17] and output in the previous period is similarly determined according to the equation:

$$Y_{t-1} = A_{t-1} / [1 - c_{t-1} - (a_{t-1}/u^n_{t-1})g^e_{t-1} - a_{t-1}u^n_{t-1}d_{t-1} \\ - (a_{t-1}/u^n_{t-1} - a_{t-1}/u^a_{t-2})] \qquad [5.21]$$

Now, for simplicity, we will denote the supermultipliers for period t and $t-1$ respectively as follows:

$$m_t = 1/[1 - c_t - (a_t/u^n_t)g^e_t - a_t u^n_t d_t - (a_t/u^n_t - a_t/u^a_{t-1})] \qquad [5.22]$$

$$m_{t-1} = 1/[1 - c_{t-1} - (a_{t-1}/u^n_{t-1})g^e_{t-1} - a_{t-1}u^n_{t-1}d_{t-1} \\ - (a_{t-1}/u^n_{t-1} - a_{t-1}/u^a_{t-2})] \qquad [5.23]$$

The value of m_t will be different from m_{t-1} purely on the grounds that u^a_{t-1} is a different value to u^a_{t-2}. Thus, for example, even supposing $g^e_t = g^e_{t-1}$, $c_t = c_{t-1}$, $a_t = a_{t-1}$, $u^n_{t-1} = u^n_t$ and $d_t = d_{t-1}$, if $u^a_{t-1} > u^a_{t-2}$, then $m_t > m_{t-1}$. We can write the equations for the determination of long-run average output in period t and $t-1$ in the simple form:

$$Y_t = A_t \,.\, m_t \qquad [5.24]$$

$$Y_{t-1} = A_{t-1} \,.\, m_{t-1} \qquad [5.25]$$

Substituting equations [5.24] and [5.25] into [5.20] allows us to express the average growth rate of output in period t as:

$$g^y_t = A_t \,.\, m_t - A_{t-1} \,.\, m_{t-1} / A_{t-1} \,.\, m_{t-1} \qquad [5.26]$$

With re-arrangement and manipulation we can get the following demand-led growth equation for period t:

$$g^y_t = g^A_t + \Delta m_t (A_t / A_{t-1}) \qquad [5.27]$$

where g^A_t is the growth rate of autonomous demand and Δm_t is the change in the supermultiplier in period t as determined by $(m_t - m_{t-1})/m_{t-1}$.[18] This growth equation shows that the growth rate of output is determined by the growth rate of aggregate demand, as determined by two elements: (i) the growth rate of autonomous demand, g^A_t; and (ii) the *change* in the value of the supermultiplier, Δm_t. It is evident that if $m_t = m_{t-1}$ so that $\Delta m_t = 0$, the growth of output will be determined wholly by the growth in autonomous demand; that is, $g^y_t = g^A_t$. While the growth in autonomous demand is conceived to be the main determinant, lasting changes in the supermultiplier can be a contributor to the determination of economic growth in this model.

The analytical limitations of our demand-led growth model should be mentioned here. In our model the *average* long-run growth of output and the *average* long-run growth in the capital stock over a period are the same, although it is supposed that output and the capital stock at any time within a period will be systematically growing at different rates associated with variations in the utilisation of capacity. This could not otherwise be the case in a 'growth equation' of the form of equation [5.27]. Our model is therefore not capable of accounting for the role of variations in the utilisation of capacity in the growth process *within any* period. Nor is it capable of accounting for interactions which occur between changes in capacity and demand along a growth path *within* a period. The complexity of these ongoing interactions in the growth process which, as discussed above, mean utilisation will be systematically different on average to normal (or desired), belongs to a separate analysis that accounts for cyclical changes in activity. Instead, our model accounts for variations between the average growth of output and capital accumulation associated with endogenous variations in average utilisation *between* different historical periods. It also endeavours to account for the process by which capacity adjusts to demand when the trend growth in demand deviates from that expected by firms and average utilisation systematically deviates from normal by reference to sequences of periods in which the trend growth rate changes from one period to the next. In this rather mechanical way the model represents interactions between demand and capacity along a path of changing *trend* rates of growth and accumulation. Notwithstanding these limitations, our model does provide an analytical framework of the fundamental Keynesian notion that the growth in output and capital accumulation is wholly determined by the growth in aggregate demand.

5.4 Conclusion

A central feature of the demand-led growth model is that unlike 'steady-state' models the growth rate depends not only on the growth rate of autonomous demand but also on the long-run change in the value of the supermultiplier. This stems from the historical periodisation incorporated into the model in which the supermultiplier will invariably be different from one period to next. Therefore, according to the model, trend growth in an historical period or epoch is explained not just by reference to factors determining the growth of autonomous demand but also by reference to factors which can cause long-run changes in the value of the supermultiplier such as changes in income distribution, technical change and the revision of expectations by firms about long-run demand. Moreover, in this model history is considered to play a central role in determining the value of the supermultiplier which, in any period, is dependent in part on events which have occurred in previous periods. The growth in autonomous demand nevertheless remains the main driving force of economic growth in the model. Hence, much of the explanation of growth will consist in identifying the key factors determining the growth in autonomous demand. These will include the longstanding fiscal policy of government, the long-running monetary policy of the central bank, developments in the financial system affecting the financing of private expenditure, market regulations and institutions affecting entrepreneurship and innovation and government policies on trade and industry, all of which influence in various ways the different components of autonomous demand. By reference to these factors, explaining growth will thus entail identifying the roles of the different components of autonomous demand: government spending, autonomous private investment, autonomous consumption and, for an open economy, exports.[19] Generated by a combination of these components it is evident that growth in autonomous demand can be driven by different sources in different historical periods. Hence, appealing to history, the role of government spending was a major driver in most advanced nations during the post-war recovery phase of the Second World War, whilst, in the last twenty years, autonomous consumption has played a more prominent role in generating demand than in previous historical periods. More intricately, the determination of the growth in autonomous demand as a whole will entail a causal interaction between its components and, therefore, should be part of any explanation of demand-led growth. In this respect, the sum will be greater than its parts. Hence, for example, large-scale government capital expenditure on transport and communications infrastructure, which improves productivity growth, is likely to augment the other components of autonomous demand by contributing to stronger growth in private investment in innovation and in autonomous consumption as well as contributing to a nation's better export performance. In this regard the causal relationship between these autonomous components of autonomous demand cannot be considered functional in the sense of the causal relationship supposed between income and demand, as defined by our supermultiplier. Instead, the causal relationship between them is conceived to be generally complex and contingent on a wide set of circumstances, such that they could only be properly explained in concrete terms in accordance with the historical approach proposed.

Notes

1. In our view this approach is consistent with the methodology employed by Adam Smith in the *Wealth of Nations* (1976 [1776]), in which a theoretical system informs an historical analysis of the major forces determining economic development. The important property of Adam Smith's theoretical system is that it is open to social and institutional factors playing a key explanatory role that can only be properly understood by reference to their history. On Smith's position, see Aspromourgos (2009, pp. 247–51).

2. Besides the Solow (1956) and Swan (1956) models, supply-side growth theory include the 'endogenous growth' models, of notably, Romer (1986; 1990) and Lucas (1988).
3. In classical analysis because quantities of commodities produced are *given* in the determination of normal prices and distribution the long-period method entails the gravitation of prices and distributive variables around their normal values but not quantities. Only in long-period marginalist economics in which normal prices and quantities are conceived to be simultaneously determined do quantities as well as prices gravitate around normal magnitudes. Thus, in classical analysis, the long-run equilibrium level of output and its composition upon which given, normal prices and distribution are determined along the growth path, can be conceived to be determined in a path-dependent manner so that it does not represent an equilibrium output in the sense that actual levels of output gravitate around it. Indeed, this methodological approach is adopted in the growth analysis which follows.
4. The notion that firms would permanently maintain excess productive capacity is originally attributable to Steindl (1952, pp. 4–14). Besides accommodating peak fluctuations in demand, Steindl argued that a 'more general reason' for maintaining excess capacity was that in competing with rivals, firms wanted to be in a position to expand their market share and establish their 'goodwill' in being able to reliably supply greater demand in the market.
5. This is simply given by the following equation:

$$s_tY_t = A_t + (a_t/u_t^n)\,(Y_{t+1}^e - Y_t)$$

   where it is assumed $Y_t > Y_t^m$, in which Y_t^m is the minimum level of income necessary given the existence of positive autonomous consumption, C^A, such that $s_tY_t^m = C^A$. This condition is necessary in a global (or closed) economy for saving net of autonomous consumption to be positive (i.e. $s_tY_t > C^A$).
6. By re-arrangement we can obtain a more comprehensible form for this term. Substituting K_t/Y_t^* for a_t into the term $s_t/(a_t/u_t^n)$ obtains $s_t/(K_t/Y_t^*/u_t^n)$ and then, by re-arrangement, to $s_t(Y_t^* u_t^n)/K_t$. This simplifies to $s_t\,Y_t/K_t$, where it is recalled that Y_t is capacity income.
7. This conception was largely proposed by Garegnani (1992) as a critique of the Cambridge conception that distribution was dependent on the rate of capital accumulation, which had been variously advanced by Kaldor (1955–56, pp. 94–100), Kahn (1959), Robinson (1962, pp. 11–13, 40–41) and Marglin (1984). In the Cambridge conception the limit to growth posed by existing capacity saving is essentially surmounted by generating growth in autonomous demand sufficient to cause a change in distribution which, in turn, induces a higher propensity to save so that additional saving is generated to accommodate the additional autonomous demand. This can be illustrated by reference to equation [5.12] above. Suppose, through government policy, g_t^A is increased to a rate which is higher than the existing g_t^y. This expansion in demand, through inflation, then brings about a redistribution of income from wages to profits, which, on the plausible assumption that capitalists have a higher propensity to save than wage earners, causes the value of s_t to increase, producing an increase in capacity saving necessary to facilitate

the expansion in autonomous expenditure. In this way the value of g_t^y will then adjust to a policy-determined g_t^A. It is supposed that in this process the resulting increase in the rate of capital accumulation will be associated with a higher rate of profit and, for a given technique, a lower real wage. In contrast, according to Garegnani's (1992) proposition, g_t^y can adjust to a higher g_t^A without any change in distribution by the degree of utilisation increasing in the long run. The increase in income (output) derived from a higher degree of utilisation of productive capacity is then able to generate the necessary saving to facilitate the additional autonomous demand. Importantly, this argument entails the rejection of a given normal utilisation and, thereby, of a steady-state growth model. For other related criticisms of the Cambridge approach to growth and distribution, see Vianello (1985), Ciccone (1986) and Garegnani and Palumbo (1998).

8. Also see Kaldor (1957, p. 601 n. 1).
9. This appears to be consistent with the conception briefly outlined by Ciccone (1987, pp. 103–6). Also see Amadeo (1986).
10. Hence, interpreted by reference to the frequency and amplitude of fluctuations in demand (and, hence, utilisation) which have occurred, the magnitude of divergence between u_{t-1}^n and u_{t-1}^a in period $t-1$ will provide important information to firms in the determination of u_t^n in period t (Amadeo, 1986, p. 155). Other major related factors which will influence the determination of normal utilisation are the technology embodied in newly installed capacity, the expected fluctuations in demand as based on historical experience and the degree of spare capacity planned for newly installed capacity (see Ciccone, 1986).
11. With respect to the third term on the right-hand side of equation [5.15], if we denote K_t^n as the capital stock with normal utilization and K_t^r the capital stock that would be realised in period t based on the average utilisation in period $t-1$, then $K_t^n - K_t^r = (a_t/u_t^n - a_t/u_{t-1}^a)Y_t$. Hence, for example, if $u_{t-1}^a > u_t^n$, this means for an existing level of demand and output, Y_t ,the capital stock that would be realised without any adjustment to induced investment in period t, K_t^r, is smaller than necessary for aggregate production to occur at a normal degree of utilisation: that is, $K_t^n > K_t^r$.
12. It should be noted that the aggregate level of investment in the model consists of induced investment determined according to equation [5.15] *plus* non-capacity creating autonomous investment which is part of and contributes to autonomous demand. On autonomous investment, see n. 19 below.
13. Note that $I_{t-1}^I - (a_{t-1}u_{t-1}^n d_{t-1})Yt_{-1} = (K_t^d - K_{t-1}^r) = (a_{t-1}/u_{t-1}^n)Y_t^e - (a_{t-1}/u_{t-2}^a)Y_{t-1}$, $I_{t-2}^I - (a_{t-2}\ u_{t-2}^n d_{t-2})Yt_{-2} = (K_{t-1}^d - K_{t-2}^r) = (a_{t-2}/u_{t-2}^n)Y_{t-1}^e - (a_{t-2}/u_{t-3}^a)Y_{t-2}, \ldots$ and so on.
14. This long-run position corresponds to equality between saving and autonomous expenditure *plus* induced investment, expressed as follows:

$$sY_t = A_t + a_t/u_t^n(Y_{t+1}^e - Y_t) + a_t\, u_t^n\, d_t Y_t + (a_t/u_t^n - a_t/u_{t-1}^a)Y_t$$

where the condition $Y_t > Y_t^m$ is met (see n. 5). Given the propensity to save, the level of saving adjusts, *via* the supermultiplier, to any given level of autonomous expenditure *plus* capacity-adjusting investment through changes in the long-run level of income (i.e. Y_t).

15. The capital stock can also be shown to be historically determined by reference to saving endogenously generated by income, as follows:

$$K_t = K_{t-1} + sY_{t-1} - A_{t-1} + (u^n_{t-1} - u^a_{t-1})\ a_{t-1}d_{t-1}Y_{t-1}$$

where A_t is the autonomous demand in period $t-1$ which absorbs part of saving but is assumed not to be adding to productive capacity.
16. Through this sequence of effects on average utilisation in period *t*, (i.e. u^a_t), the capital stock in period $t+1$ will, in turn, be affected since $K_{t+1} = K_t + I^I_t + (u^n_t - u^a_t)\ a_t d_t Y_t$.
17. In accord with equation [5.15], induced investment in period $t+1$ is:

$$I^I_{t+1} = (a_{t+1}/u^n_{t+1})\ (Y^e_{t+2} - Y_{t+1}) + a_{t+1}u^n_{t+1}d_{t+1}Y_{t+1} + (a_{t+1}/u^n_{t+1} - a_{t+1}/u^a_t)Y_{t+1}$$

Hence, while higher utilisation means less capital is required per unit of output (in period *t*), it leads to a faster rate of depreciation of capital per unit of output which tends to induce a greater level of investment in the future (i.e. $t+1$).
18. It follows from equation [5.18] above that the average rate of capital accumulation, g^k_t, will be equal to the average growth rate of output in period *t*; that is, $g^k_t = g^y_t$.
19. Autonomous private investment will mainly, though not exclusively, be connected to innovation, consisting of investment in research and development and in the installation of technically superior capital equipment or that necessary to produce new innovative products. Relaxing our assumption that all autonomous investment is non-capacity creating, then clearly capacity-creating private and public investment, which 'sets the pace' for demand growth, is included as part of, and contributes to, autonomous demand.

References

Amadeo, E. J. (1986) 'The role of capacity utilisation in long period analysis', *Political Economy: Studies in the Surplus Approach*, 2(2): 147–60.

Aspromourgos, T. (2009) *The Science of Wealth: Adam Smith and Framing of Political Economy* (London: Routledge).

Ciccone, R. (1986) 'Accumulation and capacity utilization: some critical considerations on Joan Robinson's Theory of Distribution', *Political Economy: Studies in the Surplus Approach*, 2(1): 17–36.

Ciccone, R. (1987) 'Accumulation, capacity utilisation and distribution: a reply', *Political Economy: Studies in the Surplus Approach*, 3(1): 97–111.

Garegnani, P. (1984) 'Value and Distribution in the Classical Economists and Marx', *Oxford Economic Papers, New Series*, 36, 295–325.

Garegnani, P. (1990) 'Sraffa: Classical versus Marginalist Analysis', in K. Bharadwaj and B. Schefold (eds), *Essays on Piero Sraffa* (London: Unwin Hyman): pp. 112–41.

Garegnani, P. (1992) 'Some Notes for an Analysis of Accumulation', in J. Halevi, D. Laibman and E. J. Nell (eds), *Beyond the Steady State* (London: Macmillan), 47–71.

Garegnani, P. and Palumbo, A. (1998) 'Accumulation of Capital', in H. D. Kurz and N. Salvadori (eds) *The Elgar Companion to Classical Economics*, Volume A–K (Cheltenham U.K: Edward Elgar): pp. 10–18.

Hicks, J. R. (1950) *A Contribution to the Theory of the Trade Cycle* (Oxford: Clarendon Press).

Kahn, R. F. (1959) 'Exercises in the Analysis of Growth', *Oxford Economic Papers*, New Series, 11(2): 143–56.

Kaldor, N. (1955–56). 'Alternative theories of distribution', *Review of Economic Studies*, 23(2): 83–100.

Kaldor, N. (1957) 'A model of economic growth', *Economic Journal*, 67(268): 591–624.

Lucas, R. E. (1988) 'On the mechanics of economic development', *Journal of Monetary Economics*, 22: 3–42.

Marglin, S. A. (1984) *Growth, Distribution and Prices* (Cambridge, MA: Harvard University Press).

Palumbo, A. and Trezzini, A. (2003) 'Growth without normal capacity utilization', *European Journal History of Economic Thought*, 10(1); 109–35.

Robinson, J. (1962) *Essays in the Theory of Economic Growth* (London: Macmillan).

Romer, P. M. (1986) 'Increasing returns and long-run growth', *Journal of Political Economy*, 94: 1002–37.

Romer, P. M. (1990) 'Endogenous technical change', *Journal of Political Economy*, 98: S71–S102.

Serrano, F. (1995) 'Long period effective demand and the Sraffian supermultiplier', *Contributions to Political Economy*, 14: 67–90.

Smith, A. (1976 [1776]) *An Inquiry into the Nature and Causes of the Wealth of Nations*, Glasgow edn, Vols. I–II (Indianapolis: Liberty Classics).

Solow, R. M. (1956) 'A contribution to the theory of economic growth', *Quarterly Journal of Economics*, 70: 65–94.

Sraffa, P. (1960) *Production of Commodities by Means of Commodities: Prelude to a Critique of Economic Theory* (Cambridge: Cambridge University Press).

Steindl, J. (1952) *Maturity and Stagnation in American Capitalism* (Oxford: Blackwell).

Swan, T. W. (1956) 'Economic growth and capital accumulation', *Economic Record*, 32: 334–61.

Trezzini, A. (1995) 'Capacity utilisation in the long run and the autonomous components of aggregate demand', *Contributions to Political Economy*, 14: 33–66.

Trezzini, A. (1998) 'Capacity utilisation in the long run: some further considerations', *Contributions to Political Economy*, 17: 53–67.

Trezzini, A. (2005) 'The irreversibility of Consumption and Economic Growth', *Quaderno di Ricerca del Centro Studi e Documentazione 'Piero Sraffa'*, no. 6.

Vianello, F. (1985) 'The Pace of Accumulation', *Political Economy*, 1(1): 69–87.

6
Normal Paths of Growth Shaped by the Supermultiplier

*Óscar Dejuán**

6.1 Introduction

The Keynesian principle of effective demand states that the equilibrium level of output in a given period is a multiple of the expected autonomous demand (Kalecki, 1971; Keynes, 1936). Can we extrapolate this principle to a long-run dynamic analysis and conclude that the rate of growth of output will eventually depend on the expected rate of growth of autonomous demand? A positive answer would be a significant step towards a long-period theory of output which, according to Eatwell and Milgate, provides the most solid ground for demand-led growth (Milgate, 1982; Eatwell, 1983; Eatwell and Milgate, 1983; Eatwell, 2012).

A theory of demand-led growth becomes more compelling and manageable when it is expressed by the multiplier and the accelerator mechanisms. As a matter of fact, they have been present from the first growth model (Harrod, 1939). Harrod showed that, for an economy defined by the capital/output ratio and the saving propensity, a rate of growth exists that warrants the absorption of the full-capacity output and full-capacity savings, year after year. The model, however, was so unstable that it was soon rejected as a 'knife edge':

> On either side of this line it is a 'field' in which centrifugal forces operate, the magnitude of which varies directly as the distance of any point in it from the warranted line. Departure from the warranted

* Ana González Martínez (University of Castilla – La Mancha) helped me with the search for the data contained in Section 6.2. Marc Lavoie (University of Ottawa) made useful comments on the first version of this chapter.

> line sets up an inducement to depart farther from it. The moving equilibrium of advance is thus a highly unstable one. (Harrod, 1939)

Keynes' most direct disciples tried to accommodate demand-led growth in a steady-state model where endogenous changes in distribution made possible the adjustment to a different path of growth (Robinson, 1956; Robinson, 1962; Kaldor, 1955–56; Kaldor, 1961). Kaleckian economists consider that capacity utilisation is the endogenous variable that makes possible the adjustment to the autonomous trend (Lavoie, 2010; Hein *et al.*, 2011). The same opinion is shared by most Sraffian economists (Garegnani, 1983; 1992; Ciccone, 1986; 1987; Trezzini, 1995; 1998; Palumbo and Trezzini, 2003). Apparently they have defected to the original project of building a normal theory of output parallel to the normal theory of prices.

Not all Sraffian economists share this view. John Eatwell in the Memorial of Pierangelo Garegnani (22 February 2012) defends the advisability of filling the classical-Keynesian core of economic theory with a theory of normal output with the typical gravitation properties. Cesaratto (2012) separates core-Sraffians, the group based in Rome under the leadership of Garegnani, from periphery-Sraffians which includes at least three models based on the supermultiplier (Serrano, 1995; Bortis, 1997; Dejuán, 2005). According to Cesaratto the last group shows a plausible way of reconciling demand-led growth with the independence of technology (the optimal capital/output ratio, which implies a normal degree of capacity utilisation) and the independence of distribution. This chapter aims to clarify these issues and to show the stabilising properties of the multiplier–accelerator interaction, when it is properly formulated.

Beginning (Section 6.2) with a glance at the empirical facts that the theoretical approach proposes to explain, is the Spanish case presented as an example. Section 6.3 develops an investment function based on a flexible accelerator which adjusts capacity to a growing demand and takes into account excesses of inventories and the capacity to remove them. Section 6.4 integrates the accelerator of investment with the multiplier of consumption. The resulting supermultiplier is able to explain macroeconomic equilibrium at a given moment and through time. The autonomous trend, if it lasts long enough, determines the actual growth of demand and output. It also shapes the warranted (normal capacity) rate of growth, which becomes an endogenous variable. In Section 6.5 it is concluded that the supermultiplier, if properly formulated, is a stabilising mechanism leading the economy towards the normal path of growth compatible with the autonomous trend (given technology, distribution and expenditure patterns).

6.2 A glance at a growing economy

For economists inspired by Schumpeter (1912), the railway provides a good metaphor to explain the dynamics of a capitalist economy. The trucks correspond to the productive capacity that serves induced consumption which represents 60–70 per cent of total output. The locomotive corresponds to the productive capacity that produces goods and services for autonomous demand in a broad sense. In the locomotive there are two clearly separated engine rooms. The front part, which houses the drivers, can be identified with the capacity serving autonomous demand in a strict sense. It includes: (a) autonomous consumption by households; (b) residential investment; (c) modernisation investment by firms that transforms the existing capacity, instead of expanding it; (d) real public expenditure; and (d) exports. The machines that will expand the productive capacity of the economy are produced in the rear part of the locomotive. This is identified as the 'expansionary investment' and is explained through the accelerator mechanism.

The railway metaphor conveys several lessons whose economic translation is obvious. (1) The train will run at the speed set by the locomotive. The trucks may go slower than the locomotive during the first moments after acceleration; slower, after a deceleration. But this is a transient phenomenon; eventually the rates of growth will converge. (2) Whatever the potential of this locomotive, the actual speed of the train is determined by the engine drivers. (3) When they decide to go faster, they accelerate the locomotive (the rate of capacity utilisation). Later, to avoid burning out the engine, they switch in additional power which implies a rise in the relative weight of the rear part of the locomotive. (The acceleration of investment brings about a rise in the share of the equipment sector).

With these Schumpeterian lenses, we are now going to look at the Spanish economy. Table 6.1 and Figure 6.1 show two long waves of prosperity (1964–74 and 1997–2008) and two long recessions (1975–96 and 2008 till the present day).[1] These movements can be related to the dynamics of autonomous demand in a strict sense. The car industry was the main engine of growth during the 1960s and 1970s. It produced for the home market (durable consumption goods paid for by consumer credit) and for foreign markets – cars have been the top Spanish export since the 1960s. From 1996 on, construction became the key engine of the Spanish

Table 6.1 Key economic indicators during long cycles, Spain 1964–2010

%	1964–74	1975–96	1997–07	2008–10
g(cars) (rate of growth of car production)	21.14	5.28	1.34	–7.47
g(h) (rate of growth of residential invest)	6.85	–0.43	7.45	–19.62
g_i (rate of growth of productive investment)	10.51	2.69	7.38	–5.66
g_y (rate of growth of GDP)	5.79	2.22	3.84	–3.66
i (share of productive investment in GDP)	23.96	19.23	23.22	21.05
u (rate of capacity utilization)	83.64	78.43	80.01	73.47

Source: INE, Ministerio Industria (Spain), ANFAC, AMECO.

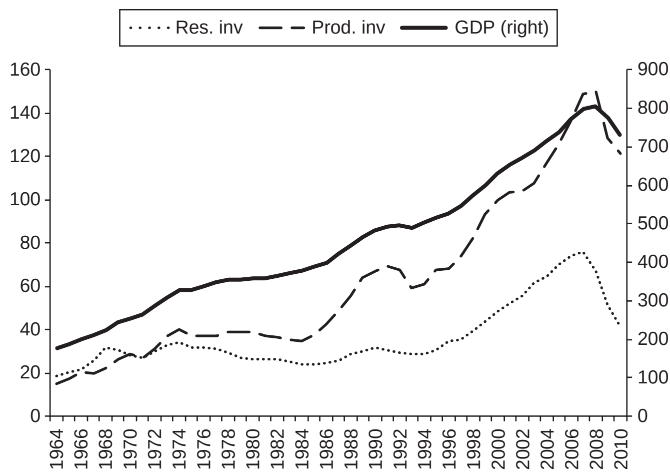


Figure 6.1 GDP, productive investment and residential investment, Spain 1964–2010 (billion Euros year 2000)
Source: See Table A (Appendix).

economy under the stimulus of easy mortgage loans. The saturation of the market was already clear in 2006. Its collapse after the financial crisis of 2007–08 plunged the economy into a serious recession.

During the long waves of prosperity accumulation accelerated, leading to a rise in the share of investment in aggregate demand (above 23 per cent) and the rate of capacity utilisation (above 80 per cent). During recessions both the investment share and the rate of capacity utilisation fell. The average rate of capacity utilisation over the 47 years under review is 80 per cent.

Table 6.2 shows the rate of capacity utilisation in the economies of the OECD for which the complete sets of data are available. We appreciate there are some level-differences that probably accord to the way each country defines normal capacity. The rate rises in boom periods but not in the steady way that would correspond with a demand expanding for several years. The low standard deviation proves that an inner force (investment) keeps the rate of capacity utilisation close to its normal position.

6.3 Expansionary investment and the flexible accelerator

The kernel of a growth model is the *accumulation of capital* that we shall identify with *productive investment of the expansionary type*. It can be called *induced investment,* provided we realise that it is different from the 'inducement' associated with consumption. Expansionary investment aims to create capacity in order to match efficiently the expected increases in demand.

According to the principle of acceleration, investment expenditure undertaken at the end of period t (I_t) can be computed as the difference between required capacity for the next and subsequent years ($K_{R,t+1}$) and installed capacity in t (K_{It}). It can be expressed by any of the following expressions which derive one from another, taking into account the definitions of the variables. D_t stands for aggregate demand in year t and g_d for its expected rate of growth. k is the optimal capital/output ratio.

Table 6.2 Rate of capacity utilisation

	Mean	**Standard deviation**
Australia	1.00	0.01
Belgium	0.78	0.03
Canada	0.76	0.11
Denmark	0.81	0.03
France	0.86	0.05
Germany	0.84	0.03
Italy	0.76	0.03
New Zealand	0.87	0.04
Norway	0.85	0.08
Spain	0.08	0.03
UK	0.79	0.07
USA	0.90	0.06

Source: Own computation with data from OECD and AMECO.

It corresponds to the normal rate of capacity utilisation that we shall normalise at unity ($u^*=1$). The actual rate is the ratio between actual and normal output: ($u_t = Y / Y^*$)[2]. $E_{kt} = K_{It} - K_{Rt}$ is the excess of capacity.

$$I_t = K_{R,\,t+1} - K_{It} \qquad [6.1a]$$

$$I_t = k \cdot D_t \cdot (1 + g_d) - \frac{k}{u_t} \cdot Y_t \qquad [6.1b]$$

$$I_t = k \cdot g_d \cdot D_t + \left(k \cdot D_t - \frac{k}{u_t} \cdot Y_t \right) \qquad [6.1c]$$

$$I_t = k \cdot g_d \cdot D_t - E_{kt} \qquad [6.1d]$$

Most of the preceding formulae are divided into two parts. The first one ($k \cdot g_d \cdot D_t$) refer to *ex ante investment*: production of the capital goods required to respond to increasing demand. The second part (that subtracts excess of capacity) refers to *ex post investment*: actual investment expenditure at the end of the year. Such a subtraction brings about undesired inventories (E_{yt}) that will be discounted from next year's demand. Actual production in year $t+1$ will be: $Y_{t+1} = D_{t+1} - E_{yt}$.

When the economy is growing along its long-period path of growth we can write $Y_t = D_t$ and $u_t = 1$ and use the pure acceleration mechanism:

$$I_t = k \cdot g_d \cdot Y_t \qquad [6.2]$$

In the general case investment depends on the expected growth of demand and the deviations from normal capacity. Both concepts have been a matter of controversy and require further clarification.

(a) Normal capacity

Competition forces firms to invest in the best (available) technology and to use it in the most efficient way, i.e., at normal capacity. After a sudden and unexpected rise in demand, firms are supposed to raise capacity utilisation to gain as many customers as possible. The actual rate of profit may also increase for a while because installed capacity is being used more hours per day. But if firms continue over-utilising capacity they risk losing customers in the next peak of demand, impairing the risk-free profitability (profits free from the risk of losing customers in the peaks of demand). The maximisation of profits in the long run (or mere survival) compels firms to adjust capacity towards its normal rate.

Ciccone is right in assuming that firms plan investment on the basis of the regular fluctuations in demand which allows for the definition of the normal rate as the average of the cycle (Ciccone, 1986; 1987). There is no reason to accelerate investment in the peak of the cycle where $u_t > 1$, since this possibility has already been taken into account in the investment decision. This conclusion does not hold, however, after permanent increases in demand like those which result from an acceleration in the autonomous trend.

It is important to draw a distinction between the normal rate of capacity and the normal rate of employment in long-period analysis. A normal rate of employment is not an equilibrium condition either in the short run or in the long run. The reason is clear: unemployment (unlike under-utilisation or over-utilisation of capacity) does not affect the profitability of firms.

(b) Expected growth of demand

Eatwell (1983) states that the proper independent variable of Keynes' *General Theory* is not investment, as such, but the long-term expectations of growth envisaged by investors (g_d). He cites the following passage:

> If we suppose a state of expectations to continue for a sufficient length of time for the effect on employment to have worked itself out so completely that there is, broadly speaking, no piece of employment going on which would not take place if the new state of expectation had always existed, the steady level of employment thus attained may be called the long period employment corresponding to that state of expectation. (Keynes, 1936, p. 48)

Keynes emphasised the importance of the 'animal spirits' of entrepreneurs to such an extent that most of his disciples have interpreted them as the prevalence of subjective factors in investment decisions. In the long-run perspective considered by investors, the 'animal spirits' of entrepreneurs are usually anchored to objective factors like the autonomous trend. In the model a persistent autonomous trend will forge expectations of demand growth. If one million houses are built in year one, 1,025 million in year two, 1,051 in year three and so on, an induced demand for machinery suggests an expected 2.5 per cent growth of demand at the factory. Even if in one year firms' production remains constant (because they were running excesses of inventories and/or capacity), they will eventually need to match the demand growth imposed by the autonomous trend. With technology constant, house-building cannot grow at a rate of 2.5 per cent if the equipment sector is not expanding production at at least the same rate.

Contrary to Palumbo and Trezzini's (2003) view, the flexible accelerator function (see equations [6.1a]–[6.1d]) does not require firms to have perfect foresight or a huge amount of information. Neither does it require the economy to be permanently in a fully adjusted path of growth. At the end of the year, firms decide investment expenditure having regard to the demand orders they receive and the ongoing imbalances.

Models that incorporate the accelerator mechanism will be stable provided demand expectations depend on objective factors (such as the autonomous trend) rather than the ups and downs of current income and capacity utilisation. The extreme instability of Harrod's (1939) model stems from the fact that the expected rate of growth depends only on psychological factors and is unduly influenced by the actual rate of capacity utilisation. Such behaviour refutes the autonomy of the autonomous trend.

6.4 Dynamic equilibrium through the supermultiplier

According to the Keynesian-Kaleckian principle of effective demand, the equilibrium level of output at a given moment does not depend on the productive capacity of the economy. Output and employment are supposed to adjust to expected aggregate demand which includes: autonomous demand (Z_t), expansionary investment by firms (I_t) and induced consumption by households:

$$Y_t = D_t = C_t + I_t + Z_t \qquad [6.3]$$

In our closed and private economy autonomous demand has been identified with residential investment. We take as given the expected demand at year t (Z_t) and its expected rate of growth (the autonomous trend, g_z). To simplify the exposition we shall assume that rentiers buy these houses with mortgage loans and let them to workers. The future stream of rents received by rentiers will be consumed.

It is well documented that induced consumption is a relatively high and stable proportion of disposable income. In a Kaleckian mood, the average propensity to consume can be computed as a weighted average of the propensities to consume of workers out of wages (c_w) and of capitalists out of profits (c_b). The proper weighting is the share of wages and profits in income (ω and β, respectively). To make the graphical representation of the model easier assume that all wages are consumed ($c_w = 1$) and all profits are saved ($s_b = 1$). In fact, workers are forced to save part of their wages in order to pay rent. But since rentiers consume all these rents we can write $c_w = 1$.

$$C_t = c \cdot Y_t = c_w \cdot \omega \cdot Y_t = \omega \cdot Y_t \qquad [6.4]$$

Let us consider, to begin with, an economy that is growing along its long-period path without excesses of capacity and inventories. The autonomous trend has prevailed long enough to anchor long-term expectations of demand, so $g_d = g_z$. Actually all variables are growing at the same rate, that we can identify with the warranted rate (g_t^*). In these circumstances we can compute investment by the pure acceleration formula given in [6.2] and determine the level of output by any of the following expressions:

$$Y_t = c \cdot Y_t + k \cdot g_t^* \cdot Y_t + Z_t \qquad [6.5]$$

$$Y_t = \frac{1}{1-c} \cdot (Z_t + I_t) = \mu \cdot (Z_t + I_t) \qquad [6.6]$$

$$Y_t = \frac{1}{1-c-k \cdot g_t^*} \cdot Z_t = \mu^* \cdot Z_t \qquad [6.7]$$

Expression [6.6] uses the simple multiplier. The equilibrium level of output is μ times the (broad) autonomous demand that includes residential investment (Z_t) and productive investment of the expansionary type (I_t). The last one can be computed *ad hoc* using any formula of the acceleration in [6.1]. Even simpler, it is g^* times the installed capacity or $(1 + g^*)$ times the previous level of investment.

Expression [6.7] combines the multiplier and the acceleration in a single expression that Serrano (1995) (after Hicks, 1950) called the super-multiplier. Output in year t is μ^* times the (strict) autonomous demand for this year (Z_t).

Such an economy will grow steadily along a fully adjusted path. Harrod (1939) called it the 'warranted path'. If the rate of growth envisaged by entrepreneurs is g^* and they increase production at this rate, the entire supply will be absorbed by demand and all the savings generated in the process of production will be invested. Dividing [6.5] through income and clearing for g^* we get the following expression:

$$g_t^* = \frac{1-c-z_t}{k} = \frac{s-z_t}{k} = \frac{i_t}{k} = \frac{s_{it}}{k} \qquad [6.8]$$

As in Harrod (1939), the warranted rate depends on technology (the optimal capital/output ratio, k) and expenditure patterns (the aggregate propensity to save, $s=1-c$). But it is also affected by the structure of the economy, which changes with the autonomous trend. Such a structure is represented by the share of residential investment in income in a given period ($z_t=Z_t/Y_t$). Alternatively it can be represented by the share in income of productive investment (i_t) that will attract, through the credit channels, an equivalent amount of savings ($s_{i,t}$).

So far the workings of an economy have been described along its long path of growth. This is not a realistic description of how capitalist economies actually grow. If the autonomous trend is truly 'autonomous', it will change from time to time. In the example in Table 6.A (in the Appendix) we suppose that after year three g_z rises from 0.025 to 0.5. Firms adjust capacity to match the increasing demand and, at the end of the year, speed up investment over and above the level corresponding to the pure accelerator process. Consider [6.1d]: $I_t=k \cdot g_d \cdot D_t - E_{kt}$. A negative E_{kt} means a shortage of capacity that is filled by depleting inventories (negative E_{yt}). Next year production will be higher than the level justified by expected demand since firms have to replenish their stocks. The actual rate of growth of this economy can be replicated by the following formula where the two adjustment mechanisms are apparent: in the short run, the actual rate of capacity utilisation (u_t); in the long run, the structure of demand (z_t):

$$g_{y,t} = \frac{(1-c-z_t)\cdot u_t}{k} \qquad [6.9]$$

Figure 6.2, derived from the data in table A (Appendix), presents the traverse from one equilibrium path to another. For the given value of the parameters the adjustment is cyclical. The rate of capacity utilisation rises to recover cyclically its normal position ($u_t=1$).[3] Overutilisation of capacity speeds up the pace of accumulation raising the share of investment in income (i) until it enables a new steady growth path ($g_y=g_d=g^*=g_z$).

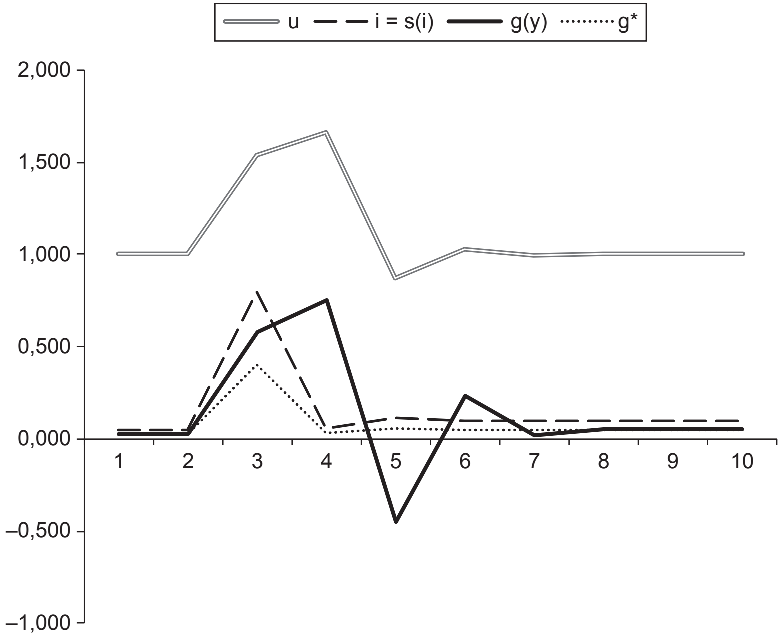


Figure 6.2 The traverse (after an acceleration of the autonomous trend)

Figure 6.3 shows the structural adjustments that make the warranted rate equal to the autonomous trend. Given technology (represented by the labour coefficient, *l,* and the optimal capital coefficient, *k*), we can draw the frontiers of distribution and growth. In the north-east quadrant we detect that the result of a rise in the autonomous trend is an increase in the share of expansionary investment in income ($i_t = I_t / Y_t$ rises). It occurs at the expense of per-worker consumption-type expenditure ($ç_t$); more specifically, at the expense of autonomous demand (z_t) because we have assumed a constant propensity to consume. The result looks paradoxical. The fall of z comes after the acceleration in autonomous demand (g_z). As a matter of fact the yearly production of houses has increased a lot but, since income has increased faster, the ratio Z_t/Y_t has fallen.

The south quadrants of Figure 6.3 look at the financial side of the economy; more precisely at the part related to business loans for expansionary investment. To simplify the exposition we shall suppose that all types of investment are financed by bank loans allocated on a demand basis.

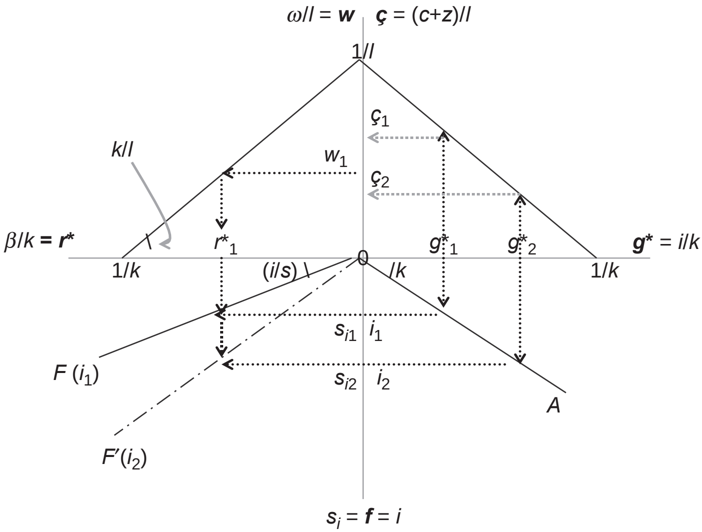


Figure 6.3 Structural adjustments towards the new autonomous trend

The demand for credit to finance accumulation (in the south-east) is a function of the rate of growth:

$$f_t^d = i_t = g_t^* \cdot k \qquad [6.10]$$

The supply of credit for expansionary investment is supposed to match this demand. It will absorb the savings generated in production that are not lent to rentiers. In relative terms it amounts to $s_{it} = 1 - c - z_t = s - z_t$. The equilibrium share of accumulated savings may also be obtained by multiplying the unit profit ($r^* \cdot k = \beta$), times the propensity to save out of profits ($s_b = 1$), times the share of expansionary investment in total savings:[4]

$$f_t^s = s_{it} = s - z_t = \beta \cdot s_b \cdot (i_t / s) \qquad [6.11]$$

The sub-index '*t*' borne by s_i, *i* and z remind us that these parameters adjust when the economy shifts to a different path of growth. The remaining ones (k, r^*, β, s, s_b) keep at their initial level. The term in brackets is the structural variable that makes the new path of growth possible and sustainable. Angle <r0F> opens whenever *i* increases.

Note that the change in the supply of savings for accumulation purposes has been driven neither by technology ($k_t = k$, $u_t = 1$), nor by distribution ($w = w_1$ and $r^* = r^*_1$),[5] nor by the saving propensity ($s = 1 - c$; $s_p = 1$). It is simply the result of the acceleration of expansionary investment in order to recover normal capacity after the acceleration in the autonomous trend.

A final word about the requirements of the supermultiplier model and the maximum autonomous trend compatible with it. The model computes the consumption and investment derived from the production of goods feeding autonomous demand. When autonomous demand is negligible we cannot compute and use the supermultiplier given in [6.7] (for the same reason that we cannot compute the simple multiplier when income is entirely devoted to consumption). We can still use the multiplier-accelerator model (our expression [6.6]) which would lead to the following equilibrium rate:

$$\hat{g}^* = \frac{1-c}{k} = \frac{s}{k} \qquad [6.12]$$

This is Harrod's warranted rate that now indicates the maximum rate of growth of autonomous demand that could be efficiently matched given technology (k) and expenditure patterns (c,s). A further increase in the autonomous trend cannot be attended by reducing the weight of autonomous demand. Firms will over-utilise capacity and u_t will be above 1, despite the effort of investors. Yet (as shown by Dejuán, 2005) the model continues to be stable and u_t converges to the following expression (it derives from [6.9], after equating $z_t = 0$):

$$u'' = \frac{k \cdot g_z}{1-c} = \frac{k}{s} g_z = \frac{g_z}{g^*} \qquad [6.13]$$

This recalls the thesis of the endogeneity of the utilisation rate (Palumbo and Trezzini, 2003; Hein *et al.*, 2011). Two provisos: (1) this applies only to the elementary model ($Y = C + I$) which is only found in text books; and (2) over-utilisation is restricted by technology.

According to Schumpeter (1912), capitalist economies overcome their natural tendency to stagnation thanks to the presence of innovative entrepreneurs who launch the production of new goods or traditional goods for new markets. The supermultiplier model we have studied is meant for this type of economy.

6.5 Conclusions

This paper extrapolates the Keynesian-Kaleckian principle of effective demand to the long run, in an attempt to build a demand-led growth theory coherent with the surplus approach. It has been shown that the autonomous trend determines not only the actual rate of growth of actual output, but also the warranted (normal capacity) rate. It 'rules the roost' of growth rates provided it is persistent enough. At least there is a tendency towards it, whenever firms try to be efficient and maximise profits in a long-run perspective.

The flexible supermultiplier nets two adjustment mechanisms. After the acceleration in the autonomous trend, firms will increase the utilisation rate to match demand. Afterwards they will speed up accumulation until the share of the expansionary investment in aggregate demand makes it possible to follow a new fully adjusted path of growth. The rate of capacity utilisation is endogenously determined in the short run. In the long run, the accumulation of capital re-establishes normal capacity utilisation and it is the warranted rate which becomes endogenous.

Contrary to the dominant opinion, it is contended that the multiplier-accelerator is a stabilising mechanism, leading the economy to fully adjusted positions (usually in a cyclical way). Three conditions are required to accomplish this task:

1. Autonomous demand and its expected rate of growth are truly autonomous.
2. Investment decisions are based on long-term expectations of demand growth that will be anchored on the autonomous trend provided it lasts long enough.
3. Investors take into account the excesses or shortages of capacity and inventories to get rid of them.

This conclusion stands as a critique to the Kaleckian and Sraffian models referred to in the introduction, which load the burden of adjustment (both in the short and in the long run) on the rate of capacity utilisation. It is felt that they do not take the restrictions imposed by technology and enforced by competition seriously enough. In a competitive economy firms are compelled to use the best technology in the best way. Firms can accommodate stationary fluctuations in demand and ordinary cycles by pushing up the rate of capacity utilisation. But if increases in demand are permanent (as happens when the autonomous trend rises), firms are bound to speed up investment to avoid the risk of losing customers in the next seasonal peak of demand.

The adjustment brought about by the supermultiplier does not require changes in distribution and/or in savings propensities, although they would make the adjustment towards a new autonomous trend easier. Accord is found with Palumbo and Trezzini's (2003) criticism of the first post-Keynesian economists referred to in the introduction: it is not sensible to assume a fall in wages during the boom, just to ensure the continuity of a mythical steady state. But it is felt that Palumbo and Trezzini are wrong when they conclude that an inbuilt tendency towards normal capacity implies perfect foresight in a steady-state pattern of growth. It has been shown that the flexible accelerator is able to work out of equilibrium. Nor should it be concluded that the endogeneity of the warranted rate implies that savings determine investment. After a change in the autonomous trend income will grow until the yearly generation of savings coincides with the yearly investments required by the new trend. The structural change occurs in the allocation of these savings between productive and residential investment (capacity-creating and non-capacity creating investment).

The opinion presented here is that the multiplier-accelerator model goes a long way in the determination of a theory of normal output. It is different but complementary and compatible with the Sraffian theory of normal prices. Both theories share the same set of exogenous data: technology and distribution (the real wage, as proposed). In addition, the theory of normal output includes certain expenditure patterns and the expected rate of growth of autonomous demand. The resulting theory of output shows (1) that the actual rate of growth of GDP is ultimately determined by the autonomous trend; and (2) that the structure of demand and productive capacity tends to accommodate to the requirements of the autonomous trend.

Palumbo and Trezzini rightly point out that the autonomous trend lacks the persistency of the traditional exogenous variables of the surplus approach (labour productivity, the real wage, workers' propensities to consume and save and so on). Certainly, in epochs without a well-established autonomous trend, the economy will suffer from macroeconomic instability. Yet the tendency towards the normal rate of growth and the proper structure of demand is always at work through the investment function. These tendencies will carve a path when the autonomous trend is persistent enough. China, an export-led economy with modern technology and low labour costs, can achieve a normal path of growth that trebles the path of the USA. So we can expect that the share of expansionary investment on income will be much higher in the Chinese economy. The rate of capacity utilisation, however, will be around its normal level in both countries. Within a given country, we can distinguish long waves of prosperity and depression. We saw in the comment about Table 6.2 that the capacity rate rises in boom periods but remains tied to the normal one.

Once the core of economic theory is well established, researchers are supposed to focus on the institutional framework that determines the exogenous data and the interactions between them. During the economic boom prior to the first great depression of the 21st century, for example, economists should have asked the following questions: For how long can residential investment act as a locomotive if the house prices are increasing three times faster than nominal wages? Can banks fill the gap with mortgage credits without leading households into a debt trap?

Notes

1. Inside these long waves we find cyclical movements of a shorter length. Within the long recession from 1975 to 2006 there was an economic boom between 1987 and 1992. In the last long wave of expansion there was a bust in the first two years of the 21st century. Car production is not included in Figure 6.1 because the data are in physical units.
2. Alternatively, it could be defined by the ratio between required and installed capacity $u_t'=K_{Rt}/K_{It}$. Both ratios coincide when there are no undesired stocks. They are accounted for in u_t, not in u_t'.
3. The table does not consider any restriction in the adjustment process. If we fix maximum and minimum levels for u_t and e_t the adjustment will be longer and smoother. If we consider that only a part of undesired inventories can be stored, the adjustment will be faster. There are many variants of the model that alter the shape of the process without changing the final results.
4. In our numerical example this ratio equals the share in profits of the savings financing expansionary investment: $S_I/B=I/S=(I/Y)/(S/Y)=i/s$.
5. The last column of Table 6.A (Appendix) shows that the actual rate of profit (r_t) is positively related to capacity utilisation.

References

Bortis, H. (1997) *Institutions, Behaviour and Economic Theory* (Cambridge, UK: Cambridge University Press).

Cesaratto, S. (2012) 'NeoKaleckian and Sraffian controversies on accumulation theory', unpublished paper.

Ciccone, R. (1986) 'Accumulation and capacity utilization: some critical considerations on Joan Robinson's Theory of Distribution', *Political Economy, Studies in the Surplus Approach*, 2: 17–36.

Ciccone, R. (1987) 'Accumulation, capacity utilization and distribution: a reply', *Political Economy, Studies in the Surplus Approach*, 3: 97–111.

Dejuán, Ó. (2005) 'Paths of accumulation and growth: towards a Keynesian long-period theory of output', *Review of Political Economy*, 17: 231–52.

Eatwell, J. (1983) 'The long period theory of employment' *Cambridge Journal of Economics*, 7: 269–85.

Eatwell, J. (2012) 'The theory of value and the foundations of economic policy', *In memoriam of Pierangelo Garegnani*. Roma.

Eatwell, J. and Milgate, M. (eds) (1983) *Keynes's Economics and the Theory of Value and Distribution* (New York: Oxford University Press).

Garegnani, P. (1983) 'Two Routes to Effective Demand', in J. Kregel (ed.) *Distribution, Effective Demand and International Economic Relations* (London: Macmillan).

Garegnani, P. (1992) 'Some notes for an analysis of accumulation', in J. Halevi, D. Laibman and E. J. Nell (eds), *Beyond the Steady State. A Revival of Growth Theory* (London: Macmillan).

Harrod, R. F. (1939) 'An essay in dynamic theory', *Economic Journal*, 49: 14–33.

Hein, E., Lavoie, M. and Treeck, T. V. (2011) 'Some instability puzzles in Kaleckian models of growth and distribution: a critical survey', *Cambridge Journal of Economics*, 35: 587–612.

Hicks, J. R. (1950) *A Contribution to the Theory of the Trade Cycle* (Oxford: Clarendon).

Kaldor, N. (1955–56) 'Alternative theories of distribution', *Review of Economic Studies*, 23: 63–100.

Kaldor, N. (1961) 'Capital Accumulation and Economic Growth', in F. A. Lutz and D. C. Hague (eds), *The Theory of Capital* (New York: St Martins).

Kalecki, M. (1971) *Selected Essays in the Dynamics of the Capitalist Economy (1933–1970)* (Cambridge: Cambridge University Press).

Keynes, J. M. (1936) *The General Theory of Employment, Interest and Money* (London and New York: Harcourt Brace Javanovich).

Lavoie, M. (2010) 'Surveying Short-run and Long-run Stability Issues with the Kaleckian Model of Growth', in M. Setterfield (ed.), *Handbook of Alternative Theories of Economic Growth* (Cheltenham, UK: Edward Elgar).

Milgate, M. (1982) *Capital and Employment. A Study of Keynes's Economics* (London: Academic Press).

Palumbo, A. and Trezzini, A. (2003) 'Growth without normal capacity utilization', *European Journal of the History of Economic Thought*, 10: 109–35.

Robinson, J. (1956) *The Accumulation of Capital* (London: Macmillan).

Robinson, J. (1962) *Essays on the Theory of Economic Growth (*London: Macmillan).

Schumpeter, J. (1912) *Theorie der Wirtschaftlichen Entwicklung* [*The Theory of Economic Development*] (Leipzig: Verlag Dunker & Humbolt).

Serrano, F. (1995) 'Long period effective demand and the Sraffian supermultiplier', *Contributions to Political Economy*, 14: 67–90.

Trezzini, A. (1995) 'Capacity utilisation in the long run and the autonomous components of aggregate demand', *Contributions to Political Economy*, 17: 33–66.

Trezzini, A. (1998) 'Capacity utilisation in the long run: some further considerations', *Contributions to Political Economy*, 17: 53–67.

APPENDIX

Data

Technology: $\boldsymbol{l}=0.8$, labour coefficient; $\boldsymbol{k}=2$, optimal or normal capital/output ratio.

Distribution: $\boldsymbol{w}=1$, real wage. $\boldsymbol{r^*}=0.1$ (normal rate of profit corresponding to $k=2$ and $w=1$). The shares in income of wages and profits are also constant: $\omega=w\cdot l=0.8$; $\beta=1-\omega=0.2$.

Expenditure patterns: $\boldsymbol{c_w}=1$, propensity to consume out of wages; $\boldsymbol{s_b}=1$, propensity to save out of profits. The aggregate propensity to consume and save are also constant: $\boldsymbol{c}=c_w\cdot\omega=0.8$; $\boldsymbol{s}=s_b\cdot\beta=0.2$.

Autonomous trend: $\boldsymbol{g_z}$. It starts at 0.025. In period 3 it shifts to 0.05. (Remember to change the g incorporated in columns Z, D and I. Also in the supermultiplier, that rises from 6.66 to 10 when g_z doubles). The initial value of autonomous demand is $Z_t=1$.

Definitions of variables

$\boldsymbol{Z_t}$: Proper autonomous demand that we identify with residential investment.
$\boldsymbol{D_t}=\mu^*\cdot Z_t$, aggregate demand to which production adjusts. $\mu^*=1/(1-c-k\cdot g)$, supermultiplier.
$\boldsymbol{Y_t}=D_t-E_{y(t-1)}$, actual production (after discounting the inherited excess of inventories).
$\boldsymbol{K_{Rt}}=k\cdot D_t$, required capital.
$\boldsymbol{K_{It}}=K_{I(t-1)}+I_{t-1}=(k/u_t)Y_t$, installed capacity at the beginning of period t.
$\boldsymbol{E_{kt}}=K_{It}-K_{R,t}$ excess of capacity.
$\boldsymbol{L_t}=l\cdot Y_t$, labour employed.
$\boldsymbol{W_t}=w\cdot L_t$, mass of wages.
$\boldsymbol{B_t}=r\cdot K_{I,t}=Y-W$, mass of profits ('benefits').
$\boldsymbol{C_t}$: $c\cdot Y_t$, induced consumption. (It coincides with wages, although a part of C corresponds to the consumption of rentiers from the rents paid annually by working households).
$\boldsymbol{S_t}=s\cdot Y_{t.}=s_b\cdot B_t$
$\boldsymbol{I_t}=K_{R,t+1}-K_{It}=\mathrm{k}^*\cdot D_t\cdot(1+g_z)-K_{It}=k^*\cdot g_z\cdot D_t-E_{kt}$, *ex post* productive investment (of the expansionary type).
$\boldsymbol{E_{yt}}=Y_t-(C_t+I_t+Z_t)$, excess of inventories or undesired stocks. A negative number implies that inventories are below the desired level. It coincides with the excess of savings: $E_{yt}=S_t-(I_t+Z_t)$.
$\boldsymbol{e_t}=E_{yt}/Y_t$, share in income of undesired inventories.
$\boldsymbol{u_t}=Y_t/Y_t^*$ rate of capacity utilisation. (It accounts for unwanted inventories that are not anticipated when capacity utilisation is defined as $u_t=K_{Rt}/K_{It}$).
$\boldsymbol{z_t}=Z_t/Y_t$, share in income of autonomous demand.
$\boldsymbol{i_t}=I_t/Y_t$, share in income of expansionary investment. It coincides with the share of the savings financing expansionary investment: $s_{it}=(Y_t-C_t-Z_t)/Y_t=1-c-z_t$.
$\boldsymbol{g_{yt}}=(Y_t-Y_{t-1})/Y_{t-1}$, actual rate of growth of output.
$\boldsymbol{g_t^*}=i/k=(1-c-z_t)/k$, warranted rate of growth.
$\boldsymbol{r_t}=B_t/K_{It}$ actual rate of profit.

Table 6.A Adjustment towards a new autonomous trend

t	Z	D	Y	KR	KI	E(k)	L	W	B	C	S	I	E(y)	e	u	z	i=s(i)	g(y)	g*	r
1	1,000	6,667	6,667	13,333	13,333	0,000	5,333	5,333	1,333	5,333	1,333	0,333	0,000	0,000	1,000	0,150	0,050	0,025	0,025	0,100
2	1,025	6,833	6,833	13,667	13,667	0,000	5,467	5,467	1,367	5,467	1,367	0,683	–0,342	–0,050	1,000	0,150	0,100	0,025	0,050	0,100
3	1,076	10,763	11,104	21,525	14,350	–7,175	8,883	8,883	2,221	8,883	2,221	8,251	–7,107	–0,640	1,548	0,097	0,743	0,625	0,372	0,155
4	1,130	11,301	18,407	22,601	22,601	0,000	14,726	14,726	3,681	14,726	3,681	1,130	1,421	0,077	1,629	0,061	0,061	0,658	0,031	0,163
5	1,187	11,866	10,444	23,731	23,731	0,000	8,355	8,355	2,089	8,355	2,089	1,187	–0,284	–0,027	0,880	0,114	0,114	–0,433	0,057	0,088
6	1,246	12,459	12,743	24,918	24,918	0,000	10,195	10,195	2,549	10,195	2,549	1,246	0,057	0,004	1,023	0,098	0,098	0,220	0,049	0,102
7	1,308	13,082	13,025	26,164	26,164	0,000	10,420	10,420	2,605	10,420	2,605	1,308	–0,011	–0,001	0,996	0,100	0,100	0,022	0,050	0,100
8	1,374	13,736	13,747	27,472	27,472	0,000	10,998	10,998	2,749	10,998	2,749	1,374	0,002	0,000	1,001	0,100	0,100	0,055	0,050	0,100
9	1,442	14,423	14,421	28,846	28,846	0,000	11,536	11,536	2,884	11,536	2,884	1,442	0,000	0,000	1,000	0,100	0,100	0,049	0,050	0,100
10	1,514	15,144	15,144	30,288	30,288	0,000	12,115	12,115	3,029	12,115	3,029	1,514	0,000	0,000	1,000	0,100	0,100	0,050	0,050	0,100

7
The Pattern of Economic Growth of the Brazilian Economy 1970–2005: A Demand-Led Growth Perspective

Fabio N. P. de Freitas and Esther Dweck

7.1 Introduction

The present work analyses the pattern of economic growth prevailing in Brazil from 1970 to 2005. The analytical framework for this investigation is based on the classical supermultiplier demand-led growth model combined with the hypothesis that the balance of payments was the main potential (and often the most effective) constraint for the expansion of the Brazilian economy during the period in focus. We adopt a demand-led growth accounting methodology to analyse the Brazilian economy, based on a multi-source database compiled for this investigation.

The chapter is organised as follows. First, the Brazilian growth experience is put into context. Growth during the period 1970–2005 is compared to previous periods, starting from 1900. The Brazilian experience is also contrasted with the growth trajectory of other reference countries and the world economy. The following section presents our theoretical framework, as well as the empirical methodology adopted for this investigation, arguing that the labour constraint is not a prevalent characteristic of the Brazilian growth experience. In Section 7.4, the analytical framework is applied to investigating the pattern of economic growth prevailing in Brazil from 1970 to 2005. The last section provides the main results and our conclusions.

7.2 The Brazilian economic growth experience in perspective

Brazil's was one of the fastest-growing economies in the twentieth century. From the early 1900s until the end of the World War II (WWII), it grew at a relatively high average rate (3.8 per cent per year approximately). However, it also experienced huge instability, which was an important feature of the primary export pattern of development that characterised the period. The Brazilian economy performed particularly well from the beginning of the state-led industrialisation process after the end of WWII, until the Latin America crises in the 1980s. Indeed, as Figure 7.1 shows, from the late 1950s until the 1970s Brazil's economy grew at an average annual rate of 6.8 per cent.

The worst periods in terms of growth performance were the 1980s and the 1990s, when the country grew at an average rate of 2 per cent per year. Prolonged stagnation due to the Latin America crises started at the beginning of the 1980s and continued in the 1990s, when intensive liberal reforms were implemented. In the 2000s, the Brazilian economy recovered and grew at higher rates (3.3 per cent per year on average), in line with the better performance of the world economy (see Table 7.2 ahead). Even so, the country was not able to recover the expansion rate pace that prevailed in the primary export period at the beginning of the last century.

It is worth comparing the Brazilian performance with that of other reference countries. First, the levels of GDP per capita of Brazil, South Korea and the USA will be compared. The US economy is the benchmark country

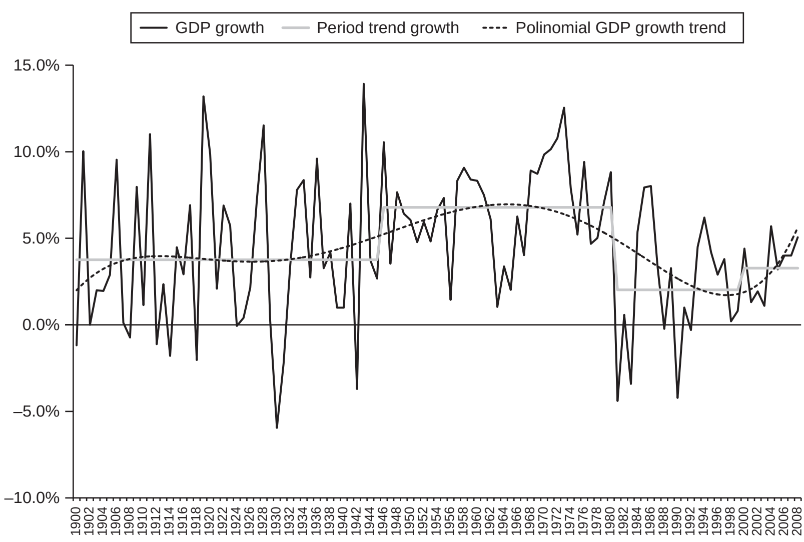


Figure 7.1 Brazilian annual real growth rates, 1900–2008
Source: Authors' elaboration based on Maddison (2010).

for the convergence analysis. South Korea has been included to allow for a contrast between Brazil and a successful economy in the Asian periphery.

Figure 7.2 and Table 7.1 show that both Brazil and South Korea engaged in a process of catching up with the US economy until the 1980s. Starting from relative shares of, respectively, 17.5 and 8.9 per cent of US per capita GDP in 1950, Brazil and South Korea achieved relative shares of 28.0 and 22.1 per cent in 1980, respectively. After that, the Latin America crises interrupted the Brazilian convergence process, and the country lagged behind until the beginning of the 2000s. Indeed, after two decades of stagnation Brazil reached a relative share of 19.4 per cent of US per capita GDP in the late 1990s. In the following decade, it managed to recover only modestly, achieving 20.6 per cent in 2008. In contrast, South Korea's convergence process continued in the period under review and, despite a brief interruption due to the Asian Crisis, the country reached a relative share of 62.9 per cent of US per capita GDP in 2008.

Another revealing comparison can be drawn between the economic growth experiences of Brazil, South Korea, the United States of America and the world economy. Indeed, in Table 7.2 one can observe the same pattern of convergence and divergence in terms of GDP per capita levels as discussed before. Additionally, one can also observe that Brazil only attained a higher growth rate than the world economy in the state-led

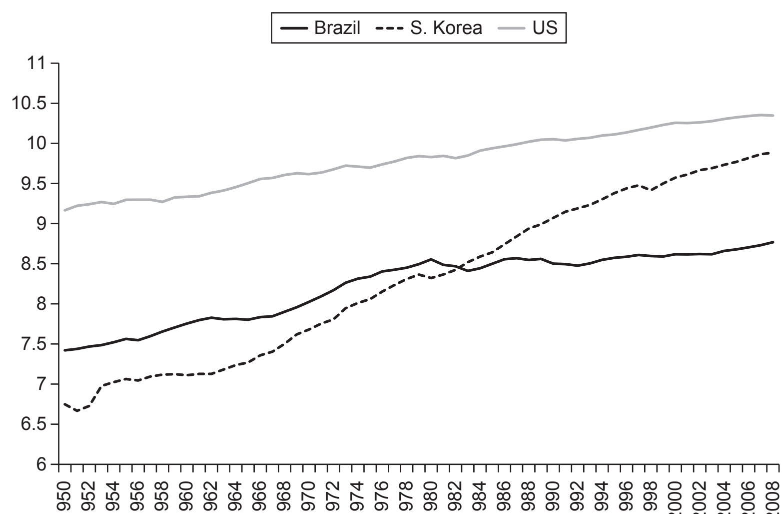


Figure 7.2 Per capita GDP levels 1950–2008, Brazil, South Korea and USA
Source: Authors' elaboration based on Maddison (2010).

Table 7.1 Relative per capita GDP (% of US per capita GDP)

Year	Brazil	South Korea
1950	17.5	8.9
1980	28.0	22.1
2000	19.4	50.5
2008	20.6	60.9

Source: Authors' elaboration based on Maddison (2010).

Table 7.2 Average real GDP growth rates (% per year)

Periods	Brazil	South Korea	US	World
1950–1980	6.8	7.5	3.6	4.5
1980–2000	2.0	7.5	3.2	3.0
2000–2008	3.3	4.4	2.1	4.2

Source: Authors' elaboration based on Maddison (2010).

development phase that lasted from 1950 to 1980, while South Korea managed to maintain an expansion rate above the world economy's for the whole period in focus (1950–2008).

The comparison drawn above provides the background to the analysis of our main subject. In the period concerned, Brazil presented its highest growth rates in the 1970s, experienced a long-lasting period of stagnation in the 1980s and 1990s, finally to present a modest recovery of GDP growth rates at the end of the 2000s. So, the present study aims to analyse the key features of the pattern of economic growth shown by the Brazilian economy in that period.[1]

7.3 The analytical framework

The Brazilian growth experience since the 1970s has been the subject of important debates. More recently, the literature[2] dedicated to interpreting and explaining the huge decline in Brazilian growth rates has been influenced by neoclassical economic theory, according to which the price system is supposed to convey information on the relative scarcity of resources, which is transmitted to consumers and producers and drives the choices they make towards the full utilisation of available resources. Hence, long-term economic growth should be characterised as a supply-constrained process, in which the rate of expansion depends on the growth of capital and labour inputs available to the economy, as well as on productivity growth. Therefore, according to that interpretative literature, the causes behind the trend break in Brazil's GDP growth rate would be: low domestic saving rate, low investment in human capital, and low (or even negative) growth rate of total factor productivity. These proximate causes have been subject to quantitative evaluation using a supply-side growth accounting empirical approach, inspired by the neoclassical growth theory. Furthermore, the literature referred to has also attempted to identify the more fundamental causes behind the decline in the Brazilian GDP growth trend. It has been suggested that 'market-unfriendly' institutions and pervasive market (and government) imperfections inherited from the post-WWII state – inducing an inward-oriented development strategy – are behind the poor growth performance of the Brazilian economy since the 1980s.

Nevertheless, contrary to the usual neoclassical viewpoint, one can argue that the Brazilian economy has not been normally constrained by the availability of resources in general and of labour in particular. As occurs in many developing economies, Brazil's economy shows a high degree of structural heterogeneity and a significant labour surplus. The great disparity in the levels of labour productivity observed across economic sectors is a quantitative expression of that kind of heterogeneity.

Table 7.3 shows significant levels of cross-sector divergence in labour productivity for the whole period. The low relative productivity of agriculture, forestry and fishing in the 1970s indicates the existence of a labour surplus in this sector. As Brazil grew and developed, labour productivity in this sector increased at a rate above the average productivity growth rate of the economy and, at the same time, the sector's employment share declined. Yet, in the same period, due to rapid urbanisation, the employment shares of wholesale and retail trade, hotels and restaurants, and community, social and personal services increased and their labour productivity grew at a lower rate than the average productivity rate of the economy. It is worth noting that, in contrast to the former's growth movement, this latter contributed to the increase in structural heterogeneity. So the Brazilian industrialisation-*cum*-urbanisation process was characterised by the reduction of labour surplus in primary activities, alongside an increase in mostly informal low-productive, low-paid services localised in urban centres. In sum, the industrialisation process was not able to produce a decline in the degree of structural heterogeneity or to eliminate the surplus labour existing in the Brazilian economy. Hence, it seems implausible that a generalised labour force

Table 7.3 Relative productivity, employment structure and labour productivity growth by sector

Variable	Sector	Agriculture, Forestry, and Fishing	Mining and Quarrying	Manufacturing	Public Utilities	Construction	Wholesale and Retail Trade, Hotels and Restaurants	Transport, Storage, and Communication	Finance, Insurance, and Real Estate	Community, Social and Personal Services (including Government Services)	Sectoral Sum
Relative Sectoral Labour Productivity levels	1970	19	269	180	133	142	83	99	460	164	100
	1980	17	205	190	250	138	69	141	357	110	100
	1990	28	372	143	470	135	41	132	377	95	100
	2000	37	646	166	1010	141	36	122	267	92	100
	2005	45	620	167	888	134	33	119	251	88	100
Sectoral Employment Structure	1970	47.9%	0.5%	14.2%	1.1%	6.1%	9.6%	3.3%	4.0%	13.2%	100%
	1980	37.2%	0.6%	13.6%	0.7%	8.5%	10.9%	3.2%	6.3%	19.0%	100%
	1990	25.5%	0.6%	15.5%	0.6%	6.7%	16.5%	3.9%	6.5%	24.3%	100%
	2000	20.7%	0.4%	13.0%	0.3%	6.2%	19.7%	4.3%	6.5%	28.9%	100%
	2005	18.7%	0.4%	13.1%	0.4%	5.6%	21.1%	4.5%	6.8%	29.4%	100%
Sectoral Labour Productivity growth	1970–1980	3.4%	2.0%	5.4%	11.7%	4.5%	2.9%	8.6%	2.2%	0.7%	4.8%
	1980–1990	3.0%	4.1%	–4.7%	4.4%	–2.1%	–6.9%	–2.6%	–1.4%	–3.3%	–1.9%
	1990–2000	4.0%	6.7%	2.5%	9.0%	1.4%	–0.4%	0.2%	–2.5%	0.6%	1.0%
	2000–2005	4.1%	–0.7%	0.2%	–2.4%	–0.9%	–1.8%	–0.5%	–1.0%	–0.6%	0.1%

Sources: Authors' elaboration based on Groningen Growth and Development Centre 10-sector Database (June, 2007), and Timmer and de Vries (2007).

constraint was a prevalent characteristic of the Brazilian economy's growth trajectory during the period under study.

The behaviour of labour productivity growth in relation to the GDP growth path that we observe supports the last conclusion. In fact, as shown in Figure 7.3, both series present a similar behaviour and move normally in the same direction. According to the neoclassical standpoint this phenomenon could be explained by assuming that the behaviour of the labour productivity growth rate drives that of GDP. Certainly, assuming that the economy has achieved almost continuous full employment, movements in the former are expected to produce similar movements in the latter. But once it is assumed that the Brazilian economy in the period under study was not characterised by the existence of continuous full employment, that kind of explanation completely loses its plausibility, and an alternative must be offered. A potential candidate is the hypothesis subjacent to the so-called 'Kaldor-Verdoorn Law' literature[3], according to which the latter empirical regularity suggests that the behaviour of GDP growth is the cause behind the observed labour productivity growth movements. This would be the case due to the existence of static and dynamic economies of scale associated with the process of economic growth.

Figure 7.3 Labour productivity and GDP growth rates, Brazil 1951–2009
Source: The Conference Board Total Economy Database (September 2010).

Consequently, conceiving economic expansion as a demand-led process, an increase in the demand-induced growth rate is thought to push labour productivity growth upward and to soften the impact of faster growth on labour requirements (and *vice versa*). Generally, the inducement we are analysing is not strong enough to prevent the GDP growth rate from being positively related to the employment growth rate. It only implies that the values of output-elasticity of employment range between zero and one. We conclude, therefore, that the availability of the labour force was not an effective constraint on the long-term expansion of the Brazilian economy during the period under review.

7.4 Theoretical background

The present chapter examines the hypothesis that the Brazilian economic expansion of 1970–2005 was a demand-led growth process subject to a balance of payments constraint. From the theoretical standpoint our analytical framework is based on the classical (or Sraffian) supermultiplier demand-led growth model proposed by Serrano (1995 and 1996).[4] A simplified, small open economy version of the model has been used as a reference. In this section we will briefly discuss the main hypothesis and the implications of the model adopted for this investigation.

First, the equilibrium between real aggregate supply and aggregate demand will be considered. By using the maximum number of components on the demand side that our database allows, the following equation is obtained:

$$Y + M = C_{HND} + C_{HD} + I_H + C_G + I_G + I_{SE} + I_{PE} + X \qquad [7.1]$$

where Y is the gross domestic product; M is imports; C_{HND} is household non-durable consumption; C_{HD} is household durable consumption; I_H is household (residential) investment; C_G is government consumption; I_G is government investment; I_{SE} is state-owned enterprises investment; I_{PE} is private enterprises investment; and X refers to exports.

Now, it is assumed that total imports are related to total aggregate demand as expressed in the equation below:

$$M = (1 - \mu)\,(C_{HND} + C_{HD} + I_H + C_G + I_G + I_{SE} + I_{PE} + X) \qquad [7.2]$$

where μ is the share of domestic content in total demand.

Additionally, it is also assumed that household non-durable consumption and private enterprise investment are induced expenditures:

$$C_{HND} = cY \qquad [7.3]$$

$$I_{PE} = hY \qquad [7.4]$$

where c is the propensity to consume non-durables, and h is the propensity of private enterprises to invest. Finally, Z designates the total autonomous expenditure, as follows:

$$Z = C_{HD} + I_H + C_G + I_G + I_{SE} + X \qquad [7.5]$$

According to the supermultiplier model, aggregate consumption expenditures contain an induced component related to the purchasing power introduced into the economy by current production decisions and is usually associated with the wage bill generated by such decisions. So, one major determinant of the propensity to consume (c) is the wage share in aggregate income. As regards aggregate private enterprise investment, equation [7.4] is intended to capture how the level of economic activity can influence the investment decisions of that particular agent. Private enterprises' propensity to invest (h) is considered an endogenous variable in the classical supermultiplier model, since the behaviour of h can be explained by the deviations between the realised output capacity utilisation rate and the normal capacity utilisation rate. In this sense, the process of capitalist competition is expected to induce an increase of h insofar as a positive deviation occurs and *vice versa*. The variation in h will then induce an adjustment of productive capacity to match aggregate demand. As a result, the model predicts that the level of capacity utilisation tends to gravitate around its normal level. Furthermore, the model also predicts that the average value of h is thought to be positively related to the GDP growth rate trend.

Yet, substituting expressions [7.3], [7.4] and [7.5] in equations [7.1] and [7.2], and then solving the resulting system for the real GDP level will result in the following expression for the real GDP equilibrium level (Y^*):

$$Y^* = \left(\frac{\mu}{1 - \mu(c + h)}\right) Z = \alpha Z \qquad [7.6]$$

where α is the supermultiplier, which, as shown in equation [7.6], depends on how the propensity to consume, the propensity to invest and the share of domestic content in demand behave.

Hence, according to the model, the behaviour of the supermultiplier (α) and the total autonomous expenditures (Z) can explain GDP behaviour. It is worth noting that the variables contained in the supermultiplier formula exert their influence on GDP growth rates exclusively by means of their rate of change – and not by their levels. This feature is significant for the analysis of the relation between functional income distribution and economic growth. According to the supermultiplier model, the main channel through which the share of wages (or the share of profits) affects the GDP growth rate is by influencing the propensity to consume and, hence, the supermultiplier itself. However, according to the model, the propensity to consume depends on the level of wage share. So, in order for this variable to exert its influence on the GDP expansion rate it must present a positive or negative growth rate. Therefore, in the classical supermultiplier model, functional income distribution can influence GDP growth insofar as it varies. In other words, a *once and for all* change in income distribution can only have a *temporary* effect on GDP growth rate (that is, it produces only a *level* effect on GDP).[5]

In our theoretical framework the balance of payments constraint is the main obstacle to economic expansion. It is a financial constraint on demand-led growth that is related to the availability of international currency, which in the period under study means basically the availability of dollars (Serrano, 2003; Medeiros and Serrano, 1999). Indeed, the accumulation of foreign reserves can keep the overall balance of payments in surplus. On the other hand, a deficit position cannot be sustained if the country in deficit is not an issuer of internationally accepted currency, such as the USA in the postwar period and, notably, after the end of the Bretton-Woods system. This is because a deficit position may set forces in train that are strong enough to limit the level of activity and the pace of economic expansion, so it could lead to an unsustainable loss of foreign reserves and/or pressure over nominal exchange rates and, therefore, over inflation rates. Eventually, and as a result of the action of those forces, authorities will be persuaded to limit the level of economic activity and the growth rate by using economic policy instruments to control aggregate demand.

It should be noted, however, that in our view the external restriction discussed above is not a permanent constraint. The literature[6] on balance of payments constrained growth, by contrast, deals with external constraints, and sees the balance of payments equilibrium as an attractor or an economic policy goal. Nevertheless, as we have seen, a surplus of the overall balance of payments is sustainable, and it is not presumed that governments of countries in that position will always push the economy to the limits defined by external constraints. Therefore, our theoretical framework conceives two possible growth regimes: a balance of payments constrained demand-led growth process, and a policy-constrained (or pure) demand-led growth process.

As Medeiros and Serrano (2001) observed, exports have a dual role in the process of economic growth.[7] First, they are a source of autonomous aggregate demand and, as such, export growth contributes directly to the expansion of GDP. Second, exports are a source of international currency and, therefore, indirectly enable the expansion of both domestic expenditures and GDP.[8] Consequently, in contrast with the literature on balance of payments constrained growth, our framework does not impose an export-led growth pattern on the interpretation of economic data. In fact, our theoretical framework is flexible enough to capture either a pattern of economic growth led by exports or a pattern of growth led by domestic expenditures. Such flexibility is particularly suitable for dealing with relatively closed economies like Brazil's, which, due to internal and/or external factors, has been alternating over time between the two regimes.

7.5 Empirical methodology

In order to examine the growth pattern of the Brazilian economy in the period under study, we adopted a demand-side growth accounting decomposition methodology rather than the supply-side growth accounting approach inspired by neoclassical growth theory.[9] It is worth noting that our empirical approach has been adapted from the structural decomposition analysis developed in the input-output literature,[10] which is largely based on Leontief's open input-output model. This model is compatible with our theoretical framework since it is a demand-oriented model.

The methodology has been adapted to capture the distinction between autonomous and induced variables that characterises the classical super-multiplier model. Indeed, it proved necessary to construct empirical proxies to distinguish between autonomous expenditures and induced expenditures. Government consumption was regarded as autonomous spending; household consumption was divided into autonomous and induced components. A distinction between durable and non-durable (including services) consumption was used as a proxy for distinguishing autonomous consumption from induced consumption, on the grounds that durable consumption is usually financed by credit, which is a source of purchasing power not directly related to current production decisions.[11]

Moreover, the amount of credit applied to durable consumption depends on institutional and social factors, such as the evolution of durable consumption patterns; income distribution patterns; institutionally determined base rates; the intensity of competition in the financial system that regulates banking; banking risk policies; tax structure; the legal framework behind the credit system; the public sector's role in the credit system; the relations between the financial sector and the central bank; and the monetary and credit policy framework adopted by the monetary authority, among others. It is acknowledged that households' debt-to-income ratio also influences the expansion of credit to durable consumption. Note, however, that such an influence is rather weak, particularly when longer periods of analysis are considered and, consequently, the institutional and social factors just mentioned are subject to significant changes, even when the analysis considers one country's experience only. On the other hand, non-durable consumption largely depends on the purchasing power introduced into the economy by current production decisions, namely payments of wages. Finally, as regards investment expenditures, since capitalist competition prompts a trend for the adjustment of productive capacity to match demand, private enterprise investment was considered induced spending. Other aggregate investment components (housing, government and state-owned enterprises investments) included in the analysis were considered autonomous, since they are neither driven by capitalist competition[12], nor directly related to the purchasing power introduced into the economy by production decisions.

To meet our methodology's requirements, a database was compiled from various sources: from the historical statistics of the system of national accounts (SNA);[13] complementary information from the work of Dos Santos and Pires (2007); and from some old input-output tables (specifically those for 1970, 1975, 1980 and 1985). The old input-output tables allowed for the division of household consumption into durable and non-durable components for the years before 1990.[14] Dos Santos and Pires' (2007) investment database allowed us to break down enterprise investment into private and state-owned components and, thus, to quantify private enterprise investment for the 2000s, which are not covered by official statistics.

The decomposition formula is derived from the national account identity between aggregate supply and aggregate demand.[15] The final formula is as follows:[16]

$$
\begin{aligned}
g = {} & \alpha(1)\left[\frac{C_{HND}(0)}{Y(0)}\right]g_c + \alpha(1)\left[\frac{I_{PE}(0)}{Y(0)}\right]g_h + \alpha(1)\left[\frac{C_{HD}(0)}{Y(0)}\right]g_{C_{HD}} \\
& + \alpha(1)\left[\frac{I_H(0)}{Y(0)}\right]g_{I_H} + \alpha(1)\left[\frac{C_G(0)}{Y(0)}\right]g_{C_G} + \alpha(1)\left[\frac{I_G(0)}{Y(0)}\right]g_{I_G} \\
& + \alpha(1)\left[\frac{I_{SE}(0)}{Y(0)}\right]g_{I_{SE}} + \alpha(1)\left[\frac{X(0)}{Y(0)}\right]g_X + \frac{\alpha(1)}{\mu(1)}g_\mu + \alpha(1)\left[\frac{E(0)}{Y(0)}\right]g_E
\end{aligned}
\qquad [7.7]
$$

where E is the value of inventory change; g is real GDP growth rate; and g_i refers to the real growth rate of variable i.

On the left-hand side (LHS) of the equation, the observed real GDP growth rate lies between a base period (denoted as period (0)) and a final period (denoted as period (1)).[17] The right-hand side (RHS) of the equation contains the factors that contribute to the expansion of the GDP. First, in order for the observed real GDP growth rate to match exactly the RHS of the equation, the last term referring to the contribution of inventory change growth rate was incorporated into the RHS. For a similar reason, all variables were deflated using the GDP price deflator, instead of using their own price deflators. This procedure has a significant implication: as for the changes in the variables' growth rates contained on the RHS of the equation, they will account changes expressed not only in volume, but also in relative prices (relative to the GDP price deflator). Second, on the RHS the contribution from each variable depends on its respective real growth rate multiplied by an accounting GDP elasticity specific to each variable being considered. All accounting elasticities depend on the value of the supermultiplier in the final period; all of them, except for the one related to the share of domestic content in demand, will also depend on the ratio of the respective variable value to the value of the GDP in the base period.

Finally, one should note that there is an important difference between the method of accounting for the contributions of the variables for autonomous expenditures, on the one hand, and the variables contained in the supermultiplier formula, on the other. The reason is that, in the latter case, the decomposition captures the contributions of the *propensity* to consume, the *propensity* to invest and the share of domestic content growth rates. That is, to be compatible with our theoretical framework, household non-durable (induced) consumption, private enterprise (induced) investment, and import contributions are related to the growth rate of those ratios. As for the autonomous expenditure variables, however, the decomposition captures their contributions in relation to their own growth rates.

The results produced by the growth accounting decomposition will be presented in two alternative forms. First, from the perspective of the sources of the contributing factors, the decomposition will quantify the contributions of both domestic and external sectors. The contribution of the domestic sector to GDP growth will be measured according to the following expression:

Domestic sector contribution

$$= \alpha(1)\left[\frac{C_{HND}(0)}{Y(0)}\right]g_c + \alpha(1)\left[\frac{I_{PE}(0)}{Y(0)}\right]g_h + \alpha(1)\left[\frac{C_{HD}(0)}{Y(0)}\right]g_{C_{HD}} + \alpha(1)\left[\frac{I_H(0)}{Y(0)}\right]g_{I_H} + \alpha(1)\left[\frac{C_G(0)}{Y(0)}\right]g_{C_G} + \alpha(1)\left[\frac{I_G(0)}{Y(0)}\right]g_{I_G} + \alpha(1)\left[\frac{I_{SE}(0)}{Y(0)}\right]g_{I_{SE}} \quad [7.8]$$

The contribution from domestic sector will then be divided into two components: private sector and public sector, which are represented, respectively, by the two following equations:

Public sector contribution

$$= \alpha(1)\left[\frac{C_G(0)}{Y(0)}\right]g_{C_G} + \alpha(1)\left[\frac{I_G(0)}{Y(0)}\right]g_{I_G} + \alpha(1)\left[\frac{I_{SE}(0)}{Y(0)}\right]g_{I_{SE}} \quad [7.9]$$

and

Private sector contribution

$$= \alpha(1)\left[\frac{C_{HND}(0)}{Y(0)}\right]g_c + \alpha(1)\left[\frac{I_{PE}(0)}{Y(0)}\right]g_h + \alpha(1)\left[\frac{C_{HD}(0)}{Y(0)}\right]g_{C_{HD}} + \alpha(1)\left[\frac{I_H(0)}{Y(0)}\right] \quad [7.10]$$

The external sector contribution, in turn, captures the influence of export growth and of changes in the share of domestic content in demand, and is measured according to the following equation:

$$\textit{External sector contribution} = \alpha(1)\left[\frac{X(0)}{Y(0)}\right]g_x + \frac{\alpha(1)}{\mu(1)}g_\mu \quad [7.11]$$ [18]

So the real GDP growth rate will be decomposed into its contributing factors according to the following equation:

$$\begin{aligned} g &= \textit{Domestic sector contribution} \\ &\quad + \textit{External sector contribution} + \alpha(1)\left[\frac{E(0)}{Y(0)}\right]g_E \\ &= \textit{Private sector contribution} + \textit{public sector contribution} \\ &\quad + \textit{External sector contribution} + \alpha(1)\left[\frac{E(0)}{Y(0)}\right]g_E \end{aligned} \qquad [7.12]^{19}$$

Second, from another perspective, the RHS terms of decomposition equation [7.7] are grouped according to whether contributions come from autonomous expenditures or from changes in the supermultiplier. In our model contributions of autonomous expenditures will be quantified as follows:

$$\begin{aligned} &\textit{Autonomous expenditures contribution} \\ &= \alpha(1)\left[\frac{C_{HD}(0)}{Y(0)}\right]g_{C_{HD}} + \alpha(1)\left[\frac{I_H(0)}{Y(0)}\right]g_{I_H} + \alpha(1)\left[\frac{C_G(0)}{Y(0)}\right]g_{C_G} \\ &\quad + \alpha(1)\left[\frac{I_G(0)}{Y(0)}\right]g_{I_G} + \alpha(1)\left[\frac{I_{SE}(0)}{Y(0)}\right]g_{I_{SE}} + \alpha(1)\left[\frac{x(0)}{Y(0)}\right]g_x \end{aligned} \qquad [7.13]$$

For its part, changes in the supermultiplier contribution will be measured quantitatively according to equation [7.14]:

$$\begin{aligned} &\textit{Supermultiplier contribution} \\ &= \alpha(1)\left[\frac{C_{HND}(0)}{Y(0)}\right]g_c + \alpha(1)\left[\frac{I_{PE}(0)}{Y(0)}\right]g_h + = \left[\frac{\alpha(1)}{\mu(1)}\right] = g_\mu \end{aligned} \qquad [7.14]$$

Hence, combining equations [7.13] and [7.14] and including the term related to the contribution of inventory change results in an alternative decomposition for the real GDP growth rate, which can be expressed as follows:

$$\begin{aligned} g &= \textit{Autonomous expenditures contribution} \\ &\quad + \textit{Supermultiplier contribution} + \alpha(1)\left(\frac{E(0)}{Y(0)}\right)g_E \end{aligned} \qquad [7.15]$$

Table 7.4 enables the decomposition quantitative results that are presented in Section 7.4 to be connected with the decomposition formulas discussed above. The first column ('Variables') lists the symbols for the variables used in the decomposition analysis. Column six ('Total')

Table 7.4 Decomposition equations in table form

Variables	Domestic Sector		External Sector	Inventory Changes	Total	Autonomous Expenditures	Supermultiplier	Inventory Changes
	Public	Private						
C_G	$\alpha(1)\left[\frac{C_G(0)}{Y(0)}\right]g_{C_G}$				$\alpha(1)\left[\frac{C_G(0)}{Y(0)}\right]g_{C_G}$	$\alpha(1)\left[\frac{C_G(0)}{Y(0)}\right]g_{C_G}$		
I_G	$\alpha(1)\left[\frac{I_G(0)}{Y(0)}\right]g_{I_G}$				$\alpha(1)\left[\frac{I_G(0)}{Y(0)}\right]g_{I_G}$	$\alpha(1)\left[\frac{I_G(0)}{Y(0)}\right]g_{I_G}$		
I_{SE}	$\alpha(1)\left[\frac{I_{SE}(0)}{Y(0)}\right]g_{I_{SE}}$				$\alpha(1)\left[\frac{I_{SE}(0)}{Y(0)}\right]g_{I_{SE}}$	$\alpha(1)\left[\frac{I_{SE}(0)}{Y(0)}\right]g_{I_{SE}}$		
I_H		$\alpha(1)\left[\frac{I_H(0)}{Y(0)}\right]g_{I_H}$			$\alpha(1)\left[\frac{I_H(0)}{Y(0)}\right]g_{I_H}$	$\alpha(1)\left[\frac{I_H(0)}{Y(0)}\right]g_{I_H}$		
C_{HD}		$\alpha(1)\left[\frac{C_{HD}(0)}{Y(0)}\right]g_{C_{HD}}$			$\alpha(1)\left[\frac{C_{HD}(0)}{Y(0)}\right]g_{C_{HD}}$	$\alpha(1)\left[\frac{C_{HD}(0)}{Y(0)}\right]g_{C_{HD}}$		
C_{HND}		$\alpha(1)\left[\frac{C_{HND}(0)}{Y(0)}\right]g_C$			$\alpha(1)\left[\frac{C_{HND}(0)}{Y(0)}\right]g_C$		$\alpha(1)\left[\frac{C_{HND}(0)}{Y(0)}\right]g_C$	
I_{PE}		$\alpha(1)\left[\frac{I_{PE}(0)}{Y(0)}\right]g_h$			$\alpha(1)\left[\frac{I_{PE}(0)}{Y(0)}\right]g_h$		$\alpha(1)\left[\frac{I_{PE}(0)}{Y(0)}\right]g_h$	
μ			$\frac{\alpha(1)}{\mu(1)}g_\mu$		$\frac{\alpha(1)}{\mu(1)}g_\mu$		$\frac{\alpha(1)}{\mu(1)}g_\mu$	
X			$\alpha(1)\left[\frac{X(0)}{Y(0)}\right]g_X$		$\alpha(1)\left[\frac{X(0)}{Y(0)}\right]g_X$	$\alpha(1)\left[\frac{X(0)}{Y(0)}\right]g_X$		
E				$\alpha(1)\left[\frac{E(0)}{Y(0)}\right]g_E$	$\alpha(1)\left[\frac{E(0)}{Y(0)}\right]g_E$			$\alpha(1)\left[\frac{E(0)}{Y(0)}\right]g_E$
Total	$\sum$	$\sum$	$\sum$	$\sum$	g	$\sum$	$\sum$	$\sum$

reproduces the decomposition equation [7.7] in tabular form, where the column total cell equals the observed real GDP growth rate. This column also divides the table into two parts, each containing one of the two alternative methods of presenting the decomposition results as previously described in equations [7.8] to [7.15]. On the left-hand side, columns two ('Domestic sector – Public'), three ('Domestic sector – Private') and four ('External sector') represent equations [7.9], [7.10] and [7.11], respectively. Column five ('Inventory changes') represents the contribution for inventory change. In its turn, the sum of the totals for columns two, three, four and five corresponds to equation [7.12]. The right-hand side of column six ('Totals') presents the decomposition in an alternative way. So, column seven ('Autonomous expenditures') gives equation [7.13] in tabular form, while column eight ('Supermultiplier') contains equation [7.14]. Finally, the sum of the totals contained in columns seven, eight and nine represents equation [7.15], where the total cell of column nine ('Inventory changes') represents the contribution from inventory change.

Before concluding this section, the analytical status of the empirical methodology previously described will be briefly addressed. Similarly to the neoclassical supply-side growth accounting decomposition, the demand-side decomposition method adopted here is not intended to provide an empirical test of the theoretical framework subjacent to it. In fact, it is assumed that such framework is valid for an interpretation of the pattern of economic growth prevailing in the Brazilian economy in the period under study. In this sense, the decomposition exercise is to be understood, at best, as an indirect way of evaluating our theoretical framework, depending on its ability to provide a plausible and consistent explanation for the observed facts.

7.6 Some results

This section analyses the growth performance of the Brazilian economy for the whole period between 1970 and 2005.[20] We begin with an overview of the external situation and how it connects with the process of economic growth in Brazil during that period.

As shown in Figure 7.4, in the 1970s the Brazilian economy experienced huge current account deficits, mostly financed by financial account surplus positions. However, a deficit situation in the overall balance of payments emerged as a consequence of the two oil shocks and their impact on the world economy. In particular, the second oil shock in 1979 and the consequent increase in US interest rates had a major negative effect on Latin America peripheral countries, leading to the external debt crises that

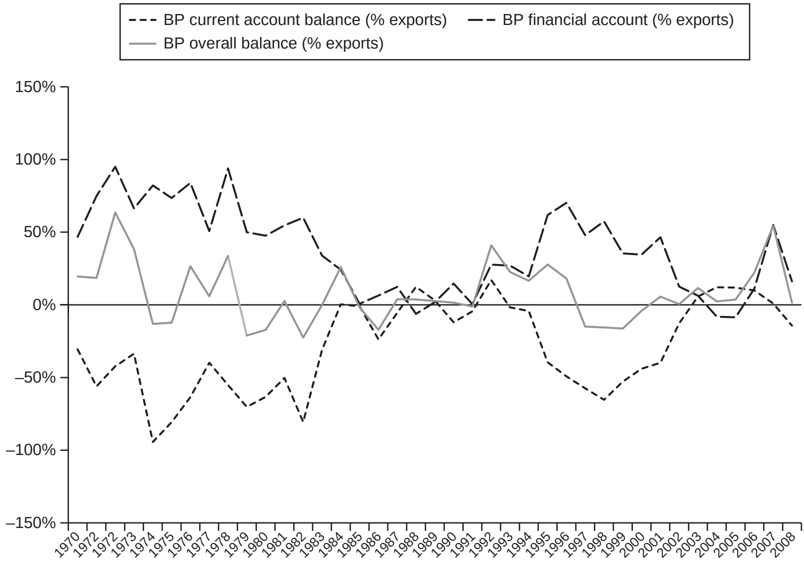


Figure 7.4 Selected balance of payments accounts, Brazil 1970–2008 (percentage of exports)
Source: Authors' elaboration based on IBGE, FGV, and Dos Santos and Pires (2007).

jeopardised the region's economic performance in the 1980s. Figure 7.4 also shows that in the 1980s, as international capital flows to Brazil were interrupted, the problem of external constraint became more acute.

Moreover, in order to keep the current account deficits under control, the Brazilian monetary authorities continued to promote nominal exchange rate devaluations. Combined with a high level of formal and informal indexation of nominal contracts and the escalation of social conflicts over income distribution, those measures also caused the inflationary process to escalate. Brazil entered a period of chronically high inflation that lasted until the early 1990s, when the return of international capital flows enabled the Real stabilisation plan to be successfully implemented in 1994, which in turn put an end to the policy of continuous nominal exchange rate devaluations and managed partially to eliminate inflation-indexed nominal contracts.

On the other hand, the return of the Brazilian economy to the international financial circuit in the early 1990s prompted a real exchange rate appreciation, which, along with the trade liberalisation policies that had been introduced in the late 1980s and intensified in the early 1990s, produced the high current account deficits that were characteristic of the late 1990s. So the balance of payments constraint continued to be a binding constraint on GDP growth. In fact, during this period the Brazilian economy faced a series of external shocks related to the Mexican, Asian and Russian crises. Finally, in 1998 Brazil experienced a balance of payments crisis and had to appeal to the IMF for help.[21]

The Brazilian crisis prompted an important change in the national macroeconomic policy regime. The new policy framework combined inflation targeting, large primary government budget surplus and floating (but still widely managed) exchange rates. In particular, the focus on price stabilisation and sound finance helped shape a conservative consensus on matters of economic policy. At least until 2003, such conservative consensus coexisted with a situation of considerable external pressure (particularly in the year 2002, due to the Brazilian presidential election). However, the balance of payments situation has been improving since 2003 mostly as a result of real exchange rate depreciation after 1999; favourable development of international commodity trading; and resumption of large gross private capital flows (Serrano and Summa, 2012). The current account deficit was converted to surplus and the overall balance of payments also showed a surplus, which enabled a policy of foreign reserves accumulation and payment of accumulated foreign debt (including IMF loans) to be implemented.[22] The continuous accumulation of foreign reserves, however, indicated that the economy was able to expand at higher rates and, therefore, that the binding constraint in the period was not the external one – a conclusion supported by the trend towards real exchange rate appreciation, observed from 2003 on, as a result of the *modus operandi* of the inflation targeting regime adopted in Brazil. The Brazilian central bank took advantage of the favourable external environment to adopt high positive interest rate differentials to control the nominal exchange rate in order to achieve the official inflation rate target (Serrano, 2010). Hence, from 2003 on, the Brazilian economy was characterised by a demand-led growth process chiefly influenced by internal factors, most of them related to the political dispute involving the maintenance or the modification of the macroeconomic policy regime inside and outside the government (Barbosa and Souza, 2010; Serrano and Summa, 2012).

Now, applying the decomposition methodology enables us to identify the pattern of economic growth prevailing in the Brazil during the period analysed. Table 7.5 shows the results.

As shown in Table 7.5, from 1970 to 2005 Brazil grew by 3.99 per cent per year. The domestic sector was the main source of GDP growth in the period, contributing 2.5 percentage points (hereafter denoted pp), while

Table 7.5 Average annual rate of growth decomposition, Brazil, 1970–2005 (constant 1980 prices)

Variables	Domestic Sector		External Sector	Inventory Change	Total	Autonomous Expenditures	Super-multiplier	Inventory Change
	Public	Private						
C_G	1.83%				1.83%	1.83%		
I_G	0.11%				0.11%	0.11%		
I_{SE}	–0.004%				–0.004%	–0.004%		
I_H		0.35%			0.35%	0.35%		
C_{HD}		0.42%			0.42%	0.42%		
C_{HND}		–0.15%			–0.15%		–0.15%	
I_{PE}		–0.04%			–0.04%		–0.04%	
μ			–0.11%		–0.11%		–0.11%	
X			1.43%		1.43%	1.43%		
E				0.16%	0.16%			0.16%
Total	1.93%	0.57%	1.32%	0.16%	3.99%	4.13%	–0.30%	0.16%

Source: Authors' elaboration based on IBGE, FGV and Dos Santos & Pires (2007).

the external sector had a contribution of 1.32 pp. Within the domestic sector, the public sector was the major source of GDP growth, with a contribution of 1.93 pp, while the private sector made a more modest contribution of 0.57 pp. These results suggest that the Brazilian economy was characterised by a domestic demand-led growth pattern or, more specifically, by a 'government demand-led' growth pattern, in which the expansion of the domestic market figured as the major source of economic growth.[23]

Next, looking at the other method of presenting the decomposition results, autonomous expenditure growth figures as the main source of economic growth, with a positive contribution of 4.13 pp, in contrast to the supermultiplier's change negative contribution of −0.30 pp. This result shows that, during the whole period under review, autonomous expenditure growth was the major determinant of the GDP growth rate trend. For some periods, nevertheless, changes in the supermultiplier made a significant contribution to GDP growth, as can be seen where the GDP growth rate and the autonomous expenditure growth rate diverge in Figure 7.5.

The average annual autonomous expenditure growth rate was 4.7 per cent for the whole period. Exports grew, on average, 6.3 per cent

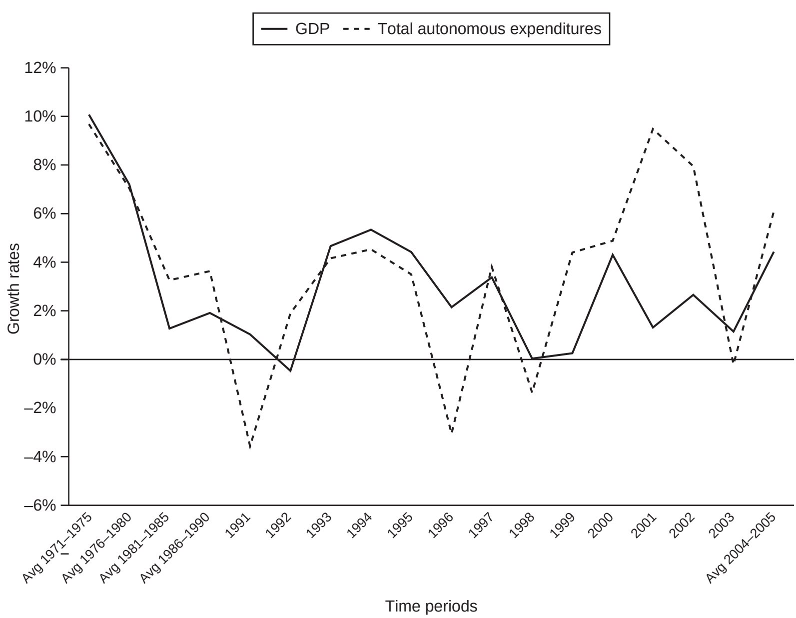


Figure 7.5 GDP growth rates and autonomous expenditure growth rates
Source: Authors' elaboration based on IBGE, FGV, and Dos Santos and Pires (2007).

annually, while autonomous private and public expenditures grew, respectively, 3.4 per cent and 4.7 per cent per year on average. This result contrasts with the one related to the growth decomposition presented above. Indeed, if we took growth rates as a reference, the Brazilian economy should be expected to follow an export-led growth pattern. Still, as the decomposition shows, this is not the case. The reason is that a higher export growth rate is more than compensated by the lower value of exports' share of GDP in the base year (that is, 7.0 per cent in 1970). On the other hand, public autonomous expenditures had a GDP share of 18.0 per cent in 1970. That share, combined with the growth rate of public autonomous expenditures in the period, explains why the public sector made the highest contribution to GDP expansion. Two explanations for the initial low share of exports in GDP are the inward-oriented state-led development process in the post-WWII period and the continental size of Brazil.[24] The shares of public and private autonomous expenditures, as well as the share of exports, are shown in Figure 7.6 below. It is worth mentioning, however, that the situation has changed since the 1970s. As shown in Figure 7.6, by the end of the period, exports' share of GDP had increased to around 15 per cent, whereas public autonomous expenditures were of the order of 22 per cent.[25] In the light of that, it can

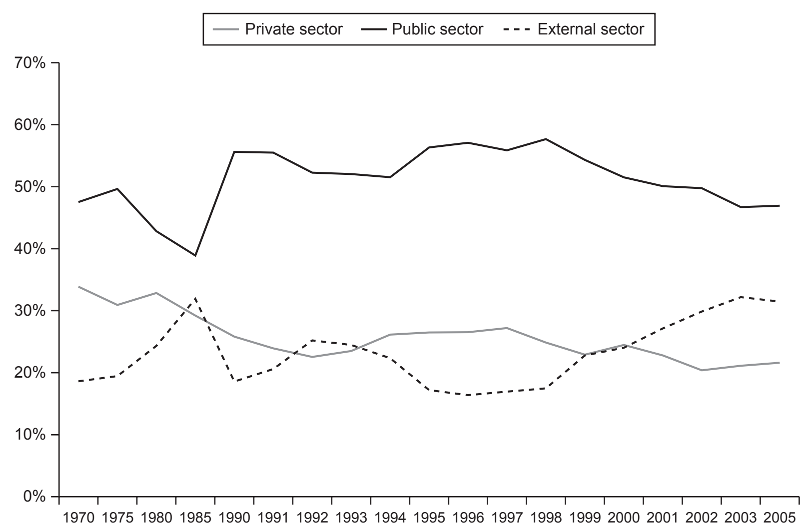


Figure 7.6 Autonomous expenditures share of GDP
Sources: Authors' elaboration based on IBGE, FGV, and Dos Santos and Pires (2007).

be suggested that the Brazilian economy has become more dependent on the dynamism of external markets.[26] Yet it is not clear whether this result means a structural change has occurred in the Brazilian economy's growth pattern.

Another important issue concerns how the components of autonomous expenditure growth behave over time. Figure 7.7 shows that exports were the most volatile component, followed by private and public autonomous expenditures. This suggests that the observed pattern of expansion of the Brazilian economy, which is mostly driven by public sector expenditures, would be more volatile if private and external sectors were dominant sources of economic growth. Therefore, should the Brazilian economy become more dependent on the dynamism of external markets (a possibility suggested above), one would also expect it to be more susceptible to fluctuations in exports.

Our analysis now turns to the behaviour of the supermultiplier. As illustrated in Figure 7.8, the supermultiplier showed a declining trend during the period with which we are concerned. In the 1970s

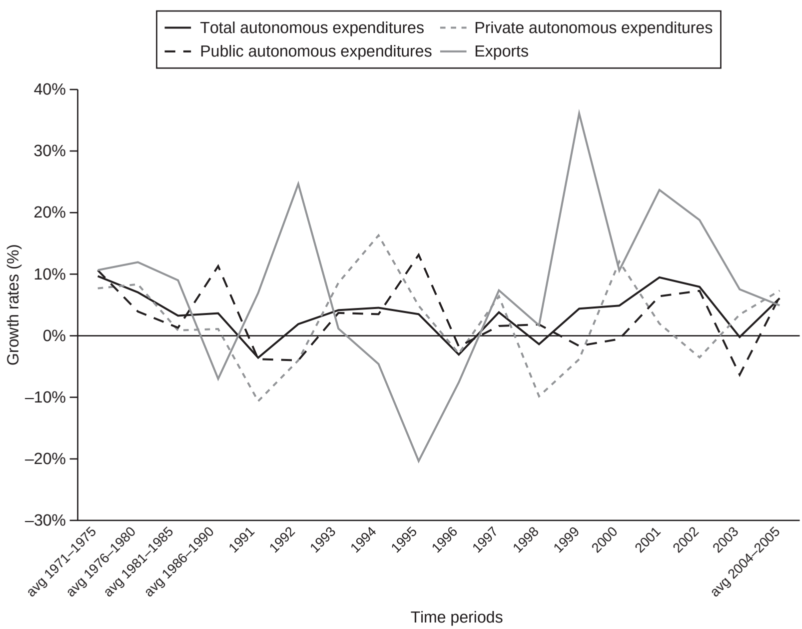


Figure 7.7 Annual average growth rates of autonomous expenditure components (percentage per annum)
Sources: Authors' elaboration based on IBGE, FGV, and Dos Santos and Pires (2007).

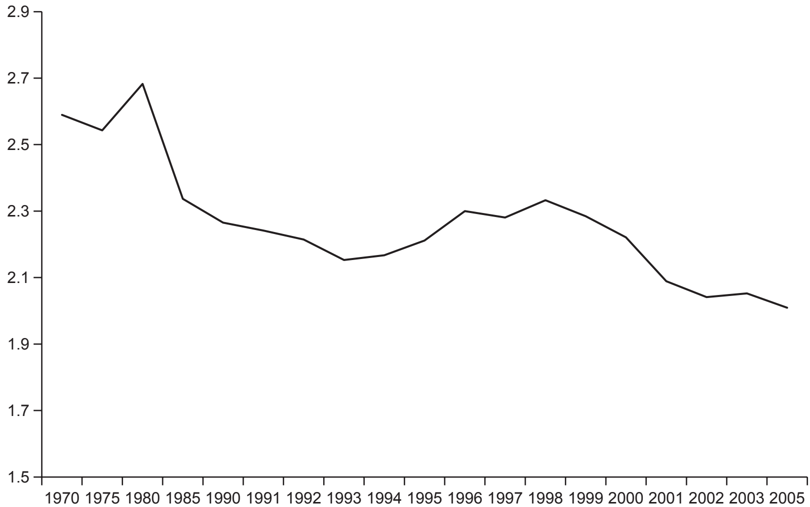


Figure 7.8 Supermultiplier
Sources: Authors' elaboration based on IBGE, FGV, and Dos Santos and Pires (2007).

its value ranged between 2.5 and 2.7, but by the end of the decade it had declined to approximately 2.0. Table 7.5 demonstrates that this decline was due to three factors: a reduction in propensity to consume, a decline in private enterprise propensity to invest, and a decrease in the share of domestic content in total demand.

As Table 7.5 shows, the growth in propensity to consume contributed –0.15 pp to the GDP growth rate (4.0 per cent), almost half the value of the total contribution of the supermultiplier to economic growth (–0.30 pp). As we have argued above, according to the supermultiplier model the wage share (of GDP) is one of the major determinants of propensity to consume. Figure 7.9 illustrates this point by showing that propensity to consume and wage share tend to move in the same direction. So the worsening of functional income distribution apparently contributed to the decline of propensity to consume.

Next, the contribution of the share of domestic content will be assessed. According to Table 7.5, that component scored a contribution of –0.11 pp to the observed GDP growth rate and explained one-third of the decline in the supermultiplier. As shown in Figure 7.10, the share of domestic content presented an upward trend until 1989, with a peak value of almost 95 per cent, followed by a significant downward

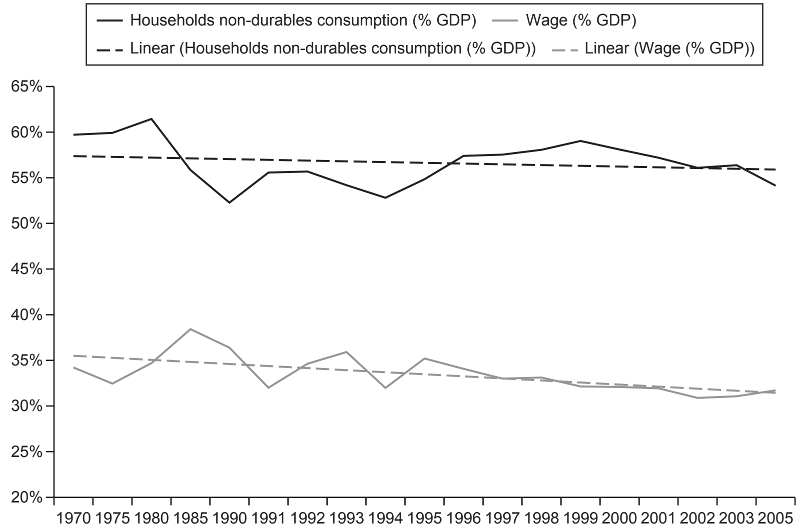


Figure 7.9 Household non-durables consumption and wages (as percentage of GDP)
Sources: Authors' elaboration based on IBGE, FGV, and Dos Santos and Pires (2007).

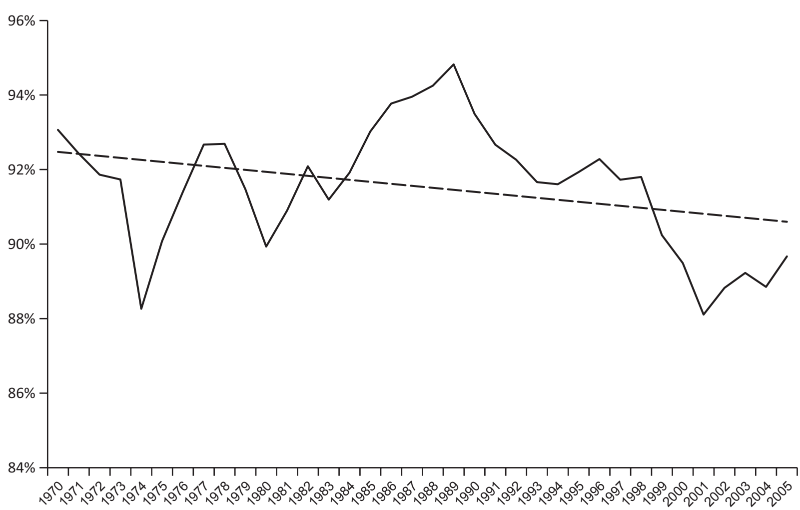


Figure 7.10 Share of domestic content in demand (percentage)
Sources: Authors' elaboration based on IBGE, FGV, and Dos Santos and Pires (2007).

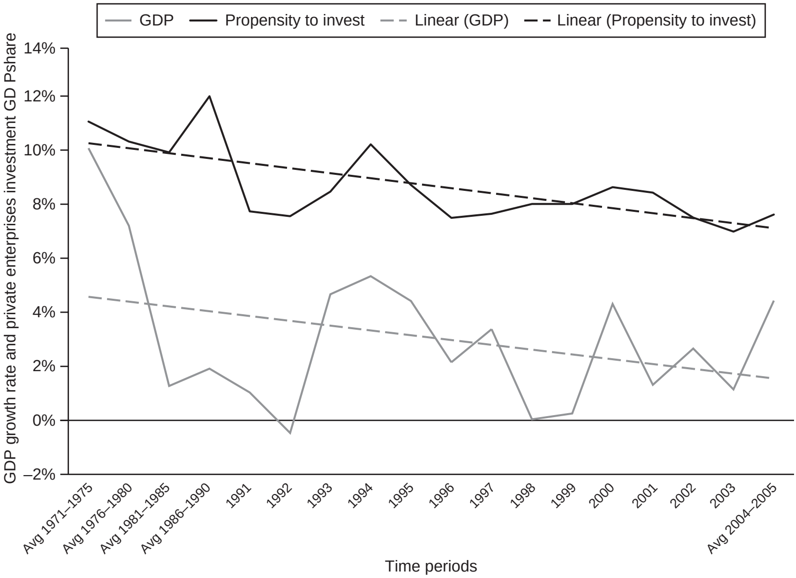


Figure 7.11 GDP growth rates and private enterprise investment share of GDP (percentage)

Source: Authors' elaboration based on IBGE, FGV, and Dos Santos and Pires (2007).

trend falling to approximately 90 per cent in 2005. This last trend resulted from a combination of trade liberalisation policies initiated in 1987 and the overvaluation of the real exchange rate.[27]

Table 7.5 shows that the contribution made by private enterprise's propensity to invest to the GDP growth rate scored –0.04 pp. Furthermore, as mentioned before, according to the classical supermultiplier model, the relation between the ratio of private enterprise investment to GDP, on the one hand, and the GDP growth rate trend, on the other, is expected to be positive. According to Figure 7.11, the observed behaviour of these two variables in our period is compatible with the relation proposed by the model. The downward trend of the GDP growth rate is followed by a decrease in private enterprise propensity to invest, as measured in this chapter.[28]

7.7 Concluding remarks

Between 1970 and 2005 Brazil experienced a huge decline in GDP growth rates, from an average rate of almost 8.5 per cent per year in the 1970s to approximately 2.3 per cent per year from the 1980s to 2005.

The neoclassical standpoint has dominated the more recent literature dedicated to studying that period. However, we are adopting a completely different perspective, based on a combination of the classical supermultiplier demand-led growth model and the hypothesis that the balance of payments was the main potential (and often the most effective) constraint on the expansion of the Brazilian economy during the period under study.

The main results of our investigation are the following. First, public sector expenditures were the major source of economic growth in the period analysed. The main influence of external conditions in general and exports in particular was the indirect effect they exerted on the growth of domestic demand. Although exports made a significant contribution, their growth path proved to be very unstable and, even when their contribution to economic growth was relatively high, they were unable to sustain alone the pace of GDP growth observed in the 1970s. This explains why in certain periods the Brazilian economy did experience, as suggested by Medeiros and Serrano (2001), a pattern of export-led 'stagnation'.

Second, our empirical investigation showed that the classical supermultiplier model is able to explain the main features of Brazilian growth in those decades. The contribution of autonomous expenditures to economic growth considered in the model is able to explain most of the observed GDP growth rate trend for the whole period. The share of GDP that wages represent seemed to be a significant determinant of propensity to consume. The propensity to invest by private enterprises was positively related to the GDP growth rate trend. Furthermore, our methodology proved flexible enough to enable a history-friendly approach to the analysis of growth experiences. In particular, the approach allowed for the consistent incorporation into the analysis of features pertaining to institutional and power relations, also proving compatible with the main theoretical and methodological aspects of the classical surplus approach to political economy.

Notes

1. Lack of space prevents the analysis here of the sub-periods between 1970 and 2005, which will be the subject of a future work.
2. See Abreu and Verner (1997), Ferreira and Malliagros (1998), Pinheiro *et al.* (2001), Gomes *et al.*, (2003), Bacha and Bonelli (2004) and Loayza *et al.* (2005).

3. See Kaldor (1989[1981]), McCombie *et al.* (2003), Thirlwall (1983) and, for evidence related to the Brazilian economy, see Marinho *et al.* (2002).
4. See also Cesaratto *et al.* (2003) and Serrano and Freitas (2007).
5. This specification must be contrasted with the results obtained in the neo-Kaleckian growth models, especially those connected with the contributions of Marglin and Bhaduri. As is known, those models allows for the existence of a regular relation between the wage share (or the profit share) and the trend GDP rate of growth.
6. See Kaldor (1978[1971]), Kaldor (1989[1983]), McCombie and Thirlwall (1994) and McCombie and Roberts (2002).
7. The balance of payments constrained growth literature does not properly acknowledge the dual role of exports, frequently ignoring the possibility or the likelihood of a pure demand-led growth regime in the sense defined above.
8. It must be noted that, as a source of international finance, exports stand out as a source of international currency without present or future international payment flows. The same is not true for other sources of external finance such as, for example, direct investment, portfolio investment, trade credits and international loans.
9. See Solow (1957) for the seminal contribution and Hulten (2010) for a recent survey of the literature.
10. For a survey of the literature see Miller and Blair (2009, chapter 13).
11. Following Serrano (1995), here we are using the Kaleckian classification of autonomous expenditures as related to the source of purchasing power introduced in the economy. For a very clear account of Kalecki's use of this classification see López and Assous (2010, chapter 2, appendix 2.1).
12. State-owned enterprises' investment was classified as an autonomous expenditure because, for most of this period, capitalist competition did not exert a major influence on its behaviour.
13. Data pertaining to periods until 1989 were provided by the Getulio Vargas Foundation (FGV in Portuguese). Data from 1990 on have been provided by the Brazilian Institute of Geography and Statistics (IBGE in Portuguese), which is now responsible for the SNA.
14. For data after (and including) 1990 the Brazilian National Accounts System adopted the United Nations 1993 revision of the SNA methodology, which requires the compilation of use tables to allow for the identification of durable consumption.
15. Note that for the decomposition to be exhaustive we have to include in the RHS of equation [7.1] an extra variable that captures the unplanned inventory accumulation. The inclusion of this variable transforms the equilibrium condition represented in equation [7.1] into an identity. For more details, see the Appendix.
16. The detailed derivation of the decomposition formula is shown in the Appendix.
17. The symbols (0) and (1) attached to a variable in level mean that the level of the variable in question is evaluated in the base period and in the final period respectively.
18. Note that this way of calculating the external sector contributions to economic growth contrasts with the usual method that uses net exports.

19. Again, the inclusion of the inventory change growth contribution is deemed necessary to guarantee equality between the RHS of the equation and the observed real GDP growth rate.
20. For a more detailed account of the Brazilian economy in this period see Serrano (1998), Serrano and Summa (2012) and Barbosa and Souza (2010)
21. Note that from 1997 to 2000 the deficit in the overall balance of payments is evidence of the severity of the external problems faced by the Brazilian economy in the late 1990s.
22. It should be noted, however, that this policy gained momentum from 2006 on and, therefore, in a period not covered by our analysis.
23. However, it does not mean that the external sector was not an important determinant of the growth process, since, as indicated above, its influence on growth can also be exerted indirectly by means of the balance of payments constraint to the domestic demand expansion.
24. One should note that the continental size of the Brazilian economy continues to contribute to the prevalence of a domestic demand-led growth pattern.
25. The figures for the shares of autonomous expenditures in GDP must be analysed with caution. One implication of the classical supermultiplier model is that a lower trend of GDP growth rate is expected to be related to a higher share of total autonomous expenditures in GDP. This occurs because lower GDP growth rates tend to lower the propensity to spend (that is, a lower ratio of private enterprise investment to GDP) and, consequently, to lower the supermultiplier value. Equation [7.6] shows that the share of total autonomous expenditures and the supermultiplier value are inversely related. Therefore, from the viewpoint of the model adopted here, the share of autonomous expenditures in GDP should not be interpreted as an indicator of the dynamism of this expenditure unless the behaviour of the variables contained in the supermultiplier formula are thoroughly analysed, particularly the behaviour of private enterprise investment.
26. More recent data (after 2005) reveal that as a share of GDP, exports have declined to values ranging between 11 per cent and 12 per cent. Nevertheless, these values are higher than the average export share observed from 1970 to 1999.
27. The Brazilian crisis of 1998 interrupted the over-valuation trend for a while and led to a depreciation trend that lasted until 2003. From then on, as mentioned before, the Brazilian central bank maintained relatively high international interest rate differentials to keep inflation under control. The management of nominal exchange rate led to the appreciation of real exchange rate or to the maintenance of a relatively overvalued real exchange rate from 2003 on.
28. It should be noted that the positive correlation between GDP growth rates and investment (in physical capital) to GDP ratio is one of the most robust empirical relations obtained in the economic growth applied literature (see, for example, Sala-i-Martin, 1997). In addition, many very influential models in the growth literature are not compatible with this stylised fact, such as the Solow model (Solow, 1956) and the neo-Kaleckian models (see Marglin and Bhaduri, 1990).

References

Abreu, M. and Verner, D. (1997) *Long-Term Brazilian Economic Growth: 1930–94* (Paris: OECD).

Bacha, E. L. and Bonelli, R. (2004) 'Accounting for Brazil's Long Term Growth', *Discussion Paper* 1018 (Rio de Janeiro: IPEA – Institute for Applied Economic Research).

Barbosa, N. H. and Souza, J. A. P. (2010) 'A inflexão do Governo Lula: política econômica, crescimento e distribuição de renda', in E. Sader and M. A. Garcia (eds), *Brasil entre o Passado e o Futuro* (São Paulo: Boitempo Editorial): pp. 57–110.

Cesaratto, S., Serrano, F. and Stirati, A. (2003) 'Technical change, effective demand and employment', *Review of Political Economy*, 15(1): 33–52.

Dos Santos, C. H. and Pires, M.C. (2007) 'Reestimativas do Investimento Privado Brasileiro: qual a sensibilidade do investimento privado "referência 1985" a aumentos da carga tributária?', *Discussion Paper* 1297 (Rio de Janeiro: IPEA – Institute for Applied Economic Research).

Elias, V. J. (1978) 'Sources of Economic Growth in Latin American Countries', *Review of Economics and Statistics*, LX(30): 362–70.

Elias, V. J. (1992) *Sources of growth – a study of seven Latin American economies* (San Francisco: ICS Press).

Ferreira, P. C. and Malliagros, T.G. (1998) 'Impactos produtivos da infra-estrutura no Brasil – 1950–1995', *Pesquisa e planejamento econômico*, vol. 28, no. 2: 315–38 (Rio de Janeiro: IPEA).

Gomes, V., de A. Pessôa, S. and Veloso, F. (2003) 'Evolução da Produtividade Total dos Fatores na Economia Brasileira: Uma Análise Comparativa', *Pesquisa e Planejamento Econômico*, 33(3), December (Rio de Janeiro: IPEA).

Groningen Growth and Development Centre 10-sector database, June 2007, http://www.ggdc.net/

Hulten, C. R. (2010) 'Growth Accounting', in B. H. Hall and N. Rosenberg (eds), *Handbook of the Economics of Innovation* (Amsterdam and Oxford: North-Holland): pp. 987–1031.

Kaldor, N. (1978 [1971]) 'Conflicts in National Economic Objectives', in N. Kaldor, *Further Essays on Economic Theory* (New York: Holmes & Meier).

Kaldor, N. (1989 [1981]) 'The Role of Increasing Returns, Technical Progress and Cumulative Causation in the Theory of International Trade and Economic Growth', in N. Kaldor, *Further Essays on Economic Theory and Policy* (New York: Holmes & Meier).

Kaldor, N. (1989 [1983]) 'The Role of Effective Demand in the Short Run and the Long Run', in N. Kaldor, *Further Essays on Economic Theory* (New York: Holmes & Meier).

Loayza, N., Fajnzylber, P. and Calderón, C. (2005) *Economic Growth in Latin America and the Caribbean: Stylized Facts, Explanations, and Forecasts* (Washington: World Bank).

López, J. and Assous, M. (2010) *Michal Kalecki* (Basingstoke and New York: Palgrave Macmillan).

Maddison, A. (2010) *Historical Statistics of the World Economy: 1–2008 AD* (Groningen: Growth and Development Centre).

Marglin, S. A. and Bhaduri, A. (1990) 'Profit Squeeze and Keynesian Theory', in S. A. Marglin and J. B. Schor (eds) *The Golden Age of Capitalism* (Oxford: Oxford University Press).

Marinho, E. L., Nogueira, C. G. and Rosa, A. L. T. (2002) 'Evidências Empíricas da Lei de Kaldor-Verdoorn para a Indústria de Transformacão no Brasil (1985–1997)', *Revista Brasileira de Economia*, 56(3): 457–82.

McCombie, J. S. L. and Thirlwall, A. P. (1994) *Economic Growth and the Balance of Payments Constraint* (London: Macmillan).

McCombie, J., Pugno, M. and Soro, B. (2003) *Productivity Growth and Economic Performance: Essays on the Verdoorn's Law* (Basingstoke and Hampshire: Palgrave Macmillan).

McCombie, J. S. L. and Roberts, M. (2002) 'The Role of the Balance of Payments in Economic Growth', in M. Setterfield (ed.), *Demand-Led Growth: Challenging the Supply Side Vision of the Long Run* (Aldershot: Edward Elgar).

Medeiros, C. A. and Serrano, F. (1999) 'Padrões Monetários Internacionais e Restrição Externa ao Crescimento', in J. L. Fiori (ed.), *Estados e Moedas no Desenvolvimento das Nações* (Petrópolis: Vozes).

Medeiros, C. A. and Serrano, F. (2001) 'Inserção externa, exportações e crescimento no Brasil', in J. Fiori and C. Medeiros (eds), *Polarização Mundial e Crescimento* (Petrópolis: Vozes).

Miller R. E. and Blair P. D. (2009) *Input–Output Analysis: Foundations and Extensions* (Cambridge: Cambridge University Press).

Pinheiro, A. C., Gill, I. S., Servén, L., Thomas, M. R. (2001) *Brazilian Economic Growth 1900–2000: Lessons and Policy Implications* (Washington, DC: World Bank).

Sala-i-Martin, X. (1997) 'I just ran two million regressions', *American Economic Review*, Papers and Proceedings of the Hundred and Fourth Annual Meeting of the American Economic Association, 87(2):178–83.

Serrano, F. (1995) 'Long period effective demand and the Sraffian supermultiplier', *Contributions to Political Economy*, 14(1): 67–90.

Serrano, F. (1996) 'The Sraffian supermultiplier', unpublished dissertation, Cambridge University, England.

Serrano, F. (2003) 'From "static" gold to the floating dollar', *Contributions to Political Economy*, 22(1): 87–102.

Serrano, F. (2010) 'Taxa de juros, taxa de câmbio e o sistema de metas de inflação no Brasil', *Revista de Economia Política*, 30: 63–72.

Serrano, F. and Freitas, F. N. P. (2007) El Supermultiplicador Sraffiano y el Papel de la Demanda Efectiva en los Modelos de Crecimiento, *Circus*, 1: 19–35.

Serrano, F. and Summa, R. (2012) 'Macroeconomic policy, growth and income distribution in the Brazilian economy in the 2000s', *Investigación Económica*, 71, October–December: 55–92.

Solow, R. M. (1956) 'A contribution to the theory of economic growth', *Quarterly Journal of Economics*, 70(1): 65–94.

Solow, R. M. (1957) 'Technical change and the aggregate production function', *Review of Economics and Statistics*, 39(3): 312–20.

Thirlwall, A. P. (1983) 'A plain man's guide to Kaldor's growth laws', *Journal of Post Keynesian Economics*, 5(3): 345–58.

Timmer, M. P. and de Vries, G. J. (2007) 'A Cross-Country Database for Sectoral Employment and Productivity in Asia and Latin America, 1950–2005', *Groningen Growth and Development Centre Research Memorandum GD-98*, (Groningen: University of Groningen).

Appendix

List of symbols

Y – Gross domestic product
M – Imports
C_{HND} – Households non-durable consumption
C_{HD} – Households durable consumption
I_H – Households (residential) investment
C_G – Government consumption
I_G – Government investment
I_{SE} – State-owned enterprises investment
I_{PE} – Private enterprises investment
X – Exports
E – Inventory change
μ– Share of domestic content in demand
c – Households' propensity to consume
h – Private enterprise's propensity to invest
Z – Total autonomous expenditures
α – The supermultiplier
g – GDP growth rate
gi – Growth rate of variable i

The decomposition methodology

We will start with the national account identity between aggregate supply and aggregate demand. Using the maximum number of components of the demand side that our database allows results in the following equation:

$$Y+M=C_{HND}+C_{HD}+I_H+C_G+I_G+I_{SE}+I_{PE}+X+E$$

Now, it is assumed that imports are related to total aggregate demand as expressed in:

$$M=(1-\mu)\ (C_{HND}+C_{HD}+I_H+C_G+I_G+I_{SE}+I_{PE}+X+E)$$

Next, it is assumed that:

$$C_{HND}=cY$$

$$IPE=hY$$

$$Z=C_{HD}+I_H+C_G+I_G+I_{SE}+X$$

Then, by substituting the above relations on the first two equations, the following is obtained:

$$Y=\mu cY+\mu hY+\mu(Z+E)$$

This equation will serve as a starting point for the subsequent GDP growth decomposition analysis. So, GDP change will be taken as described in the following equation:

$$
\begin{aligned}
Y(1) - Y(0) = {} & \mu(1)c(1)Y(1) - \mu(0)c(0)Y(0) + \mu(1)h(1)Y(1) \\
& - \mu(0)h(0)Y(0) + \mu(1)[Z(1) + E(1)] - \mu(0)[Z(0) + E(0)]
\end{aligned}
$$

Adding and subtracting the terms $\mu(1)c(1)Y(0)$ and $\mu(1)h(1)Y(0)$ to/from the RHS of the equation and assuming the fact that $\Delta Y = gY(0)$, the following is obtained:

$$
\begin{aligned}
gY(0) = {} & \mu(1)c(1)gY(0) + \mu(1)h(1)gY(0) + [\mu(1)c(1) - \mu(0)c(0)]Y(0) \\
& + [\mu(1)h(1) - \mu(0)h(0)]\ Y(0) + \mu(1)\ [\Delta Z + \Delta E] + \Delta\mu[Z(0) + E(0)]
\end{aligned}
$$

Dividing both sides of the equation by $Y(0)$ results in the following:

$$
\begin{aligned}
g = {} & \mu(1)c(1)g + \mu(1)h(1)g + \mu(1)c(1) - \mu(0)c(0) + \mu(1)h(1) - \mu(0)h(0) \\
& + \mu(1)\left[\frac{\Delta Z}{Y(0)}\right] + \mu(1)\left[\frac{\Delta E}{Y(0)}\right] + \Delta\mu\left[\frac{Z(0) + E(0)}{Y(0)}\right]
\end{aligned}
$$

By adding and subtracting $\mu(1)\,c(0)$ and $\mu(0)\,h(0)$ to/from the RHS, the following is obtained:

$$
\begin{aligned}
g = {} & \mu(1)\ c(1)\ g + \mu(1)\ h(1)\ g + \mu(1)\ \Delta c + \Delta\mu\, c(0) + \mu(1)\ \Delta h + \Delta\mu h(0) \\
& + \mu(1)\left[\frac{\Delta Z}{Y(0)}\right] + \mu(1)\left[\frac{\Delta E}{Y(0)}\right] + \Delta\mu\left[\frac{Z(0) + E(0)}{Y(0)}\right]
\end{aligned}
$$

Solving the above equation for the growth rate results in:

$$
\begin{aligned}
g = {} & \alpha(1)\Delta c + \frac{\alpha(1)}{\mu(1)}\Delta\mu c(0) + \alpha(1)\Delta h + \frac{\alpha(1)}{\mu(1)}\Delta\mu h(0) + \alpha(1)\left[\frac{\Delta Z}{Y(0)}\right] \\
& + \alpha(1)\left[\frac{\Delta E}{Y(0)}\right] + \frac{\alpha(1)}{\mu(1)}\Delta\mu\left[\frac{Z(0) + E(0)}{Y(0)}\right]
\end{aligned}
$$

First collect all the terms in which $\Delta\mu$ appears. Then, put $\dfrac{\alpha(1)}{\mu(1)}\Delta\mu$ in evidence and use the definition that $\Delta\mu = g_{\mu}\mu(0)$ to arrive at the fourth term on the RHS of the equation below. In addition, the fact that:

$$
\Delta Z = \Delta C_{HD} + \Delta I_H + \Delta C_G + \Delta I_G + \Delta I_{SE} + \Delta X
$$

is used to obtain the third term on the RHS of the following expression.

$$g = \alpha(1)\Delta c + \alpha(1)\Delta h + \alpha(1)\left[\frac{\Delta C_{HD} + \Delta I_H + \Delta C_G + \Delta I_G + \Delta I_{SE} + \Delta X}{Y(0)}\right]$$
$$+ \frac{\alpha(1)}{\mu(1)} g_{\mu} \left[\frac{\mu(0)\{c(0)Y(0) + h(0)Y(0) + Z(0) + E(0)\}}{Y(0)}\right]$$
$$+ \alpha(1)\left[\frac{\Delta E}{Y(0)}\right]$$

However, it is known that:

$$Y(0) = \mu(0)c(0)Y(0) + \mu(0)h(0)Y(0) + \mu(0)(Z(0) + E(0)).$$

So, the fourth term on the RHS is equal to $\frac{\alpha(1)}{\mu(1)} g_{\mu}$. On the other hand, the third term on the RHS can be dismembered to isolate the individual contributions of each type of expenditure involved. Further, all contributions can be expressed in terms of the real growth rate of the variables involved. As a consequence, the following equation (which appears in the text – see p. 170) is obtained:

$$g = \alpha(1)\left[\frac{C_{HND}(0)}{Y(0)}\right] g_c + \alpha(1)\left[\frac{I_{PE}(0)}{Y(0)}\right] g_h + \alpha(1)\left[\frac{C_{HD}(0)}{Y(0)}\right] g_{C_{HD}} + \alpha(1)\left[\frac{I_H(0)}{Y(0)}\right] g_{I_H}$$
$$+ \alpha(1)\left[\frac{C_G(0)}{Y(0)}\right] g_{C_G} + \alpha(1)\left(\frac{I_G(0)}{Y(0)}\right) g_{I_G} + \alpha(1)\left(\frac{I_{SE}(0)}{Y(0)}\right) g_{I_{SE}}$$
$$+ \alpha(1)\left[\frac{X(0)}{Y(0)}\right] g_X + \frac{\alpha(1)}{\mu(1)} g_{\mu} + \alpha(1)\left(\frac{E(0)}{Y(0)}\right) g_E$$

Part II

Understanding the International Economic Order

8

Continuity and Change in the International Economic Order: Towards a Sraffian Interpretation of the Changing Trend of Commodity Prices in the 2000s

*Franklin Serrano**

8.1 Introduction

In the first decade of the twenty-first century we can observe some elements of continuity and others of change in the international economic order. In terms of continuity, what is perhaps most striking is the resilience of the 'floating dollar standard', which was neither a cause of the major world crisis of 2008, nor was negatively affected by it.[1] In terms of change, there is the new tendency towards a greater relative autonomy of the (relatively fast) rates of growth of GDP of many developing economies from the (low) growth rates of the advanced capitalist countries, A second (and intimately connected) change is the increasing absolute (dollar) and relative prices for internationally traded 'commodities' (food and raw materials in general).

Both new trends are ultimately related to four interconnected features of the world economy after 2003. The first was the relatively low rates

*In addition to the presentation at the Sraffa 2010 Conference at Roma Tre University, the chapter draws on two later presentations on these topics, at CEFIDAR, Buenos Aires, in August 2011 and at the 'Congreso Internacional Crisis de la teoría económica y políticas alternativas ante la crisis global', 7–11 November 2011 at UNAM, Mexico City. The author is grateful to the organisers and participants of these three events. Special thanks are due to very useful discussions on the topic with my colleagues, Professors Carlos Medeiros, Eduardo Crespo and Eduardo Pinto and also to Fabian Amico and an anonymous referee for their comments on an earlier draft. None of them, of course, bears any responsibility for the remaining errors and mistakes.

of interest in the USA and other advanced countries and the accompanying massive gross capital flows towards the developing economies in general. The second was the shift towards policies of fast growth of internal demand in large developing countries and the acceleration of South–South trade. Third, the revival of natural resource nationalism in many developing and transition countries. Finally, there has been a marked improvement in policies for managing the balance of payments in many developing countries. These improved policies have variously included managed floating exchange rate regimes, a massive accumulation of foreign exchange reserves, the early repayment (or default in the case of Argentina) of official external debt, the setting up of sovereign funds and the selective taxation of some commodity exports (and at times subsidy of some commodity imports).[2]

All these elements taken together have led to the remarkable absence of balance of payments crises originating in the periphery and a quick recovery from the late 2008 world financial crisis.

These changes appear to be the combined result of (i) the favourable state of the world economy in terms of both growth of trade and access to capital flows; and (ii) an important reaction to economic policies adopted by developing countries to the various financial and balance of payments crises of the late 1990s up to 2002. In any case those changes in the international situation are behind the partial decoupling of the trend growth rates of developing countries as a whole (both for commodity and industrial-exporting countries) observed in the 2000s. Ultimately they were also very important for the change in relative commodity prices, which in this decade saw a reversal of the falling trend of the two previous ones.

The purpose of this chapter is to provide a preliminary attempt to understand the broad proximate causes of this marked, and to a large extent surprising, change in dollar and relative international 'commodity' prices in this new international context. For reasons of space references to the other new feature of the world economy, the tendency towards a decoupling of trend rates of growth of developing countries and of the political and policy changes in many developing countries that ultimately rendered these two trends possible, despite its importance, will be kept to the minimum required for an understanding of the question in focus, namely, the direct causes of increased dollar and relative international 'commodity' prices.[3] Again due to space limitations discussion is mainly on the analysis of commodity market behaviour and on the rapid increase of dollar commodity prices in the 2000s, which are generally not well understood instead of focusing on the much better-known causes of the low rate of increase of the dollar prices of manufactured exports during this period, which are connected with money wage and productivity trends in both the advanced and the newly industrialised Asian countries.

We consider that the modern classical surplus approach, revived by Sraffa and Garegnani since the early 1960s, provides a useful analytical framework on which to base this study. It is well known that the classical notion of price of production, revived by Sraffa, focuses on the objective material elements in the 'cost of production' of all produced goods. This should also apply to the relatively standardised foods and raw materials traded in international markets that are nowadays known as 'commodities'.[4] Moreover, the theory of prices of production foregrounds the necessary connections between costs of production' and distributive variables, and in particular the (social) rules that govern the distribution of income across between wages, profits and different types of rents (and also exchange rates), in a particular historical situation.

Most analyses of the recent increase in commodity prices have focused almost exclusively on the demand side, whether that of final users or for speculative purposes. There has been a relative neglect of the role of certain supply constraints and of more persistent cost-of-production elements. And there has been, if anything, an even greater neglect of the close connection between these changes in the costs of production and associated changes in the distribution of income.

We should perhaps also note that this neglect of cost-of-production and income distribution aspects in the recent literature contrasts sharply not only with the Sraffian classical surplus approach but also with the analyses of the pioneers of development economics, such as Singer, Prebisch and Lewis, who focused on precisely these aspects when studying the long-run trends of terms of trade.[5] In order to illustrate the close links between our own Sraffian standpoint and that of the pioneers of development economics, we shall discuss below the extent to which both the main changes in the international economic order, namely, the partial decoupling of the growth rates of developing countries and the new trend of relative commodity prices, were in part anticipated as possibilities by pioneer development economist Arthur Lewis.

The analysis will argue that changing international relative commodity prices in the 2000s reflect changes in their relative production costs. Nominal unit dollar costs of commodities have increased rapidly due to the deliberate restriction of the supply of oil (a result of both OPEC and the revival of 'natural resource nationalism'), rising costs of minerals (due in part to capacity constraints on more efficient mines) and the fast growth of real wages in major mineral and agricultural commodity-exporting countries (together with the real appreciation of many of their currencies). On the other hand, unit dollar costs of manufactured export products rose slowly due to the slow growth of real wages in advanced capitalist countries, combined with real wages which, though growing fast, are not keeping pace with growth of industrial productivity in the developing countries that export industrial goods (particularly in China and other Asian NICs).

The argument of this chapter will proceed as follows. Section 8.2 briefly describes the main characteristics of the recent boom in commodity prices. Section 8.3 addresses the *demand*-side elements of the boom in dollar commodity prices. Section 8.4 looks at some of the supply and *cost-of-production* aspects of the rise in dollar prices of specific broad groups of commodities. Section 8.5 first discusses the changes in the balance of payments position of commodity-exporting countries in the 2000s, including the partial decoupling of growth rates of developing countries in general, in order to show the important role for the cost and dollar prices of commodities in general of the trend towards a *nominal (dollar) and real appreciation of the currencies of major commodity-exporting countries*. Section 8.6 deals briefly with some better-known causes of the slow growth of non-commodity international dollar prices, which are needed to understand the question of the rise in the *relative* prices of commodities. Section 8.7 concludes by comparing and constrasting our own suggested interpretation of the recent change in relative prices with the old views of Singer, Prebisch and Lewis.

8.2 The recent boom in commodity prices

Apart from being extremely volatile, international commodity prices in general (agricultural and mineral) have increased substantially between 2000 and 2010, both in absolute (US dollar) terms and in relative terms, whether we compare them with international dollar prices of manufactured goods or in terms of the overall internal price indexes of most countries.

Dollar prices for oil started recovering, after reaching extremely low historical levels in 1999.[6] Nominal increases in metals and food prices came later. By 2003 oil prices started growing much faster and metals prices even more so, and by 2007 the rapid rise in food prices began. In the wake of the global financial crisis, prices of all types of commodities fell drastically in 2009, but quickly recovered in 2010. Prices reached a peak in mid-2011 with food and metals (but not oil) at higher nominal dollar levels than in 2008. Over the whole decade, it was dollar energy prices that increased most, followed by metals, while food prices increased by much less. According to IMF data, crude oil prices increased at an annual rate of 17.84 per cent a year from 1999 to 2002, and then by 18.14 per cent a year from 2003 to 2010. Prices of metals actually decreased by 0.2 per cent a year from 1999 to 2002, but then increased at a very fast rate of 20.36 per cent a year from 2003 to 2010. Food prices increased from 1999 to 2002 at a low rate of 0.28 per cent a year, but rose by 4.3 per cent a year from 2003 to 2010.

Over the same period, international industrial prices and world inflation did not keep pace with such fast commodity price increases, resulting in a large increase in the *relative* price of all types of commodities. Indeed, world inflation actually fell from an average of 4.44 per cent a year during 1999–2002 to 3.87 per cent in 2003–10. Consumer price inflation in the richest developed economies remained practically stable around 2 per cent over the whole decade. But inflation in developing countries fell from 8.19 per cent during 1999–2002 to 6.72 per cent in 2003–2010 in the case of Latin American countries. International dollar prices for manufactured products also increased at a much lower rate than commodity prices. WTO's index of manufacturing unit values (MUVs) actually decreased at a rate of 2 per cent a year from 1999 to 2002 and then increased at an annual rate of 4.85 per cent from 2003 to 2010. For commodities as a whole, 2003–08 was the longest phase of fast-rising dollar prices and had the largest overall through to peak price increase (131 per cent) since 1900 (World Bank, 2009, 2012).

8.3 Demand aspects

8.3.1 The 'China demand' effect for metals

Most analysts attribute the rapid increase in both dollar and relative commodity prices to the acceleration of the growth of the world economy in the 2000s, based on the very fast growth rates of some developing countries. According to this view, the increasing weight of developing countries in the world economy has been characterised by a process of heavy industrialisation and urbanisation, as well as by the spread of new Western consumption habits. This process has increased the demand for oil, metals and food in these countries and is the basis for a 'super cycle' of commodity prices that could last for a couple of decades.

Among these fast-growing developing countries, it is usually China's extremely rapid commodity import growth rate that is emphasised. Indeed, between 2002 and 2003, Chinese imports of commodities increased more than 40 per cent. In 2003 China was responsible for 26.5 per cent of the world demand for steel, 19.8 per cent for copper and 19 per cent for aluminum. This account, however, has a number of serious limitations as an explanation for rising commodity prices.

First of all, the world economy did not actually grow significantly faster in the 2000s compared with the 1990s. In fact, it grew much faster in the second half of the 1990s than in the 2000s, as the US 'dotcom boom' then coincided with China's very high growth rate. It is true that the growth rate of developing countries was, as a whole, faster in the 2000s than in the 1990s but this largely just made up for the reduction in the growth rate of the advanced economies.

The same growth pattern can also be seen in indices of world industrial production and the volume of world trade in merchandise.

Thus neither world GDP, nor world industrial production nor world volume of trade grew relatively faster in the 2000s than the 1990s, even if we exclude the big recession year of 2009 (as perhaps we should not, since periodic crises are, of course, a constituent part of the commodities and business cycles).

Another problem with explanations centred on the side of demand is that, as is well known, the income elasticity of world demand for most commodities is definitely below one, reflecting both technical change and the usual trend towards reduced intensity of commodities in GDP as income increases. Indeed, since the early 1970s, the decline of commodity intensities has been faster for food and energy commodities. In the case of metals we observe the same, though a little less pronounced, declining intensity, when we exclude China. But in this one case, the China effect on demand is so great that measures of the metals intensity of world GDP since 1995 increase, rather than decrease. Since then, indeed, there has been increasing metal intensity of world GDP. This means that since 1995 the income elasticity of the world demand for metals has definitely been substantially greater than one. In fact China's metal intensity relative to GDP was three times higher than that of the rest of the world in 1990 and became nine times higher than the world's average by 2008 (World Bank, 2009, 2012).

The third problem is that data from China confirm that the 'China demand effect' has only been highly relevant for the world economy as a whole as regards metals. While it is true that consumption and imports of all types of commodities in China have been growing rapidly, in most cases they started from a *very low base* level. Even for a few non-metal products where this was not the case and Chinese consumption has had a greater weight in world demand, China's role seems to have been mostly partially to compensate for the marked decline in the demand for commodities coming from the richer countries.

The big exception, where the Chinese demand has really had a big impact was the extraordinarily high increase in the use of metals in China. This has been associated with the very high growth rates in public and private investment, particularly in construction and infrastructure in general. In the period between 2002 and 2007, China's consumption of coffee increased 32 per cent, but as the world's total consumption actually decreased by 1.9 per cent over the same period, China's contribution to world demand growth in that period was just 0.1 per cent. Beef consumption grew 27 per cent in China over that period, during which world consumption of beef grew 7.2 per cent, to which China contributed only 2 per cent.

In the case of oil, China's consumption grew 48.7 per cent while the consumption in the rest of the word grew 6.6 per cent; the Chinese contribution to world demand growth was of 2.7 per cent over a five-year period (about 0.5 per cent a year).

The case of metals contrasts sharply with that general pattern. Chinese consumption of iron ore increased by 224.9 per cent between 2002 and 2007, while in the rest of the world it grew by 19.5 per cent, and China's contribution to world consumption growth was no less than 38.4 per cent. In the case of other metals such as aluminum, zinc and copper Chinese consumption grew at less spectacular rates over the period, between 70 and 125 per cent, but there was still a 'China effect' of between 10 and 20 per cent of the growth in world consumption of those metals over that five-year period (Jenkins, 2011).

The upshot of all this is that the China effect on the world demand for commodities, apart from metals, has been much smaller than is usually thought.

And even in the case of metals, where massive Chinese demand was of crucial importance for the fast growth of world demand, we find that the acceleration of the world demand for metals started around 1995, with a further acceleration around 2001, while, as we saw in Section 8.2, the prices of metals only start booming after 2003. Something seems to be missing in the demand-side explanation, even for metals.

8.3.2 Speculation

In part precisely because of the quite modest growth in 'requirements of use' (or final demand) for many commodities whose prices also increased sharply, many analysts have attributed the rise in commodity prices to increased speculation. Three elements have appeared, separately or together, as explanatory variables for the increase in speculation in commodity markets: falling interest rates in the USA; the devaluation of the dollar relative to other major currencies such as the euro; and financial deregulation and financial innovation in futures markets for commodities.

On this view, reduced short-term interest rates in the USA would have reduced the attractiveness of financial assets relative to commodities and cheapened the formation of speculative inventories. Moreover, the increase in dollar commodity prices would be a leading indicator that monetary policy is too expansionary. Low interest rates would later lead to a substantial increase in overall inflation and speculators would be anticipating this higher future inflation caused by monetary policy, using commodities as a hedge.

In reality, the idea that lower interest rates may to a certain extent increase the speculative demand for commodity prices is in itself reasonable. But the idea that the large increases in commodity prices that have occurred in the 2000s can be attributed to market expectations of a massive future acceleration of inflation in the USA does not make much sense and is not supported by independent evidence.. 'Core' and/or trend inflation in the USA and other developed countries has been kept low, in spite of drastic increases in dollar commodity prices. The ultimate cause of this regime of inflation moderation appears to be the weak bargaining power of American workers, reinforced by the relatively slow growth of employment in advanced countries and it seems clear that market participants understand that this trend is not likely to change any time soon.[7]

Other analysts argue that commodity speculators are hedging against the devaluation of the dollar against key currencies such as the euro or yen (i.e., currencies of rich countries that do not export commodities). It is pointed out that during some of the sharper short-term dollar commodity price increases the dollar was falling in value relative to these currencies. In fact this presumed correlation between dollar commodity prices and the value of the dollar relative to currencies of rich non-commodity-exporting countries is not at all robust and there are many periods, such as between 1984 and 1995, in which a large and almost continuous fall in the value of the dollar relative to these currencies has coincided with low and falling dollar commodity prices. It is more likely that in certain periods when this correlation can be observed, the falling dollar and rising commodity prices could be responding to a common third cause, such as low interest rate in the USA relative to those of other rich countries.[8]

It is unlikely, however, that lower American interest rates could, on their own, cause such large swings in dollar commodity prices. Indeed, while changes in the interest rate could be having some effect, it is interesting to note that this variable does not appear as a statistically significant driver of dollar commodity prices even in recent studies made by leading advocates of this view.[9] It is probable that this small effect has been totally overshadowed by the wild short-term fluctuations of expectations in the organised commodity markets.

Indeed, the process of financial deregulation and financial innovation, particularly in the American economy, brought an enormous amount of financial funds to commodities futures markets. By some estimates the value of funds directed to these markets by purely financial speculators grew from 13 billion dollars in 2003 to more than 260 billion dollars by March 2008. For many this was a major cause both of the high volatility of dollar commodity prices and, more unusually, also of their rising trend.[10]

On the other hand, those who want to deny that a speculative bubble could be amplifying the recent rise in dollar commodity prices, have two main arguments. First, the data do not show episodes of very large accumulation of physical inventories of commodities, which is considered a necessary 'signature' of speculative activity.[11] Moreover, it is argued that the volume of transactions in the futures markets merely reflects the compensating bets of different agents and does not affect the physical availability of commodities.[12] Therefore, activity in futures markets cannot affect the balance of supply and demand in the spot market, and thus has no effect on prices. Neither of these arguments, however, seems to be very solid.

A large accumulation of inventories is not really a necessary condition for speculation. Those who believe that it is argue that a large accumulation of inventories would be needed to sustain a speculative price increase, in order to compensate for the equilibrating role of the large fall in demand and large increase in supply that presumably occur when the spot price increases.

But perhaps one of the most important features of the products that we call commodities is precisely the difficulty of both decreasing significantly the quantity demanded and increasing significantly the quantity produced in the very short run. Very small imbalances between supply and demand can therefore lead to very large short-term primary fluctuations in spot prices, and divergent expectations between speculators allow the price movements to be magnified in both directions, as the same limited amount of inventories is sold and resold with speculators selling to other speculators who think the price will continue to rise, for instance.[13]

As for the argument of the irrelevance of activity in the futures markets, it fails to see the relevance of the potentially riskless gains of arbitrage between spot and future market prices of the same standardised commodity.

Both the spot and the forward prices of commodities are strongly influenced by the spot prices expected in subsequent periods. If there is a general expectation that the spot price will be higher at a subsequent date, both the spot and the forward prices will tend to increase right now. If speculators, expecting a higher price tomorrow, buy spot today, thinking of selling the commodity tomorrow, the spot price today will tend to increase. And such increase in the spot price is transmitted to the futures market, since now there is the option to reduce the supply allocated for future delivery and sell it today at the initially higher current spot price.

On the other hand, it is also true that if there is a large increase in purchases in the futures market today the spot price will rise today, because now investors have the option of selling more at the higher current forward price for future delivery. Recent financial deregulation and innovation, and in particular the extremely low margin requirements that allow extraordinarily high leverage ratios for financial speculators in commodity markets has greatly increased the availability of credit for speculative purchases in future commodities markets that are very quickly transmitted by arbitrage also to spot market prices.

There is also the counter-argument that volatility was also high for a number of commodities which have no organised futures market, such as rice and iron ore.[14] But this probably means that in these markets there were other forms of access to cheap and plentiful credit, not that no speculation occurred in these spot or the other futures markets.[15] Given all this, it seems reasonable to argue that, in spite of the controversy on the size of the effect of recent regulatory changes on future markets, overall speculation has indeed played an increasingly important role in world commodity markets over the 2000s.

But speculation usually works both ways, sometimes greatly intensifying the price increase when output is perhaps only a little lower than current final demand, and at other times causing dramatic price falls when output is greater than final demand. But then speculation cannot really explain the rising dollar and relative prices of commodities, for there seems to be no obvious reason why the massive short-term price increases have been, on average, so much higher than the, also very large, short-term decreases in commodity prices. Speculation, as we have seen, depends crucially on the expected spot market prices. Why, within such instability and volatility, were expected prices rising, rather than falling or being nearly random?

We thus see that to look only at the demand, whether final or speculative, is either misleading or, in the case of metals where world demand really boomed, incomplete. In order to understand the rise in commodity prices we must look at the amounts and specially the costs of the commodities brought to the market (supply). As Garegnani (1988, p. 254) put it: 'After all, we do explain and try to forecast the trend of the price of, say, copper, on the basis of the technical changes in its mining or the richeness of new mines, etc. – in spite of the fact that the prices of copper may fluctuate perhaps as much as 50% on either side of its trend value'.

8.4 Specific cost-of-production and supply aspects

8.4.1 Oil

Oil prices started to rise earlier and higher than other types of commodities. Here the China (and India) demand effect is quite widely known but particularly misleading. China's oil consumption amounted to less than 8 per cent of world consumption in 2008, and India's to around 3 per cent. And overall world oil consumption grew a little more than 2 per cent a year from 2000 to 2008.[16]

Clearly we must turn to the supply side for a satisfactory explanation of the rising trend of dollar oil prices. But we should be clear what exactly we mean by supply. One of the meanings of supply is the cost of production, and of particular interest is the cost of the oil production methods that have to be used to meet demand. Another meaning refers to the availability of physical quantities. And by availability we may mean either the existing inventories of already produced oil; the current productive capacity that would allow rapid expansion of output to meet demand; or the known deposits or reserves of oil under the ground (or ocean); or even the total physical or geological endowment of the resource left in the planet.

World supply of oil in the sense of existing inventories has not, on average, been scarce relative to demand over the years in which oil prices have been increasing, with world production closely matching the trend of world consumption.

Despite the popular view of increasing physical or geological scarcity and the Hubbert Curve, and the world 'peak oil' doctrine, the availability of oil in the sense of known reserves has not really been an issue (Radetzki, 2010). Not only were the world's oil reserves substantially higher in 2010 than in 2000 in absolute terms, but more importantly, the ratio of current proven reserves to current production reserves actually showed a modest increase – from 40 to 42 – between 2002 and 2007 (World Bank, 2009). In fact this fundamental overabundance of oil reserves together with the crucial role of oil as a basic good that is used directly and indirectly in the production of every other good is the key to understanding that royalties do not reflect a presumed looming physical scarcity of the resource, but mostly political power and strategic policies of states.[17]

Therefore, it is to governments' energy strategies and to productive capacity constraints and costs (including royalties charged) that we must mainly refer if we want to understand the supply side of the oil market.[18] These productive capacities and associated costs, together with American geopolitical and energy security policies, generate different types of monopoly, absolute and differential rents for public and private resource owners and producers.

A key feature of the world oil market is the fact that in the lowest cost and highest reserve regions, which are mostly OPEC countries, not only are the reserve deposits of oil vastly overabundant but the cartel also tends to produce at much less than potential full capacity.

Traditionally, Saudi Arabia has played the role of 'swing producer', keeping a sizeable planned degree of spare capacity to smooth the adjustment of current supply and demand in the world oil market and enforcing OPEC member coordination.

Saudi Arabia's ultimate objective seems to have been an attempt to restrict the supply of oil from OPEC countries. This policy has, over a long period of time, allowed market oil prices not to fall for too long below an informal and tacitly acknowledged floor: a price high enough to cover the production costs of the US and Canadian oil industries. This has been very important to ensure the long-run survival and profitability of the huge (and politically powerful) but relatively high-cost American oil industry. This fundamental aspect of American strategic energy security policy is based on the special geopolitical relationship between Saudi Arabia and the USA. The informal floor for oil supply prices creates a peculiar kind of classical *monopoly rent* in the OPEC countries which is then shared between OPEC members and the big multinational oil companies according to the (sometimes shifting) bargaining power of the two groups. Thus, OPEC's royalties are determined as a share of such specific monopoly rents.[19]

Note that the American floor price of production also includes, besides the usual elements of production costs, the royalties received by the (generally private and numerous) owners of American oil resources. These royalties are an *absolute rent* determined by the relative bargaining power of the owners in relation to the extraction industry, a bargain that is directly affected by various aspects of the American government's overall energy policy.[20]

Note that when market prices for oil are oscillating around this American floor price, production of oil in other countries and regions where the costs of production (including politically determined royalty rates, as in most cases the ownership of the subsoil resources is public) are lower than in the USA, albeit higher than those of the OPEC members, does generate classical *differential rents* for these countries.

But the American price of production is simply the oil production floor price. When world demand for oil increases sufficiently, beyond OPEC output and the productive capacity of American and other intermediate-cost regions, production becomes viable in regions of the world where costs (including absolute rent taken as state royalties) are much higher. And the actual price of oil production in these high demand conditions is given by the cost of the productive capacity that has to be activated to meet world demand.[21] This higher price of production generates further differential rents for all other lower-cost regions, even if their costs are above those that set the American floor price. Just to give a curious example, in 2008, when the demand for oil was growing fast and the market price for oil reached record levels, there were reports that in South Africa, extremely costly and highly polluting coal was being used to produce synthetic oil, something that seems to have been last done in Germany towards the end of World War II.

Equipped with this view of the structure of the world oil market, we can now turn to the increase in market prices during the last decade. The dollar market price for oil began to rise from the record low prices of 1999. Those market prices were initially substantially below the floor American price of production described above and seriously threatened both OPEC rents and royalties and the viability of the American oil industry. Then some members of OPEC, in particular Venezuela and Saudi Arabia, made an effort to coordinate all OPEC members and restrict both current output and investment in new capacity, in order to reduce the massive unplanned spare capacity both in OPEC and in other higher-cost producing countries.[22]

This tactic was quite successful and market prices began to recover rapidly. Later, as demand also began to increase more rapidly after 2003, with OPEC skillfully managing not to increase its production in line with it,[23] average market prices began to rise and most of the new production to meet rising world demand had to come from regions with higher costs for technological, geological, environmental or regulatory reasons, such as the tar sands of Canada, which nowadays seem to be setting the American floor price of production. Over time, exports to meet world demand even came from regions with very high production costs, such as offshore oil from Brazil.

This process of rising dollar prices of production and rents has been generally misunderstood as being either the consequence of the always predicted, and always missed, impending world peak of total oil production or, more plausibly, as an indication that 'all the cheap stuff has gone' with some analysts expressing doubts as to the true level of remaining reserves even in the OPEC countries.[24]

In many non-OPEC countries a very important element of the rising price of oil production seems to have been a substantial revival of 'natural resource nationalism' as a large number of developing countries, in most regions and even in Africa, seized the opportunity of rising market prices to renegotiate contracts with international private oil companies on more favourable terms. This process has increased state control of oil reserves and substantially raised the royalty rates and thus the absolute rent component of the price of production of oil. This movement stands in marked contrast to what had happened in the 1990s, where the subservient attitude of the state in oil producing countries was the international norm, as exemplified, for instance, by the contrast between contracts made by the Russian state under Presidents Yeltsin and Putin, or by the changing attitudes of Latin American and African governments in their relations with multinational oil corporations.[25]

8.4.2 Metals

Unlike oil but similarly to the case of coal, the ratio of current deposits to current production has decreased for many metals over the 2000s due to the rapid growth of world demand. Indeed, ratios of reserves to current production decreased between 2000 and 2007 for products such as bauxite, iron ore, nickel, tin and zinc, increasing somewhat for copper and lead (World Bank, 2009).

Note, however, that the mineral deposits described in these indices are those in existing mining areas, not the total fixed and finite physical availability of the metal in the earth's crust. In fact, there is not much interest in searching for new mining areas when existing ones have deposits that can last many decades at current rates of production. For instance, while the ratio of iron reserves to current production has fallen from 132 years' equivalent production to 79 since 2007, the amount of iron ore actually available on earth in the mid-2000s was estimated at about 120 million years' equivalent production, and about 2.5 billion years' for copper.[26] Thus, what really matters on the supply side is the existing productive capacity and the extraction costs. As in the case of oil, there is a long lead time for increases in mining productive capacity (some estimate more than 5 years on average). An unexpected change in demand may therefore leave the mining industry as a whole with large amounts of unplanned spare capacity, which can often lead to the complete shutdown of production in the higher-cost mines when demand reduces. Conversely, if demand accelerates unexpectedly, market prices rise to the point that it becomes viable to operate mines at high or rising extraction costs for long periods of time, generating differential rents for the producers with lower costs. Over longer periods of time this tendency towards increasing costs may be, and has historically been, checked by major improvements in extraction technology and technical progress in general. Also, similarly to oil, mining deposits often pay royalties as an institutionally and conventionally determined absolute rent for the private or more often state owners of subsoil rights.

Let us now turn to the rapid increase in metal dollar prices in the 2000s. In spite of the rapid rise in demand since the mid-1990s, driven by the acceleration of Chinese rates of infrastructure investment, and also of the increase in energy costs driven by the recovering oil price, dollar prices of metals only started to grow rapidly after 2003, and even more so than did energy prices after 2006.

The explanation for the initial period of very low dollar prices seems to be connected with the very high spare capacity for low-cost producers that came as the metals intensity of world demand fell from the mid-1980s to the mid-1990s. Over time, as demand continued to grow rapidly, spare capacity fell and at the same time new capacity has been growing slowly due to the long lead times and the further massive acceleration of demand after 2003. This process made the operation of higher-cost producers and increased differential rents viable. The market situation strengthened the bargaining power of the state in many developing countries, especially in relation to multinational private firms. This contributed to the marked revival of 'natural resource nationalism' that has been observed in mineral-producing countries in South America and Africa, and in Russia and parts of Asia. These governments have successfully increased royalties and thus the absolute rent component of the price of production of many metals.

Therefore, among the main causes of the boom in dollar metal prices after 2003 are rising energy costs, productive capacity constraints for lower-cost producers, rising extraction costs and 'natural resource nationalism' (Gopinath, 2011; Radetzki, 2012b).

8.4.3 Food prices

Dollar food prices increased much less and much later than oil or metal prices. Agricultural supply can respond to increases in demand much quicker than it is the case for minerals. On the other hand, due to its dependence on weather conditions, food output can change erratically over short periods. Thus, in the case of food, not only has demand grown relatively slowly but there has also been no clear sustained need to produce under increasing costs. Of course, the quality and quantity of specific types of land may be scarce and differential rents are generated on land with higher costs, but, contrary to what many believe, there is no inherent tendency towards decreasing returns in this sector.[27] Technical progress also appears in general to have been faster in food production than in mineral or oil extraction. For all these reasons agriculture price trends are much more clearly dominated by events on the supply side.

During the recent boom, current supply conditions as expressed by the stock-to-use ratios remained fairly stable (with the exception of wheat where the ratio fell), showing that production in general grew in line with demand.[28]

Agricultural dollar prices seem to have increased, first because of the effects of rapidly rising oil prices on energy and fertilizer costs, but also, for specific crops, because of the energy policies of the USA and the European Union.

The sudden very rapid increase in the demand for biofuels in these regions since 2006 seems to have strengthened the link between the oil price and some agricultural prices. It created an *opportunity cost* such that when the price of food that can be used as biofuel falls below a certain level determined in relation to the oil price, it is more profitable to divert the crop to meet the large and rising demand for biofuels rather than allow the market price to fall further.[29] Given this link, rising oil prices increase the floor below which the prices of these crops do not fall.[30]

8.5 General costs: the role of exchange rates

8.5.1 Decoupling growth under financial integration under the floating dollar standard

While the devaluation of the dollar relative to the key currencies of advanced economies does not seem to have played a major role in the rise of dollar commodity prices, its devaluation *relative to the currencies* *of the set of commodity-exporting countries* may have been an important element in the rise of dollar commodity prices. The revaluation of these currencies relative to the dollar increases the unit costs of production of all types of commodities, measured in dollars[31] (particularly where labour costs are more significant).[32] But in order to understand this effect we must discuss some key aspects of the balance of payments position of commodity-exporting countries during the 2000s.

After the series of balance of payments crises of the late 1990s and early 2000s many developing (or 'emerging') economies, including both industrial exporters and a large number of commodity-exporting countries, adopted a deliberate policy of reducing their external vulnerability. These countries made a huge effort to pay back their foreign debt (both private and official) and accumulate foreign exchange reserves, and many set up large sovereign stabilisation funds. Most also adopted heavily-managed floating exchange rate regimes to mitigate speculative pressures. When the world economy, international capital flows and commodity export volumes and dollar prices started growing rapidly after 2003, with international trade expansion fuelled largely by the rapid growth of both international trade and internal markets in the major industrial-exporting developing economies in Asia, these changes in macroeconomic policies allowed many of the commodity-exporting countries to grow without incurring major current account deficits and external debts. This, together with improved management of exchange rates and short-term foreign capital inflows, led to a drastic reduction in interest rate spreads for commodity-exporting countries and thus a major, and unprecedented, improvement in their balance of payments position. With a marked decrease in external vulnerability of these economies, no serious currency crises have since arisen in the commodity periphery.

This gave many developing countries (including in Africa and Latin America) scope to implement anti-cyclical macroeconomic policies and, more importantly, led to the growth of both external and internal markets, investment and productive capacities.

Led by China, the so-called South–South trade grew very rapidly in most regions. This led to a partial decoupling of the *trend* rates of growth of this large group of developing countries in the 2000s from the (slow) rate of growth of advanced countries, though trade and financial integration has made the cyclical fluctuations around these higher trends correlate very strongly with the fate of the world economy as a whole and hence with the cyclical fluctuations of advanced capitalist countries.[33] This decoupling of growth rates was anticipated in part by Lewis. In his Nobel lecture (published as Lewis, 1980) Lewis argued that it would be possible for developing countries to grow fast, even in the light of a slowdown in the growth of advanced capitalist countries, if the South–South trade among developing countries grew fast enough. This required that larger countries should rapidly expand both their internal markets and imports, acting as growth locomotives.

However, as correctly pointed out by Akyüz (2012), this decoupling of trend is only partial, because in spite of the improved policies and management of capital inflows the growth of developing countries in the 2000s still depended very much on these massive capital flows from the advanced capitalist economies. There has been a decoupling of the trend of growth[34] but certainly not a financial decoupling (see also Yeyati and Williams, 2012).

Indeed, this marked change in the international economic order has been very dependent on one particular element of continuity within it: the floating dollar standard. The fact that the US dollar remains the dominant international currency means that it is still the American central bank that sets the basic international interest rate and that American monetary policy and developments in the US (and to a much smaller extent European) financial markets play a key role in the dynamics of the pricing and quantities of international capital flows towards developing countries.[35]

Notwithstanding the much improved management of these international capital inflows by developing countries in general and by many important commodity-exporting countries in particular,[36] the relatively low public and private rates of interest in the USA were absolutely crucial in creating favourable conditions for the large amounts of capital flows of all kinds directed towards the developing world, and the experience of the 2000s cannot be understood without taking into account these elements that are, of course, completely beyond the control of the developing countries.

8.5.2 The devaluation of the dollar relative to commodity currencies

While the macroeconomic policies of commodity-exporting countries have (with a few exceptions) been successful in spurring growth while preventing large current account deficits, and the massive capital inflows have allowed the continuous build-up of foreign exchange reserves, what really matters to us here is that a tendency towards the appreciation of the real exchange rate of these countries as a whole gradually set in. In some countries it came as a gradual trend towards nominal appreciation which was accommodated in order to control domestic inflation. In a few money wages tended to grow rapidly relative to productivity due to much better conditions in the labour market and more nationalist or progressive governments (such as in Russia, Argentina and Venezuela). Some other countries had some combination of both tendencies. The consequence of this general tendency towards a real revaluation of the national currencies of commodity exporters relative to the US dollar has been, since 2003, the relatively rapid increase in the labour unit costs in dollars of most commodity-exporting countries. Thus, the revaluation of the 'commodity currencies' relative to the dollar has been another of the important neglected elements affecting the dollar cost of production and dollar prices of most commodities (particularly food but also minerals) during the recent boom.[37]

The importance of the balance of payments position (and hence of the exchange rates) of commodity-exporting countries in general as a main determinant of the terms of trade of the 'periphery' in earlier cycles has been explicitly emphasised by Patnaik (2002)[38] and by Ginzburg and Simonazzi (2004), where the connection between exchange rate devaluations and falling real wages in commodity-exporting countries is made clear. It is also most likely that the strong correlation between the growth of world demand and relative non-oil commodity prices over long periods found by Erten and Ocampo (2012) may be reflecting the fact (stressed by Ginzburg and Simonazzi) that in periods of rapid world growth, because of both trade and capital flows, the nominal and real exchange rate of many of the periphery countries tends to appreciate and periods of slow growth and crisis are marked by a series of devaluations in the commodity-exporting periphery. It is the peculiarity of the cycle of the 2000s that during the global crisis such devaluations were sharp but quickly reversed since the balance of payments position of developing commodity-exporting countries has improved drastically relative to earlier commodity price cycles.

Note that in this interpretation, although financial speculation is not seen as a direct determinant of the trend of commodity prices, international financial developments matter a lot, since the international capital inflows and their management by commodity-exporting countries are both seen as crucial aspects of the improved balance of payments position of these countries, which ultimately explains the course of their exchange rate.

8.6 The change in relative commodity prices: the 'China cost effect'

The large increases in nominal dollar prices we have discussed eventually transformed themselves into large increases in relative prices, because the international prices in dollars of non-commodity tradable industrial goods grew at modest rates and domestic inflation rates (which depend on prices of non-tradable services) were kept low in most countries.

The fact that in spite of the commodity price boom inflation did not accelerate in advanced countries allowed the boom to last much longer than had been the case in earlier cycles. On the demand side, the moderate inflation prevented governments from thinking they had to pursue restrictive policies, which if adopted would have cut back world demand for commodities and brought commodity prices down. On the supply side, the absence of persistent inflation prevented drastic increases in the prices of other goods and services as a reaction to rapidly rising dollar commodity prices, which if it had occurred sufficiently would have cancelled out the increase in *relative* commodity prices.

The ultimate reason for this low inflationary impact of the commodity boom seems to be the unusually low bargaining power of industrial workers in most advanced economies. The strong competition from low dollar labour cost industrial-exporting developing countries such as China appears to have been one of many key elements that have weakened the bargaining power of the workers in advanced economies.

We can see, then, that although the 'China effect' on demand has been greatly overestimated, another type of 'China effect', this time in terms of costs, has played an important role in the increase in the relative price of commodities that has not generally been recognised.

The low level and slow growth of dollar unit labour costs for more sophisticated manufactured goods exported from China, Mexico and many other industrialising developing countries seem to have ultimately played an important role in sustaining the great change in the relative prices of commodities. This has been due to a peculiar combination of the relatively low growth of money wages, very rapid growth in labour productivity and a deliberate attempt by governments to prevent or at least minimise the appreciation of the currencies of industrial-exporting developing countries relative to the US dollar (which in fact generally revalued much less than the currencies of most commodity-exporting countries).

8.7 Concluding remarks: Singer-Prebisch, Singer 2 and Lewis 1 and 2

It is well known that Singer (1950) and Prebisch (1949) attributed the long-run decline in relative commodity prices to the fact that real unit labour costs in the central industrial countries tended to remain fairly stable, even if the productivity rate of the manufactured goods exported by the centre was growing much faster than that of the commodities produced by the periphery. This, they stated, reflected the strong bargaining position of the workers in advanced economies. This bargaining power made both money and real wages grow more or less in line with productivity. In the periphery with its low productivity in food and virtually unlimited supply of labour, real unit labour costs of the exported commodities tended to fall as real wages in the periphery did not grow in line with the growth of productivity.

Lewis (1978) also argued that the terms of trade between the centre and the periphery depended on the levels of the relative real wage in these two regions and not really on the type of product that was exported.

As in the 1960s and 1970s, many developing countries began to industrialise and to produce and even export simpler industrial goods. Singer (1998) noted that this had not changed the terms of trade in favour of the periphery. He then formulated what he called the Singer 2 hypothesis, according to which the key point was that the periphery continued to export unsophisticated goods of low unit value and low technological content, implying a 'commodification' of those industrial goods.

However, since the 1990s a number of low-wage developing countries, especially in Asia, have started to export far more sophisticated industrial goods and even some services (as shown by the information technology services exports of India). Contrary to the models of Singer-Prebisch, Lewis and Singer 2, these developing countries now export many (but not all) of the same products as the advanced central countries. Thus, neither the Singer-Prebisch nor the Lewis nor the Singer 2 hypotheses appear to be an entirely valid explanation of the recent trend of relative commodity prices. In fact Lewis himself anticipated in 1976 that if low-wage countries, in addition to commodities, started also to export large quantities of sophisticated industrial goods, then the level of terms of trade would turn against the latter. This possible Lewis 2 hypothesis seems to have become relevant in recent years; since productivity has tended to grow much faster in industrial goods (especially in electronics) than in commodities and nowadays neither in the old countries of the centre nor in the new industrial-exporting periphery are real wages growing in line with productivity, the real unit labour costs of industrial goods would tend to fall relative to those of commodities.

In our view, what he have called the China cost effect shows that Lewis's intuition was right and that for the moment that secular declining trend of terms of trade has been reversed. For now real wages are not growing in line with productivity in the central countries, partly because of the fierce competition from low-wage workers in the periphery that now exports sophisticated industrial goods. Moreover, as we saw above, this movement towards higher relative prices for commodities is intensified by the tendency for the exchange rates of commodity-exporting countries to appreciate relative both to the dollar and to the currencies of the industrial-exporting developing countries. This seems to be leading to more rapid increases in real and dollar wages in the commodity-exporting countries and stagnant real wages in the advanced industrial economies, the latter constrained not only by the lower wages but also by the much more rapidly growing productivity of developing industrial-exporting countries such as China (and others mainly in East Asia).[39]

Notes

1. For reasons of space and scope very little will be said in this chapter about this crucial aspect of the international economic order (which in practice means that the US economy is not subjected to balance of payments constraints) apart from brief remarks in Section 8.5 below. For a more detailed discussion of the 'floating dollar standard' see Serrano (2003, 2004, 2008) and Feldman (2009).
2. On these changes see Serrano (2008), Akyüz (2012) and Frenkel and Rapetti (2011). Note that for many authors such as Frenkel and Rapetti (20011) the improved balance of payments situation is considered good because it presumably helps each developing country seen in isolation to avoid a tendency towards the overvaluation of the exchange rate of their currencies. Here (see section 8.5 below) the issue is that it allows commodity-exporting countries as a whole to avoid the opposite tendency towards competitive uncoordinated exchange rate devaluations that depress commodity prices and terms of trade.
3. A more detailed analysis of these new aspects of the international economic order can be found in Freitas, Medeiros and Serrano (2013).
4. Having clarified its precise meaning in this context the word commodities will not be written in quotes in the rest of the chapter.
5. For a classical surplus 'cost of production' interpretation of the previous period of rising commodity prices in the 1970s see Sylos-Labini (1982).
6. Most of the data in this and in the next section come from Ferreira (2012).
7. For data that confirm the limited impact of recent commodity price increases in the trend of US inflation (in contrast with what happened in the 1970s) see Tootell (2011). On the causes of the low bargaining power of workers in advanced capitalist countries see Pivetti (2011).
8. In fact, perhaps the only relevant, though indirect, effect of the nominal devaluation of the dollar relative to the currencies of the other advanced economies is that it moderates the nominal and real increase of oil and other commodities in terms of the domestic currencies of these countries.

This reduces the magnitude of the initial domestic inflationary shock of the dollar commodity price boom and perhaps helps the boom to continue, by preventing more drastic action of the central banks to fight domestic inflation. Such measures could end up slowing down the growth of aggregate demand in these rich economies, thereby cutting the world demand for commodities, which would lower current dollar commodity prices. This possibility is raised in UNCTAD (2008). This latter effect does not seem to have been decisive, for even in the USA, due to the weak bargaining position of workers, the fact that the economy faced the full impact of the rising dollar commodity and specially oil prices did not generate a price-wage-price spiral that could increase core inflation and lead to contractionary measures by the central bank (see previous footnote).

9. Jeffrey Frankel has been the most emphatic advocate of this view and has incorrectly been predicting high inflation in rich countries for many years (Frankel, 2006). In Frankel and Rose (2009) we find the admission that the lack of a reliable significant econometric effect of low interest rates of commodity prices is a 'disappointment'.
10. For the former view see Pollin and Heintz (2011) and the latter see Wray (2008).
11. This point was raised by Krugman (2008).
12. See Irwin and Sanders (2011).
13. Note that classical economists from Smith to modern Sraffians were quite aware of the importance of all this and other causes of irregular oscillations of market prices in the very short run and thus did not think it was either possible or useful to postulate either demand or supply as regular functions of prices, using instead only the twin concepts of effectual demand and natural price to analyse the long-run trends of prices and quantities (Garegnani, 1983).
14. Note that, on the other hand, for obvious material reasons there is really no organised market for *immediate* physical delivery of oil. What is called the spot market for oil is actually a 'relatively near future' market.
15. As due to the possibility of arbitrage expected spot prices affect strongly both the current spot and futures market prices, persistent diferences between current spot and future market prices can be traced to variable margins connected to perceived financial costs and risks and convenience yields attributed to assured delivery by market participants. See Kaldor (1939).
16. See Serrano (2008).
17. Note that, contrary to appearances, the fact that oil is not scarce should make environmental concerns more rather than less serious. Given that the use of oil causes many forms of undesirable pollution, its high energy efficiency and lack of scarcity mean that there is no reason why the serious negative externalities coming from its excessive use would be somehow reflected in its market price.
18. Ravagnani (2008) provides a detailed critique of attempts to use the notion of a given, known and scarce finite stock of resources and the associated 'Hotelling rule' for the determination of the price of production of oil and other non-renewable natural resources.
19. See Serrano (2004). Adelman has for decades been making the crucial point that Middle East oil is not scarce (Adelman, 2004; Adelman and Watkins, 2008). Roncaglia (2003) provides a very interesting analysis of the

relationship between OPEC members, big oil companies and the USA but unfortunately his argument that this could be seen as a case of a 'trilateral oligopoly' obscures certain issues. First, there is no such thing as oligopoly in classical competition theory, where it is the degree of capital mobility and not the number of firms that is the relevant measure of the degree of competition. Second, and more importantly, the idea of a trilateral oligopoly leads to a misunderstanding of the superficially puzzling behaviour of Saudi Arabia which reflects their special political relationship with the USA and is therefore not the result of an autonomous notion of Saudi national interest. Routledge (2003) drew attention to the crucial issue of the coordination between OPEC and the USA in order to prevent the price of oil from falling below the American 'floor supply price'.

20. As shown by Piccioni and Ravagnani (2002) and Ravagnani (2008). Note that Fratini (2009) draws attention to the fact that absolute rent, which should be taken as given as a share of the gross output or the value of the product is not the same thing as a monopoly rent (which would appear to be the difference between an exogenously given price of product and its cost of production). For an excellent discussion of the different types of rent in both the old and the modern classical surplus approach see Fratini (2008).
21. Both Schefold (2001) and Kurz and Salvadori (2009) draw attention to the importance of productive capacity constraints on the extraction of minerals being one important precondition to use differential rent theory in the case of non-renewable natural resources.
22. See Routledge (2003).
23. See Davidson (2008).
24. See Lynch (2006) and Davidson (2008) for a critique of these views.
25. On the fall and rise of natural resource nationalism, see Medeiros (2011). On the important role of the increase in the nationalisation of most of the world's oil reserves during the 2000s for the oil price increase, see Radetzki (2012a).
26. See Tilton (2009). Tilton adds for the sake of comparison that our solar system is about 5 billion years old and will not last more than 10 billion years. Again, as in the case of oil, this abundance of the resource only makes it even more unlikely that the negative environmental externalities associated with mining, which can be substantial, would ever be reflected in the market prices of minerals.
27. Pasinetti (1999) explains very well how, even in Ricardo, the assumption of increasing costs in corn production was an historically contingent empirical assumption instead of the result of a general or natural law.
28. See Baffes and Haniotis (2010).
29. Fratini (2008) mentions that such cases should be taken as further evidence of the existence of absolute rents. In the end he may well be right but we think it seems more appropriate to take these opportunity costs as yielding a type of differential rent, reserving absolute rent only to those rents that are determined directly by bargaining and political and institutional forces (and not by differences in cost or prices).
30. For empirical evidence on the boom in agricultural prices, see Baffes and Haniotis (2010) and World Bank (2009, 2012).
31. Radetzki (2008) argues that the direct effect of a devaluation of the dollar in general would be to increase the demand from commodity consumers (by

lowering the price of commodities in local currency) and increase the dollar cost for commodity producers. This observation may help us to explain why the devaluation of the dollar relative to other advanced countries' currencies had little effect on commodity prices as their weight on world demand for commodities is not that large and also why the devaluation of the dollar relative to the currencies of the commodity-producing countries has had important and persistent effects on the dollar cost of production and prices.

32. For econometric evidence of this impact of exchange rates on food prices see World Bank (2012).
33. For data on these trend and cyclical relations see Yeyati and Williams (2012) and Akyüz (2012).
34. Akyüz (2010, 2012) argues that China's growth is still very heavily dependent on exports. For a different view, which puts more emphasis on the Chinese internal market than Akyüz does, see Medeiros (2006), Anderson (2007) and Kotz and Zhu (2010).
35. See Rozada and Yeyati (2011) in particular for evidence on the overwhelming importance of American long interest rates on public debt and the yield of American high-risk domestic assets on the spreads charged on 'emerging' markets.
36. As shown by Frenkel and Rapetti (2011), due to their improved balance of payments situation emerging market spreads have fallen for the first time ever below the spread of highrisk American assets around 2004 and have remained so ever since, including in the turbulent period after the global crisis of 2008.
37. In the recent boom strong evidence that the nominal revaluation of commodity currencies (currencies of a few key commodity exporters) actually preceded increases in dollar price of commodities can be found in Chen, Rogoff and Rossi (2010) (though interpreted not as cause of the boom but as a rational expectation of future commodity prices). In fact, as it was pointed out to me by Fabian Amico it is not really important if the dollar commodity price rises begin to happen before or after the appreciation of the currencies. What matters for explaining the trend of prices is that the appreciation of the commodity currencies persists and in particular is not more than cancelled out by the usual series of large subsequent devaluations.
38. Patnaik (2002) however interprets and criticises Prebisch's views on the terms of trade as if for him those were determined in the long run by demand. As will be seen in the next section, Prebisch's theory is interpreted differently here.
39. See ILO (2013), for data on trends of real wage growth in the 2000s that seem to be consistent with the above interpretation. Real wages grew fast in high-productivity growth developing industrial exporting countries, very fast in some commodity-exporting countries and rather slowly in the advanced industrial economies.

References

Adelman, M. (2004) 'The real oil problem', *Regulation*, 27(1) (spring): 16–21.

Adelman, M. A. and Watkins, G. C. (2008) 'Reserve prices and mineral resource theory', *Energy Journal*, 29, Special Issue: 1–16.

Akyüz, Y. (2010) 'Export Dependence and Sustainability of Growth in China and The East Asian Production Network', *South Centre Research Paper*, RP 27, April.

Akyüz, Y. (2012) 'The Staggering Rise of the South?', *South Centre Research Paper* 44, March.

Anderson, J. (2007) 'Is China Export-Led?' *UBS Investment Research Asian Focus*, UBS, 27 September.

Baffes, J. and Haniotis, T. (2010) 'Placing the 2006/08 Commodity Price Boom into Perspective', *World Bank Policy Research Working Paper* 5371.

Chen, Y., Rogoff, K. and Rossi, B. (2010) 'Can exchange rates forecast commodity prices?', *Quarterly Journal of Economics*, 125(3), (August): 1145–94.

Davidson, P. (2008) 'Crude Oil Prices: "Market Fundamentals" or Speculation?', *Challenge*, 57(4), (July): 110–18.

Erten, B. and Ocampo, J. (2012) 'Super-cycles of Commodity Prices since the Mid-nineteenth Century', *UN DESA Working Paper* 109, February.

Feldman, G. (2009) 'Global imbalances and the world economic crisis: a critical appraisal', http://www.boeckler.de/pdf/v_2009_10_30_feldman.pdf, accessed 25 January 2013.

Ferreira, S. (2012) 'Condições Externas e a Dinâmica da Inflação no Brasil 1994–2010: Uma Interpretação Estrutural', PhD thesis, Instituto de Economia, UFRJ.

Frankel, J. (2006) 'The Effect of Monetary Policy on Real Commodity Prices', in J. Campbell (ed.), *Asset Prices and Monetary Policy* (Chicago: University of Chicago Press).

Frankel, J. and Rose, A. (2010) 'Determinants of Agricultural and Mineral Commodity Prices', HKS faculty research working papers, RWP10-038, Harvard.

Fratini, S. M. (2008) 'Alcune osservazioni sulla teoria classica della rendita', *QA – Rivista dell'Associazione Rossi-Doria*, March.

Fratini, S. M. (2009) 'La rendita assoluta di Marx e le equazioni di prezzo di Sraffa', Department of Economics Working Paper 0105, Roma Tre University.

Freitas, F., Medeiros, C. and Serrano, F. (2013) 'The Decoupling of the Trend of Economic Growth of the Developing Countries in the 2000s', Paper for the First World Keynes Conference, Izmir University of Economics, Izmir/Turkey 26–29 June 2013.

Frenkel, R. and Rapetti, M. (2011) 'Fragilidad externa o desindustrialización: Cuál es la principal amenaza para América Latina en la próxima década?', *Serie macroeconomía del desarrollo,* 116, CEPAL, November.

Garegnani, P. (1983) 'The classical theory of wages and the role of demand schedules in the determination of relative prices', *American Economic Review*, 73(2), (May): 309–13.

Garegnani, P. (1988) 'Actual and normal magnitudes: a comment on Asimakopulos', *Political Economy*, 4(2): 251–58.

Ginzburg, A. and Simonazzi, A. (2004) 'Disinflation in industrialized countries, foreign debt cycles, and the costs of stability', *Quaderni di Ricerca Centro Sraffa*, 4, Uniroma 3.

Gopinath, D. (2011) 'Supply responses will be delayed for hard commodities', *Trusted Sources Consulting*, September.

ILO (2013) *Global Wage Report 2012/13: Wages and Equitable Growth* (ILO).

Irwin, S. and Sanders, D. (2011) 'Index funds, financialization, and commodity futures markets', *Applied Economic Perspectives and Policy*, 33: 1–31.

Jenkins, R. (2011) 'El "efecto China" en los precios de los productos básicos y en el valor de las exportaciones de América Latina', *Revista CEPAL*, (April), 103: 77–93.
Kaldor, N. (1939) 'Speculation and Economic Stability', *Review of Economic Studies*, October.
Kotz, D. and Zhu, A. (2010) 'The Dependence of China's Economic Growth on Exports and Investment', *Review of Radical Political Economics*, XX: n.10.
Krugman, P. (2008) 'More on oil and speculation', http://krugman.blogs.nytimes.com/2008/05/13/more-on-oil-and-speculation/
Kurz, H. D. and Salvadori, N. (2009) 'On the Theory of Exhaustible Resources: Ricardo vs. Hotelling', ISER Discussion Paper, Institute of Social and Economic Research, Osaka University, 0756.
Lewis, A. (1976) 'The Diffusion of Development', in T. Wilson (ed.), *The Market and the State* (Oxford: Oxford University Press).
Lewis, A. (1978) *Evolution of the International Economic Order* (Princeton: Princeton University Press).
Lewis, A. (1980) 'The slowing down of the engine of growth', *American Economic Review*, 70(4), (September): 555–64.
Lynch, M. C. (2006) 'Crop circles in the desert: the strange controversy over Saudi oil production', http://masterresource.org/wp-content/uploads/2008/12/crop-circles.pdf
Medeiros, C. (2006) 'A China como um duplo pólo na economia mundial e a recentralização Asiática', *Revista de Economia Politica*, v. 26.
Medeiros, C. (2011) 'The political economy of the rise and decline of developmental states', *Panoeconomicus*, 58(1): 43–56.
Pasinetti, L. L. (1999) 'Economic Theory and Technical Progress', Lecture at the Royal Economic Society Annual Conference, Nottingham.
Patnaik, P. (2002) 'Globalization of capital and terms of trade movements', http://www.networkideas.org/featart/mar2002/globalisation_1.pdf
Piccioni, M. and Ravagnani, F. (2002) 'Absolute rent and the "normal price" of exhaustible resources', Centro di ricerche e documentazione 'Piero Sraffa', Quaderno di Ricerca 2, Roma Tre University.
Pivetti, M. (2011) 'On Advanced Capitalism and the Determinants of the Change in Income Distribution', Paper presented at the *Congreso Internacional Crisis de la teoria economica y politicas alternativas ante la crisis global*, 7–11 November, (Mexico City: UNAM).
Pollin, R. and Heintz, J. (2011) 'How Wall Street speculation is driving up gasoline prices today', *PERI Research Brief*, June.
Prebisch, R. (1949) 'O desenvolvimento econômico da América Latina e seus principais problemas', *Revista Brasileira de Economia*, 3(3), (July): 47–111.
Radetzki, M. (2008) *A Handbook of Primary Commodities in the Global Economy* (Cambridge: Cambridge University Press).
Radetzki, M. (2010) 'Peak Oil and other threatening peaks – Chimeras without substance', *Energy Policy*, 38(11), (November): 6566–9.
Radetzki, M. (2012a) 'Politics – not OPEC interventions – explain oil's extraordinary price history', *Energy Policy*, 46, (April): 382–5.
Radetzki, M. (2012b) 'The perseverance of the ongoing metal and mineral boom', M*ineral Economics*, 46, (August): 382–5.
Ravagnani, F. (2008) 'Classical theory and exhaustible natural resources: notes on the current debate', *Review of Political Economy*, 20(1): 79–93.

Roncaglia, A. (2003) 'Energy and market power: an alternative approach to the economics of oil', *Journal of Post Keynesian Economics*, 25(4), (Summer): 641–59.

Rozada, M. and Yeyati, E. (2011) 'Risk Appetite and Emerging Market Spreads', *Business School Working Papers*, 2011-03, Universidad Torcuato Di Tella.

Rutledge, I. (2003) 'Profitability and supply price in the US domestic oil industry: implications for the political economy of oil in the twenty-first century', *Cambridge Journal of Economics*, 27(1): 1–23.

Schefold, B. (2001) 'Critique of the corn-*guano* model', *Metroeconomica*, 52(3), (August): 316–28.

Serrano, F. (2003) 'From "static" gold to the floating dollar', *Contributions to Political Economy*, 22, (November): 87–102.

Serrano, F. (2004) 'Power relations and American macroeconomic policy, from Bretton Woods to the floating dollar standard', *http://www.networkideas.org/feathm/sep2006/pdf/power_relations.pdf*

Serrano, F. (2008) 'A economia americana, o padrão dollar flexível e a expansão mundial nos anos 2000', in J. Fiori, C. Medeiros and F. Serrano (2008), *O Mito do Colapso do Poder Americano* (Rio de Janeiro e São Paulo: Editora Record).

Singer, H. (1950) 'The Distribution of Gains between Investing and Borrowing Countries', *American Economic Review.* Papers and Proceedings, 40, May.

Singer, H. (1998) 'Beyond terms of trade: convergence/divergence and creative/uncreative destruction', *Zagreb International Review of Economics and Business*, 1(1), (May): 13ff.

Sylos-Labini, P. (1982) 'Rigid prices, flexible prices and inflation', *Banca Nazionale del Lavoro Quarterly Review*, 140: 37–68.

Tilton, J. (2009) 'Is mineral depletion a threat to sustainable mining?', http://inside.mines.edu/UserFiles/File/economicsBusiness/Tilton/Sustainable_Mining_Paper.pdf

Tootell, G. (2011) 'Do Commodity Price Spikes Cause Long-Term Inflation?' *Public Policy Briefs*, Federal Reserve Bank of Boston, n. 11-1.

UNCTAD (2008) 'Trade and Development Report, 2008', UNCTAD.

World Bank (2009) *Global Economic Prospects 2009: Commodities at the Crossroads*, World Bank.

World Bank (2012) 'Global Commodity Market Outlook', *Global Economic Prospects January 2012*, World Bank.

Wray, L. R. (2008) 'The Commodities Market Bubble: Money Manager Capitalism and the Financialization of Commodities', *Public Policy Brief*, Levy Economics Institute, 96.

Yeyati, E. and Williams, T. (2012) 'Emerging economies in the 2000s: real decoupling and financial recoupling', *Policy Research Working Paper Series* 5961, World Bank.

9

The Political Economy of the Rise and Decline of Developmental States

*Carlos Aguiar de Medeiros**†

9.1 Introduction

The spread of industries in several peripheral countries after the Second World War and the great divergence that has opened up between them since the 1980s has sparked widespread debate on economic development. Interpretations based on neoclassical and on institutional economics[1] are the major fields of historical explanations. Despite their wide differences on the determinants of economic growth, they share a common perspective on three basic aspects: first, the supposition that strategies of development are built on a set of government policies and on institutions that model private behaviours (of course, they disagree on which policies and institutions promote economic development); second, a 'methodological nationalism'[2] in which individual countries' performances are essentially explained by domestic factors. The third is a corollary of the two above perceptions and says that the state as a major initiator of positive change (in resource allocation, as in the heterodox reasoning, or in the creation of market institutions, as in orthodox thought) is responsible for the success or failure of growth strategies. For the mainstream school, the wrong policies of populist states play the dominant role, for the heterodox, they are those of liberal or neo-liberal states. For both, a meritocratic state is central to successful development

*Another version of this paper was published in *Panoeconomicus*, 2011, 1, 43–56. The author thanks CNPQ for financial support, and Franklin Serrano and Andrea Ginzburg for the comments on the original paper presented at the Sraffa Conference, Roma 2–4 December 2010, and an anonymous referee. Any errors are mine.
†Associate Professor at IE, Federal University of Rio de Janeiro.

strategies (to avoid rent-seeking cases according to neoclassical authors, to discipline large firms according to the institutionalists).[3]

Stemming from a methodological perspective based on classical political economy, on Latin American structuralism and on a Gramscian view of the state's formation,[4] this chapter takes a critical stance on these three basic aspects. To begin with it considers that modern explanations about developmental states and the role of institutions neglect the different challenges and circumstances created by initial conditions and how different economic and social structures influence institutions (a bias opposed to the classical structuralism of the United Nations Economic Commission for Latin America and the Caribbean (ECLAC), which neglected the autonomous role of institutions). The major challenge in explaining development strategies is to articulate the two dimensions. Second, it proposes that the international environment is not identical for all developing nations and is influenced by the hegemonic state's economic and political actions, creating different development opportunities for their industrialisation. Finally, it considers that the state (and its developmental historical variant)[5] may not be viewed as an agent in isolation from social classes and relations with other states, but rather, as a central institution where the dominant class or some of its sectors lead a coalition of power, building a hegemonic project compatible with a particular accumulation strategy.[6]

In addition to this introduction, the chapter has three main sections. In the next two sections there is an attempt to illustrate some of these issues using the analysis of national patterns of industrialisation and development strategies experienced by peripheral countries during both the highest convergence period and the greatest divergence phase. The third and final section explores how new challenges have been answered by developing nations and some attempts to rebuild a 'neo-developmental state'.

9.2 National development strategies in the golden age of the post-war and developmental states (DS)

The partial and limited spread of industrialisation in the post-war period (especially among industries and activities closer to the innovative processes, such as the capital goods sector) was a consequence of national strategies led by development-oriented states specifically geared to creating modern industry and its infrastructure as the main engines of economic growth.

Under US hegemony national development was basically an accumulation strategy and a hegemonic project of industrial national capital coordinated by national states favouring the formation of large industrial companies and their markets. These developmental strategies were followed in several countries and took different routes according to the size of the economy, the natural resource base, income distribution, geopolitical position and the political power underlying that strategy.

With distinct levels of success and with a much lower income per person, a few countries (in Latin America, especially Brazil and Mexico, in Asia, the Asian Tigers, especially Korea and Taiwan, the Southeast Asian countries, India and China) followed a path of industrialisation somewhat similar to what European countries had experienced in restructuring during the post-war period. The strategy was to transplant the key industries typical of the American manufacturing pattern – metal mechanics, automobiles and chemicals – and their consumption patterns, centred on durable consumption goods financed by credit. As Prebisch (2011 [1949]) remarked, the typical Keynesian post-war policies on the periphery would require structural change to offset the external constraints, and industrialisation would be the basis for a policy directed towards high growth and unemployment reduction.

Between 1950 and 1980, the steady increase in per capita income in those countries resulted from the increased pace of industrial output growth and the transfer of surplus labour in agriculture to urban activities led by industry and services. In countries where this shift was greater, as in Korea or Brazil, the growth rate was higher, and in countries where it was less intense, as in India, the rate of growth of income per capita and per employed person was lower.

Despite the diversity of initial conditions within those countries, they faced challenges caused by the large technological gap in relation to industrialised countries, the narrowness of domestic markets, the problems of coordination and financing of complementary blocks of investments in new sectors and the restrictions on balance of payments.[7]

The influence of industrial success in the Soviet Union gave major political legitimacy to long-term planning. But national industrialization strategies were not only distinguished by planning. In those most successful countries such as Brazil and Korea (and later in China, since Deng Xiaoping's reforms), strategies were the result of industrial policy and public investment. In some countries, like Korea, Taiwan or Mexico, the state has directly controlled the 'commanding heights' of the financial sector (Haggard *et al.*, 1993). Besides finance, in many countries industrial incentives were directly applied through state enterprises operating in strategic industrial activities and infrastructure.

Thus, regardless of the higher or lower share of exports in the composition of industries' final demand, the late industrialisation of the last century was led by the states. The conventional distinction between a strategy based on import substitution industrialisation (ISI) associated with the state leadership in countries like Brazil and Mexico, and an export-oriented industrialisation (EOI) associated with a pro-market strategy that would have prevailed in Korea, Taiwan or Thailand, does not, in fact, coincide with the historical evidence on industrialisation pursued in those countries.[8] All strategies originally included import substitution processes and selective opening and put greater or lesser emphasis on industrial exports according to different factors.

An essential part of national development strategies was the macroeconomic regime in which the exchange rate and the fiscal and monetary policy were subordinated to the objectives of industrial development.[9] Until the 1970s, external financing was scarce and the constraint on foreign currency imposed strict control on foreign exchange.

However, in spite of common strategies, these countries followed different patterns. Two aspects stand out: the levels of income distribution associated with the industrialisation process (inequality in Latin America was much higher than in Asia), and the share of industrial exports in countries' total exports (much higher in Asia). Besides these structural dimensions, one important difference was the role played by foreign capital, which was larger in Argentina, Brazil and Mexico than in Korea, Taiwan or India. As regards the state's power to influence the economy, one may observe that although the common base was a coalition between the military, technocratic planners, and the private industrial sector, this coalition was stronger in Korea or Taiwan than in other countries. Both countries built a 'cohesive capitalist' state (Kohli, 2004) by dislodging the landowners and enjoyed the strong support of the USA in their strategy of communist contention in Asia. It was also influenced by Japanese institutions. Except for Korea and Taiwan, the developmental state was more 'fragmented and multi-class' and the industrial sector had to establish political alliances with powerful landowners.

Different patterns

Unlike Western Europe, industrialisation in peripheral countries led by developmental states was not accompanied by social democratic coalitions aiming at the distribution of income and full employment.[10] The goal of greater equity was subordinated to growth and industrialisation goals. Income distribution was quite uneven according to the different social coalitions supported in the state. The pattern of income distribution was essentially influenced by the way land and the modernisation of agricultural production had evolved. In countries where the productivity in food production[11] was lower and internal structural heterogeneity[12] higher, peasant income was lower and the heavier was the weight of the traditional oligarchy on political power – such as in Brazil, India, the Philippines or Indonesia where the state was more 'fragmented and multi-class'. In these cases, industrialisation took place accompanied by social marginalisation and the exclusion of rural masses and suburban areas from modern consumption, leading to large income concentration. In countries where land reforms and simultaneous modernization of agriculture took place (as in Korea and Taiwan) internal structural heterogeneity and social polarization was less intense, and the state was more cohesive.

Similarly, export performance followed a distinct path.

In East Asia, import substitution was quickly followed (as early as the 1960s) by industrial exports (textiles and clothing in the early stages), generating a greater trade diversification and a consequent positive effect on the balance of payments. In Latin America, the diversification of exports happened much more slowly and less intensively. Several hypotheses explain this discrepancy. Excessive protectionism[13] or pessimism about Latin America's ability to export industrial products[14] would have prevailed in the region, in contrast to the clearly export-oriented strategies of Asian countries.

What distinguishes these approaches is the lack of connection between strategies and economic structure. Following Diamand (1978) in the case of Argentina, Mahon (1992) and more recently Bresser Pereira (2010), here we argue that among countries with highly competitive export industries based on natural resources, such as those of Latin America, there came to be an external heterogeneity or an 'imbalance in the production structure' between the productivity of the primary export sector and that of the industrial sector. This imbalance led to the formation of an uncompetitive exchange rate for industry, which contributed to the specialisation of the export sector. Industrial policies favoured domestic industry through preferential exchange rates and tariffs, but these policies achieved poor results in promoting industrial exports.[15]

In Latin America, especially in Argentina, the high share of exports-based natural resources exacerbated a distributive exchange-rate conflict between the primary exporters, the industrial sector, non-tradable support activities and the working class.

In Asia, for a small group of countries like Korea, Taiwan and certainly city-states like Hong Kong and Singapore, the scarcity of natural resources made the export of industrial products the obligatory route to industrialisation. Poor natural resource bases were favourable to a more balanced economic structure, making it possible to establish a real exchange rate more favourable to industry. In these 'invited countries' (Medeiros, 1997) industrial exports were directed to the USA, which gave them preferential access. In addition the export companies of these countries benefited from the road previously travelled by Japan which provided these latecomers with some trade networks and investment (Medeiros, 1997; Ginzburg and Simonazzi, 2003).

Thus, due to dissimilar structural and geopolitical circumstances, these strategies have evolved[16] to yield different macroeconomic and social results.

This fact had important consequence for the ease with which Asian countries reacted to the debt crisis of 1980, but it had already manifested itself in the pattern of indebted growth followed by Brazil, Mexico and Korea in the 1970s.[17]

9.3 Neo-liberalism and the crisis of the developmental nation states

The Reagan-Thatcher offensive against organised labour, the external debt crisis in the peripheral countries and the collapse of the USSR in 1991, at a time when a new technological revolution based on information and telecommunications was rising, led to significant changes in the international division of labour. Under unrestrained competition, finance and productive internationalisation greatly enlarged.

The 'Washington–Wall Street complex'[18] (and its leadership through the World Bank and the International Monetary Fund) established itself as the centre of political power and of the ideology. This brought about a new strategy of accumulation and a new hegemonic project increasing the dominance of capital, in general, and finance capital, in particular, over other factions and interests.[19] It also coincided with a new US trade offensive to open the hitherto regulated internal markets of the newly industrialised exporting nations.[20]

Among the industrialised nations, the large corporations, exposed to intense international competition, sought greater autonomy from the state, the workers and the chain of domestic suppliers demanding greater state support for the globalisation processes of production and finance in new spatial and regional arrangements. Transplanting labour-intensive activities to peripheral countries was intense, remaking the international division of labour. The firm's strategy of going global introduced a fracture between national capitalism and national capital with important repercussions for macroeconomic policies and political coalitions.

The crisis of the developmental state was similar to the end of the Keynesian national welfare state in industrialised countries. If the latter was associated with the end of the subordination of monetary and fiscal policy to full employment, the former crisis was connected with the end of the subordination of fiscal and monetary policy to industrial development.

In fact, as has happened with national Keynesianism, development strategies based on industry and on internal markets were abandoned in many countries and a new hegemonic project led by cosmopolitan capital was established. This brought about in many 'fragmented' states what can be considered (in a Gramscian expression) a capitalist passive revolution aimed at rebuilding a more cohesive state around big business and financial interests.

Despite the differences observed in time and space, the discontinuity in development strategy involved two major forces: financial openness and a big business revulsion against the developmental state.

Financial openness played an important role in the crisis and discontinuity of national developmental strategies (in both industrialised and, mainly, semi-industrialised nations) insofar as it exposed the economies to volatile capital inflows and dissolved the role of domestic credit as a mechanism for coordinating investments (Haggard and Maxfield, 1993). It was in the wake of the exchange-rate crises that Washington Consensus structural reforms were introduced on a huge scale. The institutional position of the central bank (with an exclusive focus on price stability) was strengthened due to international constraints that followed the external crisis, and the role of manufacturing interests and their institutions was diminished.

As a corollary to this macroeconomic and institutional change, there was a split – to the extent to which the nations opened their economies – between the interests of the large corporations and the national industrial strategies that had hitherto been the basis of national development. Cultivated and promoted by their developmental state, domestic businesses (challenged in their own markets) began to seek new opportunities and strategies for accumulation, especially through the formation of joint ventures with multinational corporations and through majority interests or participation in the business of privatisations. Such opportunities of going global demanded new functions and policies from the state (Medeiros, 2009).

The rebellion by large corporations against developmental states occurred everywhere.[21] It was generally accompanied by public opinion that identified developmentalism and industrial policies – such as those implemented by countries like Brazil, Korea or Indonesia – with political authoritarianism, with 'crony capitalism' and, in the case of Brazil, with income concentration. The political legitimacy of industry-based accumulation strategies and, consequently, the hegemony of this project were profoundly shaken.[22]

But a great divergence took place.

In Latin America the intense debt crisis of the 1980s brought about high inflation rates and a deep recession. This caused a structural crisis in prevailing state-led growth and created new coalitions of internal and external interests around the agenda of the Washington Consensus reforms that spread all over the region in the 1990s (Medeiros, 2008a).

In Asia, the external shock of the 1980s was not so disruptive. In large countries such as China or India, the debt ratio was too low to make any substantial negative impact (Hughes and Singh, 1991). In East and South-East Asian countries, thanks to better solvency ratios, the surge of Japanese investments and the clustering of production chains in the region (Medeiros, 1997), the majority of the economies (more or less open) experienced high growth, thereby preserving the bulk of institutions which had developed earlier. This clustering of success and collapses in space and time highlights the limitations of 'methodological nationalism' (Ocampo *et al.*, 2007; Medeiros, 1997). Only in the 1990s, but mainly after the 1997 crisis, did a strong offensive against developmental institutions occur.

Thus, there were different national answers to the liberalisation process. In Latin American countries, liberalisation took the form of a radical U-turn (Palma, 2010), from a very weak national position; in Korea and other Asian countries, liberalisation was adopted later and tied to a trajectory of high growth. Other Asian countries like China, Taiwan or India did not dismantle the main developmental institutions.

The degree and impact of these changes on national developmental strategies essentially depended on the extension and circumstances of the external crisis but the different ways in which they occurred in Asian and Latin American countries were conditioned by the production structure, the regional dynamics and the power and political cohesion of the nation states.

Different paths

Throughout the 1990s, it was possible to identify various reactions to the liberalisation and technological pressures. One common response to the new challenges was the pursuit of an 'integrationist' strategy (Amsden, 2001), or as Lall (2000) puts it, 'a passive strategy dependent on foreign direct investment (FDI)'. This was based on two pillars: on the micro side, this strategy was built by the formation of new private alliances and re-specialisation in activities with absolute cost advantages (whether in industrial commodity chains, as in Mexico, or in natural resources, as in most South American countries and in Russia throughout the 1990s).[23] On the macro side, the strategy centred on exports and on external financing and investment as the main growth machine.

In Mexico, the liberalisation process, initiated after the 1982 default in external debt and bank nationalisation, accelerated at the beginning of the nineties, moving towards the North American Free Trade Agreement (NAFTA) established in 1994. Led by small group of technopols,[24] a victorious coalition formed by large Mexican groups, mainly in the non-tradable sector, and American multinational companies, inaugurated a growth strategy based on exports of labour-intensive industrial activities in a 'shallow' trade specialisation. This export model increased Mexican dependency on US markets and investments and promoted the rise and internationalisation of domestic conglomerates. The latter endeavour fractured the early connections between Mexican capitalism and Mexican big business, and the state played the role of protagonist in privatisation deals and in providing massive finance support.

In South America, the rebellion of cosmopolitan big business against the developmental state started during the late 1970s and the 1980s as a consequence of external crisis and hyperinflation in many countries, occurring alongside the expansion of the power and influence of the holders of dollarised assets, such as the traditional exporters, banks and non-tradable activities in association with foreign capital. Argentina's external debt resulted during the 1980s in huge wealth and debt transfers from state to big business. Starting in 1989, the Structural Adjustment Programme and massive privatisation supported by Washington institutions and the Argentinean elite generated a premature deindustrialisation and denationalisation, but simultaneously, a large centralisation of capital took place headed by commodity exporters and finance.

In Brazil, it was the external crisis of the 1980s, high inflation and the eruption of an autonomous labour struggle that undermined the development coalition. With more diversified industry that partially resisted the process of trade and financial opening, some important public enterprises (including a big development bank) were preserved from the massive privatisation and denationalisation of the mid-1990s.[25] The *desenvolvimentistas* – the technocratic, intellectual and industrial leaders who had led the old economic strategy – were not completely dislodged from the state as had happened in Mexico or Argentina. But the winners from these liberal transformations were the foreign investors and big business in the finance sector and in the production of commodities.[26]

Led by the bureaucracies close to Washington's institutions, the reforms removed industry and its bureaucracies (planning and labour ministries, intermediary government agencies, etc.) from the 'commanding heights' of the economy.

An essential feature of this strategy was a macroeconomic regime based on monetary stability and financial openness. The power of the financial sector in these countries increased not only because its assets grew faster, but because of the predominance in its economic policy of its main interests – higher interest rates and low inflation. Due to high levels of external debt and the growing influence of the International Monetary Fund (IMF) on domestic policies, this financial domination found its expression through orthodox central banks that assumed in these countries the 'commanding heights' of the economy.

Some of these changes and the demise of the developmental state also occurred in Korea at the beginning of the 1990s[27] and in many Asian countries following the 1997–98 external crisis. Throughout the 1990s, several Asian nations followed a mixed strategy based on industrial incentives and on foreign direct investment and exports integrated in commodity chains in a 'flying geese' model. Korea, under pressure from America, opened up its financial system, eliminating the influence hitherto exerted by government on credit and investment. Big *chaebols* decided that government intervention was a hindrance to new economic opportunities.

The national development strategy was not completely changed in China and India (both with military power and autonomous geopolitical presence) nor in Taiwan and Singapore, who followed a path of greater autonomy (or of greater resistance), preserving the national developmental strategy and its hegemonic project in a new context. Although it relinquished some previous economic regulatory mechanisms, the developmental state in dynamic East Asian countries survived.

9.4 National development strategies at the beginning of the new millennium

At the beginning of the new millennium great changes occurred in the world economy. Higher international growth, a substantial rise in commodity prices, lower rates of interest and a continuous expansion of industrial commodities chains (mainly) located in Asia were the principal factors. The rise of China as a great trading power was at the centre of these changes. These circumstances brought about better and more widespread economic opportunities for many peripheral countries. Even for less competitive Latin American countries, the rise in commodity prices allowed the rare simultaneous occurrence of economic growth with positive current account balance and a sharp contraction of external debt. The 2008 financial crisis brought about a great recession in industrialised countries (an effect that still continues) but did not change some of these new and structural circumstances for less industrialised countries. In this context, many countries introduced Keynesian expansionist measures against the hitherto predominant orthodox opinion. Politically, the once strong IMF, World Bank and World Trade Organization (WTO) lost influence in the face of the waves of crisis hitting the countries that had followed their main prescription. Nationalism gained more legitimacy.

Faced with these circumstances, developing countries adopted three different strategies, two of which were not very different from the route taken in the 1990s. The first, a 'passive and integrationist' strategy, like the one followed by Mexico and some East European countries in the 1990s, was merely a continuation of the neo-liberal strategy and gained more support despite its weak results. The second, a 'neo-developmental' strategy adopted by China, Taiwan, Singapore, or even India, explored new opportunities to upgrade their industrial structures including new policies without radically changing their previous mechanisms of industrial and financial regulation. Finally the third, a neo-Keynesian, took place in many primary exporter countries.

As has been argued here, specific institutions and mechanisms for coordinating production were created to solve the problems of industrialisation according to the particular production sectors and technological stages. Although the institutions required for coordinating industrial sectors in agrarian economies (the post-war challenge) are different from those required for industrial upgrading (the present challenge), they continue to be necessary, as the Asian experience indicates. As Lall (2000) observed, commercial policies, credit and subsidy policies, infrastructural development, skill development, technological incentive, and the attraction and delimitation of foreign direct investment (FDI) continue to be the instruments of industrial policy-making. In fact, industrial policies (not only horizontal but also vertical ones) are necessary both for the creation of the incentives of the innovation process in activities involving rapid transformations in the international economy, and for the construction of a new infrastructure.

Thus, the emergence of new challenges to the strategies of industrialisation, resulting from new information and telecommunications technologies (ITT) and from the formation of global and regional production chains, placed new demands on national industrial policies. The construction of a new transportation and communications infrastructure, the dissemination of new technologies and the pursuit of specialisation in specific production segments became part of the ordinary agenda of national projects of industrial upgrading. This 'neo-developmental' strategy is less centred on the internal market as the prevailing dimension of accumulation.[28] The processes of productive regionalisation and internationalisation of 'national champions' strongly expanded firms' investment horizons. The strategies of buying established technology and of adaptation based on process innovations, which typically distinguished Japan and Korea, were challenged by modularised production and new strategies based on greater proximity to proprietary activities and activities related to product innovation. In China, as well as Korea and Taiwan, a 'second phase of catching up' (Chang, 2006), based on innovation and the construction of proprietary national technologies, would be the basic challenge of industrial upgrading.

Based on these new challenges, a 'neo-development strategy' aimed at continuing the 'catch-up' strategy was developed by countries that knew how (or were able) to resist external and internal pressures.

In Asia, this new strategy was mainly followed by China which strongly combined public investment in infrastructure – the main driver of overall growth productivity – with a selective industrial policy in ITT in expansionist macroeconomics, including low interest rates, an anti-cyclical fiscal policy and the maintenance of a competitive rate of exchange. The subordination of finance and enterprise to the development goals was achieved through the maintenance of political centralisation and by the leadership of state enterprises in the 'commanding heights' of the economy. In India, with a much more fragmented society, some of the old regulation instruments were also preserved, favouring a less ambitious, but nevertheless active industrial policy.

As we have been observing throughout this chapter, a strategy is not only the outcome of a decision by a state; its coherence is socially and structurally conditioned. In Korea, after the liberal reforms implemented in the midst of an ample IMF financing adjustment, the previous industrial policies and institutions were dismantled. But the extraordinary expansion of exports (partially induced by Chinese expansion) that followed the 1997 crisis permitted a substantial reduction in sovereign debt and less interference from the IMF. In this new situation, some big *chaebols* rebuilt, with the Korean state, a new coalition for industrial modernisation based on innovation. Thus, due to a developed and homogeneous structure and pragmatic economic policy, institutional change in Korea did not interrupt its high road to investment and productivity. In a more interventionist Malaysia, something similar happened. A neo-developmental strategy is being followed without or 'beyond the developmental state' (Fine, 2005).

A neo-developmental strategy, as we have described it, is today less focused on the productive sector than was the case in the past and more centred on innovation processes in new technologies through several policies and instruments.[29]

As observed in South America between 2002 and 2008, a spectacular rise in the price of commodities permitted those countries to obtain higher growth rates, a sharp contraction in external debt and an accumulation of reserves. After the evident failure of neo-liberal strategies based on Washington Consensus reforms, a more pragmatic macroeconomic policy became established. At the same time, various nationalist movements spread from Patagonia to the Andes countries (Venezuela and Bolivia, also highly critical of the market and liberal institutions supported by the USA) as a backlash to the radical liberal experiments of the 1990s. These movements created regional policies and agreements such as ALBA (Bolivarian Alliance of America) and UNASUL (South American Nations Union) with alternative goals to the free trade initiatives led by the USA.

In Russia, a similar situation occurred, turning around the tragic decade of the 1990s and enabling the new government to construct significant sovereign reserves.[30] Countries as diverse as Argentina, Brazil and Russia could achieve higher growth prompted by internal markets that were now released from the external constraints that had blocked them throughout the 1990s. Higher minimum wages, higher social transfers and employment were achieved. And this occurred, essentially, without changing the pattern of economic growth.

Thus, various countries began constructing new development strategies situated somewhere between the neo-developmental strategy based on 'a second catching-up phase', and a passive and integrationist strategy. The new strategy, a 'neo-Keynesian' one, tries to distance itself, on the one hand from the previous strategy of national development, and on the other from the pro-finance and liberal macroeconomic policy advocated by the Bretton Woods institutions.

Without the particular conditions that support a 'high road' which we observed in some Asian countries, the state, in this third way, has less power to induce structural change. The economic and social cohesiveness required for this is missing.

As we have seen, for different structural and political reasons in major Latin American countries and Eastern Europe, the major private economic groups that in the past were the main beneficiaries of industrial policy are, nowadays, much more closely associated with international commodity chains in asymmetrical regional agreements and in non-tradable activities (in the case of Mexico or Eastern European countries that adopted the integrationist strategy); or became fragmented and failed to survive the radical process of liberalisation (as was the case in Argentina); or relocated to sectors based on natural resources and related support activities in services and construction (Brazil and Russia). Of course, in these countries, there are large segments of national manufacturing industry not connected to global chains that have resisted and survived. Nowadays, they are exposed to strong competition from China and need a more active industrial policy but these interests are diffuse and have less power to exert leadership in economic policy or to build political support for a comprehensive industrial policy.[31] On the other hand, opportunities to expand investment in natural resources have greatly enlarged.

But despite those new economic coalitions that explain why only a few countries are building solid developmental strategies based on technological catch-up, the re-birth of Keynesian (or, at least, more pragmatic) economic policy in many peripheral countries since 2008 has increased the possibilities for higher rates of growth and new perspectives on structural change. This has opened up more space for social and economic varieties of accumulation and growth strategies for the days ahead.

Notes

1. For a recent classification of the institutionalist approach, see Fine (2005).
2. For an original reference to this expression, see Gore (1996); see also Medeiros (1997).
3. Skocpol (1985) is an essential reference for this Weberian approach.
4. This tradition is very influential in some contemporary analyses of globalisation, such as Jessop (2002) and Morton (2007), and broad perspectives on capitalist state and institutions like the social structure of accumulation theory (SSA), McDonough, Reich and Kotz (2010).
5. The internationally consecrated formulation about the Developmental State is Johnson's (1982). Along these lines Amsden (2001) considered that the developmental state was predicated to perform four functions: development banking; local-content management; selective seclusion; and national firm

formation. In accordance with this description a state dedicated to promoting industry through indirect and discretionary interventions on prices and investment is considered to be developmental. However, the success of this strategy depends less on design and more on implementation and this depends on internal interests, conditioned by economic structures and the action of the hegemonic state.

6. Jessop (2002) in his definition of economic domination considers two dimensions. 'The first is internal to the economy and concerns the power of one or another fraction of capital …to impose its immediate interests on other fractions… . The second dimension of economic domination … refers to the capacity of capital in general, a given fraction of capital, or particular capitals to steer the evolution of other institutional orders in line with the demands of capital accumulation' (*op. cit.*, p. 29). In his critical observations on the present chapter Andrea Ginzburg considered that Jessop solved *ex ante* the problem of 'embeddedness'. I consider that this critique is somehow exaggerated when applied to this second dimension. Jessop follows here a Gramscian perspective that concedes to the state a relative autonomy.
7. The analysis of questions regarding the implementation of modern industries in agrarian societies was the main focus of study by the 'pioneers of development economics' such as Rosenstein-Rodan, Nurkse, Lewis, Hirschman, Furtado and Prebisch, inducing a new meaning to development economics.
8. State intervention is a phenomenon that has been common across the development experience, in the successful cases as well as the failures.
9. In his critical analysis of this chapter, Andrea Ginzburg noted that no effective demand analysis was provided here to explain the developmental strategies taken in this period. To clarify the point, the perspective adopted here follows the broad idea that during this period, given the expansionist macroeconomic policies adopted by the most successful developmental states, the main constraint on growth was exerted by the balance of payments. For a discussion see Medeiros and Serrano (2001).
10. During this same period and among industrialised countries a 'Keynesian National Welfare State' (Jessop, 2002) was built.
11. The distinction developed by Lewis (1978) and Furtado (1969) between tropical agriculture and that of a temperate climate is essential for the understanding of distinct starting points of the 'primary exporter' model and for the different levels of heterogeneity of social and economic structures.
12. See Anibal Pinto (1973). Here by internal structural heterogeneity is meant the productivity gap between food production and industrial goods and by external productivity or unbalanced economic structure the productivity gap between primary export sector and industry.
13. Brazil is an exception to this pattern. See Fishlow (1991).
14. For discussion see Pazos (1984).
15. In his comments on this chapter Andrea Ginzburg considered that Mahon and Diamand and the author accepted a strong price elasticity optimism and adopted in this analysis a supply-side static approach, as in 'Dutch Disease' hypothesis. This is a misunderstanding. The point under consideration is that although many industrial policies can induce export diversification, the existence of a competitive exchange rate for industry is a necessary (but

surely not sufficient) condition for this achievement. The fact that Korea or Taiwan had a competitive exchange rate during this period was the outcome not only of an industrial strategy (that was based in several policies) but of the fact that, unlike Latin American countries, they had no other way of obtaining currency. Structure matters as well as institutions.

16. Due to a higher balance-of-payments constraint the stronger incentives for substitution of imports in LAC induced a higher horizontal diversification than in Asia where it could be balanced by a higher priority in productivity. One may consider that the creation of *maquiladora* (in-bond industry) in Mexico occurred in 1965 as in Taiwan.
17. Although the ratio of debt to GDP was not very different in Korea and Brazil, the ratio of debt service to export that better reflects the country's external solvency was much lower in Korea. For an analysis of the 1970s external indebtedness in LAC, see Medeiros (2008a).
18. Wade and Veneroso (1998).
19. For discussions on these formulations see Serrano (2004) and Glyn (2006).
20. The liberating pressure was particularly strengthened in the environment of the WTO in the Uruguay Round that started in the 1980s.
21. 'In Korea, the giant conglomerates (the *chaebols*) have aggressively campaigned during the 1990s to convince the population that the government should abandon its industrial policy and financial regulation' (Chang, 2006, p. 253).
22. Fine (2005) examines different interpretations of the crisis of the developmental state, from what he calls the 'political school' to the 'economic school'. For the former, the developmental state was a singular historical construction that lost its functionality with the success of development. For the latter authors, the State crisis derived from financial opening and the lack of adaptation of institutions to the new environment; for others, such as Chang (2006), the financial opening destroyed the basic mechanism of investment coordination.
23. 'Within the context of the above-mentioned structural heterogeneity, LA has developed two types of successful "modern-sector" regional oligopolies: those involved in large scale capital intensive commodity production for exports, and those that have mastered the technique of organizing low-value-added labour intensive production chains – sometimes for exports (most agricultural products) and sometimes in services (eg. retail)' (Palma, 2010, p. 33).
24. This expression captures the political, ideological and economic work of a small group of academics who acted as organic intellectuals, in a Gramscian meaning, of the new socioeconomic bloc. Unlike the old technocrats, they had to be good politicians.
25. This resistance was only partial. Through the 1990s there was a disarticulation of some productive chains and an abandonment of some innovation-intensive activities, creating a kind of 'regressive specialization' (Coutinho, 1997), although some modern capital goods manufactures more integrated with mineral and raw material resources were preserved.
26. 'During the 1990s, Brazil's industrial elite withdrew without much conflict from areas attractive to international investors: industrialized food production and distribution, supermarket chains and automobile spare parts in the

beginning of the decade and, later, telecommunications, advanced services and financial institutions. The industrial elite has either migrated to the tertiary sector or retreated from business altogether, investing their capital in the financial market, pension funds or real estate. Other possibilities were "support activities" such as the building sector, packaging industry, car sales concessionaires, activities related to business and law consulting, business promotion, educational and cultural events, administrative and honorific positions in the "third sector" (NGOs) and the administration of real estate' (Rocco, 2007, p. 208).

27. The dissolution of the powerful Economic Planning Board into the Ministry of Finance and Economy was a milestone for the new state.
28. This is not to say that this strategy is led by exports: in case of China and India internal markets are still the main demand source for capital accumulation, but in both countries exports are much larger than they were in the past.
29. And this, as we have argued throughout this chapter, is conditioned by economic structure, the geopolitical position of the country and political coalitions between states and social classes.
30. With the end of the socialism and the radical liberalism of the 1990s, Russia resumed a national developmentalism strategy based, however, on natural resources. Thanks to its geopolitical position and greater state control over oil and natural gas exports, transfers to other sectors of the economy increased substantially; however, they did not result in greater export diversification.
31. The industrialisation of natural resources as an alternative industrial strategy has been under discussion in many small Latin American countries.

References

Amsden, A. (2001) *The rise of "the rest": challenges to the west from late-industrializing economies* (Oxford: Oxford University Press).

Bresser Pereira, L. C. (2010) *Globalização e Competição*. (Sao Paulo: Edit Campus).

Chang, H. J. (2006) *The East Asian Development Experience, The Miracle, the Crisis and the Future* (New York: Zed Books).

Coutinho, L. (1997) 'A Especialização Regressiva: um balanço do desempenho industrial pós-estabilização', in J. P. R. Veloso (ed.) *Brasil: Desafio em um país em Transformação* (Rio de Janeiro: Jose Olympio Edit), 81–106.

Diamand, M. (1978) 'Towards a Change in the Economic Paradigm Through the Experience of Developing Countries', *Journal of Development Economics*, 5, 19–53.

Fine, B. (2005) 'Beyond the Developmental State. Towards a political economy of Development', in C. Lapavitsas and M. Noguchi (eds) *Beyond Market Driven Development* (New York: Routledge).

Fishlow, A. (1991) 'Some Reflections on Comparative Latin American Economic Performance and Policy', in T. Banuri (ed.) *Economic Liberalization: No Panacea The Experiences of Latin America and Asia* (New York: Oxford University Press), 149–71.

Furtado, C. (1969) *Teoria e Política do Desenvolvimento Econômico* (São Paulo: Companhia Editora Nacional).

Ginzburg, A. and Simonazzi, A. (2003) 'Patterns of Industrialization and the Flying Geese Model: The Case of Electronics in East Asia', paper presented at the Conference on Clusters, Industrial Districts and Firms: the Challenge of Globalization, Modena, Italy 12–13 September.

Glyn, A. (2006) *Capitalism Unleashed* (London and Oxford: Oxford University Press).

Gore, C. (1996) 'Methodological Nationalism and the Misunderstanding of east Asian Industrialization', *European Journal of Development Research*, vol. 8, no. 1.

Haggard, S., Lee, C. H. and Maxfield, S. (1993) *The Politics of Finance in Developing Countries* (Ithaca and London: Cornell University Press).

Hughes, A. and Singh, A. (1991) 'The World Economic Slowdown and the Asian and Latin American Economies: A Comparative Analysis of Economic Structure, Policy and Performance', in T. Banuri (ed.) *Economic Liberalization: No Panacea The Experiences of Latin America and Asia* (New York: Oxford University Press), 57–100.

Jessop, B. (2002) *The Future of The Capitalist State* (Cambridge: Polity).

Johnson, C. (1982) *MITI and the Japanese Miracle* (New York: Stanford University Press).

Jilberto, A. and Hogenboom, B. (2007) 'Latin America Conglomerates in a Neo-Liberal Era', in A. E. F. Jilberto and B. Hogenboom (eds) *Big Business and Economic Development* (London: Routledge), 156–87.

Kohli, A. (2004) *State Directed Development, Political Power and Industrialization in the Global Periphery* (New York: Cambridge University Press).

Lall, S. (2000) 'The Technological Structure and Performance of Developing Country Manufactured Exports, 1985–98'. *Oxford Development Studies*, vol. 28, no. 3, 337–69.

Lewis, A. (1978) *Evolution of International Economic Order* (Princeton: Princeton University Press).

Mahon, J. (1992) 'Was Latin America too Rich to Prosper? Structural and Political Obstacles to Export Led Industrial Growth', *Journal of Development Studies,* vol. 28, no. 2, 241–63.

McDonough, T., Reich, M. and Kotz, D. (2010) *Contemporary Capitalism and Its Crises, Social Structure of Accumulation Theory for the 21st Century* (New York: Cambridge University Press).

Medeiros, C. A. (1997) 'Globalização e a inserção internacional diferenciada da Ásia e da América Latina', in M. C. Tavares and J. L. Fiori (eds) *Poder e Dinheiro* (Rio de Janeiro: Vozes), 279–347.

Medeiros, C. A. (2008a) 'Financial Dependency and Growth Cycles in Latin American Countries', *JPKE*, vol. 31, no. 1, 79–101.

Medeiros, C. A. (2008b) 'Desenvolvimento Econômico e Ascensão Nacional: rupturas e transições na China e na Rússia', in J. L. Fiori, C. Medeiros, and F. Serrano, *O Mito do Colapso do Poder Americano* (Rio de Janeiro: Vozes), 173–273.

Medeiros, C. A. (2009) 'Latin American Auction', *New Left Review*, no. 55, 109–32.

Medeiros, C. A. and Serrano, F. (2001) 'Inserção Externa, exportações e crescimento no Brasil', in J. L. Fiori and C. A. Medeiros (eds) *Polarização Mundial e Crescimento* (Petrópolis: Vozes), 105–35.

Morton, A. D. (2007) *Unravelling Gramsci, Hegemony and Passive Revolution in the Global Economy* (London: Pluto Press).

Ocampo, J. A., Jomo, K. S. and Vos, R. (2007) 'Explaining Growth Divergence', in J. A. Ocampo, K. S. Jomo and R. Vos (eds) *Growth Divergences* (New York: Zed Books), 1–25.

Palma, G. (2010) 'Why has productivity growth stagnated in most Latin American countries since the neo-liberal reforms?' *Cambridge Working papers in Economics*, 1030 (Available at http://www.econ.cam.ca.uk/dae/repec/cam/pdf/cwpe1030.pdf).

Pazos, F. (1984) 'Have Import Substitution Policies Either Precipitated or Aggravated the Debt Crisis?' *Journal of Interamerican Studies and World Affairs*, 27 (Winter), 57–73.

Pinto, A. (1973) *Distribuição de renda na América latina e desenvolvimento (*Rio de Janeiro: Zahar).

Prebisch, R. (2011 [1949]) 'O Desenvolvimento Econômico da América Latina e alguns de seus principais problemas', in Adolfo Gurrieri (ed.) *O Manifesto Latino-Americano e outros Ensaios* (Rio de Janeiro: Contraponto) 95–153.

Rocco, R. (2007) 'Economic Flexibilization and Denationalization in Brazil' in Alex E. F. Jilberto and Barbara Hogenboom (eds) *Big Business and Economic Development* (London: Routledge) 191–211.

Serrano, F. (2004) 'Relações de Poder e a Política Macroeconômica Americana, de Bretton Woods ao Padrão Flexível' in José Luis Fiori (ed.) *O Poder Americano* (Rio de Janeiro: Edit Vozes), 179–225.

Skocpol, T. (1985) 'Bringing the State Back In: Strategies of Analysis in Current Research' in Peter Evans, D. Rueschemeyer, T. Skocpol, *Bringing the State Back In* (Cambridge: Cambridge University Press).

Wade, R. and Veneroso, F. (1998) 'The Asian Crisis: the high debt model versus the Wall Street Treasury IMF Complex', *New Left Review*, no. 228, 3–25.

10

Harmonic and Conflict Views in International Economic Relations: a Sraffian View

*Sergio Cesaratto**

In dealing with the European crisis and the frequent accusation of German 'mercantilist behaviour' (Cesaratto, 2011, 2012b, 2013; Cesaratto and Stirati, 2011), it seemed natural to look for analytical approaches to international economic relations (IER) that went beyond naive pro-European rhetoric and mainstream economic beliefs in the harmonic virtues of international laissez-faire. Some Sraffian contributions to the demolition of these beliefs will be recalled below. The pre-laissez-faire, mercantilist tradition was another natural candidate for attention. Mercantilism, the world of non-harmony, may be envisaged as an underground tradition, which a group of northern European economists called 'the other canon' (www.othercanon.org), parallel to the laissez-faire tradition. Kalecki's view of net exports as a way of realising profits also buttresses a conflict view of IER. A cynical view also springs from political realism, a major tradition in political science. An intellectual father of political realism was Thomas Hobbes, contemporary of many British mercantilists. In this tradition, a social contract is enforceable at domestic level by attributing authority to the Prince, but not at international level where sovereign states do not submissively recognise any higher authority. Mercantilism and political realism converge in international political economy (IPE), a field that arose in the early 1970s as an attempt to bridge the gap between the disciplines of international economics and international relations (Strange, 1970). Political realism is commonly juxtaposed with a liberal tradition that holds a more harmonic view of IER. IPE has recently been colonised by neoclassical political scientists. The Sraffian criticism of neoclassical economics therefore appears remarkably

*I thank an anonymous referee for useful comments.

precious as a response to neoclassical imperialism. The nation-state is at the centre of mercantilism and IPE. The classical and Marxist approaches are not on easy terms with the notion of nation-state, so that it was also impossible to avoid this topic.

This chapter is a preliminary exploration of the complementarity of the classical conflict view of income distribution and the disharmonic traditions of IER in opposition to the harmonic beliefs of economic and political liberalism.[1]

10.1 Mercantilism and laissez-faire

In classic treatises on mercantilism, Furniss (1920), Suviranta (1923) and Heckscher (1955) regard it as a first *systematic* attempt to understand economic phenomena.[2] All trained in neoclassical economics, these economists found mercantilism profoundly different from laissez-faire doctrines, both classical and marginalist.

Assessing Adam Smith's famous criticism of mercantilism, to which Smith devoted a quarter of *The Wealth of Nations*, Suviranta (1923, p. 160) notes that '[t]he difference between the mercantile and the liberal point of view was not accidental, arising merely from confusion in thought, but it was deep-rooted in the different character of these economic systems'. In the latter point of view, the purpose of economic activity is '[p]roducing wealth for satisfying human wants, i.e. the ultimate end is consumption'. On the other hand, according to mercantilism '[t]he logical consequence of the fact that the people were primarily thought of as a capital material, was that consumption also came primarily to be servant of production, and not a means of satisfying human wants' (*ibid.*, p. 162).

In a similar vein, Heckscher (1955 [II], p. 285) considered mercantilism to be amoral, because it broke with the ethical foundations of the Middle Ages political views, and characteristically put the raison d'état, not individualism, centre-stage: 'the welfare of society or, in actual fact, the welfare of the state was substituted in place of the amelioration of the individual. This was a perfectly simple corollary of the raison d'état, or pure Machiavellism. … In addition the raison d'état was conceived emphatically as materialistic or economic' (*ibid.*, p. 286). Mercantilism was also amoral with regard to its means. In particular, the pursuit of individual interest was seen as a function of the welfare of the state, and had to be regulated for that purpose (*ibid.*, p. 293 and *passim*). Heckscher regarded this view as being opposite to that of laissez-faire economists influenced by the utilitarian goal of improving social happiness, seen as the sum of individual welfare (*ibid.*): both the individual and the state served the 'community'. The mercantilists held a more cynical view in which both individuals and the community served the state (*ibid.*, pp. 328–9). According to Heckscher mercantilists thought in terms of nation-states, and did not regard the individuals composing a nation as equal.

Rational thought pervaded mercantilism: 'Rationalism characterized mercantilism to so high a degree. There was little mysticism in the arguments … this rationalism expressed itself in references to *nature*. Nature was conceived as a factor which also influenced the social sphere, social life being placed parallel to physical life of the individual; and society was regarded as a body with functions similar to those of the physical body' (*ibid.*, p. 308). Mercantilists, as well as laissez-faire theorists, therefore believed in the possibility of discovering natural laws. The difference with laissez-faire theorists was that mercantilists did not regard public intervention as interference with the benign working of those laws, but as their natural complement. Mercantilists did not believe in a 'immanent social rationality' (*ibid.*, p. 321) or objective economic harmony, whereas laissez-faire 'went so far in its belief in the domination of natural laws in society that it believed in an immanent reason in the free play of forces' (*ibid.*, p. 323).[3]

Heckscher regarded the victory of laissez-faire theories over mercantilism as temporary. During the nineteenth century they were submerged by historicist and nationalistic doctrines: 'Society was regarded as a growth in the highest degree naturally determined, to be changed only by slow and gently progressive treatment, bound to tradition, each individual nation containing inherent and more or less ineradicable peculiarities' (*ibid.*, p. 334). According to the Swedish economist, laissez-faire doctrines also failed in their alleged humanitarian aims. Indeed, what they did, Heckscher argues, was to support measures that protected the individual against the state, but they failed to protect him against the market, 'against the pressure of social conditions, which did not have their origins in definite measures of the state but which, on the contrary, demanded such measures if they were to be abolished. On this point, *laissez-faire* was obstructed by its belief in natural rights, i.e., its belief in a predetermined harmony, to which was added in practical policy the influence of employer and capitalist interests' (*ibid.*, p. 337). Heckscher observes that paradoxically but not incidentally, social reforms were supported and adopted by conservatives: 'economic policy being bound up with the duty of the patriarchal state to care for the welfare of its subjects', while 'the growing importance of socialism also goaded politicians into finding remedies' (*ibid.*, p. 337).

10.2 Domestic social surplus and foreign trade surplus

In *Theories of Surplus Value* Marx argued that although James Steuart, the last great mercantilist, represented the clearest expression of the theory of 'profits upon alienation', he 'does not share the illusion that the surplus-value which accrues to the individual capitalist from selling the commodity above its value is a creation of new wealth' (Marx, (1969 [1861–3], p. 41). What Marx seems to appreciate is the idea that the surplus value is the result of one side taking advantage of the other, although the 'unequal exchange' that generates a surplus for one trading side takes place, for Marx, in the labour, not the goods market. The classic treatises on mercantilism also underline the clash between the mercantilists' view of distribution and that of 'modern' marginal theory.

Furniss (1923, pp. 198–203) regards wage determination in mercantilism as affected by the interests of the dominant classes, something far removed from the marginalist notion of wages as the natural reward of labour linked to its (marginal) contribution to production. The contrast between the mercantilist view of labour as 'the source of national wealth' and the recommendation of low wages led Furniss to perceive 'the germs of the socialist doctrine' (*ibid.*, p. 25), the term he uses to denote the classical surplus approach, in mercantilism. Indeed, very few mercantilists clearly anticipated the classical concept of surplus, and almost none came close to seeing the origin of a foreign surplus in a domestic surplus of production over consumption.[4] However, the concept was there, for instance, when they regarded the unemployed as a burden on the nation, implying that they lived on a surplus produced by the employed population, and when they argued that productive utilisation of the unemployed would increase the riches of the nation (e.g. *ibid.*, pp. 41–7 and 89–95).[5] In a famous example, Petty alludes to the ability of one section of the population to sustain the rest, including those that produce export goods (e.g. Aspromourgos 1996, p. 23). Davenant is worth quoting: 'If all hands in this Kingdom that are able were employed in useful labour our manufactures would be so increased that the commonwealth could be thereby greatly enriched and the poor, instead of being a charge, would be a benefit to the Kingdom' (quoted by Furniss, p. 91).

Suppose that the social product P just consists of *necessities*: $P=N+N'$, where N and N' are the necessities of L workers and U unemployed, respectively, both receiving a real wage equal to w, so that $P=Lw+Uw$. Defining per capita output as $\pi=P/L$, we get: $L(\pi-w)=Uw$. In other words, the unemployed survive on the surplus produced by workers above their subsistence ($\pi > w$, otherwise the unemployed would already have died). If the unemployed are put to work, the social product becomes: $P' = (L + U)\pi = L'\pi$, and the social surplus: $S = P' - L'w = L'(\pi - w)$. Now a larger surplus can be used to sustain an unproductive class (not consisting of unemployed in this case) or exported.

However, only in a late mercantilist quoted by Furniss, William Hay, an author who was presumably influenced by Petty, do we find clear coordination between domestic social surplus and foreign trade surplus:[6] 'The source of wealth is from the number of its inhabitants; … the more populous a country is, the richer it is or may be … For the earth is grateful and repays their labour not only with enough but with an abundance … Now whatever they have more than they consume, the surplus is the *riches* of the nation. This surplus is sent to other nations and is there exchanged or sold, and this is the trade of the nation. If the nation to which it is sent cannot give goods in exchange to the same value they must pay for the remainder in money; which is the balance of trade; and the nation that hath that balance in her favour must increase in wealth' (1751, quoted by Furniss, pp. 19–20, italics in Furniss).

Using the same symbols as above, the coordination between the two surpluses can be summarised as follows. In a surplus-producing economy we have: $S = P - N$. The social surplus can have a number of destinations: capitalists' consumption C_c, capital accumulation I and net exports $X - M$, that is: $S = C_c + I + X - M$. If, for simplicity C_c and I are zero, we obtain: $S = X - M$.

10.3 The centrality of national output, employment and trade surplus

As we have seen, mercantilism appears to have accorded primacy to production rather than to consumption, as in later laissez-faire theories.[7] The goal of maximising domestic production and employment, while minimising domestic consumption and imports of superfluous goods, aimed at obtaining the largest possible foreign trade surplus, which was seen by many mercantilists as the origin of net wealth for the nation. As suggested by Kalecki's lesson, we may now interpret foreign trade surplus as a way for capitalists to realise the domestic social surplus they do not consume or invest.[8] Although full coordination of internal and external surpluses was to some extent approached, it may not be said that the mercantilists were successful in this regard. So on one hand we are left with hints, by later mercantilists in particular, that the social surplus is the origin of net wealth (intended as that part of the social surplus that can be consumed or accumulated without endangering reproduction of the system on at least the same scale) and on the other hand with the idea that the origin of net wealth lies in the foreign trade surplus. How did they justify this second origin of net wealth or the importance they attributed to the trade surplus?

Profits upon alienation

Supposing, like Marx, that most mercantilists held a theory of 'profit upon alienation', it follows naturally that net gains for the nation as a whole can only be obtained by foreign trade (Heckscher [II], 1955, p. 193). One of the clearest expressions of this view is in widely quoted passages by Charles Davenant:

> It is the Interest of all Trading Nations, whatsoever, that their Home Consumption should be little, of a Cheap and Foreign Growth and that their own manufactures should be Sold at the highest Markets, and spent Abroad; Since by what is Consumed at Home, one loseth only what another gets, and the Nation in General is not at all the Richer; but all Foreign Consumption is a Clear and Certain Profit. (1697, quoted by Heckscher [II], p. 115)

So there are no 'profits' if the product is sold at home, but positive profits if consumption is kept at a minimum, cheap foreign commodities are imported, and net output sold abroad. Mercantilists had no clear notion of profits as the net income accruing to capital (Ehrlich, 1955) and as we have seen, almost without exception they did not clearly perceive the existence of a domestic surplus as the basis for a foreign surplus. They seemed to have even less idea that, given the real wage, the larger the domestic surplus, the more difficult its realisation in the domestic market, and the greater the need for 'external markets'. In any case, Kalecki's theory upholds their point of view as to the importance of net exports.

Fear of goods

The importance attributed by mercantilists to the trade surplus may be linked to the importance they seemed to attribute to the maximisation of domestic output and employment by minimising imports and maximising exports, and vice versa, to the maximisation of domestic employment to reduce imports and increase exports.[9] In this regard Johnson (1937, p. 302) describes a 'balance of work' as the difference between the labour content of exports and imports (see also Furniss, 1920, pp. 13–14; Suviranta, 1923, p. 142; Heckscher, 1955 [II], p. 366). Imports were seen as reducing domestic employment, and exports as labour 'paid by foreigners'. Here one may perhaps identify the 'fear of goods', which Heckscher saw as a leitmotif of mercantilism, as fear that limitation of the domestic market was incompatible with full employment, not such a strange concern in a low-wage economy (Heckscher, 1955 [II], pp. 121, 365).

Low wages were seen by mercantilists, though not unanimously, as a way to keep domestic consumption and imports at bay (*ibid.*, p. 364) and production costs low (*ibid.*, pp. 152–3), while encouraging hard work (Furniss, 1920, Ch. 4). A low-wage economy was therefore a central objective for a typical mercantilist (Furniss, 1920, pp. 8, 40; Heckscher, 1955 [II], pp. 46, 163–5, 153–4, 364–5). This calls for reflection.

We may distinguish a policy of *import substitution* from an *export-led* economy. The mercantilist position is evocative of both. In Keynesian terms, a policy of *import substitution* that, say, decreases the marginal propensity to import, has (*ceteris paribus*) a positive effect on domestic output and employment and – for a given level of exports – on the trade balance. In addition to the initial benefit for output and employment, by relaxing the foreign balance of trade constraint an import substitution policy also leaves more scope for growth policies based *on the growth of the domestic market*. So, in this case, the improvement in foreign trade is necessary to development of the domestic market, not an objective *per se*.

A low-wage economy would also keep imports at bay – assuming that imports of foreign goods are linked to the level of real wages. This policy, however, is hardly relevant to the development of a domestic market, but conducive to an export-led model. Once a decision to depress the domestic market is taken through a low-wage policy, then an *export-led model* is the only game in town, sustained in turn by the low labour costs. This description of a low-wage, export-led economy suits the mercantilist idea of imports as paying foreign labour, and exports as domestic labour paid by foreigners. According to the Kaleckian surplus approach, in such an economy capitalists maximise the domestic surplus they extract from workers, and get rid of it by net exports.[10]

Influx of money capital

Although a popular interpretation of mercantilism regards mere accumulation of specie as the aim of a trade surplus, mercantilists seem to have maintained that a net influx of precious metals was functional to growth of the domestic money supply, lowered the interest rate and favoured economic activity (Heckscher, 1955 [II], pp. 204, 208–9, 217–18 and *passim*). Heckscher notes that the idea that savings finance investment was absent in this literature, and capital was identified with money (*ibid.*, pp. 198–9), something that would not sound strange to modern unconventional monetary economists. This was Keynes's favourite explanation of mercantilism, expressed in Chapter 23 of the *General Theory*.

Power or plenty

Finally, the goal of a foreign trade surplus can be interpreted as a way to generate relative international political power.[11] The relative importance of power and wealth was the subject of controversy between Heckscher (1955 (I), p. 24 and *passim*) and Viner (1948); both acknowledged that the two objectives mutually sustained each other, but Heckscher attributed primacy to power and Viner to wealth. The mercantilist tradition has indeed been perceived as putting the nation-state at the centre of analysis, contrary to the cosmopolitan views of laissez-faire and Marx's theories: 'the state stood at the centre of mercantilist endeavours as they developed historically: the state was both the subject and the object of mercantilistic economic policy' (Heckscher, 1955 [I], p. 21). In short, mercantilism is 'the economic system of nationalism' (*ibid.*, [II], p. 13).[12]

In discussing presumed German mercantilism (Cesaratto, 2011, 2012b; Cesaratto and Stirati, 2011), I regarded export-led growth as a growth policy choice alternative to domestic-demand-led growth. In Germany, the Keynesian perspective was rejected by the influential Ordoliberal school as conducive to social indiscipline and inflation, regarded in turn as disturbing market-led resource allocation. Although not explicitly acknowledged, in this context, export-led growth remained the only game in town – a model that is conducive to, and simultaneously supported by, social discipline and wage restraint (Crouch, 2008, p. 479). As Ludwig Erhard (1897–1977) stated in 1953: 'foreign trade is not a specialized activity for a few who might engage in it, but the very core and even precondition of our economic and social order' (quoted by Cronin, 1996). Germany is indeed a perfect mercantile economy. At the micro level it has an excellent training, educational and R&D system; at the meso level the yardstick of a trade surplus creates an ideological climate that induces cooperation and discipline; at the macro level the systems keep wage growth in line with productivity growth and fiscal policy moderates domestic demand. Foreign policy has the promotion of German exports as priority. Paternalism is a traditional attitude of the German government; a sense of national community, traditions and nature is the main component of 'German ideology'. This perfectly suits the mercantilist tradition, particularly in its German version (Cameralism, Historical School, cf. Riha, 1985, chs 4 and 5). Codetermination is a further institutional pillar, but as Voltaire said (and as the late Vianello liked to quote): 'Incantations will destroy a flock of sheep if administered with a certain quantity of arsenic'. Just in case, a watchdog role was taken on by the Bundesbank in a unique wage-bargaining process directly involving the central bank and the leading trade union IG-Metall (Franzese and Hall, 2000, pp. 182–3). As we have seen, this model perfectly fits the Kaleckian view that net exports are a way of realising a domestic surplus, and that financial capital flows from core-surplus countries are a way to finance expenditure by trade peripheral-deficit countries, as confirmed by recent European experience (Cesaratto, 2012b).

10.4 The mercantilist tradition, liberalism and international political economy

Mercantilism was deemed dead in the late eighteenth century, when Adam Smith was self-confident enough to pass silently over the contribution of the last great quasi-contemporary mercantilist, James Steuart. Nonetheless mercantilist wisdom survived as an underground stream of thought parallel to mainstream laissez-faire economics, for instance in the work of protectionists Alexander Hamilton and Friedrich List in the German Historical School, up to modern developmentalism and in some versions of IPE.[13]

International political economy arose in the early 1970s in English-speaking countries (see Cohen, 2008). Economists Albert Hirschman and Charles Kindleberger are regarded as its forerunners. Simplifying, there are two competing political-philosophical inspirations of IPE, liberalism and political realism, which hold harmonic and conflict views of IER, respectively. The intellectual fathers of political realism are Thucydides, Machiavelli and Hobbes (see, e.g., Donnelly, 2000). All held a pessimistic view of human nature as being motivated by greed, suspicion and ambition. As a result, there is a conservative element in political realism – the human soul cannot change – but also a healthy reaction against facile utopias, which are seen as an obstacle to real change.[14] The international arena is regarded as an anarchic field governed by the rules of power: no Hobbesian social contract that delegates power to a super-national authority is acceptable to sovereign states, if not for contingent convenience.[15] The modern recovery and application of this approach to international relations is due to Edward Carr (notably a Marxist), Hans Morgenthau and Kenneth Waltz. On the other side, the liberal tradition maintains that affirmation of the market economy is a solution to domestic and international conflicts through the mutual advantages of free trade (on this, cf. the critical analysis by Albert Hirschman, 1977).

In this regard, an influential exponent of IPE, Robert Gilpin, distinguishes liberal, nationalist and Marxist traditions. 'Economic liberals', he writes, 'believe that the benefits of an international division of labour based on the principle of comparative advantage cause markets to arise spontaneously and foster harmony among the states; they also believe that expanding webs of economic interdependence create the basis for peace and cooperation in the competitive and anarchic state system' (Gilpin, 1987, pp. 12–13). On the other hand, 'Economic nationalists … stress the role of power in the rise of a market and the conflictual nature of IER'; they argue that 'economic interdependence must have a political foundation and that it creates yet another arena of interstate conflict, increases national vulnerability, and constitutes a mechanism that one society can employ to dominate another' (*ibid.*, p. 13). Gilpin identifies this approach with the mercantilist tradition. Finally, the Marxist tradition regards international relations as a field of imperial conflict and exploitation of peripheral countries.

Comparing the three approaches, Gilpin points out that, like liberals, Marxists tend to regard international trade as a modernisation force against the scepticism of nationalists (*ibid.*, p. 14).[16] Nationalists support the primacy of politics over economics, Marxists the opposite, and liberals maintain that the two spheres should remain relatively autonomous (*ibid.*, p. 26). Finally liberals and Marxists share an optimistic view of the human fate, the opposite of nationalists, who at international level base their stance on a conflict view of international relations (*ibid.*, p. 43).

The realist Gilpin is sympathetic to nationalists and Marxists, despite his personal liberal beliefs (*ibid.*, p. 25). He is also critical of the neoclassical foundations of the liberal view, accusing the dominant theory of being based on unrealistic assumptions, e.g. perfect competition, rationality, perfect information and the like. Any theory must indeed make simplifications, as long as they do not alter the substance, and this is what mainstream economists have largely done by making the alleged unrealistic assumptions. By virtue of its frail criticism of mainstream international economics, the realist tradition of IPE represented by Gilpin was exposed to liberal and neoclassical counter-criticism.[17] Not surprisingly, the latest generation of American IPE students has increasingly returned to neoclassical propositions and mainstream research methods. This new trend is called open economic policy (OEP) (Lake, 2009, pp. 50, 52). Benjamin Cohen's intellectual history of IPE (Cohen, 2008) sparked off a fierce debate on the evolution of IPE. Cohen notes divergence between American and British IPE, observing that the American school has become increasingly standardised, coming to resemble nothing so much as the methodology of neoclassical economics, featuring the same penchant for positivist analysis, formal modelling, and where possible, the systematic collection and evaluation of empirical data (*ibid.*, pp. 41–2). He also notes that political scientists 'have an inferiority complex when it comes to economics' (*ibid.*, p. 42).

Like Gilpin, defenders of traditional IPE point to the *limitations* of starting with the economic choices of rational individuals (the state was the main unit of analysis in traditional IPE). Another leading old-guard international political economist points out that a high price is paid by 'making preferences and interests exogenous, assuming that interests can be derived only from a rationalist model of human behaviour, excluding from analysis the constitutive aspects of institutional life, committing to an exclusively materialist conception of preferences and interests, and importing reductionist economic theories of politics' (Katzenstein, 2009, p. 127; see also Keohane, 2009, pp. 37–8). I do not regard this criticism of the assumption of rational choices by selfish individual agents as particularly illuminating. A non-individualistic theory of society calls rather for an alternative view of the economic fabric of society.

We cannot but refer to Marx's criticism of methodological individualism based on the classical 'surplus approach'. The methodological individualism of the early classical economists did not surprise Marx: after all, it was the ideology of a new form of society in which individuals broke previous institutional ties – feudal, religious etc. – with other individuals. This of course does not imply that ties have disappeared: they have just been superseded by more anonymous, market-dominated relations, creating the *illusion* that analysis can start with isolated individuals (e.g. Marx, 1957 [1857], pp. 82–3). The production and reproduction of social life is a collective fact for Marx, although in history the manner in which social surplus is produced and distributed has changed profoundly.[18] Individuals' interests and choices are moulded by their positions in production modes (cf. Marx, 1859). Note the degree to which the forgivable 'robinsonades' of Smith and Ricardo became the very foundations of economic and social theory with marginalism. In this theory, production is the *ex post* result of the (marginal) contributions of individual endowments of production factors, an unhistorical view in which socioeconomic relations of production are not the result of evolution of the modes by which humans produce and distribute social output and surplus.[19]

The fact that traditional IPE attributes centrality to the state as the basic unit of analysis is also problematical. On the one hand, the state is an indisputable guarantee of economic activity: 'as Carr has argued, every economic system must rest on a secure political base' (Gilpin, 1987, p. 47).[20] On the other hand, most IPE less convincingly assumes that 'society and the State form a unitary identity and that foreign policy is determined by objective national interest' when 'foreign policy (including foreign economic policy) is in large measure the outcome of conflicts between dominant groups in society' (*ibid.*, p. 48; see also Cohen, 2008, p. 125; Katzenstein 1977, p. 604; and Section 10.6 below).[21]

10.5 Comparative disadvantages?

Adam Smith accused the mercantile doctrine of looking after the interests of merchants and producers, while sacrificing those of consumers (1979 [1776], pp. 661–2). He upheld the advantages of international trade for all participating nations through exchange of surplus products, market expansion and thereby extension of the division of labour (1979 [1776], pp. 446–7). Smith held a theory of *absolute* advantages from trade which is very different from the theory of *comparative* advantages attributed to David Ricardo. A theory of absolute advantages is theoretically consistent with the pursuit of mercantilist policies, e.g. trade policies aimed at developing and safeguarding national absolute advantages. Despite his attack on mercantilism, Smith's theory of international trade is therefore not inconsistent with a disharmonic view of IER. Ricardo's theory of comparative advantages is commonly regarded as the final challenge to mercantilism: a harmonic view of international relations was seen to prevail over a conflict one, although Ricardo was very clear in limiting the validity of his celebrated theorem to the case of absence of capital mobility.

The Ricardian theory identified the origin of comparative advantages in technological differences, whereas the basic marginalist Heckscher-Ohlin-Samuelson (HOS) explanation of international trade explained specialisation on the basis of countries' different factor endowments. The international specialisation of free-trading countries with full employment thus depended on relative scarcity of factors. Sraffian authors, such as Parrinello and Steedman, have taken two critical directions. On one hand, the results of the capital theory controversy have been used to show the limited validity of the HOS theory for cases in which only land and labour are used as inputs. On the other hand, modern extensions of Ricardo's analysis have confirmed the limitations of comparative advantages that Ricardo himself pointed out.

Beginning with the first aspect, the HOS theorem may be expressed rigorously in terms of endowment of non-produced production factors, such as land and labour. On this basis the theory predicts that the country with the highest land-to-labour ratio exports land-intensive commodities. The inclusion of 'capital', however, undermined the prediction that the country with the largest 'capital supply' and the lowest interest rate exports the most capital-intensive commodity. To begin with, there is the standard problem of measuring the 'given amount of capital' irrespective of its distribution. Second, results in capital theory (assuming two sectors) show that ordering of sectors by capital intensity may change with the fall in interest rate and that the price of a more capital-intensive commodity will not fall monotonically with a fall in interest rate (Steedman, 1979a, pp. 4–5).

The international mobility of 'production factors' is seen by conventional theory as alternative to international trade: it is the same whether a relatively 'capital-rich' country exports capital-intensive commodities or 'capital' directly. The idea that capital flows arise from capital-rich countries lending *savings* to capital-poor countries is subject to capital theory criticism as much as to the domestic saving–investment nexus (Garegnani, 1983; Dalziel and Harcourt, 1997). A Kaleckian view would lead us to regard financial flows to trade deficit countries as part of a mercantilist strategy whereby surplus countries lend to deficit countries. In this view, loans precede import spending by peripheral countries, and foreign saving in core countries emerges as the result of their net exports to the periphery. The recent European crisis can be interpreted along these lines (Cesaratto, 2012b).

Marginal theory focuses on countries' different factor endowments, whereas Ricardo suggested that countries may differ because of their respective technology levels and, in the absence of capital mobility, specialise in producing the commodity for which they have the greatest comparative advantage, or the smallest comparative disadvantage. Brewer (1985) and Parrinello (2009) show that once the real wages in two potentially trading countries are, respectively, taken as given[22] and there is capital mobility, *absolute* and not comparative advantages determine the location of production. Too high a wage rate, or too low a productivity level may make a country uncompetitive, inducing capital to flow to the other country: '*We would say that a whole capitalistic economy is* *not competitive if all its capital-using techniques are unprofitable at the international equilibrium prices. This result overrules the claim that "a country must always possess a comparative advantage in something"* [as argued by Krugman]' (Parrinello, 2009, p. 52, italics in the original), which vindicates 'the intuitive idea that national competitiveness can be a source of possible economic conflict among the national economies of a global economy' (*ibid.*, p. 50).[23]

10.6 Mercantilism, classical economists and Marx on the nation-state

As we have seen, mercantilism and Marxism both reject the approach, typical of laissez-faire economics, of considering the individual as the basic analytical unit, what Marx called 'robinsonades'. The social relations of production are Marx's analytical anchor, regarded as the way a society historically organises the production and distribution of social output and surplus. The mercantilist tradition does not attain the sophistication of Marx's historical materialism, but both reject the market as the place where free choices of a variety of individuals are reconciled in a harmonic way. National communities and states, rather than Marx's social relations of production and social classes, are the reference categories of the mercantilist tradition. In an important unpublished work, Marx (1975 [1845]) rejected this approach by fiercely criticising Friedrich List (1991 [1841]), an author we can classify in the mercantilist tradition of national political economy and developmental state. Marx's stance is not surprising, as Szporluk explains:

> Marx claimed that his theory, while the result of his own intellectual endeavour, was also the reflection of objectively working historical forces and would therefore be carried out as a predestined outcome of historical development. Marx further thought that the proletariat was that 'material force' whose historical task was to realise his philosophy. When one bears all of this in mind, it is easy to see why Marx found the theories of List, particularly his view of history and his program for the future, not only objectionable but aberrant … It was axiomatic to Marx that industrial progress intensified and sharpened the antagonism between the bourgeoisie and the proletariat, an antagonism that would in the immediate future explode in a violent revolution. List, in the meantime, preached class cooperation and solidarity in the building of a nation's power. Marx thought that the Industrial Revolution, and the concomitant rule of the bourgeoisie,

> promoted the unification of the world and obliterated national differences. (Communism, he thought, would abolish nations themselves.) List claimed that the same phenomenon, the Industrial Revolution, intensified national differences and exacerbated conflicts among nations. While Marx saw the necessity of workers uniting across nations against the bourgeoisie, List called for the unification of all segments of a nation against other nations. (1988, p. 4)

The belief in free trade and comparative advantages may have led classical economists to overlook the role of the nation-state (there are, of course, exceptions, for example in *The Wealth of Nations*, and also Ricardo's belief in capital immobility can be taken as a reference to a nationalist element, but not such as to refute the general attitude). According to List, classical economists' defence of free trade served the commercial interests of Britain in having open access to foreign markets (paraphrasing Joan Robinson (1966) and Carl Schmitt, it may be argued that List regarded free trade 'as the continuation of mercantilism in other forms'). According to Marx, however, classical economists were actually decoding the secular and cosmopolitan characteristics of capitalism (Marx, 1975 [1845]; Szporluk, 1988, p. 66 and *passim*), in particular the conflicting interests of capital and labour that, in his opinion, went beyond the provincial boundaries of national states. Marx's criticism of List also reveals the German national bourgeoisie's interests behind List's vivid description of national identities (there is a similarity with Adam Smith's criticism of mercantilist writers as prejudiced defenders of merchants' interests). According to Desai (2012), Marx accuses List of hypocrisy: 'As a spokesperson for a capitalist developmental state, List was not concerned with class exploitation, only with national exploitation: "However much the individual bourgeois fights against the others, as a class the bourgeois has a common interest, and this community of interest, which is directed against the proletariat inside the country, is directed against the bourgeois of other nations outside the country. This the bourgeois calls his nationality"' (Desai, 2012, p. 62, quotation from Marx (1975 [1845], p. 281).

The expectation of a forthcoming revolution in Britain and its generalisation elsewhere presumably led Marx to dismiss the importance of the development of backward nations and to regard nationalism as an impediment to revolution rather than as a necessary historical passage to developed capitalism. For Marx, the notions of political community and national identity are illusions and false consciousness, like religion (Szporluk, 1988, p. 58). Without discarding Marx's criticism of the reactionary aspects of nationalism and the global nature of capitalism (the first part of *The Communist Manifesto* is a tribute to global capitalism), we should not forget that history vindicated the factual relevance of List's arguments about nation-states in the economic and political fields.[24] For instance, most socialist revolutions overlapped with struggles for national independence, and practical examples of international labour solidarity are rare, to say the least.

In Germany, the most representative exponent of the Young Historical School, Gustav Schmoller, also distanced himself from methodological individualism in his famous 'The Mercantile System' (1897): 'The idea that economic life has ever been a process mainly dependent on individual action, an idea based on the impression that it is merely concerned with methods of satisfying individual needs, is mistaken with regard to all stages of human civilisation, and in some respects it is more mistaken the further we go back' (*ibid*., p. 4). Schmoller's perspective is, of course, very different and somehow opposite to Marx's. Having its roots in Cameralism, German Historicism and the Romantic movements, Schmoller's Historical School views the nation-state as the supreme expression of human belonging to superior organisms (cf. Riha, 1985, chs 4 and 5). Schmoller talks of 'real *political* economies as unified organisms, the centre of which should be, not merely a state policy reaching out in all directions, but rather the living heart-beat of a united sentiment' (p. 50, italics in the original). And here we find the famous definition of mercantilism, later adopted by Heckscher:

> in its innermost kernel [mercantilism] is nothing but state making – not state making in a narrow sense, but state making and national-economy making at the same time; state making in the modern sense ... The essence of the system lies not in tariff barriers, protective duties, or navigation laws; but in something far greater: the total transformation of society and its organisation, the state and its institutions, the replacement of a local territorial economic policy with that of the national state. (*Ibid*., pp. 50–1)

This is ideology, of course, but development economists, particularly Gerschenkron (1962, p. 24), later underlined the importance of nationalist ideologies 'igniting the imaginations' of people for the mobilisation of national resources in the early stages of industrialisation. The reason why developmental bourgeoisies emerge from previous social relations of production in some nation-states, i.e., how certain pre-industrial social relations of production generate progressive pro-growth dominant classes, interested in a developmental state, are questions that have not yet been much explored. Neoclassical institutionalists seem unable to go beyond the mantra of the central role played by the protection of property rights in igniting growth (e.g. Acemoglu and Robinson, 2012). This may just be part of a story, the best of which has still to be written.[25]

10.7 Conclusions

Research on the recent European financial crisis has prompted exploration of the harmonic and disharmonic views of international economic relations. The former, more liberal view is based on the Ricardian and neoclassical trade theories. The latter is derived from pre-Smithian mercantilist conflict views of international trade. Here we have investigated the contribution that Sraffian theory can offer the latter stream of thought. This contribution cannot be underestimated since it provides a rigorous analytical rebuttal of the neoclassical theory of international trade and capital flows and supports the existence of absolute advantages, which are a source of potential trade conflict between nations. Kaleckian theory may also vindicate mercantilist attention to trade surplus. Further research is needed on the topics surveyed in this chapter. What is at stake is the space in the economics profession left by the dominant theory to these promising fields of research.

Notes

1. Conflict views do not exclude harmonic domestic and international arrangements. In fact, the social-democratic Scandinavian compromise is based on a conflict view of distribution. In the conflict approaches harmony is a subjective political result. In marginalism, harmony is an objective natural outcome of laissez-faire.
2. I regard these treatises, written in an age that saw the affirmation of nationalism, as representative of the later *reception* of mercantilist ideas. I am more interested in this reception that in the controversies on mercantilism. For example, among historians of economic thought, Coleman (1957) did not regard it as a systematic body of ideas while Bob Coats (1958) defended this view.
3. Importantly, Heckscher points out that Smith, Ricardo and Malthus perceived objective disharmony, although they did not believe that public interference would improve things much (*ibid.*, pp. 328–9).
4. See the authors quoted by Furniss (1923, pp. 25–6), for instance Chamberlen (1649): 'This may be a note to all men, especially to statesmen to look no more upon the poor as a burden but as the richest treasure of a nation, if orderly and well-employed. Which is the more manifest if we consider first,

that though they multiply more than the rich they do not only feed and clothe themselves but the rich men are fed and clothed and grow rich *by what they get out of the poor's labor over and above their maintenance*. Secondly, that the poor bear a greater burden of taxes in the city and elsewhere. For the rich either abate what they get out of the poor's labor or (which is worse) permit them to starve for want of employment' (quoted by Furniss, p. 25, italics added). And Bellers: 'Regularly laboring people are the kingdom's greatest treasure and strength, for without laborers there can be no lords; and *if the poor laborers did not raise much food and manufacture than what did subsist themselves*, every gentleman must be a laborer and every idle man must starve' (quoted by Furniss, p. 25, italics added). Johnson (1937, p. 240) quotes Dudley North (1691) who also advances a clear idea of social surplus: '[Some labourers] are more provident, other more profuse…[some] raise more fruits from the earth, than they consume in supplying their own occasions; and a surplus remains with them and is property of the riches'.

5. Mercantilists oscillate between conceiving unemployment as involuntary, e.g. Davenant: 'a defect in our constitution that many continue in wretched poverty for want of employment, though willing enough to undertake it' (quoted by Furniss 1920, p. 82); and the idea of unemployment as a sin (*ibid*., ch. 4).
6. As noted above, according to Petty the necessities produced by one section of the population also sustain those employed in the export sector. We find here a clear coordination between internal and external surpluses.
7. Suviranta, 1923, pp. 122–3, 161; Heckscher (1955 [II], p. 124): 'the power of creating wealth is more important than the wealth itself'.
8. On Kalecki and the 'Sraffian supermultiplier' approach to accumulation theory, see Cesaratto (2012a).
9. 'When people had once arrived at the view that a surplus of goods was something undesirable, the connection between this and the amount of employment followed inevitably' (Heckscher, 1955 [II], p.122. Notoriously, in Heckscher a foreign trade surplus was necessary to get rid of what he named 'fear of goods' or 'fear of redundant stocks' (*ibid*., p. 59, fn. 3).
10. Serrano (2008, p. 14) criticises Kalecki for this partially unfortunate passage: 'If exports increase and at the same time there is an equal increase in imports, overall profits remain unchanged; international trade is boosted, *but production in the country does not increase*, nor will there be any inducement for expansion of investment activity' (Kalecki, 1934, p. 16, my italics; see also 1967, p. 152). No doubt there is an ultra-mercantilist element in this sentence: international trade is a zero-sum game. Indeed, although Kalecki correctly points out that only a trade surplus generates profits, he misses the point that an increase in exports, even if accompanied by a corresponding rise in imports, nonetheless provokes an equivalent increase in domestic output: a sort of 'balanced foreign trade-budget theorem'. However, Kalecki is not wrong when he argues that for a given output, the larger the profit share and the lower the share of profits that capitalists consume or invest, the larger the trade surplus necessary to realise profits.
11. Heckscher (1955 [II], p. 317) argues that the 'obsession with power also had this result, that the interest was taken not in the absolute total of commerce nor in the utility which it represented to the inhabitants of a particular

country, but only in the superiority gained over other countries, irrespective of whether there was no absolute increase at all or perhaps even an absolute decline'. The best quotation he provides is from an important German mercantilist: 'Whether a nation be to-day mighty and rich or not depends not on the abundance or scarcity of its powers or riches, but principally on whether its neighbours possess more or less than it. For power and riches have become a relative matter, dependent on being weaker and poorer than others' (*ibid.*, p. 22). Locke would hold a similar thesis (*ibid.*, pp. 22–3).

12. Heckscher distinguishes between the Romanticist notion of nationalism concerned with traditions, ethnicity etc., and the secular mercantilist identification of it with the state interest.
13. Mercantilism has also not disappeared in lay(wo)men's and politicians' preoccupation with foreign competition. Krugman (1997) devoted a book to disproving these preoccupations. Here we ignore other important traditions in international economic relations, for instance that of Wallerstein.
14. Cf. the symposium 'American Realism and the Real World' (2003) *Review of International Studies*: 29.
15. An inspiring figure of modern political realism, Reinhold Niebuhr (1892–1971), an American Protestant priest, wrote: 'Power sacrifices justice to peace within the community and destroys peace between communities' (quoted by Donnelly, 2000, p. 27).
16. '[L]iberals believe that trade and economic intercourse are a source of peaceful relations among nations because their mutual benefits of trade and expanding interdependence among national economies will tend to foster cooperative relations. Whereas politics tends to divide, economics tends to unite people' (Gilpin, 1987, p. 31).
17. This superficial criticism is also shared by many 'post-Keynesian' economists.
18. In a famous book, polymath Jared Diamond (2005 [1997]) regards the production of an economic surplus as the trigger of human civilisation. He presents a theory of human evolution that recalls the 'four stages' found in classical authors, such as Turgot and Smith (Meek, 1971). See also below n. 25.
19. Adam Smith was quite aware of the social nature of individuals (see Cesaratto, 1996, for a comparison with Schumpeter).
20. Carr (1981 [1939]) wrote: 'Economic forces are in fact political forces. Economics can be treated neither as a minor accessory of history, nor as an independent science in the light of which history can be interpreted. Much confusion would have been saved by a general return to the term "political economy", which was given the new science by Adam Smith himself and not abandoned in favour of the abstract "economics", even in Great Britain itself, till the closing years of the nineteenth century. The science of economics presupposes a given political order, and cannot be profitably studied in isolation from politics' (*ibid.*, p. 108).
21. Traditional IPE is also interested in studying how ideologies are formed, a field called 'constructivism' (cf. Cohen, 2008, pp. 131–2). In this regard, IPE Marxist scholar Robert Cox particularly emphasised the contribution of Gramsci (see Cohen, 2008, p. 90).
22. For instance, assuming that with unemployed labour wages are close to subsistence level determined on the basis of historically determined social norms.

23. Ricardo himself limited the validity of his theory of comparative advantages to the case of no capital mobility. In this light, the anti-Ricardianism of colleagues of 'the other canon', mentioned at the beginning of this chapter, appears particularly misplaced.
24. Indeed, subsequent Marxist literature took the imperialist clash between the main economic powers into great consideration; I did not consider this development here (e.g. Brewer, 1980). Rosa Luxemburg's 'external markets' as a necessity for core-capitalist countries to realise social surplus, a view recovered by Kalecki (1967) that may vindicate some mercantilist insights, was born in precisely this context.
25. Meek (1976) points out that both Turgot and Smith regarded the protection of property rights as a result of development rather than a cause of it. In a similar vein, in a recent review of Acemoglu and Robinson (2012), Diamond (2012) is also very critical of these two neoclassical institutionalists, pointing out a causal chain that goes from the emergence of a food surplus to lawful complex societies: 'The chain of causation leading slowly from productive agriculture to government, state formation, complex institutions, and wealth involved agriculturally driven population explosions and accumulations of food surpluses, leading in turn to the need for centralized decision-making in societies much too populous for decision-making by face-to-face discussions involving all citizens, and the possibility of using the food surpluses to support kings and their bureaucrats.'

References

Aspromourgos, A. (1996) *On the Origins of Classical Economics: Distribution and Value from William Petty to Adam Smith* (London and New York: Routledge).

Acemoglu, D. and Robinson, J. A. (2012) *Why Nations Fail: The Origins of Power, Prosperity, and Poverty* (New York: Crown).

Brancaccio, E. and Fontana, G. (eds) (2011) *The Global Economic Crisis. New Perspectives on the Critique of Economic Theory and Policy* (London: Routledge).

Brewer, A. (1980) *Marxist Theories of Imperialism – A Critical Survey* (London: Routledge & Kegan Paul).

Brewer, A. (1985) 'Trade with fixed real wages and mobile capital', *Journal of International Economics*, XVIII: 177–86.

Carr, E. H. (1981 [1939]) *The Twenty Years' Crisis – 1919–1939* (London: Macmillan).

Cesaratto, S. (1996) 'Alternative Perspectives on Division of Labour, Structural Change and Economic Growth', in C. De Bresson (ed.), *Economic Interdependence and Innovative Activity: An Input-output Analysis* (Edward Elgar) (rev. version: *Studi Economici*, 67, 1999).

Cesaratto, S. (2011) 'Europe, German Mercantilism and the Current Crisis', in E. Brancaccio and G. Fontana (eds), *The Global Economic Crisis. New Perspectives on the Critique of Economic Theory and Policy* (London: Routledge).

Cesaratto, S. (2012a) 'Neo-Kaleckian and Sraffian controversies on accumulation theory', University of Siena Working Paper, 650, mimeo.

Cesaratto, S. (2012b) 'Controversial and Novel Features of the Eurozone Crisis as a Balance of Payment Crisis', in O. Dejuán, J. Uxó and E. Febrero

(2013), *Post Keynesian Views of the Economic Crisis and its Remedies* (London: Routledge).

Cesaratto, S. (2013) 'The Endless Eurozone Crisis, Where Do We Stand? A Classical-Kaleckian Overview', Dipartimento di Economia Politica e Statistica, Università di Siena, Working paper 671, http://www.econ-pol.unisi.it/dipartimento/it/node/1813 (forthcoming in *Studi Economici*).

Cesaratto, S. and Stirati, A. (2011) 'Is Germany a mercantilist country? Germany in the European and global crises', *International Journal of Political Economy*, XXXIX: 56–87.

Coats, A. W. (1958) 'In Defence of Heckscher and the Idea of Mercantilism', *Scandinavian Economic History Review*, V, 173–87 (repr. in ID, *On the History of Economic Thought*, vol. I, London: Routledge, 1992).

Cohen, B. J. (2008) *International Political Economy: An Intellectual History* (Princeton: Princeton University Press).

Coleman, D. C. (1957) 'Eli Heckscher and the idea of mercantilism', *Scandinavian Economic History Review*, V: 1–25 (repr. in ID [ed.], *Revisions in Mercantilism*, London: Methuen, 1969).

Cronin, J. E. (1996) *The World the Cold War Made: Order, Chaos and the Return of History* (London: Routledge).

Crouch, C. (2008) 'What will follow the demise of privatised Keynesianism', *Political Quarterly*, LXXIX: 476-87.

Dalziel, P. C. and Harcourt, G. C. (1997) 'A note on "Mr Meade's Relation" and international capital movements', *Cambridge Journal of Economics*, XXI: 621–31.

Desai, R. (2012) 'Marx, List, and the materiality of nations', *Rethinking Marxism*, XXIV: 47–67.

Diamond, J. (2005 [1997]) *Guns, Germs and Steel: A Short History of Everybody for the Last 13,000 Years* (London: Vintage).

Diamond, J. (2012) 'What makes countries rich or poor?', *New York Review of Books*, 7 June .

Donnelly, J. (2000) *Realism and International Relations* (Cambridge: Cambridge University Press).

Ehrlich, H. B. (1955) 'British mercantilist theories of profit', *American Journal of Economics and Sociology*, XIV: 377–86

Farrell, H. and Finnemore, M. (2009) 'Ontology, methodology, and causation in the American school of international political economy', *Review of International Political Economy*, XVI: 58–71.

Franzese, R. J. and Hall, P. A. (2000) 'Institutional Dimensions of Coordinating Wage-Bargaining and Monetary Policy', in T. Iversen, J. Pontusson and D. Soskice (eds), *Unions, Employers, and Central Banks: Macroeconomic Coordination and Institutional Change in Social Market Economies* (Cambridge: Cambridge University Press): pp. 173–204.

Furniss, E. (1920) *The Position of the Laborer in a System of Nationalism: A Study in the Labor Theories of the Later English Mercantilists* (Boston: Houghton Mifflin Company).

Garegnani, P. (1983) 'Notes on Consumption, Investment and Effective Demand', in J. Eatwell and M. Milgate (eds), (1983), *Keynes's Economics and the Theory of Value and Distribution* (London: Duckworth).

Garegnani, P. (1984) 'Value and distribution in the classical economists and Marx', *Oxford Economic Papers*, XXXVI: 291–325.

Gerschenkron, A. (1962) *Economic Backwardness in Historical Perspective: A Book of Essays* (Cambridge, MA: Belknap Press of Harvard University Press).

Gilpin, R. G. (1984) 'Richness of the tradition of political realism', *International Organization*, XXXVIII: 287–304.

Gilpin, R. G. (1987) *The Political Economy of International Relations* (Princeton: Princeton University Press).

Heckscher, E. (1955) *Mercantilism*, 2nd edn (1st edn in Swedish, 1931) (London: George Allen & Unwin).

Hirschman, A. (1977) *The Passions and the Interests: Political Arguments For Capitalism Before Its Triumph* (Princeton: Princeton University Press).

Johnson, E. A. J. (1937) *Predecessors of Adam Smith* (London: P.S. King & Son).

Jonung, L. and Drea, E. (2009) 'The euro: It can't happen, It's a bad idea, It won't last. US economists on the EMU', www.ec.europa.eu/economy_finance/publications

Kalecki, M. (1967) 'The problem of effective demand with Tugan-Baranovski and Rosa Luxemburg', in ID, *Selected Essays on the Dynamics of the Capitalist Economy* (1971) (Cambridge: Cambridge University Press).

Katzenstein, P. J. (1977) 'Domestic and international forces and strategies of foreign economic policy', *International Organization*, XXXI (issue no. 4: 'Between Power and Plenty: Foreign Economic Policies of Advanced Industrial States'): 587–606.

Katzenstein, P. J. (2009) 'Mid-Atlantic: Sitting on the knife's sharp edge', *Review of International Political Economy*, XVI: 122–35.

Keohane, R. O. (2009) 'The old IPE and the new', *Review of International Political Economy*, XVI: 34–46.

Kindleberger, C. P. (1981) 'Dominance and leadership in the international economy', *International Studies Quarterly*, XXV: 242–54.

Kriesler, P. and Halevi, J. (1996) 'Asia, Japan and the internalization of effective demand', *Economies and Sociétés – Monnaie et production*, X: 301–20.

Lake, D. A. (2009) 'TRIPs across the Atlantic: theory and epistemology in IPE', *Review of International Political Economy*, XVI: 47–57.

List, F. (1841) *The National System of Political Economy* (1991 edn) (New Jersey: Augustus M. Kelly).

Marx, K. (1975 [1845]) Draft of an Article on Friedrich List's book: *Das Nationale System der Politischen Oekonomie*, *Marx-Engels Collected Works*, Vol. 4: Marx and Engels: 1844–45 (New York: International Publishers).

Marx, K. (1973 [1857]) *Grundrisse – Foundations of the Critique of Political Economy (Rough Draft)* (Harmondsworth: Penguin Books).

Marx, K. (1977 [1859] *A Contribution to the Critique of Political Economy* (Moscow: Progress Publishers).

Marx, K. (1969 [1861–3]) *Theories of Surplus Value*, Part 1 (London: Lawrence & Wishart).

Meek, R. L. (1971) 'Smith, Turgot and the Four Stages Theory', *History of Political Economy*, 3: 9–27.

Meek, R. L. (1976) *Social Science and the Ignoble Savage* (Cambridge: Cambridge University Press).

Mundell, R. A. (1957) 'International Trade and Factor Mobility', *American Economic Review*, XLVII, 321–35.

Parrinello, S. (2009) 'The Notion of National Competitiveness in a Global Economy', in D. Vint, S. Metcalfe, H. Kurz, P. Samuelson and N. Salvadori (eds), *Essays in Honour of Ian Steedman* (London: Routledge).

Riha, T. (1985) *German Political Economy: The History of an Alternative Economics* (Bradford: MCB University Press).
Robinson, J. (1966) *The New Mercantilism* (Cambridge: Cambridge University Press).
Schmoller, G. (1897) *The Mercantile System and Its Historical Significance* (New York: Macmillan).
Serrano, F. (2008) 'Kalecki y la economia americana en los anos 2000', *Circus*, I: 7–24.
Sraffa, P. (1951) 'Introduction to Ricardo's Principles', in P. Sraffa and M. H. Dobb (eds) *The Works and Correspondence of David Ricardo*, Vol. 1 (Cambridge: Cambridge University Press).
Steedman, I. (1979a) *Trade amongst Growing Economies* (Cambridge: Cambridge University Press).
Steedman, I. (ed.) (1979b) *Fundamental Issues in Trade Theory* (London: Macmillan).
Strange, S. (1970) 'International economics and international relations: a case of mutual neglect', *International Affairs*, XLVI: 304–15.
Smith, A. (1979 [1776]) *An Inquiry into the Nature and Causes of the Wealth of Nations* (Oxford: Clarendon Press).
Suviranta, B. (1923) *The Theory of Balance of Trade in England – A Study in Mercantilism* (Helsinki: Annales Academiae Scientiarum Fennicae, B-XVII).
Szporluk, R. (1988) *Communism and Nationalism: Karl Marx Versus Friedrich List* (Oxford: Oxford University Press).
Viner, J. (1948) 'Power and plenty as objectives of foreign policy in the seventeenth and eighteenth centuries', *World Politics*, I: 1–29.

11
Lessons from the Crisis: the Macroeconomics of Leverage

Paolo Leon

11.1 Introduction

Let me advance a possible explanation of the current stagnation and of the preceding crisis and crash. The 2007 crash, generally thought to be of financial origin, is imputed to speculation without regulation and/or to the inevitable eventual failure of otherwise self-fulfilling expectations: surely, there must be more to it, and an ample literature has accordingly developed. In this chapter I shall merely offer hypotheses, unaccompanied by statistical proof. Logic mistakes are likely, but my purpose is to stimulate research by others. In the following argument I will discuss events in the USA and the UK, and to a lesser extent continental Europe (perhaps Japan) – the argument is restricted to something that looks like a closed economy; and, in effect, if a China was not there, I don't know what part of the following could stand.

As anticipated, much of the argument has already been addressed by a growing literature, part of which has tried to put together inequality, leverage and crises. One such work (Kumhof and Rancière, 2010) uses a set of non-standard (i.e. Neo-Keynesian) neoclassical assumptions to obtain a basically microeconomic relationship between income distribution and the crisis, leading to macroeconomic results. Similarly, a previous attempt at reconciling household debt and inequality (Iacoviello, 2008), attributes the growth in debt to inequality, or to a decline in the bargaining power of wage earners (called 'impatient agents'). Quite apart from the method used, I consider this causation the weak link of these works, given that in the periods examined by the authors, full employment was prevalent. Many other economists have worked on the relationship between excessive leverage and the subsequent crash, but most of this literature is descriptive, and basically proposes that anything that rises too much, sooner or later must fall. To my knowledge, the only non-neoclassical work that has put forward a reasonable explanation of the crash, by linking household debt and income distribution in a macroeconomic setting, is that of Barba and Pivetti (2009), and my approach is indeed similar to theirs. However, these authors also consider that household debt results from weak bargaining power of the wage earners. I shall try to argue that it is rather precisely *the increase in household debt that determines a worsening distribution between wages and all other incomes*.

11.2 Wealth and income

One aspect of the crisis, related to the long period preceding it, is the well established fact that households (as well as enterprises) were able *to transform the increase in their wealth*, derived from ever-rising stock and residential market values, *into new revenue*: either by speculating on their assets or by increasing their debt or both, in bull times. Wealth is a stock while income is a flow: but when wealth market values increase regularly, *wealth becomes a flow*, either because it is perceived as income, or because the increase in value can be made liquid by borrowing.[1] We may call this the macroeconomics of leverage and, as we shall see below, it is not a wealth effect *à la* Pigou. When households see this flow rising, they believe it is income and increase their consumption: they do not really need to increase their savings, as their wealth continues to rise. Also, when enterprises see their wealth flow rising, they consider it as new revenue and will push both their investment and their current expenditures, because whatever liquidity constraint prevailed beforehand, it is now removed by the increase in their wealth. Banks and financial intermediaries do not need to worry about the confusion between income and wealth, because as effective demand rises, employment and GDP grow, the profitability of enterprises increases, and creditworthiness at large is not endangered. When employment grows, so will household incomes (over and above the transformed wealth) as well as their expenditures – again, their savings will not increase, because its propensity is stifled by the regular increase in the market value of their wealth. Furthermore, as soon as employment booms, home buying increases and, with it, home values, thus making everybody richer – including firms, whose propensity to invest is enhanced by the rise in the value of stocks and buildings. One might think that these circumstances are occasional, that the crisis is a one-time event, that therefore it does not entail a disruption in the established economic theory, based on automatic market correction. As we shall see, this is not so. As far as I know, the macroeconomics of leverage has not been dealt with sufficiently in economic theory (with the possible historic exceptions of Hilferding, Galbraith and, more recently, Minsky), mostly because in methodological individualism, including models with heterogeneity of agents, debt is never a macroeconomic problem, as assets and liabilities, for the economy as a whole, are always equal and wealth is always zero (in a closed economy or in the world at large). This is a truism, however, because what happens in bull times is that both assets and liabilities increase, but it is the increase in liabilities that determines the increase in assets (the market values for houses and securities), just as lending by banks determines deposits. This is, indeed, the macroeconomics of leverage.

11.3 Income distribution

In this gargantuan (i.e., milk and honey) picture, one fact stands out: the share of labour income in GDP worsens (inequality prevails). This outcome, which is considered as exogenous by the literature, is in fact unexpected, given the increase in employment and, possibly, a tighter labour market, at least following the crisis of the early 1990s. In fact, all incomes increase, but the upper deciles increase more then the lower ones: a typical Pareto improvement with worsening distribution. Some authors divide impatient agents from patient ones, with the first that need to spend more rapidly than the second, thus increasing their debt, which is financed by the second – this is, however, the ancient abstinence argument, which implies that credit and debt cancel out in the economy as a whole, leaving the growth of leverage unexplained. There are many real reasons for this long-term trend, and many derive from increased competition in the labour market: growing female participation, immigration, liberalisation in public services, flexibility in labour laws. In all these cases, as GDP and labour demand grow, more labour supply is made available, and lesser unit wages are paid (or slower growth in wages is experienced).

All the elements that produce a weakening of the bargaining power of workers were in fact present before the crisis, but nobody has explained why it was possible for such elements to continue reducing the bargaining power of workers, when full employment conditions are prevailing: *we miss a link between the bull market and labour weakness*. Galbraith proposed a simple way to explain this phenomenon before the Great Depression: speculation benefits those that own capital (the patient ones, some would say today), and during the boom distribution works against those that only earn a wage (the impatient ones). This is more a description than an explanation, because it does not account for the decline in bargaining power as the economy booms. In addition, if everyone owns some amount of capital, Galbraith's description does not square with the 2007 crash.

I propose, instead, that income distribution is linked to the transformation of wealth into income: for workers, the increase in *wage income is less needed, when the increase of wealth income takes place, and trade unions appear less powerful*. It is often forgotten that when bargaining for wages and work conditions, workers incur the risk of being fired. Individual bargaining is always risky for workers with a family, even in times of boom, because during a spell of unemployment they receive a fraction of their original income; in addition, new jobs are available, but they are not necessarily in the same location in which the family was living, with added costs. More importantly, when each worker bargains by himself, the power imbalance between him and his possible employer is strong, even in full employment, because workers are in competition with each other. Such situations should normally increase affiliation to the unions, but these organisations tend to be egalitarian in bargaining, in order to satisfy the greatest possible number of workers; as a result, the demand for affiliation becomes weaker precisely in conditions of full employment, when each worker values himself more than he values what the unions offer. It appears, therefore, that an increase in wealth that can be transformed into income is much more reassuring for the wage earner, and does not imply any confrontation with the employer. Unknowingly, enterprises *will face a lower level of conflict with their workers* for the distribution of productivity gains, which therefore do not accrue to workers in the amount needed for dynamic equilibrium. This seems a reasonable explanation of the increase in inequality, in full employment conditions. Perhaps, this phenomenon can even induce a cultural change, if workers perceive themselves as 'human capitalists' or, simply, 'rich'.

11.4 The crash

Our story can now resume. The less wages rise *vis-à-vis* the value of the workers' wealth (their home, in general), the more they must increase their debt so as to maintain their relative standard of living – but it is the maintenance of the relative distribution of income that ensures both social cohesion and sufficient effective demand. Conventional wisdom has it that whatever the generosity of the financial sector, there is always some limiting risk beyond which each household (and each firm) cannot hope to continue borrowing: an argument that can never identify precisely where the borrowing barrier stands. Rather, each historic boom shows a financial industry bent on increasing liquidity by diluting the risk of bad debts through the concoction of ever new securities reflecting some average risk, rather than the risk of any specific asset. With capital markets booming, this is a rational behaviour, because no debt is bad if the value of the corresponding asset continues to increase.

The existence of a boom barrier is possible, but it looks like no rational expectation can discover. The crash, then, must be attributed to more substantial causes. A preferable explanation has to do with the economics of leverage in a context of worsening income distribution. In a bull market, *wealth begets wealth*, until wealth market values rise: this, again, is what can be defined as leverage. Households, however, with wages that increase less than productivity and less than GDP, must maintain their relative consumption level (as well as their social status) via new debt, based on the increased value of their homes. If so, since they use part of their wealth for consumption, they can only use part of it as leverage to acquire new property: for them *wealth does not fully beget wealth*. Thus, when households can no longer buy in the housing (or in the capital) markets, housing prices and stock indices decline. As soon as the markets start dwindling, firms will also see their wealth decrease and this implies a reduction in leverage and in perceived revenues. As revenues of all kinds slow down or simply cease to increase, effective demand decreases or ceases to increase and the multiplier stops working or works negatively: an economic crisis, and not simply a financial one, sets in.

11.5 Economic policies and the crash

We have to figure out what causes the transformation of wealth into income. I think that what happened was the result of a new series of policies that, since the early 1980s, had disregarded the need to keep effective demand nearer potential output (a position similar to the pre-1929 policies), either because a self-equilibrating 'natural rate of growth' theory was used for policy purposes or, more simply, because it was believed that supply creates its own demand (Roncaglia, 2010). Simplifying, the following appear to be the tenets of the new policies:

- if a depression sets in, then labour becomes cheaper, firms' expected real incomes increase, and their demand for labour grows, restoring equilibrium. If this does not work, and will not work because the downturn in wages reduces the volume of sales, then

- liberalisation will: the transaction costs for the economy as a whole will decline, real incomes increase, and equilibrium is restored. If liberalisation does not work, and will not work because the decline in transaction costs is tantamount to a decline in employment, wages and sales, then
- labour market flexibility will: removing wage earners' rents, due to labour rigidities, will raise firms' expected incomes and profits, restoring equilibrium: a non-event, because realised profits will not increase since demand out of wages declines (or fails to rise).

All these instances are based on a well-known fallacy: if costs of production are reduced, then demand is assumed to pick up because prices should decrease and real incomes increase (this is Pigou's real balance effect); but since the cost reduction translates into lower expenditure on inputs, then whoever supplies the inputs will face lower sales and revenues. The microeconomic foundation of macroeconomics fails, and no fixed-price assumption is needed.

It is this failure to recognise the role of effective demand, and therefore the benefits of government intervention, that has prompted the alternative, through deregulation (liberalisation) policies in the financial markets:[2] if regulation on limiting the risks in the trade of securities is relaxed, banks and financial agents will provide risky paper, but will need to show that such paper can be used to leverage other paper (as shown above), producing real effects. *Speculation is the substitute for effective demand*. Moreover, deregulation is necessary, bur also weak regulation, asymmetric powers in regulation and conflicts of interest, gullible investors, simply laissez faire, will produce the toxic securities encountered in the recent crash – possibly as in 1929.

In fact, the most important deregulation that has taken place over the last thirty years is the independence of the central bank (again, a return to the pre-1929 period): a new policy occasioned by the monetarist revival of the late 1970s and early 1980s. Such independence reduced the sovereignty of governments in issuing fiduciary money: central banks were not obliged to buy government paper to cover public deficits and governments had to sell their bonds on the market. As the bond supply increased, crowding the market, interest rates rose, public deficits and debts grew and this made it more difficult for governments to sustain effective demand. At this time, the early 1980s, a serious crisis ensued, with a large increase in unemployment. Contrary to their basic tenets, governments found that their policies were wrong and that economic activity was to be sustained, but since their economic policies did not include acting on demand (or, if they did, they had to take account of their enlarged public debt), new deregulation was enacted that further reduced the role of the central bank and of the banking system in regulating money supply. Since then, nobody has really been in control of bank money, and banks lend on the basis of their own capital. Each bank is therefore pushed to create its own wealth with the help of securitisation: buying securities which can reduce the risk of underlying assets and issuing (forging, sometimes) securities bearing uncertainty and incalculable risks, banks increased their own leveraging power. The deregulation in banking spearheaded the deregulation of the financial markets. As we have seen, speculation does have real effects: the greater financial agility, by easing the issuing of purely 'fiduciary' securities (a form of endogenous money), produces wealth, leverage and effective demand – *at least until the worsening distribution of income causes a market crash.*

This story has a moral: reducing the role of the government, through deregulation, means that no economic policy aimed at preserving full employment is really feasible, with the exception of letting speculative markets grow; and the more speculative they are, the more they affect demand. This effect can only be maintained, however, if the distribution of income does not worsen: any gain in the capital and housing markets could be considered as an increase in overall productivity and, thus, could be distributed to all the workers in the economy, not just to those that operate in the financial or housing sectors. But since workers have in the meantime lost their bargaining power and the awareness of their real status (and no Wagner Act (USA, 1935) has been passed), such redistributive policy will not take place. If, as a reaction to the crash, new regulation will really discipline capital markets, then there is no possibility for speculation to raise effective demand, and for growth to resume. The economies are at or near a dead end: if they don't regulate, a crash is likely, if they do regulate, growth slows down.

11.6 Leverage and the consumption function

I hope that all this reasoning stands up. The inspiration is certainly Keynesian (not Neo-Keynesian), as witnessed by the emphasis on effective demand, on the central bank as the last resort for financing government deficits and on income distribution. The difference concerns speculation and wealth, because while Keynes thought that buying and selling existing securities has no real effects and wealth and income are separate, in my view speculation has real effects.

This is no small change to the Keynesian model: following a suggestion of Professor De Cecco, it implies that the consumption function is based not only on income, but also on leverage: and each time a bull market appears, leverage increases, consumption is influenced, and the bargaining power of wage earners declines. If these events could be applicable to the dynamic of the capitalist system, and not just to the recent boom and crash, then the rate of interest has no place in the consumption function, and therefore neither Modigliani's life cycle nor Friedman's permanent income hypotheses are relevant. Leverage is not dependent on the rate of interest, but on the deregulation of capital markets. To make this statement clearer, the financial market may show a pseudo-Pigou effect, if speculation is enhanced by lower interest rates which cause an increase in wealth, but one would need an ever-decreasing interest rate to produce ever-increasing wealth. Instead, given the rate of interest, it is the rise in stock market indices which generates first new wealth and new incomes and finally their demise: not just once, but whenever the government is kept from targeting effective demand.

Notes

1. I owe this clarification to an unknown referee.
2. It was President Clinton, a Democrat, who relaxed whatever rules were still extant in the financial sector and ran a budget surplus. As was once suggested by Gordon Brown, such policy can be defined as a mixture of Friedman and Keynes, and is one tenet of the 'third way' of the British Labour party; he forgot that it is like mixing oil and water.

References

Barba, A. and Pivetti, M. (2009) 'Rising household debt: its causes and macroeconomic implications – a long period analysis', *Cambridge Journal of Economics*, 33: 1.

Dew-Becker, I. and Gordon, R. J. (2005) 'Where did the Productivity Growth Go? Inflation Dynamic and the Distribution of Income', *Brookings Paper in Economic Activity* 2.

Iacoviello, M. (2008) 'Household debt and income inequality 1963–2003', *Journal of Money, Credit and Banking*, 40: 5.

Keys, B. J., Mukherjee, T., Seru, A. and Vig, V. (2010) 'Did securitization lead to lax screening? Evidence from subprime loans', *Quarterly Journal of Economics*, 125: 1.

Kumhof, M. and Rancière, R. (2010) 'Inequality, Leverage and Crises', IMF Working Paper, November.

Roncaglia, A. (2010) 'Le origini culturali della crisi', *Moneta e Credito*, 63(250): 107–18.

Contents

List of Contributors

Sergio Cesaratto, University of Siena

Roberto Ciccone, Roma Tre University

Óscar Dejuán, University of Castilla – La Mancha

Esther Dweck, IE, Federal University of Rio de Janeiro (UFRJ)

Fabio N. P. de Freitas, IE, Federal University of Rio de Janeiro (UFRJ)

Paolo Leon, Roma Tre University

Carlos Aguiar de Medeiros, IE, Federal University of Rio de Janeiro (UFRJ)

Antonella Palumbo, Roma Tre University

Fabio Petri, University of Siena

Franklin Serrano, IE, Federal University of Rio de Janeiro (UFRJ)

Matthew Smith, University of Sydney

Attilio Trezzini, Roma Tre University

Introduction

Antonella Palumbo

One of the fields of research to which the modern revival of the surplus approach has devoted much attention, in the fifty years following Sraffa's *Production of Commodities by Means of Commodities*, is the theory of output and the analysis of the processes of growth and development in capitalist economies. Although no explicit theorisation about the determinants of output levels is to be found in Sraffa's work, the revival of the classical approach to value and distribution and the critique of neoclassical theory prompted by his work, as shown in Garegnani's (1978–79) seminal contribution, have relevant implications for the analysis of output.

On the one hand, the critique of neoclassical mechanisms based on factor substitutability implies (i) the lack of the analytical basis on which traditional analysis affirms the automatic tendency of the system to the full utilisation of resources and (ii) a criticism of the reduction of Keynes's contribution to the analysis of the transitory incapability of the system of ensuring sufficient demand to absorb potential output. On the other hand, the fact that in the classical theory of value and distribution output levels are regarded as given when determining relative prices and the residual distributive variable (in the so-called 'core' of surplus analysis – Garegnani, 1984), implies a degree of freedom with respect to output determination and thus compatibility with different theories of output. In the words of Garegnani (1978, p. 340), 'Ricardo's theory of distribution is *open*, in the sense that it *neither* provides premises capable of justifying the tendency of investment to adjust to saving, *nor* depends on the existence of such a tendency. This "open" character sharply distinguishes the Ricardian theory from the subsequent marginalist theories, which saw distribution as the result of forces of demand for and supply of "factors of production".'

The eleven essays contained in this volume may be said to build on these original insights, by contributing new material to the general aim of explaining the determination of output levels and their evolution over time on the basis of two fundamental notions: the absence of any mechanisms whatsoever ensuring the tendency of the system to the optimal use of resources (and thus the necessity of studying the evolution of demand and the limits it may pose on the level and growth of output); and the fundamental independence between output determination and distribution, or, more accurately, the absence of any *necessary* relation between them.

This independence, which stems from the very structure of the classical theory of value and distribution, is a distinctive element of the long-period theory of output developed within the modern revival of the surplus approach, which sets it apart from other theories attempting the extension of the principle of effective demand to the long period on different bases.

One of the most influential of such attempts has been the 'Cambridge' theory of distribution (Kaldor, 1955–56; Robinson, 1956 and 1962) postulating a strict and necessary direct relation between the rate of accumulation and the rate of profit. Starting from the 'Keynesian hypothesis' (Kaldor, 1955–56, p. 195) according to which it is the level of investment, independently determined, which causes and generates the corresponding level of saving, the Cambridge approach interprets this causation as entailing that in the long period any autonomous increase in the rate of accumulation must necessarily be met by a corresponding increase in the ratio of saving to the stock of capital, which can only be obtained if the proportion of profits in national income grows, given the higher propensity to save out of profits. The Cambridge theory of distribution is thus based on the assumption that in the long period the overall level of output does not have sufficient elasticity in responding to autonomous changes in investment, and that for this reason the composition of output between consumption and saving must necessarily change to accommodate the higher rate of investment. This would happen through an increase of money prices relative to money wages (induced by the temporary excess demand) which will produce the necessary fall in real wages.

However, as shown by Garegnani (1992) and Garegnani and Palumbo (1998), such rigidity of long-run output to independent changes in demand – which in Kaldor (1955–56) takes the form of the full-employment hypothesis while in Robinson (1956, 1962) is represented by the assumption of the long-run rigidity of the output/capital ratio – is in reality unwarranted when taking seriously the conception of the growth process as driven by demand. An essential part of this conception is in fact the notion that productive resources, far from being exogenously determined, are mainly the result of the long-run level of activity itself. This happens through the processes of creation or destruction of productive capacity that a prolonged over-utilisation or under-utilisation of installed capacity entail, while other mechanisms (such as migration, changes in the participation rate, transfers between low-productivity and high-productivity sectors – see Kaldor, 1985) are at work allowing for a similar flexibility in the supply of labour force in response to changes in the demand for labour.

Such mechanisms typically operate over long periods, so that some discrepancies between available resources and their utilisation are always observable; at the same time, the fact that unused resources tend slowly to disappear from sight implies that there is no need to observe great such discrepancies in order to recognise the wide margins of elasticity with which output may accommodate changes in demand. The necessary link between accumulation and distribution that the Cambridge theory of distribution postulates thus disappears, and the necessary trade-off between investment and consumption which characterises that theory disappears as well. Not only does this allow for a separate determination of distribution on the basis of a complex of social and institutional forces, as entailed by the classical surplus analysis (see the essays in Volume I); it also implies that the fundamental conclusion of Keynes's analysis as regards the short period, namely that a *direct* relation of causation may be postulated between investment and consumption, is confirmed and extended to the long period. All the authors contributing to this volume share this common theoretical ground and interpret the long-run principle of effective demand as entailing that the variations in the level of output act as the adjusting mechanism between investment and saving, exactly as postulated in Keynes's short-run analysis.

Similarly to what happens in the determination of distribution, the ultimate determinants of growth are to be sought in a complex of socio-historical circumstances which defy any simple and mechanical analysis of causation. Growth cannot be considered in merely quantitative terms as the increase over time of an aggregate indifferentiated amount of product; it has rather to be regarded as 'a complex process, entailing structural change of the economic system, such that it can only be plausibly explained in concrete terms by reference to social, politico-institutional and technological factors' (as Smith maintains in his essay in this volume). This also entails that the clear-cut distinction traditionally to be found in economic literature between the theory of growth and the theory of development loses part of its meaning.

This theoretical attitude with respect to the analysis of growth is not only an expression of the above-mentioned separability and mutual independence between the determination of distribution and the determination of output, which allows a plurality of determining factors to be taken into account in each: it is also an expression of the particular method of analysis which, following the lead of classical economists, is adopted in the modern revival of the surplus approach when dealing with such questions. According to Garegnani's (1984) reconstruction (see the Introduction to Volume I), outside the purely deductive and analytical 'core' of their theories (devoted to the study of the general and necessary relations between relative prices and distribution), the classical economists used to address such questions as the determinants of distribution or of accumulation (the 'intermediate data' of the core) by means of an analysis made up of a mixture of deduction, observation and generalisation from experience; where no precise and mechanical relations are postulated between variables, but complex systems of influences and interrelations are considered (this also implies the possibility of considering mutual interrelations between accumulation and distribution, provided they are part of this complex system of influences and are not regarded as necessary and mechanical relations).

This very methodological characteristic implies that in the theory of output and in the study of the determinants of growth, similarly to what happens in the theory of distribution, the classical approach requires a method of analysis in which institutional and social factors play a relevant role, so that no abstract general theorisation is possible without reference to the historical conditions in which economies develop. In principle, this implies that no neat distinction can be traced, when analysing the determinants of distribution or those of output, between pure theory and applied analysis.

To different degrees, the essays contained in this volume (as well as in Part II of the first volume, 'The Revival and Development of the Classical Theory of Distribution') bear out this methodological attitude; and this is also the reason why no sharp line has been drawn, in conceiving this volume, between a theoretical and an applied part. Rather, the two parts into which the volume is divided deal respectively with 'Demand-Led Growth in the Classical Approach' (the first seven essays) and 'Understanding the International Economic Order' (the last four essays).

In the first essay of Part I Ciccone addresses a problem which currently has a relevant impact on collective choices and policy debates, i.e., the role and management of public debt. The aim of his chapter is to show how deeply the attitude towards public debt and its policy implications may change when regarded from the theoretical angle assumed in this volume.

According to the 'view of total output as governed by aggregate demand independently of the availability of resources', public debt cannot be seen as subtracting savings to private accumulation, so that no trade-off arises in principle between public and private expenditure. The chapter highlights the potential expansionary effect of public deficits, which act on output both directly and indirectly by inducing higher capacity-building investment, and critically reviews the conclusions of mainstream literature on fiscal multipliers. On these premises, Ciccone shows analytically the potential perverse effects of restrictive fiscal policies aimed at reducing the debt/GDP ratio by comparing the long-run effects of different policy regimes, and maintains that lower values of the debt ratio might be rather achieved, under certain conditions, through expansionary fiscal policies.

Petri's chapter is conceived as a critique of the neoclassical theory of investment. Apart from the decisive critiques that have been directed at it following the capital debates, which show the fallacious bases on which the decreasing demand for factors is obtained when heteregeneous capital is considered among them, Petri finds a second weakness in the neoclassical theory of investment, which would be present even conceding the working of the factor substitution mechanism, and has to do with the necessary dependence of investment on output.

He shows that factor substitution, once the durability of fixed capital is taken into account, necessarily operates only on the investment flows. This implies that factor substitution operates slowly and that during the very long adjustment process its effects interact with the effects of other forces (like the influence of investment on aggregate demand and the influence of demand on the level of production), producing in the end the possibility that the relationship between the rate of interest and investment is in some circumstances opposite to what theory prescribes. More accurately, the rate of interest proves insufficient to determine investment as long as the quantity of labour employed at each stage of the adjustment process is indeterminate. The analysis also implies that no mechanism can be immediately envisaged, even conceding factor substitution, that automatically increases the demand for labour in response to falling wages.

The following five chapters all deal, in a more or less direct way, with the controversial methodological issue of the analytical tool to be used in studying the long-run tendencies of the system, which has been the object of much debate within the approach.

Controversy has particularly centred around the notion of 'supermultiplier', originally developed by Hicks (1950), which has been proposed by Serrano (1995) as a useful representation of the long-period relationship between the level of output and autonomous demand that in his view can be postulated in the analysis of growth within the modern revival of the surplus approach (other supermultiplier models are to be found in Bortis, 1997 and Dejuán, 2005). The supermultiplier is determined on the hypothesis that capacity-building investment is to be regarded, in the long run, as an entirely induced variable, while other components of demand (especially current public expenditure and exports) may be regarded as completely autonomous. This kind of 'unproductive' (in the sense of being unable to modify productive capacity) or 'final' expenditure would be the driving force of growth, since its variations induce in the long period changes in output both through consumption and by inducing changes in investment aimed at adjusting capacity to demand.

The advocates of the supermultiplier model stress its capability of replicating the simple Keynesian formula of the multiplier in an equally simple relation valid for the long run, where the capacity effects of investment are considered; the possibility of restating the autonomous role of demand in driving the growth process, since variations in the growth rate of autonomous demand are accommodated by changes in the relative shares of autonomous expenditure and induced investment in total output; and the independence of distribution from accumulation entailed by the model. The critics (especially Trezzini, 1995, 1998; Palumbo and Trezzini, 2003) question the assumption of average normal utilisation of capacity, by showing that any change in the rate of growth of autonomous demand implies a process of adjustment whereby the degree of capacity utilisation proves different from normal, even on average; consequently they believe that the validity of the model is limited to the extremely artificial case of autonomous expenditure growing for an unlimited long period at a constant rate (Trezzini, 1998, p. 66), and regard such an assumption as in contrast with the assumed autonomy of demand. More generally, Palumbo and Trezzini (2003) show that, from a logical point of view, the variability (both in the short and the long run) of the degree of utilisation of productive capacity is the way in which the autonomy of demand manifests itself, by producing the very discrepancies between demand and capacity that imply the need for adjustment of the latter. Analysis of the actual way in which the process of adjustment operates shows its lengthy and uncertain nature, due to the fact that the process itself is bound to generate offsetting forces which may render the complete adjustment between production and capacity an unrealistic occurrence, even without considering the variability of capacity requirements induced by technical progress. While not challenging the tendency to a uniform rate of profit and the gravitation of market prices towards normal prices, which occur on the flow of newly produced capacity (on the issue of gravitation, see the Introduction to Volume III), the systematic incapability of the system of attaining 'fully adjusted positions' (the definition is due to Vianello, 1985) challenge their role as reference points in the analysis of the evolution of quantities.

The chapter by Trezzini contained in this volume addresses this methodological question from a general point of view, by investigating the double relationship, in the analysis of growth, between actual and theoretical magnitudes, on the one side, and cyclical and trend positions of the economy, on the other side. His main thesis is that the prevailing method of studying the trend of produced quantities 'through growth paths whose characteristics ... are defined regardless of the quantities actually produced' derives from an 'unjustified *transposition* of the relation between theoretical and actual magnitudes developed in the theory of prices into the theory of the evolution of quantities'. On the basis of the different nature of the forces at stake in the determination of produced quantities, he advocates a different method of analysis capable of taking into account the fact that the pattern of cyclical fluctuations may affect the economy's trend by inducing irreversible changes (such as the presence of new capacity or its destruction, or the acquisition of consumption standards) which may not be considered as merely accidental. Both the absence of any theoretical reason to consider the determinants of trend magnitudes as persistent and independent of cyclical phenomena, and the unrealism of the hypothesis that the components of autonomous demand should grow along a smooth regular path, imply that there is no solid foundation to the attempt at establishing a relationship between normal and actual quantities analogous to the relationship between normal and actual prices.

The main thrust of Palumbo's chapter is the critique of the notion of 'potential output' which is to be found in mainstream literature and the critical analysis of the methodologies used by international institutions to obtain empirical estimates of potential output. Estimation methods are shown to be deeply influenced by the dominant theoretical framework in which potential output is exogenously determined by supply factors; on the other hand, the actual content of empirical estimates is often heavily dependent on the time series of actual output, in the theoretical (unproven) belief that actual output cannot have deviated too markedly from potential, the latter being a strong attractor of the former. Some puzzling results of this literature, especially as regards the output–inflation relationship assumed by mainstream theory, are examined. Against the limitations of the supply-side approach, potential output is regarded within the demand-led growth framework as endogenous and dependent on the economy's actual realisations. It is in this context that Palumbo examines the debate about the analytical role of the fully adjusted positions within the demand-led growth approach and concludes that a meaningful definition and quantification of potential output can only be referred to the short period, the long-run evolution of potential being a path-dependent phenomenon.

In his contribution, Smith is also critical of the limitations of the assumption of normal utilisation in the demand-led growth approach. After having exposed the limitations of a rigid supermultiplier model, he proposes an alternative growth model with variable utilisation both in the short and the long run. Though in a simplified way, the model contains an attempt at historical periodisation of the relationship between autonomous demand and aggregate output. This is achieved by letting the value of the supermultplier change period by period (each historical time period characterised by specific relations between variables is called an 'epoch'; each epoch is assumed to be long enough for fixed productive capacity to fully adjust to expected demand). Unlike in 'steady-state' models, the growth rate depends in Smith's model not only on the growth rate of autonomous demand but also on the long-run change in the value of the supermultiplier which may be influenced by such phenomena as changes in income distribution and technical change. Average utilisation of capacity over the whole timespan in the end proves to be different from normal, although the model does not allow for changes in utilisation within each period. The simplicity of the model cannot, however, obscure, in Smith's opinion, the complex interactions that exist among the different variables influencing the growth process, but rather aims at representing the basic relation between autonomous demand and output as historically contingent.

Differently from the chapters just described, in the last two chapters of Part I the supermultiplier model, together with the assumption of normal utilisation of capacity, is considered as a correct and useful representation of the actual relationship between output and autonomous demand, and a fundamental element of the theory of output in the modern classical approach.

Dejuán's paper contains a restatement of the supermultiplier model with a number of qualifications and clarifications. He particularly discusses the issue of stability, by attributing the extreme instability of Harrod's (1939) model to the excessive influence of the actual rate of capacity utilisation on the expected rate of growth. In his view, 'models that incorporate the accelerator mechanism will be stable provided demand expectations depend on objective factors (like the autonomous trend) instead of the ups and downs of current income and capacity utilization'. In order for the autonomous trend to represent an objective factor to which expectations may be anchored, however, the same rate of growth of autonomous demand has to last long enough to be perceived as persistent by entrepreneurs: 'Certainly, in epochs without a well established autonomous trend, the economy will suffer from macroeconomic instability'. A second qualification is that the model must be formulated in terms of a flexible supermultiplier if the latter has to result in a stabilising mechanism. This is obtained by making investment depend not only on the expected growth of demand but also on the deviations from normal capacity.

In their paper Freitas and Dweck use the supermultiplier demand-led growth model to analyse the pattern of Brazilian economic growth from 1970 to 2005. The basic model is integrated with the idea that at least in some sub-periods the balance of payments was a binding constraint on Brazil's growth. However, the demand-led growth structure allows different possible growth regimes to be considered. In the case of Brazil, the authors identify both a balance of payments-constrained demand-led growth process, and a policy-constrained (or pure) demand-led growth process (in which the availability of foreign exchange is not the binding constraint on growth) as relevant for different sub-periods.

The authors also propose an exercise in 'demand-side' growth accounting. Their decomposition of the rate of growth aims both at assessing the role of domestic and external demand in the different phases of the growth process (with the conclusion that Brazilian growth was mainly led by the public sector contribution), and at showing the role in the different phases, respectively, of changes in autonomous demand and changes in the supermultiplier. The authors conclude that the latter has to be regarded as a changing magnitude, which implies that the identification of the historical factors capable of determining its changes must be an essential part of the analysis of growth determinants.

The four chapters in Part II of this volume address the changing features of the international economic order and use the theoretical insights provided by the approach as interpretative tools of observed phenomena.

Serrano focuses in particular on the behaviour of the commodities market and on the sustained increase of dollar commodity prices in the 2000s. He maintains that this trend cannot be explained by the great rise in Chinese demand, and that demand factors (whether final demand or demand for speculation) seem not to be strong enough to account for the intensity and the timing of the phenomenon. Nor do any explanations in terms of natural scarcity of the various commodities seem relevant. Rather, on the basis of the classical analysis of the cost of production as the main determinant of price, he draws attention to the various factors which have contributed to an increase in the relative cost of production of commodities *comprehensive of rents*: the deliberate restriction of the supply of oil, obtained through the convergent action of OPEC policies and 'the revival of natural resource nationalism', the fast growth of real wages in commodity-exporting countries, the real appreciation of their currencies with respect to the dollar, the stagnation in unit costs of manufactured exports due to slow growth of real wages in advanced capitalist countries and fast growth in industrial productivity in the emerging countries which are new exporters of industrial products. He also emphasises the historically contingent nature of the new international economic order of the 2000s, and stresses some elements of continuity with the preceding phase such as the continuing heavy dependence of developing countries on American monetary policy and American financial markets.

Medeiros's contribution aims at assessing the decline of the developmental state and the transformations that governments and public agencies are undergoing in emerging countries as actors in the process of development. The perspective adopted is 'based on classical political economy, on Latin American structuralism and on a Gramscian perspective on the state's formation'. On these bases, he emphasises the role of different economic and social structures in influencing institutions, the uneven structure of the international environment with the hegemonic state's policies playing a crucial role, the nature of the state as 'a central institution where the dominant class or some of its sectors lead a coalition of power, building a hegemonic project compatible with a particular accumulation strategy'.

He identifies three different phases in the post-war period, the first based on national capital coordinated by national states (state-led industrialisation); the second based on the internationalisation of productive and finance capital and the opening of developing countries' markets to foreign capital (the neo-liberal phase); the third characterised by the return to expansionary Keynesian policies in developing countries in a more internationalised production context.

Cesaratto's contribution is centred on the different views of international relations that it is possible to find in the history of economic thought. He classifies the various theories into two main categories, those believing 'in the harmonic virtues of international laissez-faire' and those implying a conflictual view of international relations. He especially deals with this second group of theories, in search of insights that may contribute to a deeper understanding of current international arrangements, and particularly the current relationship, within the euro area, between Germany and Southern European countries. He reviews the contributions of various strands of thought (Mercantilism, the German Historical School, International Political Economy), and maintains that the Sraffian approach (together with Kalecki's theory of exports) may provide the useful insights scattered in these approaches with the sound analytical base they seem to lack.

In analysing some aspects of the great crisis of 2007 and the following recession, Leon's chapter addresses the issue of recent trends in distribution which is also dealt with in other contributions (Serrano in this volume, and Stirati and Pivetti in Volume I). His main thesis is that the direction of causation that some authors (among whom Pivetti) envisage between deteriorating distribution and increase in household debt should be reversed, with the latter potentially being a relevant force in determining the former. This is based on the idea that the continuous increase in wealth market values has produced an easy transformation of wealth into income, thus relaxing the liquidity constraint, and weakening the bargaining power of trade unions by rendering wage income less necessary for workers. This would explain, in Leon's view, why the lowering of real wages happened in a context of very high employment, supposedly entailing a relatively strong wage-bargaining position for workers.

References

Bortis, H. (1997) *Institutions, Behaviour and Economic Theory* (Cambridge: Cambridge University Press).

Dejuàn, Ó. (2005) 'Paths of accumulation and growth: towards a Keynesian long-period theory of output', *Review of Political Economy*, 17: 231–52.

Garegnani, P. (1978) 'Notes on consumption, investment and effective demand: I', *Cambridge Journal of Economics*, 2: 335–53.

Garegnani, P. (1979) 'Notes on consumption, investment and effective demand: II', *Cambridge Journal of Economics*, 3: 63–82.
Garegnani, P. (1984) 'Value and Distribution in the Classical Economists and Marx', *Oxford Economic Papers*, 36: 291–325.
Garegnani, P. (1992) 'Some Notes for an Analysis of Accumulation', in J. L. Halevi, D. Laibman, and E.J. Nell (eds), *Beyond the Steady State. A Revival of Growth Theory* (London: Macmillan).
Garegnani, P. and Palumbo, A. (1998) 'Accumulation of Capital', in H. D. Kurz and N. Salvadori (eds), *The Elgar Companion to Classical Economics* (Cheltenham: Edward Elgar).
Harrod, R., (1939) 'An essay on dynamic theory', *Economic Journal*, 49: 14–33.
Hicks, J. R. (1950) *A Contribution to the Theory of the Trade Cycle* (Oxford: Clarendon).
Kaldor, N. (1955–56) 'Alternative theories of distribution', *Review of Economic Studies*, 23: 63–100.
Kaldor, N. (1985) *Economics without Equilibrium* (Armonk, New York: Sharpe).
Palumbo, A. and Trezzini, A. (2003) 'Growth without normal capacity utilization', *European Journal of the History of Economic Thought*, 10: 109–35.
Robinson, J. (1956) *The Accumulation of Capital* (London: Macmillan).
Robinson, J. (1962) *Essays in the Theory of Economic Growth* (London: Macmillan).
Serrano, F. (1995) 'Long period effective demand and the Sraffian supermultiplier', *Contributions to Political Economy*, 14: 67–90.
Trezzini, A. (1995) 'Capacity utilisation in the long run and the autonomous components of aggregate demand', *Contributions to Political Economy*, 17: 33–66.
Trezzini, A. (1998) 'Capacity utilisation in the long run: some further considerations', *Contributions to Political Economy*, 17: 53–67.
Vianello, F. (1985) 'The pace of accumulation', *Political Economy*, 1: 69–87.

Part I
Demand-Led Growth in the Classical Approach

1
Public Debt and Aggregate Demand: Some Unconventional Analytics

Roberto Ciccone

1.1 Introduction

This chapter explores some basic questions about the effects on the economic system of financing public expenditure by issuing debt. It develops within a theoretical framework differing from that which is currently predominant both in pure theory and in applications to specific problems such as those addressed here. In particular, the approach in this chapter rests on two basic and closely related premises.

The first is the 'classical' explanation of income distribution in terms of circumstances of a social and institutional character, and hence the rejection of the view of distribution as determined by the forces of demand and supply held by the dominant theory in its various formulations. The second is the application of the Keynesian principle of effective demand to the long-run or trend levels of total output, which are regarded here as dependent on the size of the aggregate demand ultimately determined independently of output in conditions of full employment (or 'natural unemployment'). Aggregate demand therefore sets an upper limit on the levels of activity obtainable in the long run (and, needless to say, in the short run too), which would generally be insufficient to allow for the full employment of available resources.[1]

The close relationship between the two premises stems from the natural compatibility of the classical theory of distribution with a determination of output levels into which the available quantity of labour and other resources do not enter as directly relevant circumstances. The analytical structure of classical theory, characterised by the absence of relations of a *necessary* character between the determination of distribution and the determination of output levels, makes it possible to study changes in variables such as the aggregates of output, consumption, investment, public expenditure and so on with no need to consider the effects that those changes may have on distribution and the price system. The analysis will therefore be carried out on the assumption of a given distribution of income and associated price configuration.

Within the theoretical framework outlined above, the work seeks to single out some basic relations and propositions concerning the effects of public deficits and public debt that might, if valid, serve as a foundation to address the issue under the conditions possibly set by specific contexts.

1.2 Public debt as additional private wealth

Let us start by defining some relations between the relevant magnitudes. In a closed economy the necessary equality between aggregate expenditure and aggregate output, i.e., between total savings and total investment, entails the identity

$$S_p = I + D$$

where S_p is private savings, I private investment and D public deficit.

In our theoretical framework, the level of income is determined by the level of aggregate demand, i.e., by the sum of private and public expenditure, which implies that that identity is satisfied, through changes in the level of income, by the adjustment of the flow of private savings to the sum of private investment and public deficit. An increase in public deficit ΔD caused by a given increase in public expenditure is therefore counterbalanced by an equal increase in the flow of private savings ΔS_p. In other words, through the influence of public expenditure on aggregate demand and hence on the level of income, an increase in public deficit generates *additional* private savings, savings that would have not been formed in the absence of the increase in the deficit.

The flow of private savings is in turn equivalent to the addition to the stock of private sector wealth in the period, gross of capital depreciation:

$$\Delta W_p = S_p = I + D$$

where ΔW_P stands for the gross variation in private wealth. Our point is that public deficits add to private investment by generating savings, and therefore wealth, in the hands of the private sector. As a result, the size of private wealth at any given time would be larger, the larger the sum of the public deficits built up in the past.

It should be evident how this conclusion relates to the view of total output as governed by aggregate demand independently of the availability of resources. In neoclassical theory, where the level of income is determined (apart from temporary deviations) by the supply of resources with consequent adjustment of aggregate demand, the size of private wealth is instead basically independent of the flows of public deficits. In that theoretical framework, the condition of full employment, or 'natural' unemployment, entails a corresponding level of income with the associated amount of savings. Private savings and the formation of private wealth are therefore set independently of the sum of public deficits and private investment, with a trade-off arising between the latter two 'uses' of private savings – unless it is maintained, as in the 'Ricardian equivalence' argument, that public deficits cause an increase in the share of savings out of full-employment income, thus displacing private consumption rather than private investment.[2]

We can say therefore that within the theoretical framework adopted here, no constraint is imposed on the size of public debt by the size of private wealth, since the latter increases to the same extent as the increase in public debt. Nor is any particular significance to be attached to a variation in the proportion of private wealth constituted by government debt, as an increase in that proportion, for example, would be no more than the arithmetical consequence of the fact that the stock of public bonds held by private agents has grown proportionally more than other types of assets constituting private wealth, such as real capital, and should not be mistaken for a negative influence on capital accumulation.

While public expenditure and deficits may play a role in raising the level of aggregate demand and income, no negative consequences need therefore arise for the economic system from the related accumulation of public debt as such. It is, however, precisely the capacity of public expenditure to produce relevant and lasting effects on aggregate demand that is questioned in much of the current literature, often by claiming that, at least in the long run, the values of 'fiscal multipliers' are quite low, if not indeed zero or negative, due to the depressing effect of fiscal expansion on some components of private expenditure.[3] Examination of the arguments asserting this allegedly negative influence shows, however, that they are based either on relations peculiar to neoclassical theory and unwarranted in the analytical framework adopted here or on quite arbitrary premises or procedures.

The most traditional case of displacement of private expenditure is obviously the 'crowding out' of private investment due to the higher interest rate caused by an increased public deficit. To start with, the dependence of the level of investment on the interest rate is a typical neoclassical concept and one that finds no analytical foundation, at least in the guise of a relation of general character (as implied by the functional form it is usually given), in our theoretical context. Moreover, the debates on capital theory have shown that the idea of an inverse relation between the interest rate and investment is questionable even within the boundaries of neoclassical theory. In turn, the supposed rise in the interest rate is itself part of the same conceptual picture, since it would be generated by the 'competition' between public deficit and private investment as alternative uses for scarce private savings, the amount of the latter being established by the full-employment level of income. This scarcity has instead no place in our framework, where private savings adjust to the sum of public deficit and private investment through changes in the level of income, which implies, as pointed out above, analogous adjustment between the corresponding magnitudes of stock. A different and theoretically more 'neutral' reason is sometimes put forward to explain the rise in the interest rate following an increase in public deficit (or even in tax-financed public expenditure), namely that the expansionary effect of public spending feeds the demand for transaction money and thus pushes the interest rate up if the supply of money remains unchanged. While the analytical premises for the result are not strictly neoclassical in this case, and can indeed be found in a purely Keynesian setting, it is the very notion of the supply of money remaining unchanged in the presence of an increasing demand for money that is wholly arbitrary. The particular stock of money in existence is in no sense 'natural' and rather reflects the targets of the monetary authority, among which the level of the interest rate is certainly paramount. The assumption that the central bank would not adjust the money supply and thus let the interest rate rise therefore amounts to the assumption of an unjustified switch to a more restrictive monetary policy. It follows that a reasonable *ceteris paribus* condition would rather be to take the central bank's targets as given and hence assume an adjustment of the supply of money designed to keep the interest rate at the same level.

Another argument in support of the depressive influence of public deficit on the expenditure of private agents concerns the latter's consumption. The grounds for maintaining that private consumption would be reduced by an increase in public deficit lie in the so-called 'intertemporal budget constraint' imposed on the government, whereby the present value of future debt of the public sector (as well as the future debt of any agent) must be zero. If this constraint is to be met, each present flow of public deficit must have as a counterpart an equal present value of future budget surpluses and hence of tax increases, if we refer for simplicity to unchanged future levels of expenditure. On the assumption that households are forward-looking, the expected future change in the fiscal regime would induce them to increase savings to the same extent as the present value of the additional future taxes, coinciding with the present increase in public deficit. It is thus claimed that due to this behaviour on the part of households, the increase in public deficit is offset by an equal fall in private consumption with no net effect on aggregate demand and the level of income.

Closer examination reveals, however, that the irrelevance of aggregate demand with respect to the determination of income is actually a *presupposition* rather than a consequence of the behaviour of households outlined above. In order to have a rational foundation, this behaviour must be based on the 'Ricardian equivalence',[4] the substance of which can be seen as consisting of two elements: a) households perceive that wealth in the form of public bonds is bound to be cancelled out by the additional future taxes imposed by the government's intertemporal budget constraint; b) they wish to leave a given amount of wealth to their heirs. As a result of a), households are prompted to increase their savings (reducing consumption) to the same extent as they subscribe new public debt, so that the latter is bequeathed *in addition to*, and not in the place of, the (real) wealth they wish to hand down to the next generation. As should already be evident, the need for households to reduce consumption in order to increase savings reveals that the level of income is conceived as independent of aggregate demand from the very outset. The role played by the reduction of private consumption is in fact to prevent the public deficit from crowding out private investment and therefore the real wealth to be bequeathed. If the level of income is instead seen as depending on aggregate demand, an increase in deficit spending would generate additional income and private savings, and the invariance of private investment would require no reduction in private consumption, which would, on the contrary, be free to increase together with the increase in income. Therefore, even with regard to the argument just discussed, the ineffectiveness of fiscal policy proves to be embedded in its very premises.

In conclusion, there appears to be no general reason within the theoretical framework adopted here to presume that the influence of public deficit on aggregate demand is offset by corresponding contractions of private expenditure, either for investment or consumption. On the contrary, recognition that deficit spending helps to keep aggregate demand, and therefore the level of income, higher than it would otherwise be makes it possible to attribute it with a positive influence not only on private consumption but also on investment levels, especially in a long-run perspective. Once demand is seen as the factor governing levels of activity also in the long run, it is reasonable to suppose that at least the 'dimensional' share of investment, namely the part connected with the size of productive capacity, is affected by the prospective levels of demand and therefore by the trajectories of fiscal policy (and economic policy in general) involved in determining them.

1.3 Debt–output ratio and public expenditure reductions: the case of a single period

If, as is usual, the ratio of public debt to domestic product is taken as a 'measure' (albeit not necessarily a meaningful measure) of debt, the influence of deficit spending on both terms of the ratio renders its relation with that measure more complex than is commonly acknowledged.

Let us consider alternative fiscal policies, characterised by different levels of public expenditure, while assuming unchanged tax rates. Let the index A denote magnitudes that would obtain under the current policy and indexes R and E magnitudes that would obtain respectively under a restrictive policy (with lower flows of public expenditure) and an expansionary policy (with higher flows). We assume a closed economy in which public expenditure and private investment do not depend in any predetermined way on the level of output and private consumption is a constant proportion of the private sector's current disposable income. Moreover, the share of the latter consisting of interest on public bonds is entirely saved, a simplifying assumption that does not affect our results qualitatively because it in fact cuts off one of the channels through which public deficits can feed aggregate demand and therefore acts *against* our argument.

Apart from its theoretical premises, a further difference of our argument with respect to conventional treatment of the subject is that the effects of a change in fiscal policy will be studied by comparing the alternative values that the ratio of debt to domestic product would take under one policy or the other in the same period or series of periods. This way of reasoning is at variance with the more usual procedure whereby the debt ratio resulting, or expected to result, over a period of time from a policy change is compared with the value of the ratio calculated at some date prior to the policy change. As we shall see later (Section 1.7), the reason for adopting this approach is that policy changes apply to some previously planned policy and the counterparts of the results obtainable by the switch are therefore the effects that the older policy would continue to produce if still in force rather than the past values of relevant variables.

If we assume for the moment that private investment flows are not affected by fiscal policy and can be therefore taken as given in the face of changes in the latter, the level of income at time t under the currently planned fiscal policy would be

$$Y_t^A = m(I_t + G_t^A)$$

where I and G stand respectively for private investment and public expenditure, and m is the income multiplier. Given the assumptions stated above,

$$m = \frac{1}{1 - c(1 - z)}$$

where c is the (marginal and average) share of consumption out of private disposable income and z the (marginal and average) tax rate.[5] The multiplier is therefore assumed to produce all its effects on income within the same period in which any given level of 'autonomous' expenditure, either private or public, takes place. The admitted arbitrariness of this assumption loses much of its relevance when the analysis extends to several periods, as it does here, and its global results do not therefore depend on how the direct and indirect effects of individual flows of expenditure are actually spread out over time.

With a restrictive policy entailing lower public expenditure $G_t^R < G_t^A$, the level of income at time t would instead be

$$Y_t^R = m(I_t + G_t^R)$$

with a difference in the level of income equal to

$$\Delta Y_t^R = Y_t^A - Y_t^R = m(G_t^A - G_t^R) = m\Delta G_t^R$$

On the other hand, the lower public expenditure reduces the deficit at time t with respect to that which would result with policy A:

$$\Delta D_t^R = D_t^A - D_t^R = (1 - mz)\Delta G_t^R$$

As expressed by the term $(1 - mz)$, the cut in expenditure reduces the deficit less than its amount, as each unit of lower spending also reduces public revenues by its multiplied effect m on the level of income times the tax rate z.[6]

If we assume for simplicity that interest on public debt starts to be paid one period after issue, the stock of debt at t with policy R is lower than with policy A by the same amount as the reduction in the deficit:

$$\Delta B_t^R = B_t^A - B_t^R = (1 - mz)\Delta G_t^R$$

where B stands for the stock of debt.

Using the previous relations, we can now express the ratio of public debt to domestic product that would obtain at t with policy R:

$$\frac{B_t^R}{Y_t^R} = \frac{B_t^A - (1 - mz)\Delta G_t^R}{Y_t^A - m\Delta G_t^R}$$

The purpose of the adoption of the restrictive policy is to lower the ratio of debt to domestic product below the level that would be obtained with the actual policy, i.e., to obtain

$$\frac{B_t^R}{Y_t^R} < \frac{B_t^A}{Y_t^A}$$

The latter condition evidently requires the reduction in the domestic product to be proportionally smaller than the reduction in the debt, i.e.,

$$\frac{m\Delta G_t^R}{Y_t^A} < \frac{(1 - mz)\Delta G_t^R}{B_t^A}$$

which, after substituting for m and simple algebraic manipulation, reduces to

$$\frac{B_t^A}{Y_t^A} < (1 - c)(1 - z)$$

The quantity $(1 - c)(1 - z)$ thus sets the upper limit of the ratio of debt to domestic product within which a reduction in public expenditure in a given period would be effective in reducing the debt ratio below the level that would obtain in the same period under the planned spending policy. For values of the ratio above that limit, a lower level of public expenditure would result in a higher ratio, whereas a higher level would result in a *lower* ratio. The idea behind this result is quite simple: the larger the size of debt with respect to domestic product, the smaller the effect that a given reduction of public expenditure would produce on the former in relation to the latter. Beyond a certain threshold, the negative effect of less public spending on domestic product would be proportionally larger than on the stock of debt and the ratio would consequently prove higher than it would have been with an unchanged public deficit (and vice versa). In other words, the magnitude $(1 - c)(1 - z)$, which represents the marginal (and average) 'propensity to save' of the private sector in relation to domestic product in our analysis, also provides the ratio between the difference in the stock of private wealth and the difference in domestic product, both differences resulting from the implementation of policy *R* instead of policy *A* in period *t*:

$$(1 - c)(1 - z) = \frac{\Delta W_{pt}^R}{\Delta Y_t^R}$$

where ΔW_{pt}^R represents the amount by which private wealth at time *t* would be less under policy *R* than it would under the current policy *A*.[7] The lesser formation of private wealth would, however, be precisely the counterpart of the lower accumulation of public debt engendered by policy *R*, i.e. $\Delta W_{pt}^R = \Delta B_t^R$, and the last equality entails

$$(1 - c)(1 - z) = \frac{\Delta B_t^R}{\Delta Y_t^R}$$

The marginal rate of private savings is therefore equal to the marginal difference in the ratio of debt to domestic product generated by policy *R* as against policy *A*. It is obvious that the debt ratio obtainable with policy *R* will prove lower than that obtainable with policy *A* only if the latter is lower than the marginal difference generated by policy *R*, i.e., only if

$$\frac{B_t^A}{Y_t^A} < \frac{\Delta B_t^R}{\Delta Y_t^R}$$

which explains why the magnitude $(1 - c)(1 - z)$ functions as the upper limit of the *A* debt ratio for the restrictive policy to be effective in reducing the debt ratio.

The quantity $(1 - c)(1 - z)$ can thus be taken as the threshold for the current debt ratio, above which the effects of changes in the level of public expenditure on the debt ratio would be opposite to those commonly expected. It should be noted that, making allowances for our simplifying assumptions, the values that the threshold can plausibly take are quite small. For $c=0.70$ and $z=0.30$, for example, its magnitude would be as small as 0.21, which is much lower than the ratios of public debt to domestic product generally observable in real economies.

1.4 Expenditure reductions over a number of periods

So far we have considered the question of the effects of a change in the level of public spending that takes place in a single period. As relations in the longer run are obviously what matter most, we shall now go on to study the effects of variations in public expenditure extending over a series of periods.

As regards the absolute amount of debt, a series of lower flows of deficit spending obviously reduces the stock of debt below the level it would otherwise have found. This effect, which increases with the length of the series of reductions and the scale of their amounts, is enhanced by the increasing reduction in interest outlays caused by the cumulative effect on the stock of debt. We shall now consider a policy *R* that cuts the levels of public expenditure from time 1 to time *t* with respect to the previously planned policy *A* and express the stock of public debt at the final time *t* in relation to the debt which would have obtained at *t* with policy *A*. If we assume a zero interest rate for the moment, the relation between the two stocks of debt at time *t* would be:

$$B_t^R = B_t^A - (1 - mz)\sum_{s=1}^{s=t} \Delta G_s^R$$

where $\Delta G_s^R = G_s^A - G_s^R$ is the reduction of public spending in period *s* which policy *R* entails in comparison with policy *A*.

If we now introduce a positive interest rate *i* on public debt, each reduction of debt in comparison with policy *A* generates further reductions as time goes on through the lower flows of outlay on compound interest entailed.[8] In greater detail, one unit of lower debt at time *s* generates a lower debt at time *t* by an amount of $(1+i)^{t-s}$, and the relation between the debt stocks therefore becomes

$$B_t^R = B_t^A - (1 - mz)\sum_{s=1}^{s=t} \Delta G_s^R (1+i)^{t-s}$$

In order to identify the condition that must be fulfilled in order for the restrictive policy R to lower the ratio of public debt to domestic product at time t, let us now express the ratio as

$$\frac{B_t^R}{Y_t^R} = \frac{B_t^A - (1-mz)\sum_{s=1}^{s=t} \Delta G_s^R (1+i)^{t-s}}{Y_t^A - m\Delta G_t^R}$$

and impose the condition that it must prove smaller than the debt ratio with policy A:

$$\frac{B_t^A - (1-mz)\sum_{s=1}^{s=t} \Delta G_s^R (1+i)^{t-s}}{Y_t^A - m\Delta G_t^R} < \frac{B_t^A}{Y_t^A}$$

This condition requires the reduction in the domestic product to be proportionally smaller than the reduction in the stock of debt:

$$\frac{m\Delta G_t^R}{Y_t^A} < \frac{(1-mz)\sum_{s=1}^{s=t} \Delta G_s^R (1+i)^{t-s}}{B_t^A}$$

from which, by rearranging and expressing the value of m, we obtain:

$$\frac{B_t^A}{Y_t^A} \frac{\Delta G_t^R}{\sum_{s=1}^{s=t} \Delta G_s^R (1+i)^{t-s}} < (1-c)(1-z)$$

We thus find once again a condition containing the 'inversion threshold' of the debt ratio in relation to changes in the level of public expenditure. In comparison with the result obtained by considering a single period, the magnitude that must remain below the threshold if policy R is to bring about a lower debt ratio is still the value that the debt ratio would assume at t with policy A, this time multiplied by a quantity evidently lower than 1. (It should be remembered that the reductions ΔG_s^R are regarded as *positive* values.) As might be expected, the condition to be fulfilled is therefore less restrictive than in the case of a single period as a result of the cumulative effects of the series of reductions in public spending.

The relation we have just obtained gains in transparency if the magnitude that expresses the effects of the succession of reductions in expenditure is transferred to the right-hand side, thus reinstating this side in the position of an 'inversion threshold' for the current – policy debt ratio as in the single-period case:

$$\frac{B_t^A}{Y_t^A} < (1-c)(1-z)\frac{\sum_{s=1}^{s=t} \Delta G_s^R (1+i)^{t-s}}{\Delta G_t^R} \qquad [1.1]$$

The quantity on the right-hand side of **[1.1]** can be seen as the value of the inversion threshold of the debt ratio at time t 'corrected' by the cumulative effects of the lower flows of public spending. The 'correction' operates in such a way as to make the threshold variable in relation to the series of reductions in expenditure with respect to policy A, increasing its value, from the second period of the sequence, above the value $(1-c)(1-z)$ initially assumed by the threshold. In actual fact, the fraction on the right-hand side generally increases with the number of periods and therefore the series of expenditure reductions, even though it might show some decreasing stretches after a sufficient number of periods and depending on the actual values of ΔG_s^R, but in any case remaining above the initial value $(1-c)(1-z)$.[9] Figure 1.1 shows the behaviour of the threshold, which we shall indicate as $T(t)$, leaving aside any possibility of a drop and therefore assuming its monotonicity.

Let us then suppose that over time policy A would produce the debt ratio $\frac{B_t^A}{Y_t^A}$ shown in Figure 1.2, starting from a value higher than the initial level of $T(t)$ and then proceeding along a rising trend, which is assumed to be linear for simplicity. We shall also assume that the spending reductions entailed by policy R are such as to make the threshold $T(t)$ increase more rapidly than the debt ratio resulting from policy A, so that the threshold cuts the trend of the debt ratio after a certain number of periods and the latter finds its way below the threshold. The intersection point in τ therefore marks a divide between the two regions in which the values of the R debt ratio would stand: the northwest region H, where the R debt ratio would exceed the A ratio, and the southeast region L, where the R ratio would be lower than the A ratio.

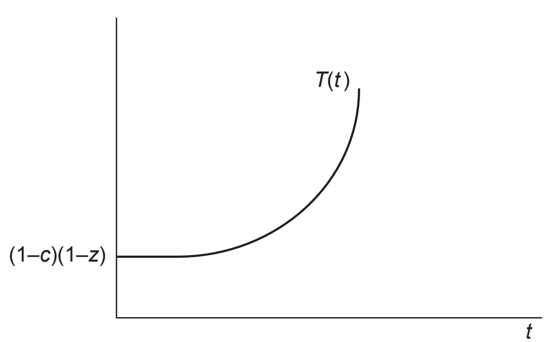


Figure 1.1 The threshold

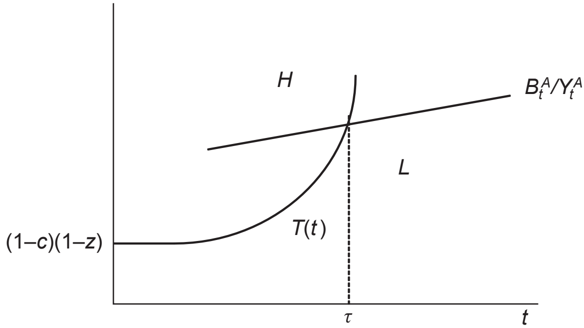


Figure 1.2 Threshold and present-policy debt ratio

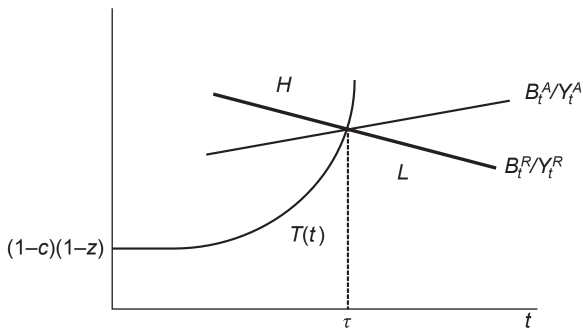


Figure 1.3 Effects of a restrictive policy (investment given)

The specific trend of the *R* debt ratio would of course depend both on the values of the spending flows entailed by policy *R*, which are those defined by the assumed reductions from the levels planned by policy *A*, and on the assumed trend of the *A* debt ratio. In order to simplify the graphic exposition, let us represent also the trend of the *R* debt ratio in linear form and assume that it is downward sloping, as shown by the thick line in Figure 1.3. Neither the decreasing course of the curve nor its monotonic behaviour is therefore a necessary characteristic of the trend of the *R* ratio, not least because of the influence of public spending on aggregate demand and output. The conclusions here are, however, not strictly dependent on the shape of the *R* ratio curve, and the assumption that the slope of the latter is negative is no more special than any other.[10]

We can thus see that as long as the debt ratio associated with policy A stands above the threshold, the debt ratio produced by policy R would in turn stay *above* the A ratio, and up to that point the effects of the restrictive policy could therefore be said to be 'perverse', as they are opposite to those expected.

On the other hand, the above exercise shows that after a number of periods the debt ratio produced by policy R would fall below what would be generated by policy A. This result might be taken to mean that the restrictive policy is ultimately effective in bringing about a debt ratio lower than otherwise, the delay – and indeed the opposite outcome – that the process goes through before reaching the target being conceivably not so relevant if the main concern is the long-run behaviour of magnitudes. Furthermore, the reductions in the levels of public spending seem to be capable, other things remaining equal, of making the $T(t)$ curve quite steep, and therefore rapidly increasing.[11]

Appealing to the temporary nature of the 'perverse' effects of a restrictive policy in order to diminish their importance, a reaction often observed when such consequences take place in actual economies, implicitly rests, however, on a belief in an inherent tendency of aggregate output to move towards its potential level and hence denies the autonomous role of demand in determining output. In the above analysis, the obstacle that prevents the lower levels of public spending from reducing the debt ratio, even to the point of raising it, is in fact the negative effect they have on levels of output. This depressing influence is permanent and the comparative diminution that the ratio eventually undergoes is obtained *despite* such influence, rather than because of its disappearance. As a matter of fact, the reduction of the debt ratio with respect to what it would be under policy A is obtained through a comparative reduction in *both* the stock of debt and domestic product, with the former decreasing proportionally more than the latter after a number of periods due to the cumulative action of spending cuts and lower interest outlays. The 'perverse' effects on the debt ratio are therefore the manifestation of a change in the trend levels of output caused by the restrictive change in fiscal policy and should be kept distinct from any sort of short-run consequences due to temporary circumstances or agents' reactions to that policy. This is reinforced by the consideration that, as will be argued shortly, it is precisely the persistence of the depressive effects of the restrictive policy that is liable in the long run to cause further changes enhancing those effects and consequently extending the number of periods for which the policy could generate a comparatively *higher* debt ratio.

1.5 Further on public expenditure reductions: negative feedbacks on private investment

The effects of given reductions in public spending have been examined so far on the assumption that the tax rate, the propensity to consume of the private sector and the interest rate on public bonds all remain constant. It has also been assumed that the other component of 'autonomous expenditure', namely private investment, does not alter the series of its magnitudes in relation to the adoption of a different fiscal policy and that therefore the effects on output levels of sequences of different levels of public spending are limited to those produced by the latter differences alone (the action of the income multiplier being of course taken into account).

The last of these assumptions appears to be the most restrictive. As already pointed out, there is no need in the theoretical framework adopted here for the interest rate to depend either on the size of the public debt, however it may be measured, or on measures of fiscal policy or more generally on changes in the levels of activity, and it is therefore legitimate to take it as given in the event of changes in those circumstances. The same holds all the more for the tax rate, which is completely under government control. As regards the private propensity to consume, while there could be reasons in support of its dependence on long-run levels of income, taking this dependence explicitly into account would only make the value of the multiplier variable without altering the principle of which the multiplier is the quantitative expression, namely that the level of aggregate output is governed by demand. The general character of the analysis undertaken here does not require us to go beyond the principle as such, and treating the propensity to consume as given constitutes a simplifying device that imposes no substantial restrictions on the results obtained.

Different considerations instead apply to the assumption that the levels of output generated by a given fiscal policy do not affect the investment plans of the private sector. Given our key principle that output is limited, in the long run no less than in cycle fluctuations, by the level of demand, it follows as a natural corollary that demand also influences the level of private investment. At least as regards the share of gross investment that could be described as 'dimensional', i.e., related to the desired scale of productive capacity, strict dependence is quite obvious on the levels of output that firms expect to be profitably attained or maintained in the periods ahead and hence on future levels of demand. It is quite reasonable, on the other hand, to presume that the trends of expected demand are influenced by the levels of demand that have prevailed in the past periods, so that lower levels of public spending will negatively affect private investment through their depressing impact on aggregate demand. Moreover, private investment may be influenced even by the *planned* levels of public spending if, as is conceivable, firms take government planning into account in forming their views about future market conditions and therefore output trends.

The conclusion that private investment is negatively affected by lower levels of public expenditure has important consequences for our analysis. Let ΔI_s^R indicate the reduction (taken as algebraically positive) in gross private investment that would be obtained in period s under the restrictive fiscal policy R compared with what it would be in the same period under policy A. Due to the influence of the lower levels of investment on levels of income and therefore on public revenues, the absolute amount of public debt that would result at time t from the application of policy R since time 1 is now larger. As above, this can be expressed in relation to what it would have been under policy A:

$$B_t^R = B_t^A - (1-mz)\sum_{s=1}^{s=t} \Delta G_s^R (1+i)^{t-s} + mz\sum_{s=1}^{s=t} \Delta I_s^R (1+i)^{t-s}$$

where the last term on the right-hand side represents the effect on the final stock of debt of the lower revenues (with consequent higher outlay on interest) caused by the lower flows of private investment expenditure. Correspondingly, the level of income at time t falls by comparison with what it would be under policy A because of the lower levels of public spending and private investment in the period:

$$Y_t^R = Y_t^A - m(\Delta G_t^R + \Delta I_t^R)$$

For the debt ratio at time t under policy R to be lower than it would be under policy A, the decrease in the level of income must be less than proportional to the decrease in the stock of debt, that is

$$\frac{Y_t^A - Y_t^R}{Y_t^A} < \frac{B_t^A - B_t^R}{B_t^A}$$

and, therefore, using the above relations:

$$\frac{m(\Delta G_t^R + \Delta I_t^R)}{Y_t^A} < \frac{(1-mz)\sum_{s=1}^{s=t} \Delta G_s^R (1+i)^{t-s} + mz\sum_{s=1}^{s=t} \Delta I_s^R (1+i)^{t-s}}{B_t^A}$$

Expressing the value of the multiplier m, the latter inequality may be written in the form:

$$\frac{B_t^A}{Y_t^A} < (1-c)(1-z)\frac{\sum_{s=1}^{s=t} \Delta G_s^R (1+i)^{t-s}}{\Delta G_t^R + \Delta I_t^R} - z\frac{\sum_{s=1}^{s=t} \Delta I_s^R (1+i)^{t-s}}{\Delta G_t^R + \Delta I_t^R} \qquad [1.2]$$

The expression on the right-hand side of relation **[1.2]** is the value now assumed by the 'inversion threshold' $T(t)$ of the debt ratio. Comparison with inequality **[1.1]**, where no changes in the levels of private investment are considered, reveals two differences. In addition to the reduction in public spending in period t, the denominator of the first fraction contains the reduction in private investment in the same period, which obviously reduces the value of the fraction (besides reinforcing the possibility of some decreasing stretches in its behaviour). The magnitude of the threshold is further reduced by the negative term in z, which could, if its values were sufficiently large, even give the threshold a decreasing course, in this case not necessarily restricted to limited stretches alone.[12] The size of this negative effect on the value of the threshold evidently depends on the size of the decreases in the flows of private investment that policy R generates by comparison with policy A. Even though this kind of influence does not lend itself to shaping in a general form, it is reasonable to expect that both the amounts and the duration of reductions in public spending would, through their impact on aggregate demand, directly affect the amounts by which private investment flows would be lower under policy R than policy A.

In conclusion, by augmenting the negative influence of the restrictive policy on output levels, reductions in private investment engender two effects. The first is a drop in the values of the threshold curve, which switches to the right and may become *decreasing* from a certain point on.[13] The second is an increase in the values of the debt ratio ensuing from the restrictive policy in any given period due to the simple fact that lower income levels generate lower public revenues. This effect would be seen graphically in an upward movement of the line representing the trend of the R debt ratio.

In Figure 1.4, $T_0(t)$ indicates the threshold curve ensuing with unchanged investment levels and T_1 the threshold associated with a series of investment reductions (by comparison with what would happen under policy A) starting from period t_1, while the two downward-sloping lines intersecting them represent the trends of the R debt ratio that would obtain respectively under the first and the second hypothesis about investment behaviour. As we can see, with no change in

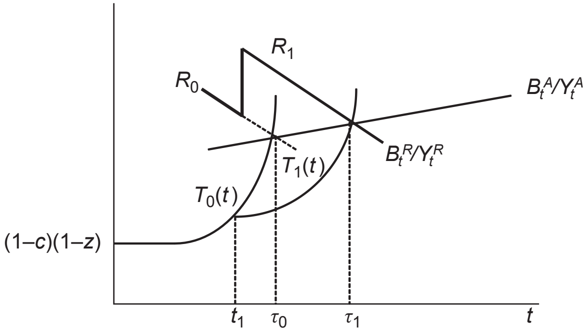


Figure 1.4 Effects of a restrictive policy (investment reacting once)

investment flows, the *R* policy would potentially bring the debt ratio below the level entailed by policy *A* in period τ_0 (line R_0). As a result of the curtailment of private investment flows at t_1, however, the threshold curve becomes $T_1(t)$ from then on with the corresponding upward movement in the trend of the *R* debt ratio (line R_1). The moment at which the latter would fall below the policy *A* ratio thus changes from period τ_0 to period τ_1 and the course of the *R* debt ratio would be represented by the bold broken line.

There is, however, nothing to rule out the possibility of the continuation of lower levels of demand and output, and their additional reduction entailed by the lower investment levels, subsequently generating even larger reductions in private investment flows. In Figure 1.5, the decreases in investment, again by comparison with the investment flows under policy *A*, are supposed to have increased still further since not only period t_2 but also period t_3 with consequent successive switches of both the threshold curve and the trend of the *R* debt ratio. The latter would thus remain above the trend of the *A* debt ratio for a number of periods, falling below it only if and when the reduction of private investment flows (by comparison with what would occur under policy *A*) stops for a sufficient length of time. Moreover, this result would eventually be obtained with levels of domestic output and private investment that could be lower by any degree than those resulting from policy *A*.

Even more drastically, policy *R* could prove totally incapable of bringing the trend of the debt ratio below that of policy *A* if the reduction in private investment flows were to become sufficiently large eventually to cause a decreasing course of the *T*(*t*) curve. This case is represented in Figure 1.6, where the decrease in private investment induced by policy

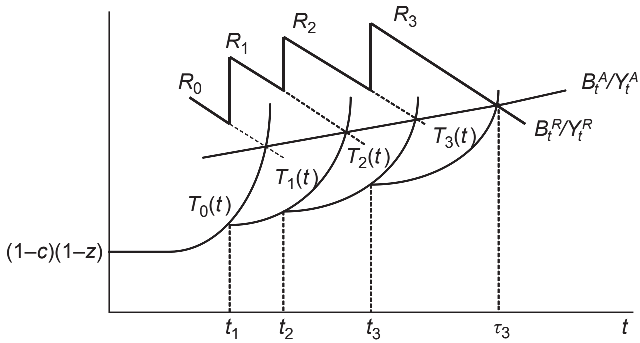


Figure 1.5 Effects of a restrictive policy (investment reacting over time)

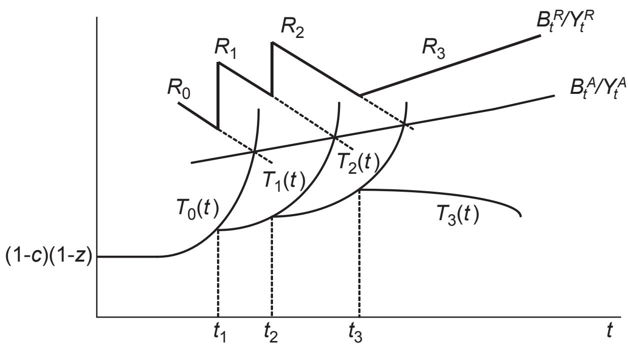


Figure 1.6 Effects of a restrictive policy (threshold eventually downward sloping)

R in period t_3, compared with the investment levels that would obtain under policy *A*, is supposed to be such as to generate a downward-sloping threshold curve. In this case, the latter would not intersect the trend of the *A* debt ratio and consequently, despite the continuation of a restrictive policy, the trend of the *R* debt ratio would continue to stand above that of the *A* ratio.[14]

1.6 The case of increases in public expenditure

We have thus seen that the attainment of a comparatively lower debt ratio is not guaranteed by the extension of the number of periods for which a policy of reduced public spending is implemented. In the absence of compensatory changes in other elements of autonomous demand, which there would be no reason to expect, precisely the prolongation of such a policy, with its negative influence on long-run demand trends, would possibly if not indeed inevitably lead to a drop in flows of private investment. In the case represented in Figure 1.5, this shifts the goal of reducing the debt ratio further into the future with no guarantee that the process will not fall into a tailspin, thus moving the position of the threshold curve $T(t)$ further and further to the right. The already uncertain reduction of the debt ratio becomes impossible in the case represented in Figure 1.6, where sufficiently strong reductions in flows of private investment cause the slope of the $T(t)$ curve to become negative and the restrictive policy gives rise to a permanent increase in the debt ratio. Furthermore, this uncertainty of results, even as regards their sign, is accompanied by a certain and permanent relinquishment of higher levels of income and employment by the community, something that could be described in the terms of the dominant theoretical framework as an instance of 'Pareto inferiority'.

These conclusions prompt examination of the situation that would instead arise with higher levels of public spending. Let E be an expansionary policy entailing increases in expenditure with respect to the current policy A. If we take into account the increases in private investment that would be induced by the higher levels of demand, the stock of public debt at time t under policy E, as against the debt that would result from policy A, is:

$$B_t^E = B_t^A + (1 - mz)\sum_{s=1}^{s=t} \Delta G_s^E (1+i)^{t-s} - mz \sum_{s=1}^{s=t} \Delta I_s^E (1+i)^{t-s}$$

where ΔG_s^E and ΔI_s^E respectively represent the increases in public expenditure and private investment in the generic period s by comparison with what would obtain under policy A. The level of income in period t would correspondingly be higher by the amount

$$\Delta Y_t^E = m(\Delta G_t^E + \Delta I_t^E)$$

The condition required for policy E to produce a lower debt ratio at t than policy A is once again that the proportional increase in the level of income exceeds the proportional increase in the absolute stock of debt:

$$\frac{m(\Delta G_t^E + \Delta I_t^E)}{Y_t^A} > \frac{(1 - mz)\sum_{s=1}^{s=t} \Delta G_s^E (1+i)^{t-s} - mz \sum_{s=1}^{s=t} \Delta I_s^E (1+i)^{t-s}}{B_t^A}$$

which, rearranging the terms and expressing the value of the multiplier m, provides the value that the debt ratio with policy A must be *above* in order for the said result to be obtained:

$$\frac{B_t^A}{Y_t^A} > (1-c)(1-z)\frac{\Delta G_s^E (1+i)^{t-s}}{\Delta G_t^E + \Delta I_t^E} - z\frac{\sum_{s=1}^{s=t} \Delta I_s^E (1+i)^{t-s}}{\Delta G_t^E + \Delta I_t^E}$$

The right-hand side of the above inequality represents the 'inversion threshold' of the $\frac{B_t^A}{Y_t^A}$ ratio, in this case a *lower* boundary, in order for the expansionary policy to generate a lower debt ratio at t. As can be seen, the expression of the threshold is strictly analogous to what was found in considering a restrictive policy, and it therefore behaves in the same way. On the assumption that private investment would respond positively to the higher aggregate demand triggered by the expansionary policy, in Figure 1.7 $T_0(t)$ represents the threshold curve assuming no change in investment flows as against what would obtain with policy A, and the curves $T_1(t)$, $T_2(t)$, and $T_3(t)$ instead exemplify the successive forms of the threshold curve resulting from increasingly large comparative increases in the flows of private investment. In the same figure, the line E_0 represents the course that the E debt ratio would take if investment flows were the same as under policy A, and the lines E_1, E_2, E_3 stand for the paths of the ratio generated by the increasing levels of investment entailed by the corresponding $T.(t)$ curve. As the figure shows, each further increase in investment flows moves both the threshold curve and the E debt-ratio line to the right, thus shifting the time

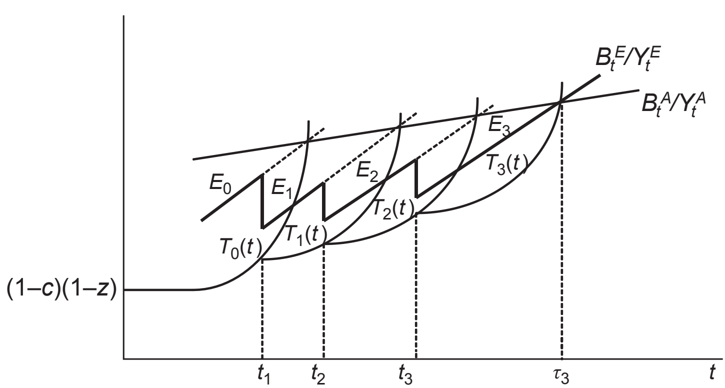


Figure 1.7 Effects of an expansionary policy (investment reacting over time)

at which the two curves jointly intersect the trend line of the *A* debt ratio and the *E* debt ratio overtakes the latter progressively from τ_0 to τ_1, then τ_2 and finally τ_3. As a result of the expansionary policy, the trend followed by the *E* debt ratio would therefore lie *below* the one generated by the current policy for a number of periods which would be larger, the larger and more persistent the effects produced directly and indirectly on aggregate demand.

In symmetry with the findings of our examination of a restrictive policy, one result of our analysis is that the trend of the expansionary policy debt ratio would lie permanently below that of the current policy if the resulting increases in private investment flows reached levels high enough to generate a downward-sloping $T(t)$ curve. This is the case represented in Figure 1.8, where the rise in private investment from period t_3 is supposed to be such as to make the $T_3(t)$ curve negatively sloped and prevent it from intersecting with the $\frac{B_t^A}{Y_t^A}$ line. In turn, the $\frac{B_t^E}{Y_t^E}$ line would never overtake the $\frac{B_t^A}{Y_t^A}$ line, and the expansionary fiscal policy would succeed in keeping the trend of the debt ratio below that of the current policy indefinitely.

1.7 Final remarks and conclusions

As stated at the outset, unlike the more usual procedure of comparing the results of a policy change with magnitudes calculated at some earlier

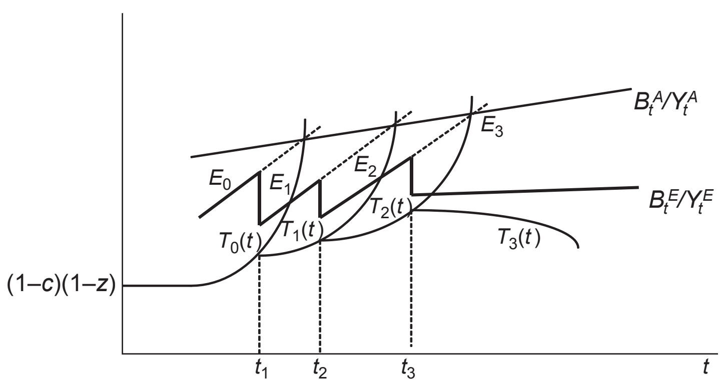


Figure 1.8 Effects of an expansionary policy (threshold eventually downward sloping)

time, our investigation compares a given trend of the debt ratio generated by the currently planned fiscal policy with the trend that would be generated by either a comparatively restrictive or a comparatively expansionary policy. The justification put forward for this procedure is the fact that a change in fiscal policy, and particularly in public spending, is generally formulated in terms of variations to be made on current policy. It is therefore more appropriate to evaluate the effects produced by the policy change over a given period through comparison with what would have been the result of current policy during the same period rather than with values prior to the change. This holds all the more if, as is evidently the case for the debt ratio and for objectives of fiscal policy in general, the level of overall output constitutes a relevant variable and is seen as demand determined by and hence dependent on the economic policies in force.

Nevertheless, the analysis can be easily supplemented if desired through the introduction of a specific value of the debt ratio that fiscal policy can be expected – according to some views of greater or lesser conceptual validity – to take as a benchmark for the actual debt ratio. On the obvious assumption that this target ratio is lower than the one obtainable with the current policy, it could be represented by a horizontal line drawn below the $\frac{B_t^A}{Y_t^A}$ line used in our graphs. The addition of such a line in Figures 1.5 and 1.6, which show how difficult if not indeed impossible it is for a restrictive fiscal policy to bring the trend of the debt ratio below the level resulting from the current policy, would then enable us to deduce analogous and actually stronger conclusions about the ability of this kind of policy to bring the debt ratio into line with an even lower target value. Adoption of the same procedure in Figures 1.7 and 1.8 would suggest that this adjustment would be more efficiently pursued by means of an expansionary policy.

Summing up, the chapter has endeavoured to show analytically that once a positive influence of public expenditure on total output is acknowledged, an *inverse* relation may hold, also in the long run, between levels of deficit spending and values of the ratio of public debt to domestic product. The results suggest that if a lower debt ratio constitutes a target of economic policy, as is currently the case in several countries, then regardless of the validity or otherwise of that objective, restrictive fiscal policies may prove ineffective and even self-defeating by generating higher rather than lower values of the debt ratio. On the contrary, expansionary policies may produce comparatively lower values of the ratio, and deficit spending would therefore be superior to fiscal retrenchment with regard not only to levels of activity and employment but also to what is often referred to as an index of solidity of the financial position of governments.[15]

The conclusions drawn here apply to policies implemented over series of periods and hold in this sense for the long run, not just the short. This is due to the fact that, contrary to what we find in much of the mainstream literature, the theoretical framework adopted here entails no need to counterbalance changes in fiscal policy with opposite changes in private expenditure, no matter how 'lagged'. The key idea of this chapter, namely that output is determined by demand, naturally implies instead that gross investment, the primary 'autonomous' component of private spending, directly depends in turn on the trend in levels of demand. Flows of private investment would thus respond to and hence assist changes in aggregate demand rather than counterbalancing them. By further enhancing the influence that fiscal policy can have in the long run on output levels, this behaviour of private investment reinforces the result that contrary to the 'orthodox' view, lower values of the debt ratio can be obtained more easily, and less painfully, through expansionary fiscal policies.

Notes

1. An application of the Keynesian principle of effective demand to the role of public spending and, more generally, to fiscal policy was put forward in a fairly large body of literature mainly dating from the 1940s and 1950s, which found perhaps its most 'radical' expression in Lerner's *functional finance* (Lerner, 1943). A major element distinguishing those contributions from this chapter is the fact that while the former often fail to specify the nature of the limit that aggregate demand imposes on total output, i.e., whether it is purely cyclical in character or extends to the trend of activity levels (cf. Ciccone, 2002, pp. 12–13, fn. 6), aggregate demand is explicitly regarded here as determining overall output in both the short and the long run.
2. For the premises implicit in the assumption that agents behave in this way, see Section 1.2 below.
3. A vast literature has developed focusing specifically on quantitative estimation of the change in the level of total output attributable to a change either in public expenditure or in tax rates or levels, a ratio generally labelled the 'multiplier of fiscal policy' or simply 'fiscal multiplier' (for a recent survey of this literature, see Ramey, 2011; see also Barba, 2001 for a critical discussion of the theoretical foundations of the claim that fiscal contractions can also be expansionary, which is still relevant to much of the work at issue). Many of these studies appear to be inspired by the assumption, implicitly or

explicitly derived from orthodox theory, that a change in fiscal policy can have no significant or in any case lasting influence on the level of demand, as it is bound to cause opposite changes in some components of the latter (private investment, private consumption or exports), so that the size of its effect on output is expected to be relatively small or even opposite in sign (but see Fazzari *et al.*, 2012 for an argument developed in a different direction). This is so by construction in the 'simulations' carried out through the use of models, the results of which obviously depend on the characteristics of the models adopted, which often reflect the set of relations peculiar to neoclassical theory. The size and sign of fiscal multipliers are alternatively investigated by means of pure empirical work relying on no theoretical structure and based exclusively on historical data. This is not without cost. Due to the dependence of fiscal magnitudes on several variables, including domestic output, a major problem encountered by this kind of research is the identification of changes in fiscal quantities (e.g. the primary balance) that can be attributed to deliberate policy actions and therefore treated as exogenous rather than resulting from the prevailing economic conditions (see the objections raised in this regard in IMF, 2010). The mass of work on 'fiscal multipliers' has in fact produced estimates that differ widely according to the assumptions and procedures adopted, e.g. ranging from −3.8 to +3.8 for public expenditure (cf. Van Brusselen, 2009, p. 18; estimates for tax-cut multipliers are reported to range between −4.8 and +3.0). The issue therefore still remains a matter for theory rather than something that can be decisively settled on empirical grounds. In the approach adopted here, which regards demand as determining the level of output, the influence of public spending on output can accordingly be expressed by the full action of the income multiplier with no need for counterbalancing variations in other constituents of demand.

4. As is well known, it was Barro (1974) who originally put forward the idea that households would reduce consumption by the same amount as the issuing of public debt, so that debt- and tax-financing of public expenditure would supposedly be equivalent in their effects on agents' behaviour. This article made no reference to Ricardo and the connection with an analogous point made by Ricardo was later noted by Buchanan (1976), since when the concept has often been referred to as 'Ricardian'. While claiming that the way in which public spending is financed should be a matter of indifference to the individual on logical grounds, Ricardo maintained, however, that this equivalence is not effectively perceived, as future taxes are in fact assigned less importance than present taxes. Unlike Barro, Ricardo thus maintains that debt financing would reduce savings by comparison with the situation in the case of tax financing (cf. Ricardo, 1951, pp. 186–7).
5. The rate z is here taken as representative of any kind of deduction from produced income accruing to the public sector as a whole as well as taxes and fees paid on purchases of goods and services. The constancy of z also rules out therefore any changes in public charges (e.g. for health services), which could furthermore come into conflict with the assumed invariance of income distribution.

6. Taking into account the value that the multiplier m assumes under our simple assumptions (see above in the text), the value of mz is lower than 1 for $z<1$ and $c<1$.
7. The difference in private wealth at t amounts to the difference in the flows of savings accruing to the private sector as a result of the difference in the level of domestic product under one policy and the other:

$$\Delta W_{pt}^{R}=(1-c)(1-z)\Delta Y_{t}^{R}$$

from which the equality stated in the text follows.
8. Interest outlays on public debt are considered net of taxes, any tax on them being no more than a clearing entry for the public sector.
9. If $T(h)$ and $T(h+1)$ are the values of the threshold at two subsequent times, their difference is

$$\begin{aligned}T(h+1)-T(h) &= (1-c)(1-z)\left[\frac{\Delta G_1^R(1+i)^h+\Delta G_2^R(1+i)^{h-1}+\cdots+\Delta G_{h+1}^R}{\Delta G_{h+1}^R}\right.\\ &\quad\left.-\frac{\Delta G_1^R(1+i)^{h-1}+\Delta G_2^R(1+i)^{h-2}+\cdots+\Delta G_h^R}{\Delta G_h^R}\right]\\ &= (1-c)(1-z)\frac{1}{\Delta G_{h+1}^R}\left\{\left[(1+i)-\frac{\Delta G_{h+1}^R}{\Delta G_h^R}\right]\right.\\ &\quad\left.\times\sum_{s=1}^{h}\Delta G_s^R(1+i)^{h-s}+\Delta G_{h+1}^R\right\}\end{aligned}$$

The algebraic sign of this quantity is determined by the quantity in braces, which can be negative if ΔG_{h+1}^R exceeds ΔG_h^R by a sufficient amount and the sum of $\Delta G_s^R(1+i)^{h-s}$ has attained a large enough value.
10. In point of fact, even with a linear form, the R debt ratio might be increasing with a slope lower than that of the A ratio (which is needed by the required intersection with the latter) or be flat. Being less favourable to our argument, the assumption of a decreasing trend is the one adopted here.
11. This can be easily seen under the simplifying assumption that the spending reductions are constant over time, i.e., $\Delta G_s^R = G^R$, in which case $T(t)$ reduces to

$$\begin{aligned}T(t) &= (1-c)(1-z)\frac{\sum_{s=1}^{s=t}\Delta G^R(1+i)^{t-s}}{\Delta G^R}=(1-c)(1-z)\sum_{s=1}^{s=t}(1+i)^{t-s}\\ &= (1-c)(1-z)\frac{(1+i)^t-1}{i}\end{aligned}$$

the values of which, interestingly enough, are independent of the magnitude of the reduction in spending. This function can increase quite rapidly. For example, with $c=0.70$, $z=0.30$ and $i=0.03$, it takes the value 0.65 for $t=3$ and is already as high as 1.12 for $t=5$.

12. The difference between two subsequent values of the $T(t)$ is now

$$T(h+1) - T(h) = (1-c)(1-z)\left[\frac{\Delta G_1^R(1+i)^h + \Delta G_2^R(1+i)^{h-1} + \cdots + \Delta G_{h+1}^R}{\Delta G_{h+1}^R + \Delta I_{h+1}^R} - \frac{\Delta G_1^R(1+i)^{h-1} + \Delta G_2^R(1+i)^{h-2} + \cdots + \Delta G_h^R}{\Delta G_h^R + \Delta I_h^R}\right] - z\left[\frac{\Delta I_1^R(1+i)^h + \Delta I_2^R(1+i)^{h-1} + \cdots + \Delta I_{h+1}^R}{\Delta G_{h+1}^R + \Delta I_{h+1}^R} - \frac{\Delta G_1^R(1+i)^{h-1} + \Delta G_2^R(1+i)^{h-2} + \cdots + \Delta G_h^R}{\Delta G_h^R + \Delta I_h^R}\right]$$

By positing that this magnitude is negative and that $\dfrac{\Delta G_{h+1}^R + \Delta I_{h+1}^R}{\Delta G_h^R + \Delta I_h^R} = 1 + \sigma_{h+1}$,

we obtain,

$$(1-c)(1-z)\left\{\left[1 - \frac{1+\sigma_{h+1}}{1+i}\right][\Delta G_1^R(1+i)^h + \Delta G_2^R(1+i)^{h-1} + \cdots + \Delta G_h^R(1+i)] + \Delta G_{h+1}^R\right\}$$
$$< z\left\{\left[1 - \frac{1+\sigma_{h+1}}{1+i}\right][\Delta I_1^R(1+i)^h + \Delta I_2^R(1+i)^{h-1} + \cdots + \Delta I_h^R(1+i)] + \Delta I_{h+1}^R\right\}$$

The above inequality, which implies the decreasing course of $T(t)$, can be satisfied by requiring no more than sufficient reductions in private investment (compared with the investment flows that would obtain under the currently planned fiscal policy). See also notes 13 and 14 below.

13. If the assumption that the reduction in public spending is constant over time (see n. 4 above) is combined for algebraic convenience with the assumption that the reduction in private investment is also constant, i.e. $\Delta I_s^R = \Delta I^R$, the expression of the threshold on the right-hand side of inequality [1.2] becomes

$$T(t) = (1-c)(1-z)\frac{\sum_1^t \Delta G^R(1+i)^{t-s}}{\Delta G^R + \Delta I^R} - z\frac{\sum_1^t \Delta I^R(1+i)^{t-s}}{\Delta G^R + \Delta I^R}$$
$$= \frac{(1+i)^t - 1}{i}\left[\frac{(1-c)(1-z)}{1+\Delta I^R/\Delta G^R} - \frac{z}{1+\Delta G^R/\Delta I^R}\right]$$

where it is evident that the size of ΔI^R relative to ΔG^R affects the values of $T(t)$ negatively.

14. Under the specific assumptions of a constant decrease in public expenditure and a constant decrease in private investment, which allow us to use the formula of the threshold already indicated in n. 8, the table below exemplifies the progressive movement of the $T(t)$ curve and the eventual change of its slope into the negative as represented in Figure 1.6. The calculations are made assuming c=0.70, z=0.30 and i=0.03, and that the reduction in investment (expressed as a proportion of the public spending reduction) starts in period 3 and increases in steps consisting of intervals of three periods while remaining constant within each interval. The table shows that when the decrease in investment reaches

t \ $\Delta I^R/\Delta G^R$	$T_0(t)$ 0.00	$T_1(t)$ 0.20	$T_2(t)$ 0.40	$T_3(t)$ 0.70	$T_4(t)$ 0.90	$T_5(t)$ 1.00
1	0.21					
2	0.43					
3	0.65	0.65				
4		0.78				
5		0.90				
6		1.04	1.04			
7			1.10			
8			1.17			
9			1.24	1.24		
10				1.24		
11				1.24		
12				1.24	1.24	
13					1.21	
14					1.18	
15					1.14	1.14
16						1.10
17						1.05
18						1.00
19						0.95

0.70 of the decrease in public spending, the threshold curve becomes flat and decreasing for further increments of the decrease in investment.

15. The analysis developed here is not immediately susceptible of transposition in terms of growth rates of variables, as is instead often the case in the literature concerning public debt and its ratio to domestic product. It is, however, easy to see that in such a reformulation of our basic relations (see Ciccone, 2002, ch. II) the rate of growth of public spending would be conceived as a policy variable positively affecting the rate of growth of overall output (see also Aspromourgos, 2007, section 2 for an analogous standpoint) and an *inverse* relation could therefore hold between the former rate and the ratio of public debt to domestic product.

References

Aspromourgos, T. (2007) 'Interest as an artefact of self-validating central bank beliefs', *Metroeconomica*, 58(4): 514–35.

Barba, A. (2001) 'Rethinking expansionary fiscal retrenchments', *Contributions to Political Economy*, 20: 57–86.

Barro, R. J. (1974) 'Are government bonds net wealth?', *Journal of Political Economy*, 82: 1095–117.

Buchanan, J. M. (1976) 'Barro on the Ricardian Equivalence Theorem', *Journal of Political Economy*, 84: 337–42.

Ciccone, R. (2002) *Debito pubblico, domanda aggregata e accumulazione* (Roma: Aracne).

Fazzari, S.M., Morley J. and Panovska I. (2012) 'State-Dependent Effects of Fiscal Policy', *Australian School of Business*, Research Paper 2012 ECON 27.

International Monetary Fund (2010) 'Will It Hurt? Macroeconomic Effects of Fiscal Consolidation', *World Economic Outlook*, ch. 3: 93–124.

Lerner, A. P. (1943) 'Functional finance and the federal debt', *Social Research*, 10: 38–51.

Ramey, V. (2011) 'Can government purchases stimulate the economy?', *Journal of Economic Literature*, 49(3): 673–8.

Ricardo, D. (1951) 'Funding System', in Piero Sraffa (ed.) with the collaboration of M. H. Dobb, *The Works and Correspondence of David Ricardo*, Vol. IV (Cambridge: Cambridge University Press): pp. 149–200.

Van Brusselen, P. (2009) 'Fiscal Stabilisation Plans and the Outlook for the World Economy', *NIME Policy Brief* 01-2009.

2

The Inevitable Dependence of Investment on Expected Demand: Implications for Neoclassical Macroeconomics

Fabio Petri

2.1 An internal critique of the neoclassical justification of Say's Law

The purpose of this chapter is to draw attention to a weakness, so far unnoticed, of the neoclassical argument in support of Say's Law – that is, of the thesis that investment *is determined by* savings,[1] and that therefore aggregate demand poses no obstacle to selling at cost-covering prices the aggregate supply of goods whatever the forces determining the latter. The neoclassical argument, relying upon an assumed negative interest elasticity of investment derived from the demand-for-capital function, neglects the problems with the marginalist or neoclassical conception of capital: as pointed out by the late Pierangelo Garegnani (1983, 1990), the discovery of reverse capital deepening undermines the foundations of Say's Law, because it undermines the belief in a negative interest elasticity of the demand for (value) capital, but then also the belief in a negative interest elasticity of aggregate investment; Garegnani concluded that the 'neoclassical synthesis' criticism of Keynes could not be accepted, and that aggregate demand had to be considered the determinant of employment and growth not only in the short period but also in the long run. In Petri (2004, ch. 7) I reinforced Garegnani's contention by showing that the attempts, after Keynes, to derive a negative interest elasticity of investment without relying on the traditional neoclassical conception of capital are all indefensible.[2] Here I intend to point out that the neoclassical argument meets grave difficulties even conceding the traditional neoclassical conception of capital–labour substitution[3] – and for a reason different from the ones adduced so far (such as 'malfunctioning' of financial intermediaries or irreducibly subjective expectations and animal spirits).

The reason is a fact hidden from sight in most current presentations of investment theory, the *inevitable* dependence of investment on desired capacity (and hence on expected demand and its variations) even in a neoclassical framework. I will point out that the marginalist, or neoclassical, approach to investment needs the continuous full employment of labour in order to arrive at a determinate influence of the interest rate upon investment; since the full employment of labour cannot be *assumed* as a starting point of the analysis (it can only be, if at all, a *result* of the analysis), labour employment must be considered variable, but then a given rate of interest leaves investment indeterminate even accepting the marginalist conception of capital–labour substitution, because the rate of interest can only determine the desired *ratio* of capital to labour, which leaves the capital to be invested in new plants indeterminate and to be determined by desired productive capacity. As I will show, some neoclassical economists (Jorgenson, and Dornbusch and Fischer) implicitly admitted it by having investment determined by expected demand, but did not see that then extremely serious problems arise concerning the capacity of wage flexibility to bring about the full employment of labour. What follows explains and expands on these points.

2.2 The neoclassical connection between demand for capital and investment

After Keynes it has become common to consider investment a function of the rate of interest *only*. Even when the influence of other variables is admitted, e.g. of income, it is generally seen as *additional* to the influence of the rate of interest, in the sense that the latter would suffice to determine investment if the other influences were very weak or absent. On the contrary a correct grasp of the neoclassical conception of capital–labour substitution implies that the rate of interest alone is unable univocally to determine investment, unless the full employment of labour is *assumed*. To see why, it is necessary to remember the traditional derivation of investment from the schedule of the demand for capital (conceived in the traditional marginalist way as a single factor, an amount of value).[4] The connection – often only implicit – between investment and demand for capital in J. B. Clark, Böhm-Bawerk, Wicksell, Marshall, Knight etc. has been described by Garegnani with a clarity that can hardly be surpassed (Garegnani, 1983: 34–7; 1990:

59–60). Investment was seen by these authors as the *flow* corresponding to the *stock* demand for capital, given that capital wears out and therefore needs a continuous flow of new capital goods for its stock to remain equal to the demand for it:

> The traditional theory implies that the delayed adjustments in the wages, rents, and prices of products do not fundamentally alter the terms of the question ... Hence the significance of the demand and supply functions for capital as a stock, which would exhibit the basic tendencies destined to emerge from the multiplicity of forces acting at any given moment in the savings investment market ... the traditional analyses of the demand and supply for capital were in effect intended to be an analysis of the demand and supply for savings, abstracting from the complications likely to operate at each particular moment of time in the savings–investment market. (Garegnani, 1990: 59–60)

Those authors had to admit of course that in a concrete economy any adjustment to a change in the data of equilibrium (e.g. labour immigration, or technical progress, or changes in the propensity to save) would also present the 'complications' Garegnani mentions, 'complications' due, for example, to differences in the age structure of fixed capital and connected irregularities of the need for replacement of scrapped plants; redistributions of purchasing power among social groups due to changes in the interest rate; possible interference of financial intermediaries; possible convenience of anticipated scrapping of fixed plants; mistaken expectations; slowness in adjustments of factor rentals; and so on. The effects of these 'complications' were to be studied if necessary at a second level of approximation; the demand-for-capital curve was believed to supply 'the basic tendencies', those emerging once the irregularities of the behaviour of prices owing to accidental or transitory disequilibrium phenomena had time to be sufficiently compensated or corrected, and therefore product and factor prices had become sufficiently close on average to their new normal levels, a process enforced by competition. For example, even without any change in optimal technologies a reduction in the rate of interest cannot but push freely competing firms to try and undercut their competitors by lowering product prices relative to money wages since average costs have decreased; if they don't, it will be new firms – whose birth will be stimulated by the persistence of prices higher than average costs – that will do it to gain market shares.

2.3 Capital–labour substitution requires capital to change 'form'

It is opportune to stress a number of important aspects of the determination of the marginalist *long-period investment function*, as I call the investment function generated by such an approach. The demand for capital is determined as the persistent demand for capital goods – aggregated in value terms – implied by the persistent demand for a given net product, this net product being the one produced when labour is fully employed, and production methods, output composition and prices being the normal ones associated with the income distribution determined by the full-employment marginal productivities of labour and capital (following general practice, I assume land is free, because land is not important for the issues to be discussed). Since at each given moment the endowment of 'capital' is crystallised in specific capital goods adapted to a specific productive method, a change (induced by a change in income distribution) in the desired (i.e., normal) capital–labour ratio in an industry can only be realised by replacement of the old durable capital goods with new ones of a different type, or for brevity, can only be realised in new plants (only in new plants can the marginal productivity of capital be determined, since only there can the normal *K/L* ratio be varied); if industry output is unchanged, the new plants will only be built to replace older plants that reach the end of their economic life and are scrapped, plants that are not so old continuing in operation as long as they earn non-negative residual quasi-rents. Changes in the output of an industry, whether due to changes in consumer choices or to changes in production methods in industries using that output as an input, will mostly be accommodated, in the short period, by changes in the degree of utilisation of existing plants, but if perceived as persistent will induce a desire to change productive capacity, and this will be the other main influence on gross investment (per unit of labour employed in new plants), affecting its composition through the desired composition of new capacity. The composition effect due to change of methods is part, in the traditional marginalist approach, of the overall operation of the so-called direct factor substitution mechanism, which changes the desired K/L ratio in the subsystem producing a given final good; the composition effect due to changes in consumer choice constitutes the indirect factor substitution mechanism. In either case, since in most cases it is impossible to utilise existing productive capacity for the production of goods different from those for which that capacity had been planned, the change in industry capacity can generally only be realised through the building of new plants where demand expands, and non-renewal of the scrapped older plants where demand contracts. Thus both the direct and the indirect substitution mechanisms between capital and labour can only operate by affecting the type and sector allocation of the *new* durable capital goods to be combined with the flow of labour gradually released by the scrapping of the durable capital goods that reach the end of their economic life. It is only through the replacement of the existing capital goods with capital goods adapted to produce different goods or adapted to a different technical method, that is, it is only through a change of the 'form' of capital, that the average economy-wide capital–labour ratio can change and a sufficiently elastic demand curve for factors can be obtained. For this reason Hicks (1932: 18–21) expressed strong doubts on the meaningfulness of a *short-period* demand curve for labour, and considered the notion of a demand curve for labour to be meaningful only if one allowed the 'form' of the given capital endowment of the economy to have the time to adapt to the changed real wage.

An implication of this view is that the process of change of the 'form' of capital brought about by a change in the rate of interest is necessarily *slow*, taking – in order to operate completely – the many years required for renewal of the entire stock of fixed plants of the economy: much longer, generally, than necessary for the new rate of interest to determine a gravitation to the new normal relative prices, a gravitation that only requires that the first plants built according to the new optimal factor proportions be capable of imposing a price equal to their lower average cost, obliging the older plants to accept that price and be content with residual quasi-rents. But economic conditions will seldom remain unchanged long enough for complete adaptation of *all* plants to an unchanging rate of interest; therefore it is implicit in this approach that in any concrete economy the rate of interest must be seen as determined, not so as to equalise the capital–labour ratio in factor demand to the ratio of the existing *total* endowment of capital to *total* labour supply, but rather so as to equalise the desired average ratio of capital to labour *in new plants* to the ratio between the flow of 'free' capital (savings) and the flow of labour 'freed' or 'released' by the gradual shutting down of old plants – a ratio that would coincide with the 'total' one only when and if the *entire* labour force were employed in plants embodying methods optimal for that income distribution, and generating productive capacities adapted to the composition of consumer demand. But since most gross investment would be generally motivated by unaltered replacement of used-up circulating and fixed capital, the implicit view of traditional marginalist economists was that the demand-for-capital schedule and its shifts would give a sufficient indication of the tendencies of the investment function.

Any non-negligible difference of actual investment from the long-period investment function would be part of those transitional or irregular 'complications' mentioned by Garegnani.[5] The most important aspect of the long-period investment function, its negative interest elasticity, could in any case be argued certainly to hold for the actual investment function too, since the *K/L* ratio employed on average in new plants would certainly tend to decrease with rises in the interest rate, while the flow of gradually 'freed' labour as well as the speed of completion of changes in industry dimensions would hardly increase.

2.4 The need for the full employment of labour

It should now be clear that the long-period investment function crucially depends on the assumption of full employment of labour. If for simplicity we assume 'investment' *I* to indicate only the part of total investment whose ratio to labour and to output will respond to changes in distribution, that is, to consist only of gross investment going to new plants,[6] and if *L^* stands for the flow of labour employed in new plants, the optimal *K/L* ratio determined by the rate of interest determines *I/L^*, but *I* remains indeterminate unless *L^* at the denominator is given. The long-period investment function assumes *L^* to correspond to the flow of labour gradually 'freed' by the closure of the oldest plants in a situation of continuous full employment of labour. It is this given *L^* that allows the *K/L* ratio corresponding to the given rate of interest to determine investment.

If the possibility of labour unemployment is admitted, then a given average *K/L* ratio in new plants does not suffice to determine investment, because new plants can employ less, or (if there already is some unemployment) more, labour than the flow 'released' by the closure of old plants, correspondingly gradually reducing or increasing employment. A given rate of interest, without some assumption fixing *L^*, leaves investment indeterminate.

2.5 A simple model with putty-clay capital

A very simple model, that stresses only the direct factor substitution mechanism, can give concreteness to the above considerations. Assume an economy where a single good is produced by labour and putty-clay capital; production within each period adapts to the demand forthcoming in that period (the analysis is in discrete time); the output can be consumed, or it can be invested, i.e., costlessly transformed into capital, but the newly produced capital becomes productive only at the beginning of the following period. The capital–labour ratio must be chosen at the moment of transformation of output into capital, from the possibilities determined by an *ex ante* production function $Y=F(K,L)$, and the resulting capital good allows only one output–labour ratio, which is constant as labour employment per unit of capital varies from zero up to a maximum corresponding to the capital–labour ratio originally chosen. (Thus there may be less than full-capacity utilisation of some or all capital goods.) Capital goods last 10 periods with constant efficiency, independently of the K/L ratio chosen at the time of their creation and of the level of utilisation of the capital good during its life. I assume no technical progress.

The economy is initially in stationary full-employment equilibrium with capital goods fully utilised: at the end of every period the oldest one-tenth of the capital goods are scrapped and replaced by new capital goods of the same type, produced during the period; the newly installed capital goods utilise in the following period the one-tenth of the labour force which is 'freed' by the scrapping of the oldest capital goods. The real wage equals the marginal product of labour in new plants; once the real wage is fixed, the real rate of interest (I neglect risk) is univocally determined (by – owing to the presence of fixed capital – rather complicated equations into which we need not enter).

Then, let us assume, at the beginning of one period the real wage unexpectedly rises (trade unions or political decisions impose this rise, without a change in labour supply) and it is expected to remain at the new level for many periods, and the real interest rate adapts rapidly, so the optimal K/L ratio in new plants rises; the quantity of output destined to investment, let us assume, does not change (this allows us to consider the quantity of capital as not changing); from the subsequent period onwards, part of the one-tenth of the labour supply 'freed' by the scrapping of the oldest plants remains unemployed; the other nine-tenths of the labour force remain employed by the already existing plants, which I assume still yield positive quasi-rents because the wage increase is small. Assume (i) that savings keep translating without difficulty into investment; (ii) that the amount of output destined to gross investment does not decrease in subsequent periods in spite of the decrease in labour employment, so the stock of capital (in the physical sense of total amount of output from which it was created) does not change and (iii) that the real wage does not change. Then after 10 periods the total physical capital K_{Tot} of the economy, measured in physical terms as the sum of the given-up consumption that allowed its creation, has not changed, and labour employment (which is less than labour supply) corresponds to the new lower L/K ratio multiplied by the aggregate capital measured as indicated. All employed labour now produces output at the new Y/L

ratio. The final labour employment as a function of the real wage is indicated by a labour demand curve that traces the marginal productivity of labour when the given physical supply K_{Tot} of capital is introduced into the economy's production function *F(·)*. This is the labour demand function that, as Hicks requested, allows the 'form' of the given quantity of capital to become adapted to the real wage.

(It would not be unrealistic to interpret the length of the period of this analysis as at least a year – fixed plants often last much longer than 10 years –, so the wage change would take at least 10 years to exert its full effect on employment. The slowness of the adjustment implicit in this theory is seldom fully perceived, so its important consequences escape general recognition. One consequence is that, in order to avoid implausible enormous falls of wages whenever unemployment were to arise, the theory must admit the presence of social forces that render wage decreases very slow (Petri 1991: 272–3). But then it is unclear why those same social forces – custom, solidarity, feelings of fairness, aversion to accepting reductions of wage relativities as stressed by Keynes, bargaining power of trade unions, threat of violence, etc. – might not be capable of totally preventing wage falls even in the presence of unemployment, thus constituting the basis for a determination of wages alternative to the neoclassical tendency toward a supply-and-demand equilibrium, and very much in line with the views of the first attentive observer of capitalism, Adam Smith. Another consequence is that even the neoclassical economist has little reason to presume that the *negative* effect, to be discussed later, of a decrease of real wages on employment through its negative effect on aggregate demand will be slower and weaker in its action than the positive effect on the demand for labour coming from capital–labour substitution.)

The assumption that production takes one period (with all productions started at the beginning of a period and ending at the end of the period) means that in each period t the output $Y_t = C_t + I_t$ cannot include the output of plants created by I_t. So Y_t is the result of the full utilisation of the plants that the economy has *at the beginning* of the period, each vintage producing and employing labour depending on the amount of capital good of that vintage and on the K/L ratio chosen for that vintage. Thus in order to determine the demand for labour the reasoning takes Y_t in each period as given, determined by the full utilisation of beginning-of-period capacity. (Changes of the real wage at the beginning of the period have no effect on labour employment in that period, at least as long as the wage change does not cause anticipated scrapping of plants.)

Let us now remember the considerable elasticity of the output of the several industries in real economies, in response to variations in demand (the elasticity that makes the working of the Keynesian multiplier possible). Variations of demand will be met at first by variations of inventories and then by variations of output levels tending to bring inventories to normal – and, in manufacturing industry, generally with little or no change in prices. The premises of this elasticity are not represented in the above model, which lacks inventories, but this elasticity should nonetheless be admitted. And it is well known that firms plan productive capacity for a level of utilisation which is considerably less than the technical maximum level (and is nonetheless esteemed to be optimal for the reasons pointed out in the literature on optimal capacity utilisation: Marris, Betancourt and Clague, Winston, Heinz Kurz etc.), so that not only under-utilisation of plant, but also above-normal utilisation is a possibility. Therefore what we described above (see p. 50) as the *maximum* output/capital ratio corresponding to the chosen *K/L* ratio must more realistically be reinterpreted as the *normal* output/capital ratio, which can be exceeded if demand is particularly high. And 'full-capacity output' must be interpreted to mean *normal*-utilisation output, not an upper limit to actual output.

Once this elasticity of output in response to demand is admitted, then there is no obstacle to admitting an autonomous influence of investment upon output, in either direction. An investment less than normal-capacity savings will encounter no obstacle in causing *Y* to be less than normal-capacity output even if initially there was full employment of a rigid labour supply. An investment greater than normal-capacity savings will cause *Y* to be greater than normal-capacity output as long as an increase in labour employment is possible.[7] Savings will adjust to investment via the variation of *Y* induced by the multiplier.

2.6 Implications of the possibility of unemployment

Once the basic intuition is grasped, it is convenient to abandon the picture of production as consisting of rigidly separate cycles and to admit, more realistically, continuous production and continuous scrapping. The scrapping of old plants causes a flow of 'freed' labour; new plants absorb a flow of new employment. The moment the possibility of unemployment is admitted, even with a constant employment in the already existing plants that continue to be utilised, the second flow can be smaller than the first, causing a gradual rise in unemployment, or greater than the first, with a gradual reduction in unemployment. The indeterminacy of labour employment in new plants implies that a given *K/L* ratio in new plants leaves investment indeterminate, as we saw above (see p. 49). This confirms the conclusion that even conceding the neoclassical conception of capital–labour substitution, income distribution is insufficient to determine investment, and therefore employment, too, is not determined. A given real wage (and corresponding real interest rate and normal relative prices) determines only the ratio *K/L* in new plants, it does not suffice to determine investment (and labour absorption) in new plants. As for employment in already existing plants, the rigid output–labour ratio implies that an assumption of decreasing marginal product of labour cannot be accepted, hence employment cannot be considered as determined by the real wage;[8] more realistically, employment will depend on output which will be determined by sales, and therefore, through the multiplier, by investment. And since the desired *K/L* ratio in new plants leaves investment indeterminate, there seems to be little alternative to considering investment as determined by the desire to reach normal-capacity utilisation, i.e., by the expected level and variations of demand.[9]

But before examining some implications of this view of investment, let us note how the above considerations destroy the neoclassical demand curve for labour. What emerges is that no incompatibility exists between a rise of real wages and a constant or increasing labour employment, *even accepting the neoclassical conception of capital–labour substitution*. Capital–labour substitution can operate only in new plants, and a greater *K/L* ratio in new plants implies a lower absorption of labour in new plants and no increase in employment elsewhere only if investment and the other autonomous components of aggregate demand remain constant or decrease. There is no reason why they should: the elasticity of output makes an increase of employment in existing plants perfectly possible if, for example, public expenditure, or investment, increases (in fact, as I argue later and have written elsewhere (cf. Petri, 2004: 320), investment *will* increase):

> the flexibility of production in response to changes in demand implies that there is no necessary influence, in the short as well as in the long period, of changes in real wages on the demand for labour. In existing plants, where capital already has a given 'form', higher real wages will bring about little or no change in output per unit of labour: employment will depend on capacity utilization which will depend on aggregate demand. In new plants, the flexibility of production of

> capital goods industries will generally pose no problem with obtaining the inputs required by the adoption of the new most profitable methods of production on the scale suggested by the expected level of aggregate demand, even if the latter is increasing considerably. Thus (apart from political reactions) there generally is no incompatibility between more employment and higher wages, all that is required is that the higher wages be accompanied by a stimulus to aggregate demand. This will be so even when it were the case that a higher wage implied a shift to more value-capital-intensive techniques and therefore required more savings: the increase in savings will be brought about by the increase in aggregate output. (Thus one might say, in neoclassical language, that owing to the adaptability of production to demand, relative factor proportions adapt to income distribution rather than the other way round.) (Petri 2011a: 411; 416, footnote 36)

Because of the above, empirical enquiries confirming that in most industries wages equal the marginal revenue product of labour would be no confirmation that the marginal product of labour *determines* real wages, because the causality must be understood to go the other way: owing to the adaptability of production to demand it will be prices and methods of production (i.e., the capital goods utilised by firms) that will adapt to a given real wage, so as to render the marginal revenue product of labour equal to the wage.

2.7 A decrease of real wages reduces investment

But – the neoclassical economist will object – all the above is based on *not* assuming the full employment of labour, and this can be at most a transitory state if the labour market is competitive: the decrease of real wages will increase the demand for labour. But will it really? I have just argued that the neoclassical decreasing demand curve for labour is destroyed by the analysis developed so far. So the effect on employment of a tendency of real wages to decrease in the presence of non-frictional unemployment must be examined anew;[10] and a readiness of workers to accept wage decreases as the normal answer to the emergence of unemployment will not be credible if, as I will argue, such decreases do not generally bring about an increase in employment even accepting neoclassical capital–labour substitution.

In existing plants, I have argued that labour employment depends on sales, not on the real wage; an increase of employment requires an increase of sales i.e., leaving government intervention and changes in the propensity to consume aside, an increase of investment. In new plants, conceding the neoclassical conception of capital, the decrease of real wages reduces the *K/L* ratio. Assume that investment is motivated by desired productive capacity and that the economy has been stationary for some time so initially entrepreneurs have little reason to expect anything but the same demand also for the next few years. Assume initially that already existing plants continue to be utilised normally. Then the new plants can only aim at satisfying the same demand that was satisfied by the scrapped plants they are replacing. Let us initially consider only the direct-substitution mechanism. The decrease of the *K/L* ratio in new plants planned for a given output corresponds to a shift on a given isoquant toward using more labour and less capital, hence *it reduces investment*. If aggregate demand did not decrease, this would not prevent an increase in the demand for labour in new plants (although a smaller increase than if *I* remained constant, see below) and a constant employment of labour in existing plants, and hence some increase in the total demand for labour. But *the decrease of investment reduces aggregate demand,* and then the assumption that the already existing plants keep being normally utilised turns out to be illegitimate, because the reduction of sales has a negative effect on employment in existing plants; thus even though the *flow* of employment in new plants increases, the overall *stock* (the level) of employment *decreases*.[11] Furthermore sooner or later the planned investment in new plants will be further revised downwards as expectations of unchanged sales turn out to be too optimistic; this further reduction of investment may well be small or even absent initially, but since *Y* remains lower than initially (its rise would need a rise in investment, while there is no incentive to such a rise) this will gradually persuade firms that they do indeed need a smaller productive capacity. Thus the decrease in wages starts a reduction in investment *and employment* that may continue for a long time.[12]

Now let us consider the indirect factor substitution mechanism. It is well known that this mechanism may not work in the direction needed by neoclassical theory, but neglecting for the sake of argument the possibility of 'perverse' income effects, the decrease in real wages changes the composition of consumption demand in favour of labour-intensive goods. The traditional derivation, from this change, of an increased demand for labour rests on an assumed unchanged total employment of capital, which in our framework where capital is putty-clay must mean an unchanged total investment. As in the direct-substitution mechanism, this assumption has no justification in view of the freedom with which investment can be decided. As in the other case, there is no reason for firms to expect future aggregate demand to be the greater one connected with *more* labour employment and an *unchanged* capital stock, since current aggregate demand is forthcoming from the income of the given capital and the not yet increased labour employment; only its composition is changing. The more plausible assumption is that the total value of expected demand for consumption goods is equal to the total current expenditure on them, and its changed composition corresponds therefore to a greater demand for labour and less demand for capital, that is, as in the direct-substitution case, less investment. Then the effect is the same as in the other case, a reduction of aggregate demand that causes a reduction of labour employment, with a likely subsequent further discouragement of investment.

2.8 Investment according to Dornbusch and Fischer

I am not the first to argue that even neoclassical theorists should admit an influence of expected sales on aggregate investment (in other words, a role for the accelerator broadly intended). This influence was indeed admitted in the first (1963) version of Jorgenson's 'neoclassical' approach to investment,[13] and it became the basis of the theory of investment in the popular macroeconomics textbook by Dornbusch and Fischer.

The basic idea of the approach of these economists was precisely, as I have argued, to take as given (expected) aggregate demand instead of labour employment in order to determine the desired capital stock and hence investment. Output is treated as if homogeneous and homogeneous with capital; then only the direct-substitution mechanism can be assumed to operate. The rate of interest selects the average capital–labour ratio on the aggregate isoquant corresponding to the planned level and composition of aggregate output; the desired capital stock changes if either the rate of interest, or planned output (i.e., expected demand), or both, change. Thus the desired capital stock is determined by the neoclassically determined capital/output average ratio, *and by the level of aggregate output*. A lower interest rate raises the desired K/Y ratio; with expected Y initially unchanged, the desired capital stock increases, although by less than if L, rather than Y, were kept fixed; the increase of the desired capital stock causes an increase of investment. Thus in the third edition (1984: 206–8) of their macro textbook Dornbusch and Fischer argue that, assuming a Cobb-Douglas aggregate production function $Y=L^{1-\gamma}K^{\gamma}$, the rental of capital (indicated as *rc*) causes a demand for capital K^* that depends on expected sales *Y*:

$$K^* = g(rc, Y) = \frac{\gamma Y}{rc} \qquad [2.1]$$

The role both of income distribution, and of *Y*, explains Dornbusch and Fischer's use of the term 'flexible accelerator' as an alternative denomination for what they also call the 'neoclassical approach' to investment. The approach of course requires the traditional and unacceptable marginalist conception of capital–labour substitution, and furthermore it is left with the problem of the speed with which the desired capital stock is reached when it changes discontinuously owing to a jump in the rate of interest;[14] but at least it avoids the frequent serious error, found in many current textbooks, of a derivation of the negative interest elasticity of the investment function from a given downward-sloping marginal-productivity-of-capital curve, forgetting that the marginal-productivity curve of capital needs a *given* labour employment, while the investment function is needed for the IS-LM model where labour employment is *variable*.[15] However, there is a price to be paid for avoiding this error: the consequences were alluded to in Section 2.7 and will now be explored further.

2.9 Criticism of the 'Keynes effect' mechanism

The view of Dornbusch and Fischer appears to have been that, since (if expected *Y* is given) the negative interest elasticity of desired capital and hence of investment obtains in their approach too, the 'neoclassical synthesis' criticism of Keynes is valid, a flexibility of money wages would ensure a tendency toward full employment. The well-known 'Keynes effect' mechanism at the heart of the 'neoclassical synthesis' relies on decreases of money wages in the presence of unemployment; this, according to Keynes, brings about some increase of employment in firms that initially expect to be able to sell more at a negligibly lower product price; this causes an excess of aggregate supply over aggregate demand since investment for the moment has not increased; the consequent decrease of the price level causes a decrease in the demand for money, hence a decrease in the rate of interest, hence an increase in investment. The same picture of how the tendency toward full labour employment operates if money wages are flexible downwards is obtained from Dornbusch and Fischer's textbook.

But their different approach to investment opens the way to a number of objections even if the neoclassical conception of capital–labour substitution is accepted without question.

First, the presence of an accelerator influence upon investment makes consideration of what has been happening to *Y* important. If, starting from a situation of desired capital–output ratio equal to the actual one, *Y* decreases for any reason (e.g. because of a decrease in exports, or in state expenditure) and remains low, then desired *K* is lower than actual *K*, and investment is discouraged; and this, through the multiplier, causes *Y* to decrease further, stimulating further decreases in desired *K*. The decrease in the rate of interest brought about by the 'Keynes effect' must then supply a *very* strong stimulus to investment to reverse this downward process. Such a strong stimulus cannot be expected, for two reasons. The first is that the increase in desired *K* is smaller than the one derived from the standard demand-for-capital curve, because the latter determines desired capital on the basis of a given employment of labour, while here firms move along a given (*K,L*) isoquant: this is shown in Figure 2.1 where the isoquant corresponding to a given *Y* is drawn, and a change in distribution that changes the optimal *K/L* ratio from α to β causes an increase of desired capital from K_1 to K_3 if labour employment is fixed at L_1, but only from K_1 to K_2 if output is fixed. The second reason is that the increase in the *K/L* ratio can be realised only in new plants, so it concerns only a very limited portion of productive capacity in every year. (The slowness of the change in the *K/L* ratio pointed out in Section 2.5 should not be forgotten: it is generally underestimated, owing to a mistaken

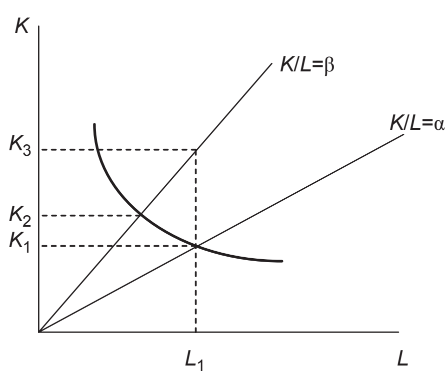


Figure 2.1 The change in the demand for *K* if *K/L* rises from α to β

tendency to conceive capital as putty-putty. Therefore the influence of Y on desired productive capacity and hence on investment has sufficient time to manifest itself.) Therefore even a neoclassical economist has little reason to expect the 'Keynes effect' to be more powerful than destabilising multiplier–accelerator interactions.

Second, the Dornbusch and Fischer approach implicitly recognises – in accordance with standard microeconomics – that the marginal products of the two factors labour and capital are tied together in such a way that if one marginal product increases, the other decreases, and that factor prices adjust to marginal products so that normal competitive extra-profits net of risk must be assumed to be (close to) zero when one studies investment.[16] This means that an increase in the desired *K/L* ratio will be associated with a change of relative factor prices consisting of a decrease in the real interest rate and an *increase* in the real wage. In order for the marginalist factor substitution mechanism to stimulate investment by raising the *K/L* ratio in new plants, the real rate of interest must decrease, i.e., *the real wage must increase*. On the contrary, the first stage of the 'Keynes effect' mechanism supposed to raise employment if money wages decrease consists of a *decrease* of real wages: firms raise employment and production because money wages decrease relative to prices that have not decreased yet; once prices start decreasing, since plausibly they decrease with some lag relative to the decrease of money wages, the real wage perhaps stops decreasing but remains lower than initially for all the deflationary period. As pointed out in Section 2.7, then investors have an incentive to adopt a *lower K/L* ratio in new plants, and this causes investment to *decrease*. To avoid this result, it would seem necessary that the decrease in real wages be strictly temporary, soon reversed by an *even greater* decrease in the price level (caused by prices rapidly adjusting to average costs including not only lower money wages but now also a lower rate of interest); then because of the rise in real wages the desired *K/L* ratio in new plants increases; this will hopefully stimulate investment, and *Y*, to the point of raising the demand for labour *in spite of* the rise in real wages. But note how one will then be admitting the possibility and indeed *necessity* of, at the same time, raising employment and real wages! Then it becomes difficult to deny that it must be the task of public intervention to secure such a result without the slowness and uncertainties of leaving it to the spontaneous workings of the market, which would in any case not be at all guaranteed to work in the required direction, because there is little reason to expect the necessary greater decrease in the price level to be sufficiently fast – firms are notoriously hesitant to decrease prices – and furthermore it is well known that price decreases raise the debt burden with possible negative effects on production and investment. (To all this one can add the well-known negative effect on the propensity to consume, and hence on the multiplier, associated with a decrease in real wages.) It is in any case striking that the rise in employment will have to be associated with a rise, not a decrease, in real wages. (Is this perhaps the reason why the Dornbusch and Fischer approach was not more widely adopted and was subsequently totally forgotten?)

2.10 A conclusion

These considerations should suffice to show how little one can trust that downward flexibility of money wages will reduce unemployment, even neglecting the Cambridge capital-theoretic criticisms, the moment one more consistently develops (full employment of labour not being initially assumed) the implications i) of the importance of durable capital; ii) of the inevitable influence of expected demand on investment; and iii) of the multiplier, and of possible multiplier–accelerator interactions broadly conceived.

To the above one must then add (i) the empirical evidence that consistently contradicts the presumption of a significant interest elasticity of investment; and (ii) the Cambridge results in capital theory, in particular the possibility of reverse capital deepening, that undermine the neoclassical conception of capital–labour substitution and show that the *theoretical* presumption of a negative interest elasticity of the demand for value capital per unit of labour has no solid foundation, so the lack of *empirical* support for such a presumption is not surprising. The conclusion must be that there is no reason at all to believe in a spontaneous tendency of market economies toward the full employment of labour.

2.11 Implications for wage flexibility and for growth theory

Two important implications of this conclusion can be pointed out.

The first is that the assumption that in the presence of unemployment money wages will decrease becomes implausible, and the Friedmanite thesis, that if in the presence of unemployment wages do not decrease then unemployment is voluntary, loses its analytical foundations. If reductions of wages have little or no effect on labour demand and can even have a negative effect, cumulative historical experience will have taught this fact to the labouring classes, ways will have been found to teach this knowledge and the consequent appropriate rules of conduct to the young, and it is then perfectly understandable that an unemployed worker will not, apart from exceptional circumstances, try to obtain a job by undercutting others. The generalised reduction in wages that wage undercutting would bring about would not reduce unemployment; it would only lower the incomes of employed workers – who often are the relatives of unemployed workers, and on whose income the living of the latter may depend. In such a situation it would be mistaken to define *unemployment* as voluntary: the absence of wage reductions is voluntary, but not unemployment. The unemployed worker, in refusing to accept a lower wage is *not* choosing the alternative 'no wage reduction, no job' over the alternative 'wage reduction, job'.

The second implication is the need to reconsider the theory of growth. The elasticity of output with respect to demand pointed out in Section 2.6 (see p. 53) strongly suggests a view of economic growth and capital accumulation as dependent on the evolution of the autonomous components of aggregate demand, because it implies that *aggregate* production can quickly adjust not only to decreases of aggregate demand, but also – within limits rarely approached – to *increases* in aggregate demand, so that it is generally possible, even in economies very close to full employment, to raise consumption *and* investment at the same time, if aggregate demand increases.[17] Hence investment is hardly ever constrained by savings; capital accumulation will result from the demand for additions to capital stocks due to increases in desired capacity, in turn due to increases of aggregate demand. A growing literature is developing these insights.

2.12 Implications for DSGE macro models

I conclude by briefly pointing out the relevance of the above analysis to the currently fashionable foundation of macro theory upon Dynamic Stochastic General Equilibrium (DSGE) models, where the problems for Say's Law pointed out in this chapter are pushed out of sight by an *assumption* of continuous full employment of the labour supply and of investment determined by savings. This assumption is generally justified by reference to the 'rigorous' microfoundations supplied by general equilibrium (GE) theory: the models are argued to be simplified renditions of the results one would derive from completely disaggregated intertemporal GE models, possibly made more realistic by the admission of adjustment costs, imperfect competition, and so on. The claimed premise of these models is therefore that intertemporal GE theory is a robust descriptive theory.

The curious thing is that the claimed consistency of this type of macro models with infinite-horizon GE theory is announced with pride, as supporting the trustworthiness of these models, while on the contrary more and more often general equilibrium specialists advance strong reservations on the descriptive validity of GE theory. One can mention Michio Morishima, Stephen Marglin, Duncan Foley and Alan Kirman as at one time convinced neoclassical theorists who have decidedly rejected GE theory. An implicit rejection or at least an agnostic attitude also emerge in the fact that the problems with uniqueness and stability have

> led many microeconomists to forsake the general equilibrium conceptualization altogether. As a result, microeconomic theory has, by and large, been reduced to a collection of techniques and tricks for resolving narrow, isolated microeconomic problems and the study of, also narrow and isolated, strategic behaviors. (Katzner, 2006, p. ix)

One can also mention the frequent denunciations by, for example, Frank Hahn or Franklin Fisher, of the sterility of stability studies based on adjustments that do not allow the implementation of disequilibrium decisions; but if time-consuming adjustments are allowed, the equilibrium becomes indeterminate because the data relative to the endowments of the several capital goods are no longer data, being altered by production. Also, many theorists are very uneasy about the utterly unrealistic assumption of complete futures markets or correct foresight; but the alternative of temporary equilibria without correct foresight, explored in the 1970s and early 1980s, is nowadays totally discredited (as evidenced by its complete disappearance from advanced micro textbooks), owing to the problems it encountered, which explains why Lucas, real business cycle theories, or DSGE models only refer to intertemporal equilibria as their 'rigorous' microfoundation.

One might then reject the DSGE approach in macroeconomics simply as a consequence of the rejection of intertemporal general equilibrium theory as a positive theory, a rejection motivated by this theory's need for the untenable assumption of complete markets or perfect foresight, by its lack of uniqueness or stability even granting the auctioneer, and by its inability to say anything on the distance between equilibrium paths and the behaviour of economies not continually perfectly in equilibrium (Petri, 1999: 50).[18]

But, as I have argued elsewhere (Petri, 1999: 53–4), it is difficult to understand the acceptance of intertemporal equilibria as descriptively valid without a more or less conscious belief that the undeniable occurrence, in actual economies, of disequilibrium and time-consuming adjustments does not destroy the neoclassical theses as to the *trend* the economy follows, which is reasonably approximated by the intertemporal equilibrium path. Only an idiot would deny that in actual economies there is no auctioneer and no complete futures markets, but rather time-consuming trial-and-error adjustments, mistakes, disequilibria, imperfect foresight; so DSGE theorists must believe that there are persistent forces that cause these disequilibria to be sufficiently corrected or compensated so that the trend the economy actually follows is not too far from the path described by their models.[19] But then the reference to disaggregated intertemporal equilibrium with perfect foresight as the 'rigorous' microfoundation of the models is only a smokescreen; behind it there is in fact a belief in the time-consuming adjustment mechanisms on whose basis the marginal approach was born and accepted, and that after Keynes were rehabilitated by monetarism carrying forward from the criticism of Keynes initiated by the neoclassical synthesis. Without some such belief the reference to intertemporal equilibria would be devoid of any justification, given that by themselves neo-Walrasian equilibria and their sequences tell us *nothing at all* about the actual path a market economy not continuously in equilibrium will follow.

For this reason, the arguments of the present chapter are relevant criticisms of DSGE models too, as well as of the whole development of neoclassical macroeconomics after Keynes. The characterisation of contemporary neoclassical macro models as simplified intertemporal general equilibrium models would, if taken seriously, deprive these models of any pretension to descriptive validity; such a pretension can only rely on traditional neoclassical macroeconomic tendencies, that is, on the same time-consuming adjustment mechanisms on which J. B. Clark or Wicksell or Pigou or Hayek, or the 'neoclassical synthesis' and monetarism, based their analyses. The relevance of the argument presented here, then, lies in its pointing out that, the moment the continuous full employment of labour is not assumed to start with, those adjustment mechanisms will not work as normally presumed, and Say's Law loses credibility even before one questions the neoclassical conception of capital–labour substitution: this was not realised because the correct implications of that conception for investment theory when labour employment is not given were not grasped. The recuperation of pre-Keynesian views initiated by the neoclassical synthesis and carried forward by monetarism, which is what lies behind the current faith in DSGE models, was made possible by a theory of investment which was mistaken not only in its foundation on an untenable conception of capital, but even if that conception is not questioned.

Notes

1. Of course I am not denying the possibility of a mutual influence (e.g. the propensity to save might depend on the rate of interest); but the basic idea of the theories I intend to criticise is that the aggregate amount of saving may well be given independently of investment, and then investment will adapt to it, while the converse (given investment, and savings adapting) does not happen; to insist on mutual determination misses this fundamental asymmetry.
2. A brief summary of my arguments is available in Petri, 2003, section 5.
3. I will not consider the attempts to defend a tendency of aggregate demand to adjust to aggregate supply on the basis of the Pigou (or 'real balance') effect: even Patinkin conceded that this effect is uncertain and anyway too weak. Cf. Petri, 2004, Appendix 7A2, pp. 292–5 for a confirmation of Patinkin's scepticism, based on recent estimates of the wealth effect.
4. I briefly recall and criticise here the two main attempts *before* Keynes to derive a negative interest elasticity of investment without having recourse to capital–labour substitution. In Walras future rentals of capital goods are treated as independent of the rate of interest, so the demand price of new capital goods (the discounted value of the given future rentals) rises when the rate of interest decreases, and this stimulates their production; but these given future rentals are an obviously indefensible assumption since the interest rate is one of the distributive variables and its changes alter the rentals of all other factors; as standard microeconomic theory teaches, product prices tend to minimum average costs, but then if the rate of interest decreases the rate of return on investments will tend to decrease too: the rentals of capital goods will decrease relative to their supply prices. The same objection applies to Irving Fisher, who assumes for each saver/investor a given series of alternative income streams among which the investor chooses – for each rate of interest – the one with the highest present value; prices are treated as given independently of the rate of interest, as in Walras. Thus Alchian (1955: 942) writes that Fisher's 'exposition … is based on the supposition that one merely changes the rate of interest and holds other prices fixed', and correctly accuses such a procedure of logical inconsistency. Actually Fisher admits that changes in the rate of interest alter relative prices, but he dismisses the need for further discussion of the issue by writing that this influence is 'a factor which, after all, is more intricate than important' (*The rate of interest*, p. 168), a statement for which no support is supplied. On Keynes on investment cf. Petri (2004: 260–2).
5. Investment must anyway be determined over not too short a period, to avoid its being relevantly influenced by transitional phenomena such as, for example, a *decrease* of the rate of interest inducing an expectation of a further decrease in a few months' time, and inducing therefore a postponement, i.e., a *decrease*, of investment; or anticipated scrapping of new plants induced by the change in prices due to a *rise* in the rate of interest, that may induce a temporary *increase* in investment.

6. The rest of gross investment – partial replacement, without alteration, of durable capital components of existing plants that are not scrapped, plus purchase of raw materials to be used in already existing plants – will be generally rigidly determined by intended outputs and by the technology embodied in the plants, and will therefore be independent of changes in income distribution except in so far as these affect the composition of demand; accordingly, it can be taken as given (and for this reason it is permissible to neglect it) as long as normal utilisation of existing plants is assumed.
7. That is, as long as normal-capacity output, the one associated with the normal utilisation of existing productive capacity, is less than is necessary to employ the entire labour supply.
8. On this issue it is worth noticing the agreement between the empirically based criticism by Dunlop and Tarshis of the decreasing marginal productivity of labour in Keynes, and the earlier theoretically based rejection of a short-period decreasing demand curve for labour by Hicks (1932).
9. Of course innovation will be another fundamental determinant of investment, but its effects do not seem relevant for a discussion of the validity of Say's Law.
10. Keynes objected that, unless investment increases and absorbs the increased saving associated with the increased output brought about by the greater employment, the decrease in real wages and increase in employment will not happen, because workers can only reduce *money* wages, and the insufficient aggregate demand will cause prices to decrease in step with money wages. This argument rests on an assumed decreasing marginal product of labour in the several plants, so if real wages were capable of permanently decreasing, the demand for labour *would* rise; I wish to question the robustness of the neoclassical argument *even conceding* a decrease of real wages. Below (Section 2.9) I discuss money wages.
11. Consider the following rough example. Suppose *I/Y* is constant, the average life of plants is 10 years, and the reduced *K/L* ratio causes *L/Y* in new plants to rise by, say, 5 per cent; the first year the increased hiring of labour in new plants as a percentage of previous labour employment is 5 per cent of 10 per cent, i.e., one half of one percentage point. If investment decreases by, say, 4 per cent, this causes a reduction of Y (and plausibly of employment in existing plants) by 4 per cent, i.e., about a 3.5 per cent reduction of labour employment.
12. This conclusion is reached without considering the negative effect on Y due to the generally admitted rise in the average propensity to save caused by the decrease of the share of wages in national income, an effect which, if admitted, would strengthen the argument.
13. On the difference between Jorgenson's 1963 and 1967 approaches cf. Petri (2004: 287–90).
14. This speed is determined by Dornbusch and Fischer through a 'gradual adjustment hypothesis' that states that the larger the gap between the existing capital stock and the desired capital stock, the more rapid a firm's rate of investment. Empirical evidence is then referred to in order to estimate the speed of adjustment.

15. This is just one of the many grave mistakes to be found in the arguments for a negative interest elasticity of aggregate investment after Keynes (cf. Petri, 2004: 271–81).
16. Without this reciprocal adjustment of factor prices, the desired *K/L* ratio would not be given by equation [2.1]. Obviously the extra-profits to be considered are the ones on new plants, existing plants earn quasi-rents. (I use 'extra-profits' to mean what standard microeconomics calls 'profits', in order to avoid confusion with the classical meaning of 'profits'.) On the need for such assumptions for the study of investment cf. Garegnani, (1983: 36; 46, fn. 25), also Petri (2011b: 67).
17. Labour constraints are usually not binding in the short run because of visible or hidden unemployment and underemployment, and over the longer run there are migrations, and structural social adaptations, e.g. changes in the participation of women.
18. A survey of these and other criticisms is in Petri (2011b).
19. Some such view is, for example, implicit in the numerous admissions by Lucas, Sargent and others that rational expectations make sense only for situations sufficiently persistent for agents to have had the time to learn how correctly to form their expectations – with the implication that during the learning mistakes are inevitable; but some learning is going on all the time, because of the continuous emergence of novelties (in each industry there may be technical progress, changes in tastes, etc.), hence those admissions imply that most markets are most of the time in disequilibrium, which can only be neglected if one looks at trends of the averages and one assumes that time-consuming adjustments operate which cause the trends to be sufficiently close to the equilibrium path.

References

Alchian, A. A. (1955) 'The Rate of Interest, Fisher's Rate of Return over Costs and Keynes's Internal Rate of Return', *American Economic Review*, 45 (December): 938–43.

Dornbusch, R. and Fischer, S. (1984) *Macroeconomics*, 3rd edn (New York: McGraw-Hill).

Garegnani, P. (1983) 'Notes on Consumption, Investment and Effective Demand', in J. Eatwell and M. Milgate (eds), *Keynes's Economics and the Theory of Value and Distribution* (New York: Oxford University Press); repr. with modifications of a paper with the same title published in two parts in *Cambridge Journal of Economics*, 1978, 2(4): 335–53; and 1979, 3(1): 63–82, in turn a translation of 'Note su consumi, investimenti e domanda effettiva', *Economia Internazionale*, 1964–5.

Garegnani, P. (1990) 'Quantity of Capital', in J. Eatwell, M. Milgate and P. Newman (eds), *The New Palgrave: Capital Theory* (London: Macmillan).

Hahn, F. H. (1981) 'Review of M. Beenstock "A neoclassical analysis of macroeconomic policy" (Cambridge University Press, 1980)', *Economic Journal*, (December): 1036–9.

Hayek, F. A. (1932) 'Money and capital: a reply', *Economic Journal*, 42: 237–49.

Hicks, J. R. (1932) *The Theory of Wages* (London: Macmillan).

Jorgenson, D. V. (1963), 'Capital theory and investment behavior', *American Economic Review*, 53(2), (May): 247–59.

Katzner, D. W. (2006) *An Introduction to the Economic Theory of Market Behavior. Microeconomics from a Walrasian Perspective* (Cheltenham, UK: Edward Elgar).

Petri, F. (1991) 'Hicks's recantation of the temporary equilibrium method', *Review of Political Economy*, 3(3): 268–88.

Petri, F. (1999) 'Professor Hahn on the neo-Ricardian criticism of neoclassical economics', in G. Mongiovi and F. Petri (eds), *Value, Distribution and Capital: Essays in Honour of Pierangelo Garegnani* (London: Routledge): pp. 19–68.

Petri, F. (2003) 'Should the Theory of Endogenous Growth be based on Say's Law and the Full Employment of Resources?', in N. Salvadori (ed.), *The Theory of Economic Growth: A 'Classical' Perspective* (Cheltenham, UK and Northampton, MA, USA: Edward Elgar): pp. 139–60.

Petri, F. (2004) *General Equilibrium, Capital and Macroeconomics* (Cheltenham: Edward Elgar).

Petri, F. (2011a) 'On the Likelihood and Relevance of Reswitching and Reverse Capital Deepening', in N. Salvadori and C. Gehrke (eds), *Keynes, Sraffa and the Criticism of Neoclassical Theory. Essays in Honour of Heinz Kurz* (Abingdon and New York: Routledge): pp. 380–418.

Petri, F. (2011b) 'On the Recent Debate on Capital Theory and General Equilibrium', in V. Caspari (ed.), *The Evolution of Economic Theory. Essays in Honour of Bertram Schefold* (Abingdon and New York: Routledge), ch. 4: pp. 55–99.

3
The Meaning of Output Trends in the Analysis of Growth

*Attilio Trezzini**

3.1 Introduction

This chapter discusses the meaning of the output trends examined in the analysis of growth and their relation to actual levels of output and income. These issues are seldom explicitly examined in the literature, where the relation between actual and theoretical magnitudes assumed in the theory of prices appears to be automatically extended to the magnitudes addressed by theories of growth. In field of economic growth, however, the distinction between actual and theoretical magnitudes overlaps with the distinction between cyclical and trend positions of the economy. This gives rise to crucial misunderstandings, at least from the standpoint adopted here.

Any particular notion of a *trend level of output* assumed in the theory of growth implies a corresponding notion for all the components of output, including the trend level of aggregate consumption, investment and saving. Specific meanings are then attached to the ratios between aggregates: trend values for the ratio of saving to income, of consumption to income or investment to income.

The chapter opens with a discussion concerning some features of the relation between theoretical and actual magnitudes assumed in the theories of prices and distribution. This introductory section prepares the ground for the critical discussion then developed of the method widely used in the literature, whereby the trend of produced quantities is studied through growth paths whose characteristics – level and rate of growth – are defined regardless of the quantities actually produced. As

* I am grateful to R. Ciccone and P. Trabucchi for useful and interesting discussions and to an anonymous referee for the comments on a first version of the chapter.

argued in the second section, this method derives from the simple and unjustified *transposition* of the relation between theoretical and actual magnitudes developed in the theory of prices into the theory of the evolution of quantities.

The viewpoint of the classical and Keynesian approach to the analysis of growth is then assumed in the third part of the paper and solid grounds are found for rejection both of this transposition and of the assumption that the trends are independent of actual fluctuations in output.

3.2 Theoretical and actual magnitudes in the analysis of prices and distribution

Our starting point is the relation between theoretical and actual magnitudes in the theory of prices and distribution, where this relation has been developed and discussed and where its general terms are therefore sufficiently clear. The relation between theoretical and actual magnitudes is addressed here with a distinction drawn between short- and long-period magnitudes or between actual and natural or normal magnitudes.[1]

The magnitudes examined by theoretical analysis, i.e., normal prices, are studied with reference to ideal situations in which it is assumed that no transitory and purely accidental phenomena affect prices and that the process of competition has been completed. Normal prices imply a uniform rate of profit on capital invested in various industries.

The determining forces behind these theoretical magnitudes are more persistent than those determining temporary deviations of actual from theoretical magnitudes. They are thus sufficiently lasting to allow, in principle, the adjustment of actual prices to normal prices. The latter are therefore seen as centres of gravitation of actual prices and hence able to account almost completely for the determination of relative prices.

Actual prices differ in fact from normal ones due to numerous circumstances affecting the composition of output, which proves inadequate with respect to the composition of effectual demand.

The circumstances are transitory in nature and more or less accidental. They can in fact be determined by an inadequate composition of productive capacity inherited from the past, by the persistence in some firms of old techniques and other phenomena that competition tends to eliminate. Further divergence of the quantities brought to the market from the corresponding effectual demand can be determined by more accidental phenomena such as the effect of a stormy night on the day's fishing. All these phenomena are transitory, however, being automatically cancelled out either by the adjustment of production levels through the process of competition or by an accidental phenomenon with the opposite effect.

Relative persistence is an essential property of the causes that determine natural prices, allowing the necessary adjustments to take place in the case of phenomena such as the composition of productive capacity or the presence of obsolete techniques, and allowing purely accidental phenomena – a stormy night – to be cancelled out by accidental phenomena with the opposite effects. This property is thus essential if the natural prices determined by these persistent causes are to be seen as centres of gravitation for actual prices and these theoretical magnitudes are therefore to be considered significant in the interpretation of actual economies.

3.3 Relevant features of theoretical prices

Some *features* of the relation between theoretical and actual magnitudes in this theoretical field should be stressed.

The conditions determining normal prices are largely *independent* of the possible levels of actual prices. They can thus be studied separately from the causes determining the tendency of actual prices toward natural prices and separately from analysis of the ways in which this tendency takes place.

Prices and distribution are determined in classical political economy on the basis of the technical conditions of production, the level and composition of normal real wages, the level and composition of the quantities produced (the conditions of reproduction), and the tendency towards a uniform rate of return on capital. This set of conditions is independent of the possible levels of actual prices and rates of return. Normal prices remain the same regardless of whether some actual prices are temporarily higher or lower than normal and regardless of whether a divergence lasts for a longer or shorter period.

This property also implies that during any possible process of adjustment or gravitation of actual to normal prices, the latter do not change. A process of adjustment is thus conceivable and normal prices are independent of it.

Normal prices have an *independent existence* with respect both to the phenomena determining divergences of actual prices and to the paths followed by actual prices in adjustment or gravitation.

From another viewpoint, this independence of the determinants of natural prices from actual prices can be understood as the fact that the effects of the causes of divergence of actual prices from natural prices are *reversible*. When the quantity of a commodity brought to the market is greater than the corresponding effective demand, the actual price proves lower than the natural price. This has no effect on the determination of the natural price, however, and the effects of the causes determining the divergence (an error in the quantity produced) are therefore reversible. Once these causes are cancelled out, they no longer affect prices.

Similar considerations can be developed with respect to marginalist theories, according to which long-run prices are determined by consumer tastes, technical conditions of production and the quantities of factors of production available. These circumstances determine equilibrium prices (and quantities) independently of actual prices.

The marginalist determination of prices and distribution determines equilibrium levels and composition of *output* simultaneously. The study of *gravitation* is, in this context, more complex and less independent of actual quantities with respect to the corresponding theoretical process of classical political economy. The independence of normal prices from the corresponding actual magnitudes and the paths of adjustment therefore requires additional assumptions, such as the existence of an auctioneer and that actual exchanges take place only at equilibrium prices.[2]

This similar conception of the relation between theoretical and actual magnitudes in classical and marginalist analyses of prices is combined, however, with radical differences.

As is known, a major difference between the marginalist and classical approaches is the simultaneity of the former as opposed to the method based on separate logical stages that characterises the latter. In addition to the fundamental implication of divergent positions with respect to the tendency towards full employment, this difference also has relevant implications of a methodological nature.

The simultaneity of marginalist analyses inevitably prompts the adoption of a uniform method in different fields of economic analysis. All the relations are regarded as sufficiently simple and general to be represented – at least in their fundamental aspects – by means of simple and general mathematical relations.

Conversely, a meaningful distinction between the method used in analysis of the relations determining prices and distribution – the *core* – and the one used in the analyses of output and accumulation – like all the relations *outside the core* – was envisaged by Garegnani (1984) in classical political economy. Being automatically enforced by the working of competition, the relations *inside the core* are comparatively simple and sufficiently general to be represented by mathematical relations – price equations – that are by nature necessarily simple and general.

The relations between magnitudes outside the core appear to be dramatically affected by circumstances of a historical, institutional and political nature that are at the same time *fundamental* and susceptible of being different in different periods or different economies. This makes the use of quantitative relations to represent their influence highly implausible if not indeed misleading. The use of such relations is limited to the construction of *models* or *examples* for strictly *heuristic* purposes.

3.4 Theoretical and actual quantities vs. cyclical fluctuations and trend magnitudes

The lack of an explicit definition of the meaning of the theoretical magnitudes addressed – generally output trends – in the theory of growth is particularly problematic because, in the field of the evolution of produced quantities over time, the distinction between long-run (persistent and fundamental) and short-run (transitory and/or accidental) phenomena affecting the determination of prices overlaps with another relevant and distinct phenomenon, namely the fact that economies generally develop over time through fluctuations in the levels of produced output that describe the trend of the economy.

Unlike the situation with prices, a large amount of data is available for aggregate output. Figure 3.1 shows the actual data for Italy from 1960 to 2012. Real aggregate output clearly displays irregular evolution over time with an increasing trend.

Also evident is the existence of recessions of a more or less lasting and intense nature that interrupt the uneven growth. Each 'cycle' thus has specific different characteristics: average rates of growth and different 'patterns of fluctuations'.

Let us first consider this picture independently of the literature. Treating the entire span as one long, uniform period seems only to prevent us from explaining why the economy grows more intensely in one period than another. This appears to be precisely the object of theoretical analysis of the determinants of economic growth.[3] It therefore seems reasonable to divide the period of observation into sub-periods marked by recessions, which can thus be defined as cycles. A trend can then be extracted from each of these cycles, as presented in Figure 3.2 (by means of simple linear regression or any other more sophisticated statistical procedure).[4]

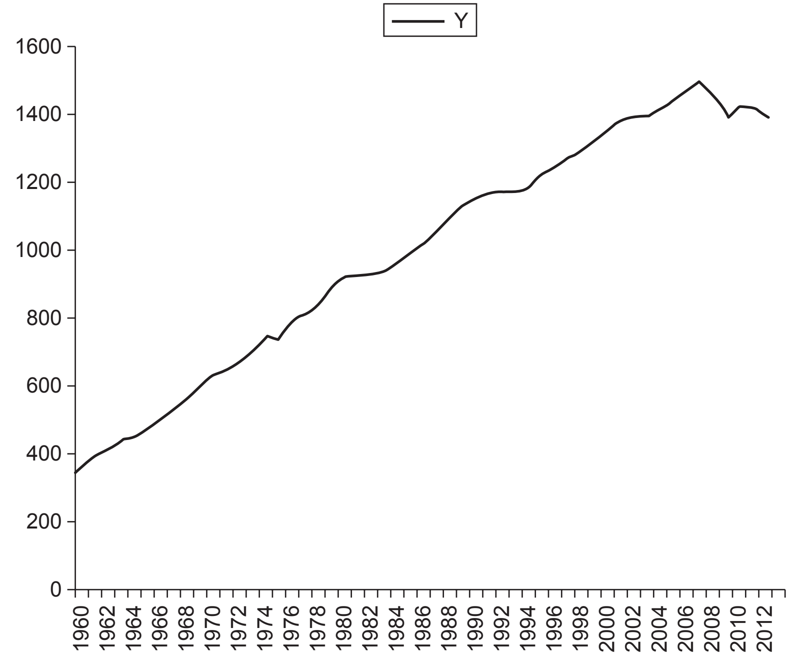


Figure 3.1 Real GDP (2000 prices), Italy, 1960–2012
Source: Ameco dataset.

These trends seem to be the phenomena requiring explanation. Why is economic growth more or less intense in one period than another? Why did the economy not grow more?

Figure 3.3 shows the actual data and their (linear) trend for real output of the first period into which it is plausible to divide the period of observation.

Any theory seeking to explain economic growth should be able to detect the major determinants of the phenomenon described in Figure 3.3. This necessarily entails a theoretical position as regards the determinants of the actual magnitudes represented by points like A in the figure, the determinants of the trend magnitudes represented by points like A*, and the relation between the two kinds of magnitude.

There is no explicit discussion of this methodological issue to be found in the literature on growth. What we are about to develop are therefore considerations about what appears to be the rationalisation of a method widely used in the literature. Apart from some exceptions

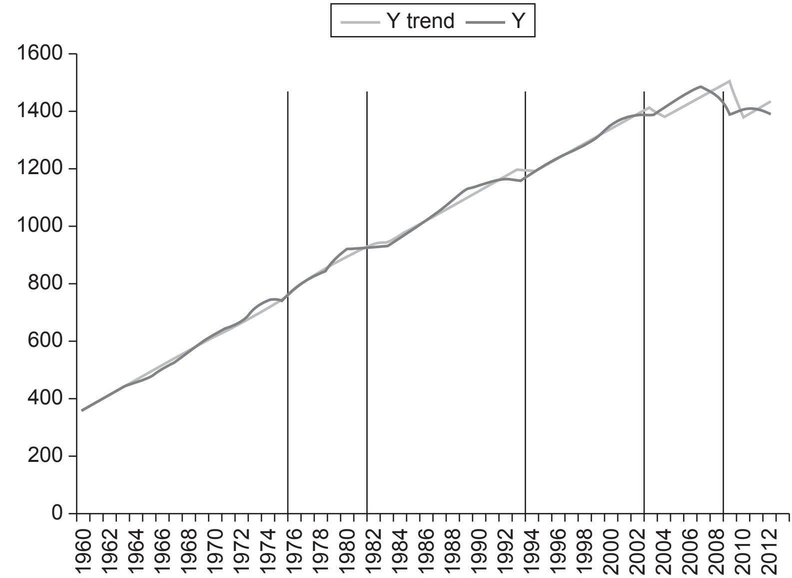


Figure 3.2 Actual data and linear trends for sub-periods: real GDP (2000 prices), Italy, 1960–2012
Source: Ameco dataset.

in both the mainstream (the Real Business Cycle) and Keynesian approaches,[5] what can be seen in most of the literature on growth is a tendency to study the determinants of the trend *independently* of the determinants of cyclical fluctuations. The causes of fluctuations are, at the same time, implicitly assumed as either *accidental* or *structural* phenomena that can in any case be studied separately and independently of the trend around which fluctuations occur.

This position seems to be based on automatic transposition of the method used in the theory of prices and distribution. The relations between the fluctuations of produced quantities (points A) and their trend levels (points A*) may in fact appear to be similar to the relations between short-period prices and the corresponding normal levels.

The assumption of the existence of trend levels of output independent of fluctuations in output is identical to the assumption of the existence of long-run or trend values for some *crucial* relations between quantities such as the capital–output ratio, K/Y, average propensity to save, S/Y, or investment–output ratio, I/Y, which can again be studied, and are actually determined in the models, independently of their actual values at different stages of the cyclical fluctuation.[6]

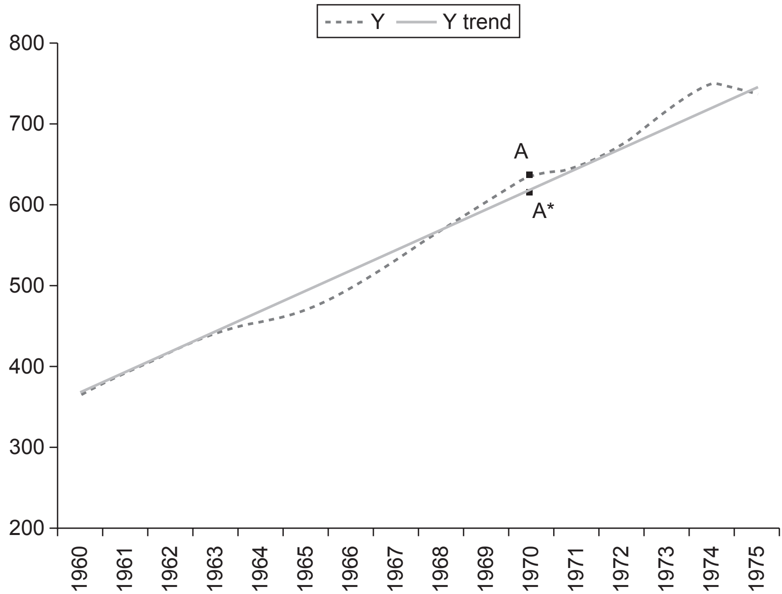


Figure 3.3 Actual data and linear trend: real GDP (2000 prices), Italy, 1960–1975
Source: Ameco dataset.

Our aim here is to consider this position, which is certainly plausible and widely accepted in the literature but may reveal some weaknesses on in-depth analysis.

3.5 The classical and Keynesian approach: the background

It is essential to state specifically that the argument presented in this chapter is developed within the framework of an approach to the analysis of growth in which the determining role of the evolution of aggregate demand is properly acknowledged. In particular, the approach adopted originates both from criticism of the marginalist idea of capital as a factor of production and from the classical surplus approach, defined as the classical and Keynesian approach to the analysis of growth.[7]

The critique of the marginalist concept of capital has been shown to be sufficient to deprive the mainstream theories of the most solid foundations for the alleged tendency towards full-employment equilibria. As such, this criticism is thus regarded as a 'second route' to the Keynesian principle of effective demand (Garegnani, 1978–79, 1983), the second in chronological terms but more solid from the theoretical viewpoint than the first opened up by Keynes himself. It is in fact completed by reference to the surplus theory of value and distribution, which is compatible with the principle of effective demand and thus makes it possible to extend the Keynesian premises to analysis of the long-run tendencies of economies.

These premises thus become the basic element for a classical and Keynesian approach to the analysis of growth and accumulation. The level of aggregate demand is conceived as the primary determinant of the level of output in any one period. It also determines the degree of utilisation of existing capacity and through it the evolution of capacity itself. Changes in aggregate demand generate changes in output through immediate changes in capacity utilisation, which lead in turn to the creation or destruction of capacity through higher or lower (nil or even negative) flows of investment. The latter in turn have effects on both demand and capacity. One central concept formulated is the *high elasticity* with which output responds to changes in aggregate demand in both the short and the long run. While this elasticity is determined in the short run by changes in the utilisation of capacity that allow different levels of production, the elasticity of output is made possible in the long run by changes in the degree of utilisation of capacity and by the consequent creation of new capacity.

3.6 The rigidity of output to demand and definition of the methodological issue

The concept of the elasticity of output with respect to changes in aggregate demand has a crucial bearing on the issue discussed here. The independence of the trend of output from cyclical fluctuations is in fact generally asserted on the basis of the assumption of the rigidity of output with respect to changes in aggregate demand. It is possible to identify two kinds of rigidity assumed in the literature.

The first is the rigidity asserted in the mainstream theories of growth due to the principle of the tendency towards full-employment equilibria. This principle leads to denial of any elasticity of output with respect to changes in demand and with it any role of aggregate demand in the growth process. In the context of mainstream analyses of growth, the tendency towards full employment is also sufficient to select a single path of evolution of output independently of any phenomenon determining fluctuations (whether accidental or systematic) superimposed on the tendency towards full employment. The full-employment trend of produced quantities therefore has an existence independent of fluctuations. This trend is also unaffected – at least in principle – by the actual positions of the economy during the cycle.[8]

The principles leading to the tendency towards full employment then lead to the *transposition* of the relation between theoretical and actual magnitudes developed in theories of prices into the analyses of growth. This transposition is consistent with the characteristic of marginalist theories that they are theories of prices, distribution and equilibrium quantities all at the same time. This simultaneity naturally and necessarily leads to extension of the same conception of the relation between long- and short-period prices to the analysis of relations between trend levels of output and their cyclical fluctuations.

The second is the rigidity assumed in most of the Keynesian growth models together with the assumption of steady-state conditions. In the analyses based on the Cambridge equation, the constancy of the ratio between capital and output is assumed as a condition defining the analysis of growth, and this assumption leads to steady-state paths being adopted as the object of analysis. This assumption is closely connected with the need to address the Harrod's instability, as discussed below, but means at the same time that the trend of the economy is conceived as independent of cyclical fluctuations in output.

The principle of the long-run elasticity of production with respect to changes in the levels of demand on which the classical and Keynesian analysis is based lies at the very root of the rejection of post-Keynesian analyses of growth. According to the authors of the classical and Keynesian approach, the denial of the elasticity of output to demand is incorrect and the role attributed on the basis of it to changes in distribution in the adjustment of savings to investment is to be rejected as misleading.[9]

3.7 Output trends and cyclical fluctuations in the classical and Keynesian approach to the analysis of growth

In the literature based on the classical and Keynesian approach to the analysis of growth, the definition of the theoretical magnitudes involved is discussed no more explicitly than in the other Keynesian approaches. This definition and the relation between the output trend and cyclical fluctuations thus appears controversial.[10]

Here we attempt to consider critically the possibility of assuming that the trend of output is independent of its cyclical fluctuations, as part of this literature does. Some elements of an alternative way of addressing the issue are then put forward. These are of course very much preliminary considerations, which may be regarded as elements for debate and discussion.

Our first step is to put forward what may be considered the most plausible rationalisation of a method based on the independence of the trend of output from its cyclical fluctuations.[11]

With respect to Figure 3.3, we first assume that the aim of the theory of growth is to explain the determination of points like A*. It is therefore necessary to see whether the relation between the actual magnitudes (point A) and the corresponding trend magnitudes (point A*) is such as to permit separate study of the determination of the trend and of divergences from it. As argued, this is the case for the actual/short-period levels of a price $\boldsymbol{p}$ and its theoretical value $\boldsymbol{p^*}$ when the viewpoint of classical political economy (and the original formulations of marginalist analyses) is adopted.

In order to separate the determination of trend, it is necessary to assume the existence of forces causing the economy to grow smoothly. The existence of long-run growth rates of the determining variables must be assumed. Theoretical mechanisms of adjustment permitting corresponding – and almost necessarily uniform – growth of the other variables should then be envisaged. This set of circumstances determines the trend of output.

Actual fluctuations around this trend should then be regarded as due in part to accidental phenomena (either temporary obstacles such as the composition of productive capacity or purely accidental phenomena represented by random variables with zero-mean values). These fluctuations appear to be of the same kind as the divergences of the actual levels of a price $\boldsymbol{p}$ from its theoretical value $\boldsymbol{p^*}$ and can therefore be overlooked completely in analysis of the long-run tendencies of the economy.

Fluctuations in output are generally seen, however, as the joint result of random phenomena and *structural* phenomena inherently endowed with cyclical behaviour.

The determinants of fluctuation patterns have in fact been examined by the theory of business cycles, in which many different theoretical explanations have been put forward. Monetary phenomena, inventory cycles, fixed investment cycles and political cycles are just some of the numerous principles suggested in this broad field of analysis to explain the fact that economies develop through fluctuations.

In order to separate the analysis of cyclical fluctuations from that of the trend of output, it is necessary to argue that the phenomena determining the irregularity of the evolution of output are in any case independent of those determining the trend. This appears to be the conception implicit in most of the large number of multiplier and accelerator models of the 1950s and 1960s and, more generally, in all the Keynesian analyses that adopt a steady-state path to represent the output trend.

The relation between the trend (A*) and cyclical/actual magnitudes (A) is assumed to possess the property identified as essential in the relation between normal and actual prices. The phenomena that determine those normal values on the one hand and trend magnitudes on the other are assumed to be *independent* of the phenomena determining the divergences from their theoretical or long-run counterparts of *actual* magnitudes on the one hand and *cyclical* magnitudes on the other. Though not absolute, the independence of the two kinds of phenomena must be sufficient to justify separate determination of the two sets of causes.

The trend is thus regarded as having independent existence and can be described as an autonomous entity with respect to actual fluctuations.

While this view can probably be challenged in any theoretical context, it does not appear to be sufficiently solid in the one assumed here, which correctly acknowledges the determining role of aggregate demand in the growth process.

3.8 An initial argument in favour of path dependence

When the role of aggregate demand in the growth process is acknowledged, the trend level of produced quantities does not appear to exist independently of the way in which it actually occurs.

As regards normal prices, it may be argued that they are not affected by the actual levels of market prices insofar as they are independent of the causes determining the divergence of actual prices. Longer periods of adjustment and greater differences (5 per cent rather than 1 per cent) between actual and normal prices are not regarded as capable of affecting normal prices, which continue to be determined by the same persistent and structural causes. A storm of two or three nights leading to a longer and more marked divergence of the actual price of fish does not alter the actual conditions of reproduction of fish.

The same reasoning does not seem to hold for the trend level of aggregate produced quantities.

Longer and more intense expansion appears to involve a number of phenomena determining the average rate of growth. Actual levels of aggregate demand that are more or less high for longer or shorter periods determine incomes that give rise to expenditure affecting the average level of aggregate demand. The same phenomena imply investments that affect not only demand but also capacity. These phenomena cannot be considered either transitory or irrelevant for the determination of output, which is the phenomenon addressed here. It therefore appears that longer and more intense expansions and shorter and less intense recessions are the way in which a steeper trend of growth is determined.

Moreover, the pattern of fluctuations affects the trend of the economy by generating phenomena that more or less *persistently* affect the determination of the levels of demand in the following periods. During a more intense or longer expansion, larger amounts of private wealth are accumulated and the incomes of households remain higher for a longer period, during which higher standards of consumption may be irreversibly acquired. A shorter period of expansion would prove insufficient to consolidate such standards. Symmetrically, longer and more intense investment booms imply that new additional productive capacity comes into existence, and this certainly affects both the possibility of producing and the decisions to invest in the following periods. This appears to be radically different from what happens when actual prices are higher to some degree over a more or less long period than their corresponding normal level.

Many phenomena generating *irreversible* effects become crucial when analysis focuses on the determination of output and its evolution. This makes it difficult to assume that a longer or more intense expansion has no effect on the trend of the economy.[12] A more detailed analysis of reality could result in a very long list of these irreversible effects.

Simple transposition of the relation between actual and normal prices into the theory of growth and accumulation does not appear to provide a solid basis for the reconstruction of an analysis correctly recognising the role of aggregate demand in growth.

In this context, the trend of output should probably be regarded simply as concisely representing a phenomenon that occurs through irregular periods of expansion and contraction. In this sense, it is similar to a simple average providing a concise indicator of how a phenomenon manifests itself in a population with different features. Neither the trend of produced quantities nor the statistical average exists independently of the variability through which the phenomena actually occur.

Conversely, this obviously implies that the features of the cycle are not independent of the growth process. They cannot be understood simply as random phenomena of greater or lesser intensity or a more or less intense manifestation of structural phenomena such as the inventory cycle or fixed investment cycle and so on. On unprejudiced observation of reality, they appear as the only way in which higher or lower *trend* rates of growth can occur, i.e., as the manifestation of the very causes of growth.

It should be stressed that the reference to classical political economy makes the classical and Keynesian approach more flexible by comparison with approaches that assume the tendency towards full employment in the long run. In the reappraisal of classical political economy, as stated above, particular importance is attached to the distinction between the method used in the *core* of the surplus theories and the one used in the analysis of relations *outside the core*. This distinction entails no necessary symmetry in the conception of the relation between theoretical and actual magnitudes in contexts of analysis that are actually different and separate.

3.9 The question of capacity adjustment

The transposition of the relation between actual and natural prices into the relation between the trend and cyclical fluctuations of output has also been fostered by another and closely connected question. The recognition of the tendency of productive capacity to adjust to aggregate demand found in most of the Keynesian approaches leads to the adoption of *fully adjusted positions* as the only ones of any relevance for long-run analyses.[13]

Another apparent symmetry is that the tendency of capacity to adjust to demand appears similar to the tendency towards a uniform rate of profit. Just as natural prices can only be associated with conditions in which the forces leading to uniformity of the rate of profit are 'at rest', the theoretical trend of the economy must be characterised by a normal degree of capacity utilisation. Any different degree of utilisation would in fact result in changes determined by the tendency to adjust capacity to demand, which would make a trend characterised by a non-normal degree of utilisation implausible as a representation of the primary forces of the growth process taking place through cyclical fluctuations.

This consideration lies ultimately at the origin of the crucial role played by Harrod's instability in the theory of growth.[14]

In the context of the Keynesian approaches to the analysis of growth, the principle of the tendency of capacity to adjust to demand has played a role similar to the one played by the tendency towards full employment in mainstream theories.

At the cost of drastic simplification, it is possible to state that in mainstream theories, due to the principles entailing the full employment of resources, given the initial amount of productive resources, the propensity to save and the production function (with all the possible changes determined by exogenous or endogenous technological changes), the path of growth around which fluctuations can occur is almost completely determined.

In the Keynesian approaches, the principle of the tendency of capacity to adjust to demand, insofar as it is interpreted as prompting us to associate normal utilisation with the positions representing the trend of output, leads to the conclusion that given the initial endowment of capital stock (productive capacity) and the ratio of saving to income, the path of growth is almost entirely determined. Different theories assert different roles of *aggregate demand* or its components in affecting growth by determining different ratios of investment to income and the corresponding ratio of actual savings to income.[15] The common assumption of normal utilisation of capacity (either constantly or on average over the business cycles) justifies the study of the trend independently of cyclical fluctuations in these approaches.

As is known, however, it has been suggested that the tendency of capacity to adjust to demand does not necessarily mean that normal utilisation is a condition to be presupposed in analysis of the long-run output of the economy.

An initial result in this direction reached by Ciccone (1986) is that the gravitation of market prices towards their normal levels does not necessarily entail full adjustment of capacity to the *level* and composition of effective demand. Normal prices may prevail when produced *output* has the right composition with respect to effective demand but aggregate productive capacity may be not normally utilised, even on the average over the fluctuations.[16] This principle was developed by challenging the validity of the Cambridge theory of distribution. Since then, however, other more general arguments have been developed in favour of the view that the economy does not necessarily tend to attain normal utilisation of capacity even on average over business fluctuations.[17] It would thus follow that it is misleading to study growth by means of paths characterised by normal utilisation even when it is implicitly assumed that this is not a constant condition, as in the steady-state models, but a condition arising on average over cyclical fluctuations, as in the supermultiplier analyses. This method would in fact lead to assuming as general relations between variables that may hold solely in very particular conditions, e.g. when demand just happens to grow constantly and homogeneously at a rate compatible with the normal utilisation of existing capacity.

The arguments against the idea that economies tend towards growth in conditions of normal utilisation include the fact that the process of adjustment of capacity to demand has effects on aggregate demand and its evolution. This adjustment occurs through investments that create both productive capacity and additional aggregate demand. It is thus very hard to argue that the demand effects of investment during this adjustment do not affect the position towards which the economy should converge. As discussed above, it is hard to maintain the necessary property of *independence* of theoretical positions with respect to the fluctuations involved in adjustment.

The process of adjustment of capacity to the evolution of aggregate demand therefore appears to be longer[18] and more uncertain than that of the adjustment of prices to their normal level.

Particular attention should be focused on the analyses based on the classical and Keynesian approach that share with the other Keynesian analyses this view of utilisation and the independence of trends from cyclical fluctuations. The analyses based on the supermultiplier assume that given the initial level of productive capacity, the capital/output coefficient, the propensity to save and the – regular and constant – growth rate of 'autonomous demand', it is possible to select a path of growth that will prove to be the trend around which cyclical fluctuations occur. These analyses regard the rate of growth (which is assumed to be constant) of autonomous demand $\mathbf{g_a}$ as the determining force and the whole process of growth is pegged to it. Capacity, output and all the components of aggregate demand are assumed to grow in accordance with this rate of growth.

The debate on the supermultiplier showed that this rate of growth of autonomous demand must be implicitly regarded as warranted (compatible with the existing productive capacity) and must remain the same forever. Any value of the rate determined independently of the level of existing capacity and/or any change in this (warranted) rate of growth would imply under/over-utilisation and a process of the adjustment of capacity to demand whereby the warranted rate adjusts to the autonomously determined (or altered) rate $\mathbf{g_a}$.[19]

It has also been pointed out just how essential the notion of 'relative persistence' is in this context.[20] In the process of gravitation of prices, it is possible to assume that the conditions determining normal prices are comparatively persistent with respect to the speed of the adjustment process. This is another feature of the relation between theoretical and actual magnitudes in the theory of prices that cannot be attributed to the relation between the trend and actual fluctuations of produced quantities. As noted above, the adjustment of capacity to aggregate demand appears to be so long, uncertain and complex that it becomes necessary, in order to argue in favour of a complete adjustment, to assume that the rate of growth $\mathbf{g_a}$ remains constant for an implausibly long period.

3.10 The intrinsic irregularity of the evolution of final demand components

A further point to be taken into consideration is that the view of the relation between trend and cyclical fluctuations under discussion here hinges on the possibility of assuming the existence of rates of growth of autonomous demand components that are themselves independent of the actual/cyclical levels of output. This is another assumption that appears far from plausible.

Examination of the evolution of the components of this final demand reveals that, by their very nature, they display an irregular pattern over time. For most of them, it is difficult to imagine a regular long-run rate of growth around which mainly random fluctuations occur. The asymmetric behaviour of this magnitude has been noted in studies on the expansion of aggregate consumption and taken as a crucial feature determining the evolution of the part of consumption that cannot be regarded as directly dependent on current income.[21] An analysis of government expenditure would certainly lead to the identification of cyclical and anti-cyclical components together with components for which a regular trend could be postulated. Different considerations may plausibly lead to similar conclusions as regards the components of aggregate demand related to international trade. The components of autonomous demand connected with the introduction of technical innovation would require yet another set of considerations, which could lead in any direction. Even though the analysis in this direction is far from being sufficiently developed, when all these components are considered without presupposing a methodology based on the assumption of a long-run constant rate of growth (necessarily uniform with the rates of growth of all the other magnitudes), it seems hard to argue that they tend to grow at a regular rate around which random fluctuations occur and thus generate a trend that can be studied independently of cyclical fluctuations. A growth rate of each of these magnitudes that proves to be more or less high on average over a period of time seems to be determined by longer and more intense expansions and shorter and less intense contractions of the corresponding magnitude.

3.11 The need to reconsider Harrod's instability

The critical considerations about the obstacles to full adjustment seem to make the assumption of an economic trend independent of the actual positions unnecessary and even misleading. They also clash, however, with the pressing need to focus on the theoretical 'adjusted' positions deriving from the idea of Harrod's instability. It is therefore necessary to examine the logical origins of this concept in order to evaluate its actual solidity. While this issue cannot be discussed here in detail for reasons of space, some relevant points can be recalled.

The concept rests on the assumption that any divergence between actual and desired utilisation generates an adjustment of capacity. It has been pointed out, however, that there may be different ways in which over- or under-utilisation can occur without creating any need to adjust capacity.[22] Moreover, it is possible to argue that the desired degree of utilisation can be affected by the same circumstances that generate increases in the actual utilisation;[23] developing this argument it is possible to argue that any divergence between the two tends to be much less relevant in determining instability.

This notion of instability is based on a conception of capital as qualitatively homogeneous and continuously divisible. This makes it possible to conceive the tendency of capacity to adjust to demand through a process of growth or decline of this homogeneous material, i.e., by increases or decreases in the rate of capital accumulation. Unprejudiced observation of reality would show that an adjustment of capacity to demand through an increase or decrease in the size of existing plants is an extremely rare phenomenon. Capacity generally adjusts through the radical destruction of entire plants and/or firms or the creation of new plants and/or firms. An individual plant may be under- or over-utilised all through its economic life and then disappear without its under- or over-utilisation ever being corrected by decreasing or increasing its size.

For these reasons it is possible to imagine under- or over-utilisation that lasts for many years (and even increases) and is corrected during a recession without making any difference – or only a limited difference – to the trend rate of accumulation.

Our scepticism about this idea of the adjustment of capacity to demand is strengthened when technical change is taken into account. The inadequacy of given productive capacity with respect to the demand for which it was installed may manifest itself when the same capacity is economically obsolete. In this case too, the inadequacy of capacity would not be corrected by an increase or decrease in size but by the creation of new plants.

The assumption of a general tendency of productive capacity to be determined by aggregate demand evolution does not ultimately appear to entail actual adjustment realisation (even on average over fluctuations). Theoretical output trends characterised by degrees of utilisation differing from normal on average and implying increasing or decreasing capacity utilisation over a period of time thus appear to be much less inconsistent than the concept of Harrod's instability makes them appear.

3.12 Conclusions: the outline of a possible alternative method of analysis

When the classical and Keynesian approach to the analysis of growth is assumed, the validity of grounding the analysis on paths of growth determined independently of actual positions of the economy seems to be challenged by the considerations developed here. This seems to be true despite the broad consensus on the use of this method. In the classical and Keynesian approach, this consensus may be based essentially on two elements: a) the transposition of the relation between theoretical and actual magnitudes assumed in the theory of prices and distribution into the analysis of the evolution of the quantities over time; and b) the symmetry between the tendency towards a uniform rate of profit in the theory of price and the tendency of productive capacity to adjust to the evolution of aggregate demand.

It has been argued that these elements are not sufficiently solid. In particular, in the context of demand-led growth, it does not seem to be possible to attribute the theoretical trend magnitudes with the property of independence from the actual magnitudes occurring in fluctuations, a property of independence similar to that attached to normal prices with respect to actual ones in classical theories of prices and distribution. The adjustment of capacity to demand has been shown to be not only longer and more complex and uncertain than the tendency towards uniformity of profits but also insufficient to guarantee the independence of the theoretical magnitude from the path of adjustment.

Let us now outline an alternative position emerging from these criticisms. Economic growth is a phenomenon consisting in the evolution of produced quantities that is likely to occur in an irregular manner, i.e., through fluctuations with different patterns. The study of the trend or of any average path of growth may not be separated from analysis of the way in which the relevant variables fluctuate in generating the trend, which is generally the object of theories. In this context, the trend has no *existence* independent of the actual irregular evolution of the variables. It is therefore a concise representation of a more complex occurrence. Growth theory should therefore investigate the change in variables and explain why, in a growing economy, output levels tend during expansion to reach levels higher than those achieved in the previous expansion and/or why recessions are interrupted at levels of output higher that those achieved in previous troughs. Symmetrically, growth theory should be able to explain why this does not occur and the opposite can even take place when the economy is not growing or is in decline.[24] Among other things, the actual evolution of produced quantities determines the actual rate of utilisation of productive capacity.[25]

Notes

1. Classical political economy and marginalist theories adopted different terms for the same concepts, partially as a reflection of differences in theoretical principles. We choose not to repeat or clarify the different definitions of the terminology used here – natural vs. normal or short-run vs. actual magnitudes – as this would only make the exposition more cumbersome without improving the completeness or solidity of the arguments presented.
2. This assumption is needed because the occurrence of exchange at other than equilibrium prices would entail a change in actual endowments and hence in the data determining equilibrium positions. See Bilancini and Petri (2008).
3. Though paradoxical, the approach of regarding a very long and heterogeneous series of time intervals as a single long period has actually been adopted in the literature more than once. It can be detected in some influential works of the 1950s, when Friedman (1957) and Modigliani and Brumberg (1954) studied the long-run propensity to save on the basis of Kuznets's data. It can also be seen in Kaldor (1957), who attempted to extract some *stylised facts* from studies of economic history to serve as driving features for the theory of growth.
4. A linear trend has been extracted through the linear interpolation of actual data. It is worth stressing that exponential interpolation should be used in order to study paths characterised by constant rates of growth.
5. References for these exceptions are given in footnotes 8 and 24.
6. This is another automatic transposition of a feature of the analyses of prices, where the relations between the variables determined by the theory – between normal prices and rate of profits or between quantities and prices – are in fact assumed as those prevailing in the long run. These relations are independent of the relationships between the same variables that occur transitorily during fluctuations of prices and rates of return.

7. This approach was originally sketched out in Garegnani (1962) (a work for Svimez, an Italian research centre for the development of southern Italy) and then developed in Garegnani (1978–9 and 1992). Various authors have since continued along the path opened up by Garegnani.
8. It is worth stressing that the assumption of the independence of the trend from actual positions of the economy has been also questioned in the *marginalist* debate on the business cycle. Particular significance attaches in this connection to Haberler (1938), esp. pp. 31–53. An important exception to the identification of full employment with the assumption of a trend determined independently of cyclical fluctuations can be found in the more recent versions of full-employment theories known as Real Business Cycle theories, where cycle and trend are studied as related phenomena both determined by stochastic shocks that affect full-employment equilibria. Even though the rigidity of output with respect to changes in demand is then asserted, this does not lead to the assumption of a trend independent of cyclical fluctuations. See Bronzi (2012).
9. See Ciccone (1986), Garegnani (1992), Vianello (1985). For the particular way in which the issue of steady state and the elasticity of output with respect to changes in aggregate demand is present in the neo-Kaleckian literature, see Trezzini (2011b).
10. The debate on the supermultiplier can be interpreted as the result of implicitly different conceptions of this issue. No explicit statement of the point is to be found, however, on either side.
11. As pointed out above, in the classical and Keynesian literature as well as most of the analyses of growth, it is virtually impossible to find an explicit treatment of the meaning of the trend of output. It is therefore unfair to attach this rationalisation of the method used directly to supermultiplier analyses. This chapter may also help to make explicit any other implicit rationalisation of the method by the authors involved.
12. The existence of irreversibility certainly matters when the viewpoint of demand-led growth is assumed and is not demonstrably irrelevant in supply-side analyses of growth. See n. 8 for traces of this problem in marginalist theories of the cycle.
13. *Fully adjusted positions* are to be understood as those in which the tendency of capacity to adjust to demand has been fully realised and the degree of capacity utilisation proves normal.
14. Once any divergence of the actual rate of growth from the warranted rate (of actual utilisation from normal) is seen as generating cumulative expansion or deflation, a path characterised by normal capacity utilisation appears the only one that can be taken as the theoretical trend of an economy. All the theories of growth have in fact attempted since Harrod's contribution to develop mechanisms capable of making the warranted rate of growth equal to the rate determined by the specific theory. In any case, however, the coinciding of the actual and warranted rates of growth is seen as a necessary condition. The exception is the case of the neo-Kaleckian models briefly discussed above.
15. At the cost of the simplification involved in any synthesis of different theories with important elements in common, we can say in post-Keynesian theories, a different rate of accumulation directly affects the ratio of actual

saving to income by altering the distribution of income. In neo-Kaleckian theories, different rates of accumulation directly affect the long-run degree of utilisation and thus determine differences in profits and savings. In the analysis based on the supermultiplier, finally, different rates of growth of autonomous demand indirectly affect the share of the amount of saving corresponding to normal capacity utilisation (capacity saving) absorbed by autonomous demand and directly affect the share of it capable of financing investment and capacity growth.

16. The composition of output may be attained with productive capacity adjusted in composition but not in level or in conditions in which new capacity has the right composition but works together with old capacity that produces output by means of obsolete techniques and under- or over-utilisation, thus generating 'quasi-rents'. Competition stops working when the rate of profit is uniform over the flows of new investment and not necessarily over the whole stock of capital.
17. Garegnani (1992, p. 58) argued in this sense as follows: 'What becomes clear in this connection is something which was perhaps not immediately evident, namely that the possibility that investment should generate a corresponding amount of savings through changes in aggregate productive capacity (the Second Keynesian Position) is inconsistent with the assumption of an economy working over time at the desired level of capacity utilisation, even only on an average taken over booms and slumps.' While this implication is evident in the extremely simplified case studied by Garegnani in which there is no autonomous component of aggregate demand, it is *only apparently* overcome – but does not completely disappear – when autonomous demand is assumed to exist (see Trezzini, 1995).
18. The length referred to is a theoretical one. What matters is the persistence of the determining forces (the data) with respect to the speed of the adjustment of the actual magnitudes to the theoretical ones. In the case of prices, the relevant adjustment of capacity to demand is limited to the new capacity created by investment flows. In the case of quantities produced, it is the whole stock of capital that has to change in composition and level. In this sense it not only appears longer in absolute terms but also is longer in relation to the persistence attributable to the determining forces: the rate of growth of accumulation or of autonomous demand. These determining forces do not appear to be persistent with respect to the length of the former adjustment, not least because they are not independent of the process of adjustment itself.
19. Trezzini (1995).
20. Trezzini (1998).
21. See Garegnani and Trezzini (2010) and Trezzini (2011a).
22. Trezzini (2011b).
23. Garegnani (1992) p. 56 and White (2006) pp. 156–58.
24. This 'path dependence' is essentially the same as that theorised in the 1950s and 1960s by authors working in a Keynesian framework (see Kaldor, 1951; Kalecki, 1968; Smithies, 1957; and Goodwin, 1967). The approach was essentially abandoned by the Keynesian tradition, unfortunately, and its revival is proving hard to achieve. It has been used more recently in some contributions based on the classical and Keynesian approach (see Garegnani and Trezzini, 2010).

25. These considerations do not necessarily mean that it is completely useless to construct models based on the assumption both of the independence of the trend of output with respect to its actual evolution and of the attainment on average of normal utilisation of capacity. These models must, however, be seen for what they are: abstractions that may help clarify some potential properties of a growing economy but involve numerous relations between variables that can neither represent nor fully explain those that occur in reality. On the other hand, this limited role is precisely what must be attributed to any type of *model* in the classical and Keynesian approach to the analysis of growth.

References

Bilancini, E. and Petri, F. (2008) 'The dynamic of general equilibrium: a comment on Professor Gintis', *Quaderno del Dipartimento di Economia Politica*, no. 538, Università di Siena.

Bortis, H. (1997) *Institutions, Behaviour and Economic Theory – A Contribution to Classical-Keynesian Political Economy* (Cambridge: Cambridge University Press).

Bronzi, M. (2012) 'Il nesso di causalità tra ciclo economico e trend: le teorie del ciclo reale e le teorie della domanda', Doctoral thesis in political economy, Università di Roma Tre.

Ciccone, R. (1986) 'Accumulation and capacity utilization: some critical considerations on Joan Robinson's Theory of Distribution', *Political Economy*, 1: 17–36.

Friedman, M. (1957) *A Theory of the Consumption Function* (Princeton: Princeton University Press).

Garegnani, P. (1962) *Il problema della domanda effettiva nello sviluppo economico italiano* Roma: Svimez.

Garegnani, P. (1978–79) 'Notes on consumption, investment and effective demand: I and II', *Cambridge Journal of Economics*, 2(4) and 3(1): 335–53, 62–82.

Garegnani, P. (1983) 'Two Routes to Effective Demand', in J. A. Kregel (ed.), *Distribution, Effective Demand and International Economic Relations* (London: Macmillan): pp. 69–80.

Garegnani, P. (1984) 'Value and Distribution in the Classical Economists and Marx', in *Oxford Economic Papers*, 36(2): 291–325.

Garegnani, P. (1992) 'Some Notes for an Analysis of Accumulation', in J. Halevi, D. Laibman and E. Nell (eds), *Beyond the Steady State: A Revival of Growth Theory* (Basingstoke and London: MacMillan).

Garegnani, P. and Palumbo, A. (1998) 'Accumulation of Capital', in N. Salvadori and H. D. Kurz (eds), *The Elgar Companion to Classical Economics* (Aldershot: Edward Elgar).

Garegnani, P. and Trezzini, A. (2010) 'Cycles and growth: a source of demand-driven endogenous growth', *Review of Political Economy*, 22(1): 119–25.

Goodwin, R. M. (1967) 'A Growth Cycle', in C. H. Feinstein (ed.), *Socialism, Capitalism and Economic Growth* (Cambridge: Cambridge University Press).

Haberler, G. (1938) *Prosperity and Depression: A Theoretical Analysis*, League of Nations, Geneva.

Harrod, R. F. (1939) 'An Essay in Dynamic Theory', *Economic Journal*, 49(193): 14–33.

Kaldor, N. (1951) 'Mr. Hicks on the Trade Cycle', *Economic Journal*, LXI, December: 833–47.

Kaldor, N. (1954) 'The relation of economic growth and cyclical fluctuations', *Economic Journal*, 64: 53–71.

Kaldor, N. (1957) 'A model of economic growth', *Economic Journal*, 67(268): 591–624.

Kalecki, M. (1968) 'Trend and the business cycle', *Economic Journal*, 78: 263–76.

Keynes, J. M. (1936) *The General Theory of Employment, Interest and Money* (London: MacMillan).

Modigliani, F. and Brumberg, R. (1954) 'Utility Analysis and the Consumption Function: An Interpretation of Cross-Section Data', in K.K. Kurihara (ed.), *Post-Keynesian Economics* (London: Allen & Unwin).

Palumbo, A. and Trezzini, A. (2003) 'Growth without normal capacity utilization', *European Journal of the History of Economic Thought*, 10: 109–35.

Serrano, F. (1995) 'Long-period effective demand and the Sraffian supermultiplier', *Contributions to Political Economy*, 14: 67–90.

Smithies, A. (1957) 'Economic fluctuations and growth', *Econometrica*, 25(1): 1–52.

Trezzini, A. (1995) 'Capacity utilization in the long run and the autonomous components of aggregate demand', *Contributions to Political Economy*, 14: 33–66.

Trezzini, A. (1998) 'Capacity utilization in the long run: some further considerations', *Contributions to Political Economy*, 17: 53–67.

Trezzini, A. (2011a) 'The irreversibility of consumption as a source of endogenous demand-driven economic growth', *Review of Political Economy*, 23(4), 537–56.

Trezzini, A. (2011b) 'Steady State and the Analysis of Long-Run Tendencies: The Case of Neo-Kaleckian Models', in R. Ciccone, C. Gehrke and G. Mongiovi (eds), *Sraffa and Modern Economics*, Vol. II (London and New York: Routledge): pp. 129–51.

Vianello, F. (1985) 'The Pace of Accumulation', *Political Economy*, 1(1): 69–87.

White, G. (2006) 'Demand-led growth and the Classical Approach to Value and Distribution: Are they Compatible?', in N. Salvadori (ed.), *Economic Growth and Distribution: On the Nature and Cause of the Wealth of Nations* (Aldershot: Edward Elgar): pp. 148–78.

4
Potential Output and Demand-Led Growth

*Antonella Palumbo**

4.1 Introduction

Potential output is generally understood as the desirable level of output, i.e., the highest which may be attained in any given situation without putting inflationary pressures upon the economy.

The political and practical relevance of the notion can hardly be overestimated. Potential output is frequently and increasingly used in the empirical analyses and policy prescriptions of such institutions as the International Monetary Fund (IMF) and the Organisation for Economic Cooperation and Development (OECD). Time series of the variable are reconstructed, both to evaluate the economy's performance and to project a potential growth path over a number of future periods,[1] against which the possible inflationary effects of policies are evaluated. It is used in the construction of derived indicators, such as the 'output gap' (the percentage deviation of actual from potential output, frequently used both in estimating inflation and for determination of the proper monetary policy);[2] and the 'cyclically adjusted budget balance' (a measure of the fiscal policy stance which is used in policy contexts such as the EU fiscal surveillance framework).[3]

Despite the frequent mention and use of potential output and related concepts, however, empirical estimation of the notions seems beset by difficulties and uncertainties ('a tricky business', according to Van Ark – see Gros *et al.* 2010, p.1).

*I wish to thank Anna Carabelli, Andrea Ginzburg and an anonymous referee for helpful comments, while retaining all responsibility for any remaining mistakes.

This chapter aims to examine the theoretical foundations of empirical research on potential output, especially that currently produced by international economic institutions and central banks. The purpose is twofold: first, to show that the notion of potential output changes significantly according to the theoretical framework within which it is defined, and how deeply empirical estimation methods are influenced by this. Second, it is maintained here that the current mainstream empirical literature, which starts from the theoretical presumption that the long-run path of potential output is supply-determined, provides empirical estimates which have generally little content apart from the representation of the theoretical belief that potential output must be a strong attractor of actual output, so that the unobservable variable is assumed (but not proved) to be some sort of moving average of actual output.

The related theoretical belief that there is a well-defined relationship between output gaps and changes in inflation, and the difficulty of detecting any systematic output–inflation relationship in actual data, is at the root of the variability and uncertainty of the estimates and of their poor performance as predictors of inflation.

Against these fundamental limitations, the notion of potential output will be analysed within the demand-led growth approach. The latter will here be defined in its broadest terms as that conception of growth according to which no automatic mechanisms ensure the tendency to full utilisation of resources, in the long as well as in the short run, so that the path of actual output is demand-constrained and determines in turn the evolution of potential output over time. This has relevant implications for empirical research. It will be maintained that not only is measurable potential output to be considered as a short-period notion, but no regular connection should be established or searched for between the output gap and inflation.

We start from the mainstream analysis of potential output, based on the idea of a permanent tendency of actual production to gravitate towards potential production, in accordance with the theories that describe growth as supply-constrained. Particular attention will be paid, respectively, to the notion of potential output in natural-rate and NAIRU models (Section 4.2) and in real business cycle and dynamic stochastic general equilibrium models (Section 4.3). Section 4.4 provides an overview of the results of current empirical estimates, almost exclusively based on supply-constrained growth models, and focuses on their limitations. In search of an alternative notion of potential output, Section 4.5 will be devoted to a brief analysis of the estimation method originally proposed by A. Okun in the 1960s, in which potential output was defined and measured on the basis of the Keynesian theory of effective demand, while Section 4.6 addresses the conception of potential output in the theories of demand-led growth and its implications for empirical research. Section 4.7 concludes.

4.2 Potential output in natural-rate and NAIRU models

The theoretical and empirical analyses on potential output which are currently conducted by central banks and international institutions[4] are largely based on the idea that potential output is determined by supply factors – the evolution of resources and productivity over time – while actual production, which can be affected by the state of demand and short-term shocks, tends to gravitate permanently towards it. This derives from the prevailing theoretical conception of the working of the economic system, which, based on the equilibrating mechanisms of neoclassical theory, regards growth as a supply-side phenomenon with no permanent limits posed by demand to the production of the economy's full potential.

Within this general view, however, different theories and models imply different specific notions of potential output and give rise to different methods for its estimation.

As a first fundamental qualification, the potential utilisation of labour is usually defined as implying natural unemployment, that is, according to Friedman's definition, the level of unemployment which is compatible with constant inflation. The behaviour of inflation on the two sides of the non-inflationary unemployment rate is regarded as symmetrical, with accelerating or decelerating inflation (and the related changes in expectations) playing the role of adjusting the system towards the natural (equilibrium) rate. Thus, monetarist and new classical models imply a strict relation between output and inflation. Demand shocks (especially monetary shocks) are regarded as causing deviations of actual from potential output, which may be more or less durable according to the rapidity of error correction in expectations, but are, however, of a temporary nature. Actual output fluctuates around the potential path, the latter depending exclusively on long-run (real) supply determinants.

The New Keynesian NAIRU models (Carlin and Soskice, 1990; Layard *et al.*, 1991) have partially different theoretical bases, relying on the hypothesis of imperfect competition or nominal and real rigidities which prevent the system from attaining efficient results. The NAIRU (Non-Accelerating Inflation Rate of Unemployment), though implying involuntary unemployment, represents equilibrium because it is the only rate of unemployment at which inflation is constant (the competing claims of workers and firms being such as to exactly exhaust the product). Spontaneous mechanisms based on accelerating or decelerating inflation[5] ensure the tendency of actual unemployment to the NAIRU. Involuntary unemployment at the NAIRU is not caused by insufficiencies of aggregate demand, but by the imperfections and market power characterising the labour (and product) market. It cannot therefore be corrected by means of demand policies, but rather through labour reforms or other competition-enhancing measures. Demand may cause deviations of the actual unemployment rate from the NAIRU.

In what follows, we will refer to the general type of NAIRU models, in which the NAIRU is an attractor of the actual unemployment rate, neglecting the possible hysteresis effects. In much of this literature hysteresis is in fact regarded at best as a medium-term phenomenon with the system converging in the end to the unique long-term NAIRU (see Layard *et al.*, 1991); see however below (Section 4.4.1) for discussion of the case of full hysteresis.

Notwithstanding their different theoretical bases, natural-rate models and NAIRU models have very similar empirical implications. First, in both types of models the equilibrium unemployment rate – whether natural rate or NAIRU – is regarded as an attractor of the actual rate, and potential output as an attractor of actual output.

Second, both approaches postulate a clear distinction between the long-run determinants of trend (long-run supply factors and labour market institutions) and the determinants of short-run fluctuations, which depend on demand shocks (policies included) and other factors, including temporary supply shocks (ECB, 2011, p. 80). Adjusting mechanisms ensure that the deviations induced by demand are only temporary[6] (see below for qualifications of this statement in the case of NAIRU models with full hysteresis).

Third, in both approaches a definite relation is postulated between the sign of output gaps and changes in inflation. Actual output being above potential necessarily implies accelerating inflation, while inflation must decelerate for negative output gaps.

On these theoretical bases, three distinct classes of methods can be identified for the empirical estimation of potential output (for a classification of methods, see Cotis *et al.*, 2005; Laxton and Tetlow, 1992):

a) On the basis of the assumed gravitation of actual output towards potential, it is possible to consider *trend output*, i.e., the growth path of actual output over time neglecting short-period fluctuations, as an empirical approximation of potential output. This defines the class of 'statistical methods' for potential output estimation, which are based on the extraction of the trend from the time series of actual output by means of various statistical techniques (see Section 4.3).

b) The theoretical presumption that the trend of potential output is exclusively determined by long-run supply factors and is totally independent of demand variables gives rise to an attempt to directly estimate the growth of such factors over time. This defines the class of 'economic' methods for potential output estimation, the most widely used of which is the *production function approach* (Giorno *et al.*, 1995; De Masi, 1997; Denis *et al.*, 2002; Billmeier, 2004; CBO, 2004). It consists in estimating, through specific hypotheses, the growth in time of the potential capital and labour inputs (the latter depending in turn on the calculation of the equilibrium rate of unemployment) and of a 'potential total factor productivity' supposedly representing the pace of technological innovation; the growth in time of potential output is then estimated by applying a production function to these hypothetical potential factors (see Section 4.4.3).

c) The supposed theoretical relationship between output gaps and inflation variations gives rise to an attempt to derive the estimates of potential output indirectly from the series of the inflation rate, by specifying a definite quantitative relationship between changes in the inflation rate and output gaps, i.e., a Phillips curve.[7] This produces the so-called 'multivariate' statistical methods for the estimation of potential output (to be distinguished from the univariate statistical methods seen above), which derive estimates of potential output from the time series of actual output but with corrections deriving from the behaviour of the inflation rate. The multivariate methods thus represent in fact a mixture between the statistical methods and the economic methods.

The output–inflation relationship is also essential for estimation of the NAIRU in the production function approach.

The essential characteristics of the notion of potential output in this literature can be summarised as follows:

- Potential output tends to be identified with the average level of actual output, so that output gaps should prove tendentially symmetrical.
- The behaviour of inflation is symmetrical on the two sides of potential output.
- The long-run path of potential output is exogenous and can be defined when knowing the laws of motion of resources and productivity.

As regards this latter point, it is worth noting that the possibility of projecting the growth path of potential output over time also depends on the fact that the majority of empirical analyses of this kind represent the economy by means of simple aggregative models with an extremely simplified definition of technical progress.[8]

4.3 Potential output in RBC and DSGE models

A different definition of potential output is to be found in real business cycle (RBC) models and gives rise to partially different estimation methods, or at least to different interpretations of the estimated values.

According to the RBC approach (Kydland and Prescott, 1982; Prescott, 1986), the economy's cyclical fluctuations are produced by the agents' optimal reaction to random real shocks. Trend and cycle have the same economic determinants and depend exclusively on supply factors, while money disturbances have negligible effects. Actual production thus always coincides with potential production: observed fluctuations in output should not be interpreted as *output gaps,* but as the fluctuating path of equilibrium itself. As for policy implications, 'costly efforts at stabilization are likely to be counterproductive. Economic fluctuations are optimal responses to uncertainty in the rate of technological change.' (Prescott, 1986, p. 21)

While according to this approach potential output should be regarded as the same thing as actual output, it is possible to obtain an estimate of *trend output* by applying a statistical filter to the time series of actual output. Actually, the technique of statistical filtering for estimating trend output was originally proposed by RBC authors (see especially Hodrick and Prescott, 1981), and has been largely utilised since.

Statistical filtering requires a (necessarily arbitrary) hypothesis of what has to be defined as *cycle* and what as *trend* in a time series. Different hypotheses give rise to different results: the time path of trend output could even be represented by a single number, expressing an average rate of growth over the whole estimation period. However, the Hodrick-Prescott filter (like other filters)[9] does generally assume that the trend itself varies in time and has a stochastic component.

The Hodrick-Prescott filter consists in the minimization of both deviations of actual output from potential and deviations of the rate of growth of potential output from a regular trend:

$$Min\left\{\sum_{t=1}^{T}(\ln Y_t - \ln Y_t^*)^2 + \lambda[(\ln Y_{t+1}^* - \ln Y_t^*) - (\ln Y_t^* - \ln Y_{t-1}^*)]^2\right\}$$

The estimation results are strictly dependent on the arbitrary value assigned to the parameter λ, determining the variability in time of the estimated series and its greater or lesser adherence to the series of actual output (Billmeier, 2004, pp. 18–19).

While in RBC models statistical filtering is regarded simply as defining the *trend* component of the actual time path of output, in natural-rate and NAIRU models the same method of estimation is regarded as approximately identifying the potential output path.

One relevant policy implication of RBC models is the virtual absence of any effects whatsoever and any possible scope for monetary policy. This extreme conclusion has been partly modified within a different class of models, the dynamic stochastic general equilibrium (DSGE) models, which stem from RBC models in assuming an intertemporal general equilibrium framework as the basis for modelling the economy. As in RBC models, fluctuations are described in terms of the optimal responses of agents to random shocks and a structure of lags implies a mechanism of propagation whereby stochastic disturbances have durable effects. However, in the so-called New Keynesian dynamic stochastic general equilibrium (NK-DSGE) models, it is assumed that nominal or real rigidities and market inefficiencies result in transitory deviations of actual output from potential output, implying scope for the active role of monetary policy in correcting output gaps (Woodford, 2003).[10]

In empirical evaluation and for policy purposes, policymakers need in the first place 'to assess the degree to which fluctuations in observed output reflect the optimal response to shocks that hit the economy, versus undesirable fluctuations' (Basu and Fernald, 2009, p. 2). Optimal fluctuations are defined as the fluctuations that would occur in case of full flexibility of prices and wages and zero mark-ups. In NK-DSGE analyses, a model is usually built representing the whole economy and then used to estimate, through simulation exercises, the following three notions of potential output (Vetlov *et al.*, 2011):

a) *Trend output*, which represents the time path of the economy net of business cycle fluctuations, the latter being partly induced by deviations of potential output itself with respect to a regular steady growth path. In fact, the identification of long-period trend does not amount, in these models, to estimating potential output but only its long-period component.
b) *Efficient output*, which is the path that output would follow if goods and labour markets were perfectly competitive (that is to say, if prices and wages were fully flexible and mark-ups were zero). Deviations of

actual from efficient output measure the losses of welfare associated with imperfect competition and nominal rigidities.

c) *Natural output*, which is the path that output would follow under flexible prices and wages but imperfectly competitive markets (so, differently from the efficient output, with different-from-zero mark-ups and mark-up shocks). Deviations of actual from natural output measure the relevance of nominal rigidities.

Efficient and natural output can deviate from actual output because the latter, unlike the former two, is determined by assuming sticky prices and/or wages. Policy should aim at identifying and correcting deviations from natural output, for the purpose of controlling inflation, and deviations from efficient output, which imply a loss of welfare, while optimal fluctuations should not be corrected but rather accommodated (Basu and Fernald, 2009; Galì *et al.*, 2011). The results of some empirical analysis carried on along these lines will be described in Section 4.4.4 below.

What is worth remarking is that the absolute majority of empirical estimates of potential output currently performed by central banks and international economic institutions are based on the kind of models which have been just described in this and the previous section and are thus based on the theoretical presumption that only supply factors matter in the determination of potential output.[11] As will be shown in the next section, this presumption deeply influences the results of estimation.

4.4 Problems and puzzles in the empirical estimates of potential output

In reviewing the results of empirical literature it is worth starting from the attempts at estimating the NAIRU, given the role that it plays also in estimation of potential output.

4.4.1 Chasing the NAIRU

There are numerous attempts in the literature to give empirical content to the notions of natural rate and NAIRU. It is worth noting that in applied research the two concepts are frequently identified with each other, notwithstanding the theoretical differences: empirically, they are both defined as the rate of unemployment that is compatible with constant inflation and to which actual unemployment tends (Gordon, 1989, p. 220; Cross, 1995).

It does not seem too far-fetched to maintain that empirical literature has generally failed to identify such equilibrium rate of unemployment

(Jenkinson, 1988). The first attempts assumed that the unobservable natural rate or NAIRU was a constant; the impossibility of detecting any such constant when analysing real economies has given rise to the notion of *time-varying NAIRU* (Gordon, 1997).[12]

Such difficulties became particularly evident in the course of attempts to explain the European experience of the 1980s and 1990s, when unemployment persistently remained much higher than in the preceding decades. This has created a dilemma: either the NAIRU tends to be quite stable, but in this case one should explain why wide and persistent deviations of the actual unemployment rate from the supposed equilibrium value do not tend to be corrected; or the NAIRU itself is very variable, but one should observe in reality some of the phenomena which according to theory justify such variability and especially its permanent changes of level (Pesaran and Smith, 1995, p. 219). The increase in actual average unemployment in the 1980s and 1990s in most European countries seems rather to have been accompanied by an increase and not a decrease in labour flexibility due to policy measures (see Blanchard and Katz, 1996, trying to justify the variability of the NAIRU, and the critical remarks in Solow, 2000, p. 5; Stirati, 2001, p. 428). Equally difficult to explain are the relevant differences in average unemployment among the different European economies (Gordon, 1989, p. 220).

Instead of identifying a single equilibrium rate of unemployment, data seem rather to show that in many cases different unemployment rates are compatible with stable inflation (Schreiber and Wolters, 2007; Karanassou *et al.*, 2003; Franz, 2005, p. 23; IMF, 2006, ch. 3).

Within the NAIRU models, one theoretical response to these empirical difficulties has been the notion of hysteresis, i.e. the supposed influence of actual unemployment rates on the equilibrium one (Blanchard and Summers, 1987; Jenkinson, 1988). This is justified in various ways, for example by assuming that high unemployment raises the share of long-term unemployment, determining loss of efficiency and competitiveness in part of the labour force and thus their inability to influence the bargained real wage (Layard *et al.*, 1991, pp. 74–5).

Hysteresis introduces a series of problems and contradictions in NAIRU models and opens up the possibility of completely overturning the very basic idea of the existence of an equilibrium unemployment rate. An equilibrium variable which is heavily influenced by the actual one is in fact a rather weak attractor of actual values, and thus an almost irrelevant theoretical concept. As for policy, full hysteresis, by subverting the relationship between actual and potential output, implies that demand policies may make the potential growth path shift permanently (Ball, 1996).

In the majority of NAIRU models, however, these radical conclusions are avoided by assuming that hysteresis is a short/medium-period phenomenon and that in the long period the system converges to a unique rate (Layard *et al.*,1991; Carlin and Soskice, 1990).

4.4.2 Estimates of the time-varying NAIRU

NAIRU is currently estimated – like potential output – by means of two different types of method, the statistical and the economic, or by some mixture of the two (Richardson *et al.*, 2000; Boone, 2000; De Masi, 1997; Billmeier, 2004; Denis *et al.*, 2002; Fabiani and Mestre, 2001). Statistical methods amount to defining the NAIRU as the trend component of the actual unemployment rate, while economic methods try to infer the NAIRU from equations representing economic relations, and particularly the Phillips curve.

The most widely used statistical methods are the Hodrick-Prescott filter, based on the double minimization both of unemployment gaps and of deviations of NAIRU from its regular growth path, and the Kalman filter, in which the actual series of unemployment rates is decomposed in a trend component, a cyclical component and error. Univariate filters produce estimates which automatically imply the tendency of actual unemployment rates to the NAIRU, but as a result the estimated NAIRU has no connection with the actual behaviour of inflation. Moreover, they have very scarce predictive powers as regards the value of the unemployment rate.

Multivariate statistical filters are based on the idea that univariate filters' estimates must be corrected by means of theory-based relationships (they are thus a combination of statistical and economic methods). The estimated NAIRU must be such that the estimated unemployment gaps are correlated with inflation changes. Such filters, however, are based on the assumption that deviations from estimated NAIRU are symmetrical over a sufficiently long period, thus amounting to calculation of a moving average of the series of actual unemployment rates, in which measures are taken to prevent estimated unemployment gaps from moving in a direction opposite to that predicted by theory.

Economic methods may use either structural-form equations or reduced-form equations to estimate the NAIRU. The former, which are usually based on a wage equation and a price equation (according to the specification proposed by Layard *et al.*, 1991) and try to infer the NAIRU from the equilibrium solution, should in principle provide a better (in the sense of more theory-based) estimate (Richardson *et al.*, 2000, p. 34). In practice they are scarcely used, because of both theoretical uncertainties surrounding the correct specification of the equations, and the instability of the resulting estimates and their lack of robustness due to the high number of unobservable parameters (Franz, 2005). A further problem has to do with the difficulty of estimating such variables as the desired wage, so that mainly estimation models based on reduced-form equations are used in applied research.

The latter usually consist of an expectation-augmented Phillips curve, where current inflation is explained by past inflation (approximating expected inflation on the assumption that inflation is a random walk – Layard *et al.*, 1991, p. 15), the unemployment gap ($u_t - u_t^*$) and other variables representing supply shocks (Gordon, 1997):

$$\pi_t = \pi_{t-1} + \alpha_1(u_t - u_t^*) + \alpha_2 z_t + e_t$$

where u_t^* is the unknown NAIRU, z_t summarises the supply shocks (z_t is normalised so as to have zero mean, since supply shocks are symmetrical by construction), while e_t is a white noise. The equation, which is typically more complex due to various lags (z_t may also be a vector of variables), is estimated by making use of two different filters, the multivariate Hodrick-Prescott filter or the multivariate Kalman filter.

The multivariate HP filter (Laxton and Tetlow, 1992) implies imposing the condition that the NAIRU is the trend component of the actual unemployment rate, according to:

$$u_t = u_t^* + \varepsilon_t$$

where ε_t is a white noise process; while the NAIRU itself follows a random walk:

$$u_t^* = u_{t-1}^* + v_t$$

where v_t is a white noise. With these restrictions, this 'economic' method of estimating the NAIRU is perfectly coincident with the application of a statistical multivariate filter to the series of actual unemployment. The quantity to minimize is:

$$Min \sum [(u_t - u_t^*)^2 + \lambda_1 (u_t^* - u_{t-1}^*)^2 + \lambda_2 e_t^2]$$

(where e_t is the residual of the Phillips curve). With this 'economic' restriction the resulting NAIRU, though being by construction the moving average of the actual unemployment rate, is much more variable than the one obtained through univariate filters (Richardson *et al.*, 2000). Resulting estimates are influenced by the two weighting parameters λ_1 e λ_2: in particular, the higher the second, the higher the weight given to inflation data and the more variable the resulting NAIRU.

The multivariate Kalman filter estimates are obtained in the hypothesis that the trend (the NAIRU) is a random walk process with or without drift:

$$u_t^* = \mu_t + u_{t-1}^* + \varepsilon_t$$

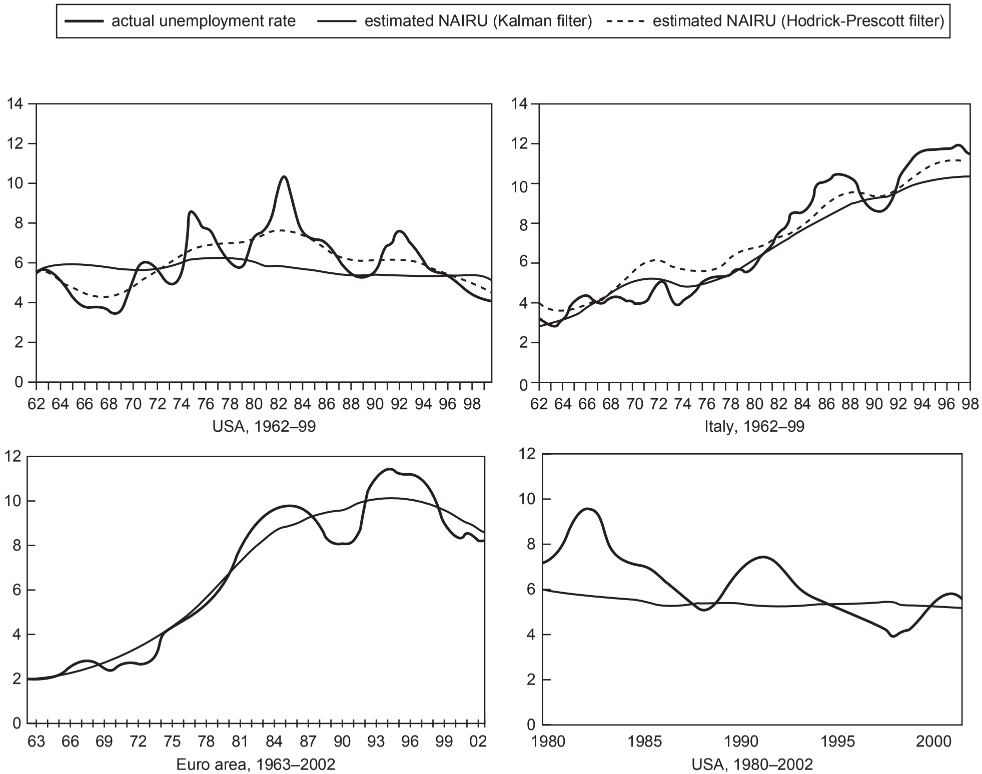
(the μ_t drift may either be absent, or deterministic, or a random walk process itself), with no economic restrictions on the trend; while the cyclical component is constrained by a Phillips curve-type relation. The Kalman filter is an iterative process in which the estimated NAIRU is progressively corrected on the basis on new information on the observed variable (see Richardson *et al.*, 2000, p. 42–3; Boone, 2000, p. 5; Denis *et al.*, 2002, p. 9ff. for details).

The whole procedure thus consists in attributing all the variability in actual unemployment rates which cannot be explained on the basis of inflation data to supposed changes in the unobservable NAIRU (apart from casual errors). Since the resulting estimates may show excessive variability, they are further corrected with smoothing procedures (Gordon, 1997; Richardson *et al.*, 2000, p. 44; Fabiani and Mestre, 2001, p. 11). As noted by Franz (2005, p. 17), 'If there is no limit on the ability of the NAIRU to fluctuate each time period, the time-varying NAIRU may jump up and down and soak up all the residual variation in the Phillips curve'.

In some of the estimation models, but not in all, further restrictions are imposed on the cyclical component to explicitly constrain it to have zero mean (Denis *et al.*, 2002, p. 10; Fabiani and Mestre, 2001, p. 10). Even if there is no such explicit hypothesis, however, the estimated NAIRU is built so as to absorb all the changes in the average level of the actual unemployment rate, automatically attributing them to supposed changes in the supply factors determining the NAIRU (Boone, 2000, p. 11; Franz, 2005, p. 18).

Figure 4.1, which reproduces some results of the estimates from different sources, confirms that the series of the estimated NAIRU always proves to be the trend component (though differently defined in different estimation models) of the actual unemployment rate series.[13]

A series of problematic aspects of NAIRU estimates are stressed in the literature: estimates are uncertain and not robust, i.e., change drastically


actual unemployment rate
estimated NAIRU (Kalman filter)
estimated NAIRU (Hodrick-Prescott filter)
USA, 1962–99
Italy, 1962–99
Euro area, 1963–2002
USA, 1980–2002

for a small change in specification (Staiger *et al.*, 1997, p. 34; CBO, 2004; Gordon, 1997, p. 21 and 24); most importantly from a theoretical point of view, they seem to be very inefficient in predicting inflation (Staiger *et al.*, 1997; Jenkinson, 1988). As seen above, inflation actually plays a relatively unimportant role in building NAIRU estimates, especially after smoothing procedures are applied. The resulting correlation between unemployment gaps and changes in inflation is indeed very low (Billmeier, 2004).

The dominant role is played instead by the actual unemployment rate: as noted by Gordon (1997, p. 28), 'fluctuations in the NAIRU seem too large to be plausible and seem mainly to mimic movements in the actual unemployment rate'. According to Galbraith (1997, p. 101), 'In general, the estimated NAIRU in a variety of studies has tracked the actual unemployment rate sluggishly. When unemployment rises, analysts tend to discover that the demographic characteristics of workers are deteriorating, or that the job-wage and wage-price dynamic has become unstable. And then the unemployment rate drifts down again, those flaws mysteriously begin to disappear, and a lower NAIRU is estimated.'[14]

4.4.3 Estimates of potential output: the production function approach

While the estimates of potential output which are obtained through statistical filters, both univariate and multivariate, are by definition the trend component of actual output, the 'production function approach' (see Section 4.2) should be based on economic relations and should identify the supply factors allegedly determining potential output. Thus in principle actual output should not automatically gravitate around the estimated potential, unless the theoretical presumption of gravitation were empirically confirmed.

In this approach, potential output is estimated through the following Cobb–Douglas production function with constant returns, applied to the time series of potential inputs:

$$Y_t^* = (N_t^*)^{\alpha} (K_t)^{1-\alpha} TFP_t^*,$$

Potential capital is identified with actual; potential labour input N_t^* is estimated by applying a statistical filter to the actual time series of labour force, so as to obtain the 'trend labour force' LF^*, and by correcting such trend labour force for the estimated NAIRU

$$N_t^* = LF_t^* (1 - NAIRU_t);$$

while 'potential' total factor productivity is obtained by applying a statistical filter to the series of the total factor productivity (TFP) which is estimated through application of a Cobb-Douglas of the same specification to actual data over the same period. Elasticity α is the observed average value of the wage share.

For what has just been seen as regards the estimates of the NAIRU, potential labour is in fact the trend component of actual employment, while potential TFP is the trend component of TFP (the residual) estimated on actual data.

Thus, despite the complication of the procedure, what the production function approach indirectly calculates is in fact a sort of moving average of the time series of actual output. Like estimates obtained through statistical filtering, these estimates are, disappointingly, very slightly correlated to inflation changes (Billmeier, 2009; ECB, 2011, p. 82–85). The attempt to define output gaps as more correlated to inflation gives rise to volatility of the estimated series of potential output, which then has to be smoothed.

It can be concluded that the tendential symmetry of output gaps which can be observed in these estimates is in reality built into the hypotheses,[15] which makes the estimation of potential output through the production function approach not significantly different from statistical filtering. Once again, the empirical content of the estimates is merely the theoretical presumption (but not the empirical proof) that actual output must necessarily gravitate, on average, around potential.

4.4.4 Estimates of potential output within NK-DSGE models

Attempts at estimating potential output within the NK-DSGE framework are still preliminary and do not offer a body of consolidated results (Vetlov *et al.*, 2011, p. 22).

As seen above, estimation in these models proceeds through simulation exercises once the 'correct' models for describing the economy have been identified. This implies 'calibrating' a general stochastic equilibrium model on the economy (thus finding, through econometric techniques, the value of the parameters which make the model mimic the true economy's fluctuations as closely as possible) and defining on this basis the nature of shocks which have hit the economy in the sample period (Basu and Fernald, 2009, p. 23). One way of obtaining the various measures of potential is to simulate the path of the economy in the counterfactual hypotheses of full flexibility of prices and wages and: a) presence of the low-frequency persistent shocks only (trend output); b) presence also of the high-frequency efficient shocks (efficient output); c) presence also of shocks on wage and prices mark-ups (natural output). Actual output may differ from all three measures due to price and wage stickiness, which, however, is regarded as a short-period phenomenon. Thus, with this estimation procedure, since the path that the economy has actually followed is theoretically defined as optimal (except for short-term fluctuations), the three definitions of potential output are in fact three different ways of smoothing the fluctuating path of actual output, and all three kinds of output gaps tend to have zero mean if the whole period is taken into consideration. Efficient and natural output are much more volatile than trend output for the fact of embodying a greater number of shocks in the definition of potential output, and thus account for a greater part of observed output variability; while trend output estimates prove to be very similar to those obtained through statistical filters (Vetlov *et al.*, 2011, p. 14).

This result may partly change depending on the hypotheses of the model and its specification. Galí and Gertler (2007, p. 30), for example, maintain that actual output gravitates around a natural inefficient level, due to the presence of mark-ups (caused by imperfect competition) systematically causing a loss of output and welfare if compared with the theoretical optimal path. This would imply an estimated efficient output systematically higher than natural and non-symmetrical output gaps: demand policy, however, would prove useless in closing them, while only supply-side policies aimed at enhancing competition would be effective.

In NK-DSGE models monetary policy is required to close the gap between actual and natural output, which should empirically prove to be highly correlated with inflation variations. However, the results of estimation in this respect are at best uncertain and mixed: besides being very volatile, natural output gives rise to output gaps whose correlation with inflation is often quite low (Vetlov *et al.*, 2011, p. 5–6).

Results of simulations are not robust and may vary widely for small changes in specification (Sims *et al.*, 2010, p. 211). The nature of shocks cannot be unequivocally defined empirically, making some arbitrariness in interpretation inevitable. Moreover, the basic model has been complicated with the introduction of many frictions, which makes the interpretation of the relationships among variables rather complex. The only reason to pursue these kind of empirical exercises rather than the traditional ones would be, according to Vetlov *et al.* (2011, p. 7), their allegedly superior theoretical consistency.

4.4.5 Estimates of potential output and the crisis

It may be concluded that the estimation of potential output according to its supply-side definition in general produces disappointing results.

The main theoretical hypothesis, i.e., that the sign and size of the output gap should be related to inflation, does not show in the data. If the estimated output gaps are constrained to be compatible with observed inflation changes, the resulting series is extremely volatile (the estimated potential oscillates wildly), which seems to contrast with the idea of potential output as a comparatively smooth series. On the other hand, estimated potential output is often nothing more than the statistical trend of actual output, or some elaborated variant of it. Its correlation with inflation changes is disappointingly low, which leads some authors to conclude that other factors, apart from the output gap, have become more relevant in influencing inflation (oil prices, prices of imports in general, expectations of inflation – see ECB, 2011, p. 85; Billmeier, 2009, p. 402).

One consequence of the way potential output is usually estimated shows in the effect on it of the deep contraction that has afflicted advanced economies in recent years. Although the recession was initially regarded mainly as a pronounced deviation from the trend of potential output, after some time it produced downward revisions of the potential itself (Gros *et al.*, 2010; ECB, 2009, p. 45; 2011, p. 73). This result, which is obviously implied in the very way estimates are built, is somehow puzzling for the mainstream theoretical conception of potential output. If the crisis is interpreted as a deep adverse demand shock, according to the theories of supply-side growth it should not be considered as capable of affecting the potential output. Explanations thus vary from regarding the whole change which has been detected in the potential level as something which will probably not affect the 'longer-term growth rate of potential output' if the economy reacts flexibly enough (ECB, 2011, p. 73), to looking for changes in supply-side determinants. These can be due either to low levels of investment in research and development, or to reduction of the labour force due to discouragement or similar effects (European Commission, 2009; ECB, 2011, p. 77).

In the end, the very content of the empirical estimates of potential output has forced applied research to admit the influence of actual on potential output, which was probably unexpected in principle.

We will now turn our attention to the theories of demand-led growth, in which the causation between actual and potential output is reverted. Although these theories have not given rise to a systematic body of applied research on the estimation of potential output, the insights they offer on the working of the system may serve to interpret some of the results and puzzles of mainstream empirical literature on the matter. As a premise, we briefly deal with the analysis of potential output proposed by A. Okun in the 1960s, which represents the first systematic attempt, along Keynesian lines, to estimate potential output for policy purposes.

4.5 Okun's definition of potential output

On the basis of macroeconomic Keynesian theory, Okun (1962; see also Council of Economic Advisors, 1962) regards potential output as the maximum attainable level of output in the short period, with given equipment and capacity. In giving empirical content to the notion, he assumes, however, that unemployment cannot be compressed below a certain threshold without causing unsustainable inflationary pressures, and arbitrarily (but following what he regarded as the consensus view of the US economy) puts such threshold at 4 per cent (Okun, 1962, p. 1). The 'unemployment gap' is then calculated as the difference between the 4 per cent non-inflationary unemployment rate, regarded by Okun as a lower bound, and the actual one. It must be noted that no specific hypothesis is introduced on the behaviour of inflation at unemployment rates above 4 per cent. In other words, 4 per cent is not regarded as the turning point which separates inflation from deflation (or accelerating inflation from decelerating inflation).[16]

A second element which is used in Okun's estimates is the so-called 'Okun's Law', establishing an empirical proportion (3:1, by Okun's own calculations) between changes in actual real output (real Gross National Product) and changes in unemployment.[17] The actual real GNP of a period in which unemployment has been 4 per cent may thus be defined as the initialising level for the series of potential output, which is then constructed by assuming that each percentage point unemployment gap implies a 3 per cent gap between potential and actual output. This allows both the level of potential output and its rates of growth over the estimation period to be defined. Some smoothing procedure is then used on the resulting estimated series in order to reduce its variability.

The results of Okun's definitions and hypotheses are shown in Figure 4.2, which is derived from his original data.

It is worth stressing the differences between Okun's estimates and the estimation methods which subsequently prevailed and are now used. First, while in natural-rate models potential output is a long-run concept, Okun (1962, p. 2) maintains that it can only be defined in the short period, on the basis of given technical conditions and given plants: 'to the extent', he maintains, 'that low utilization rates and accompanying low profits and personal incomes hold down investment in plant,

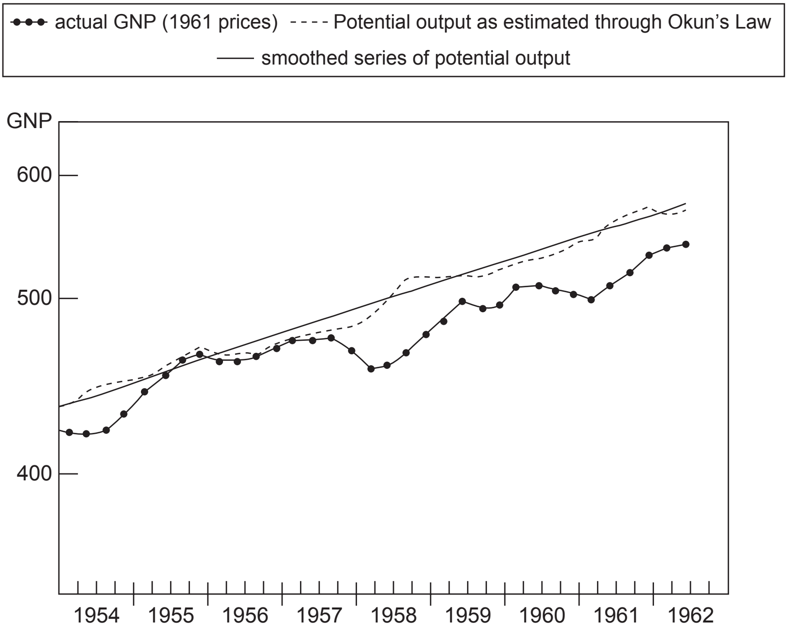


Figure 4.2 Estimates of potential output in Okun's definition, USA 1954–62
Source: Data from Okun (1962).

equipment, research, housing, and education, the growth of potential GNP will be retarded'. According to Okun, 'today's actual output influences tomorrow's productive capacity'.

Second, apart from exceptional periods of boom, actual output is systematically lower than the estimated potential; output gaps tend consequently to be mainly negative and non-symmetrical. Unlike in natural-rate and NAIRU models, potential output is theoretically characterised as the ceiling and not the average level of actual production. This implies that in Okun's view the non-inflationary rate of unemployment is not the *equilibrium* rate of unemployment.

Third, apart from the definition of the non-inflationary rate, Okun's procedure requires no use of inflation data in potential output estimation, thus implying no necessarily strict correlation between output gaps and inflation.

Okun's Law has also been used for estimation of potential output in later works. Even if no additional hypothesis is introduced on the level of potential output for initialising the series, it is possible to estimate a series of rates of change of potential output exclusively on the basis of Okun's Law (an exogenous definition of the potential rate of unemployment is, however, necessary). Such exercises are proposed for example by DeLong (2002) and Solow (2000).

4.6 Potential output in demand-led growth theories

Once set in the long period and taken as the basis for a theory of growth, the Keynesian principle of effective demand can be characterised as implying independence of investment from saving and the absence of any spontaneous tendency of the system to grow along a path with full utilisation of resources.[18]

However, some sort of mechanism must guarantee that the growth paths of actual output and potential output do not totally diverge over time. Although adjustment need be neither instantaneous nor complete, even in long periods, an adjustment mechanism must exist and this is represented by the endogeneity in the formation of resources.

Persistently under-utilised resources tend slowly to disappear – fixed capital shrinks if net investment is negative, labour may emigrate or participation rates may fall – while the pressure of high demand may, on the contrary, accelerate their formation (with the same phenomena of opposite sign; see Kaldor, 1985, pp. 35–6). Thus it may be maintained in the first place that for the principle of demand-led growth to be true it is not necessary to observe in reality wide and persistent gaps between actual production and full-capacity production (the latter being a variable quantity which can be slowly influenced by the former). At the same time, this does not imply that there are small margins of elasticity for production and capacity to adjust to demand: on the contrary, such margins are even wider because not only can the utilisation of existing resources vary, but they can be created or destroyed altogether.

Following Garegnani (1992), the loss of potential over a number of periods, due to under-utilisation of capacity in one period, is represented not only by the current under-utilisation, but also by the capacity which failed to materialise and to produce yet higher capacity. This implies that such loss is mostly invisible: borrowing an expression from Landes (1990, p. 6), we may say that 'the gap between what is and what can be is enormous'.

Equally wide are the margins for *increase* in the size of capacity, given the definition of normal utilisation as lower than technical full utilisation, and thus the possibility of over-utilisation leading to the installation of new capacity.

It may thus be maintained that it is the endogeneity of potential output and its dependence on actual that characterise demand-led growth theories. While output gaps would measure in each period the distance between actual output and capacity output, over a succession of periods they would not be sufficient to measure the distance between what has been actually realised and what could have been realised: they may at most represent the *visible* waste of capacity.

As regards inflation, no necessary and strict relation is postulated between output gaps and price changes. In the first place, the definition of normal utilisation of resources contains no reference to inflation, so that, while it can be maintained that in general prices rise more quickly on approaching full utilisation, no general and mechanic rule can be established. Inflation will be influenced by the distributive conflict (Stirati, 2001), the intensity of which need not bear any definite relation to the level of activity. As regards the long run, once admitting that high demand may produce pressures on capacity only for a limited period, before inducing the endogenous creation of new capacity, the conclusion likely follows that no permanent pressure can be exerted on costs and prices by demand (Serrano, 2006). It would thus be illegitimate to infer the measure of potential output from the actual behaviour of inflation.

4.6.1 Studying the long-run path of potential output in the demand-led growth approach

A matter of debate, within the demand-led growth approach, is whether the long-run trend of the system can be studied independently of its short-run behaviour. This has relevant implications for the possibility of defining potential output as a long-run or an exclusively short-run concept, or, more accurately, the possibility of projecting a long-run potential output path starting from an arbitrary point in time.

Some theories and models within the demand-led growth approach assume that the mechanism of capacity adjustment described above operates in such a way as to allow full adjustment between demand and capacity to be considered representative of the actual growth path of the economy (at least on average), so that the long-period relations between variables may be studied with exclusive reference to such situations of full adjustment. This is the case not only for steady-state growth models,[19] but also for the so-called 'supermultiplier' models (Serrano, 1995) in which the system's long-period growth is determined by the growth in time of autonomous demand, while investment is regarded as completely induced by the capacity adjustment mechanism. In principle, these models allow a long-period path of potential output to be defined, at each point in time, on the basis of the exogenous growth of autonomous demand, in case a very simple form of technical progress is assumed.

However, the very definition of growth as demand-determined implies that in any specific period demand may change autonomously and prove different from the previously installed capacity. Though capacity is continually built or destroyed on the basis of the strength of expected demand, normal utilisation of installed capacity is not guaranteed. The absence of automatic mechanisms for adjusting demand to capacity reproduces in each period the possibility of different-from-normal utilisation, while at the same time duration of fixed capital implies that the adjustment of capacity to demand may act slowly (Ciccone, 1986; Palumbo and Trezzini, 2003). In each period in which capacity is given, differences between actual output and capacity output may well occur and need not be symmetrical over any relevant observation period.

This implies the impossibility of defining exogenously a growth path of potential output. At each point in time, given installed capacity, it is always possible to define a hypothetical growth path that the economy would follow if, *assuming no technical change*, investment decisions allowed normal utilisation of existing capacity and if future demand were such as to allow normal utilisation of all capacity installed in each subsequent period. Such a 'normal utilisation path', however, would by no means represent the path towards which actual output gravitates, even neglecting innovation. Any under-utilisation would in fact determine a displacement of the whole path. Nor would the normal utilisation path even represent the highest possible growth path, given the upward flexibility of output and capacity.

In other words, at each point in time potential growth is in fact represented by a plurality of possible growth paths, even neglecting the role of technical progress. The path which the economy actually follows identifies the range of future growth possibilities (as noted by Setterfield, 2002, p. 5, this approach necessarily requires the notion of 'path dependence' in order to analyse long-run tendencies).

We may conclude that in the demand-led growth approach potential output may be a meaningful, quantifiable notion only with reference to the short period, in which it can be defined on the basis of installed capacity, while the long-period path of potential growth is an undetermined notion from an *ex ante* point of view. No exogenous path can be identified representing the long-run tendencies of the economy: trend and business cycle must be analysed together, with the latter influencing and determining the former.

4.7 Conclusions: policy implications

Although their uncertain and problematic nature is often recognised, empirical estimates of potential output play an extremely relevant role in shaping and binding policies. As has been shown, in the supply-side framework the actual content of such empirical estimates is essentially the *ex post* record of the economy's average actual growth path, which has led analysts to somehow recognise the adverse effect of the recent great crisis on potential output.

What matters are policy implications. Since in the supply-side framework demand policies may at best be aimed at stabilising the cycle, while playing no role whatsoever in the determination of potential, it follows that a lower estimated level of potential output, and thus a smaller negative output gap, implies less need and less scope for expansionary demand policies (ECB, 2011, p. 82). As maintained in the Economic Synopses of the Federal Reserve Bank of St Louis (2012, no 11, p. 2), 'the gap may be closing faster than we thought because potential GDP is lower than we thought. And, if potential GDP is lower than expected, then interest rates may have to rise sooner than expected to prevent an acceleration of inflation.'

As maintained by Solow (2000, p. 10), this *ex post* definition of potential output, whereby what the economy has actually realised is assumed to be not too far from potential, seems ultimately to respond only to the theoretical preconception according to which expansionary demand policies are normally to be avoided for their inflationary consequences, while all that is needed are a flexible labour market and structural reforms (ECB, 2011, p. 85).

The endogeneity of potential output and potential growth which are proper to the demand-led growth approach have entirely different policy implications. In the first place, if output is liable to be demand-constrained both in the short and the long run and the economy is devoid of self-regulating mechanisms, demand policies may prove useful not only for short-run stabilisation but also for growth. In the second place, the path dependence of the growth trajectories and the plurality of potential output paths at each moment in time imply wide scope for all kinds of industrial and capacity-building policies aimed at shaping and guiding the economy's growth along a socially chosen path.

In principle it is possible, in the demand-led growth context, to identify an estimation method of potential output which may provide a measure of the economy's performance and a guide for policy action.

Like Okun's method, it should aim at quantifying the 'visible' difference between what can be produced and what is actually produced in each short period with given capacity, while both theory and experience show the role of actual realisations in shaping the economy's future possibilities and so the impossibility of quantifying even approximately the growth of potential in time as an *ex ante* notion.

Unlike Okun's method, it should avoid any reference to inflation data in the estimation of potential output, if only in the definition of the single non-inflationary rate of unemployment as the potential one. In addition to the failure of mainstream models to establish a clear empirical relation between output and inflation, historical experience shows a variety of output–inflation relationships, such as phases of increasing unemployment at constant or even increasing inflation, decreasing inflation at persistently high unemployment rates, seemingly horizontal Phillips curves in some specific sub-periods (see IMF, 2006), and so on. It seems we must conclude that no general and systematic relationship is observable in reality,[20] and take this into account when attempting to quantify potential output.

Notes

1. See for example the Congressional Budget Office's annual *Update to the Budget and Economic Outlook* which contains 10-year projections for the US economy.
2. Reference is to the so-called 'Taylor rule', used extensively both in New Keynesian models and as a policy procedure by a number of central banks.
3. See the European Commission's bi-annual publication 'Cyclical Adjustment of Budget Balances'.
4. See, for some instances, the OECD's *Economic Outlook*, the IMF's *World Economic Outlook*, the ECB's *Monthly Bulletin*.
5. Increasing (decreasing) inflation is assumed to have contractionary (expansionary) effects on aggregate demand through the Keynes effect and the Pigou effect, provided it is not accommodated by monetary policy.
6. 'Business cycle effects and autonomous demand shocks of various kinds should wash out if we take a long enough period' (Nickell, 1997, p. 71–2).
7. In coherence with the hypotheses of the natural rate models, the Phillips curve used for estimation is assumed to be vertical in the long run while admitting only a short period trade-off between unemployment and inflation.
8. Technical progress is often exogenous and usually defined merely in terms of the growing ability of the economy to produce increasing quantities of the same single commodity with no attention paid to qualitative changes.
9. Some examples are the Baxter-King (1995) filter, which identifies high-frequency, medium-frequency and low-frequency changes in the time series;

the Beveridge-Nelson (1981) filter, which assumes a particular correlation between cycle and trend determinants; the Kalman filter based on decomposition of the series in trend, cycle and erratic component.

10. Woodford states that this class of models responds to the need for central banks to have a theoretical basis in justification and support of their action.
11. As already noted, this statement must be qualified for NAIRU models with full hysteresis. See below, Section 4.4.1.
12. 'When Milton Friedman first proposed the natural rate hypothesis ..., it sounded like royal edict had established the natural rate as another one of the universe's invariant constants. Today, there is general recognition that if a NAIRU exists, it must be changing over time' (Stiglitz, 1997, pp. 5–6).
13. The change in the level of Italian unemployment in the 1980s and 1990s compared to the previous period, for example, is entirely attributed to an exogenous change in the NAIRU.
14. Still more drastically, Franz (2005, p. 30) maintains: 'To put it differently, yet taking the risk of an oversimplification: What has been estimated with a highly sophisticated machinery is simply the trend unemployment rate. If so, this can be carried out much easier and within a few minutes, just by using an HP filter.'
15. 'Descriptively, the mean of the gap measure should be close to zero over longer time horizons' (Billmeier, 2009, p. 396). It must be further noted that such institutions as the IMF and the OECD usually revise the estimates according to specific knowledge of the economies considered: 'In all cases, estimates of potential output incorporate a substantial amount of judgement and country-specific expertise of desk officers' (De Masi, 1997, p. 10). 'Such measures may be qualified, sometimes heavily, by the judgement of country specialists' (Giorno *et al.*, 1995, p. 168).
16. It is thus a very different notion from the NAIRU. According to Tobin (1995, p. 39), 'in Keynes open-end inflation results from an "inflationary gap" in aggregate demand, while a "deflationary gap" leads to comparative stability of prices or price trends'.
17. The original formulation of Okun's Law is based on observation on US data of the period 1947–60. Okun (1962) explains the empirical relationship on the idea that increases in working hours, productivity and participation accompany the expansionary phases, while the reverse occurs in slumps.
18. For a survey of different theories within the demand-led growth approach see Setterfield (2002).
19. The absolute majority of demand-led growth models, among which those of the Cambridge tradition, are cast in terms of steady-state growth paths. In the Kaleckian models (see Blecker in Setterfield, 2002), however, the steady state is deemed compatible with a constant but non-normal degree of utilisation of capacity, thus implicitly contradicting the tendency to capacity adjustment.
20. Contrary to the widespread interpretation of the Phillips curve, it can be maintained that also in Phillips's (1958) original article the relationship between money wage variations and unemployment is described not as a definite function, but as a broad relationship which can change due to various different influences. See Palumbo (2010).

References

Ball, L. (1996) 'Disinflation and the NAIRU', *NBER Working Papers* 5520.

Basu, S. and Fernald, J. G. (2009) 'What Do We Know and Not Know about Potential Output?', *Federal Reserve Bank of San Francisco Working Paper* 2009-05, March.

Baxter, M. and King, R.G. (1995) 'Measuring Business Cycles Approximate Band-pass Filters for Economic Time Series', *NBER Working Papers* 5022.

Beveridge, S. and Nelson, C. R. (1981) 'A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the "business cycle"', *Journal of Monetary Economics*, 7: 151–74.

Billmeier, A. (2004) 'Ghostbusting: Which Output Gap Measure Really Matters?', IMF Working Papers 146, updated version in (2009), *International Economics and Economic Policy*, 6: 391–419.

Blanchard, O. and Katz, L. F. (1996) 'What We Know and do Not Know about the Natural Rate of Unemployment', *NBER Working Papers* 5822.

Blanchard, O. and Summers, L. (1987) 'Hysteresis and the European Unemployment Problem', *NBER Working Papers* W1950.

Boone, L. (2000) 'Comparing Semi-structural Methods to estimate Unobserved Variables', *OECD Economics Department Working Papers* 240.

Carlin, W. and Soskice, D. (1990) *Macroeconomics and the Wage Bargain* (Oxford: Oxford University Press).

Ciccone, R. (1986) 'Accumulation and capacity utilisation: some critical considerations on Joan Robinson's Theory of Distribution', *Political Economy*, 2: 17–36.

Congressional Budget Office (2004) 'A Summary of Alternative Methods for Estimating Potential GDP', *Background Paper*, March, http://www.cbo.gov/.

Cotis, J. P., Elmeskov, J. and Mourougane, A. (2005) 'Estimates of Potential Output: Benefits and Pitfalls from a Policy Prospect', in L. Reichlin (ed.), *The Euro Zone's Business Cycle: Stylised Facts and Measurement Issues*, CEPR.

Council of Economic Advisors (1962) *Economic Report of the President* (Washington: US Government Printing Office).

Cross, R. (ed.) (1995) *The Natural Rate of Unemployment* (Cambridge: Cambridge University Press).

De Masi, P. R. (1997) 'IMF Estimates of Potential Output: Theory and Practice', *IMF Working Papers* 177.

DeLong, B. (2002) 'Forecasting growth and unemployment', July, http://www.j-bradford-delong.net/

Denis, C., McMorrow, K. and Röger, W. (2002) 'Production Function Approach to Calculating Potential Growth and Output Gaps – Estimates for the EU Member States and the US', *European Commission, Directorate-General for Economic and Financial Affairs, Economic Papers* 176.

European Central Bank (2009) *Monthly Bulletin*, July.

European Central Bank (2011) *Monthly Bulletin*, November.

European Commission (2009) 'Impact of the Current Economic and Financial Crisis on Potential Output', *European Economy Occasional Papers* 49, June.

Fabiani, S. and Mestre, R. (2001) 'A System Approach for Measuring the Euro Area NAIRU', *ECB Working Papers* 65.

Franz, W. (2005) 'Will the (German) NAIRU Please Stand Up?', *Zentrum für Europäische Wirtschaftsforschung* (*Centre for European Economic Research*), Discussion Paper n. 03-35.

Galbraith, J. K. (1997) 'Time to Ditch the NAIRU', *Journal of Economic Perspectives*, 11: 93–108.

Galí, J. and Gertler, M. (2007) 'Macroeconomic Modeling for Monetary Policy' *Journal of Economic Perspectives*, 21: 25–46

Galí, J., Smets, F. and Wouters, R. (2011) 'Unemployment in an Estimated New Keynesian Model', *NBER Working Papers* 17084, National Bureau of Economic Research.

Garegnani, P. (1992) 'Some Notes for an Analysis of Accumulation', in J. Halevi, D. Laibman and E. Nell (eds), *Beyond the Steady-state* (Basingstoke and London: Macmillan).

Giorno, C., Richardson, P., Roseveare, D. and van den Noord, P. (1995) 'Potential Output, Output Gaps and Structural Budget Balances', *OECD Economic Studies* 24.

Gordon, R. J. (1989) 'Hysteresis in history: was there ever a Phillips curve?', *American Economic Review*, 79: 220–5.

Gordon, R .J. (1997) 'The time-varying NAIRU and its implications for economic policy', *Journal of Economic Perspectives*, 11: 11–32.

Gros, D., Alcidi, C., Biggs, M., Mayer, T., Van Ark, B. (2010) 'Forum: The Impact of the Financial Crisis on the Real Economy'; *Intereconomics*, 45(1), (January/ February): 4–20.

Hodrick, R. J. and Prescott, E. C. (1981) 'Post-war U.S. Business Cycles: An Empirical Investigation', *Northwestern University Discussion Paper* 451, re-published in *Journal of Money, Credit and Banking*, 29 (1997): 1–16.

International Monetary Fund (2006) *World Economic Outlook*, April.

Jenkinson, T. J. (1988) 'The NAIRU: Statistical Fact or Theoretical Straitjacket?', in R. Cross (ed.), *Unemployment, Hysteresis and the Natural Rate Hypothesis* (Oxford: Blackwell).

Kaldor, N. (1985) *Economics without Equilibrium* (Armonk, NY: Sharpe).

Karanassou, M., Sala, H. and Snower, D. (2003) 'The European Phillips curve: Does the NAIRU exist?', *Applied Economics Quarterly*, 49(2): 93–121.

Kydland, F.. and Prescott, E. (1982) 'Time to build and aggregate fluctuations', E*conometrica*, 50: 1345–1370.

Landes, D. (1990) 'Why are we so rich and they so poor?', *American Economic Review*, 80: 1–13.

Laxton, D. and Tetlow, R. (1992) 'A simple multivariate filter for the measurement of potential output', *Technical Report* 59, Bank of Canada.

Layard, R., Nickell, S. and Jackman, R. (1991) *Unemployment, Macroeconomic Performance and the Labour Market* (Oxford: Oxford University Press).

Nickell, S. (1997) 'Unemployment and labor market rigidities: Europe versus North America', *Journal of Economic Perspectives*, 11: 55–74.

Okun, A. (1962) 'Potential GNP: its measurement and significance', *Cowles Foundation Paper* 190.

Palumbo, A. and Trezzini, A. (2003) 'Growth without normal capacity utilisation', *European Journal of the History of Economic Thought*, 10: 109–35.

Palumbo, A. (2010) 'La teoria del salario monetario di Phillips e la curva di Phillips', in G. Bonifati and A. Simonazzi, *Il ritorno dell'economia politica. Saggi in ricordo di Fernando Vianello* (Roma: Donzelli): pp. 79–102.

Pesaran, H. and Smith, R. (1995) 'The Natural Rate Hypothesis and its Testable Implications', in R. Cross (ed.), *The Natural Rate of Unemployment: Reflections on 25 years of the Hypothesis* (Cambridge: Cambridge University Press): pp. 203–30.

Prescott, E. (1986) 'Theory Ahead of Business Cycle Measurement', *Federal Reserve Bank of Minneapolis Quarterly Review*, Fall.

Richardson, P., Boone, L., Giorno, C., Meacci, M., Rae, D. and Turner, D. (2000) 'The Concept, Policy Use and Measurement of Structural Unemployment: Estimating a Time-Varying NAIRU across 21 OECD Countries', *OECD Economics Department Working Papers* 250.

Schreiber, S. and Wolters, J. (2007) 'The long-run Phillips curve revisited: is the NAIRU framework data-consistent?', *Journal of Macroeconomics*, 29: 355–67.

Serrano, F. (1995) 'Long-period effective demand and the Sraffian supermultiplier', *Contributions to Political Economy*, 14: 67–90.

Serrano, F. (2006) 'Mind the Gap: Hysteresis, Inflation Dynamics and the Sraffian Supermultiplier', Universidade Federal de Rio de Janeiro, mimeo.

Setterfield, M. (ed.) (2002) *The Economics of Demand-led Growth* (Cheltenham: Edward Elgar).

Sims, C. A., Edge, R.M., Gürkaynak, R. S. and Reis, R. (2010) 'How Useful Are Estimated DSGE Model Forecasts for Central Bankers?' *Brookings Papers on Economic Activity*, Fall, 209–59.

Solow, R. M. (1986) 'Unemployment: getting the questions right', *Economica*, 53, suppl., S23–S34.

Solow, R. M. (2000) 'Unemployment in the United States and in Europe. A contrast and the reasons', *CESifo Working Paper Series* 231.

Staiger, D., Stock, J. H. and Watson, M. W. (1997) 'The NAIRU, unemployment and monetary policy', *Journal of Economic Perspectives*, 11: 33–59.

Stiglitz, J. (1997) 'Reflections on the natural rate hypothesis', *Journal of Economic Perspectives*, 11: 3–10.

Stirati, A. (2001) 'Inflation, unemployment and hysteresis: an alternative view', *Review of Political Economy*, 13: 427–51.

Tobin, J. (1995) 'The Natural Rate as New Classical Macroeconomics', in R. Cross (ed.), *The Natural Rate of Unemployment* (Cambridge: Cambridge University Press).

Vetlov, I., Hlédik, T., Jonsson, M., Kucsera, H. and Pisani, M. (2011) 'Potential Output in DSGE Models', *ECB Working Paper* 1351.

Woodford, M. (2003) *Interest and Prices: Foundations of a Theory of Monetary Policy* (Princeton, NJ: Princeton University Press).

5
A Historical Approach to Demand-Led Growth Theory

*Matthew Smith**

5.1 Introduction

The central purpose of this chapter is to construct an analytical framework for explaining growth in concrete terms by reference to history consistent with the view that economic growth is fundamentally determined by the growth in aggregate demand.[1] As based on the Keynesian principle of effective demand, a demand-led theory of growth supposes that the level of aggregate output is determined in the long run by aggregate demand in which saving endogenously adjusts to autonomous demand through changes in income and output associated with the adjustment of productive capacity to aggregate demand. In this approach it is the growth in demand which determines the growth in output and the rate of capital accumulation in which it is supposed there is no technological constraint on output adjusting to demand growth. Key factors in explaining growth, notably, technical progress, are therefore conceived to contribute to economic growth through their effect on the growth in demand. From this standpoint, growth is a complex process, entailing structural change of the economic system, such that it can only be plausibly explained in concrete terms by reference to social, politico-institutional and technological factors. All these factors are seen to have an historical dimension in explaining growth.

The demand-led theory constructed here proceeds by building upon the Keynesian theory of demand-led growth consistent with the

*I am grateful to Tony Aspromourgos, John King, Heinz Kurz, Rod O'Donnell, Attilio Trezzini and an anonymous referee for some valuable comments on draft versions of this paper, without implicating them in the present product.

surplus approach to value and distribution of classical economics as reconstructed by Sraffa (1960). A critical feature of this theory, which clearly distinguishes it from supply-side growth theory developed on the basis of marginalist principles,[2] is the absence of any functional relationship between the quantity of inputs to be employed productively and the (relative) prices of those inputs which could provide a tendency for the economic system to gravitate toward full employment output. This feature stems from the fundamental analytical separability that exists in classical economics between, on the one hand, the determination of long-period normal prices and distribution and, on the other hand, the determination of outputs and the aggregate level of output as well as employment (see Garegnani 1984; 1990, pp. 122–32). In the framework presented this means that at any long-period positions along a growth path the determination of normal distribution and prices is conceived to correspond with the determination of long-run equilibrium levels of aggregate output at which demand is not necessarily sufficient to bring about the full employment of productive inputs.[3]

To begin with, in Section 5.2 we build a 'supermultiplier' growth model in which utilisation of an economy's productive capacity is assumed to always correspond to a *given* normal utilisation of capacity. On the basis of this assumption we derive a familiar 'steady-state' growth model in which output and the capital stock are conceived to grow constantly at the same rate for a given technique. We shall see, however, that such a growth model is not truly consistent with the fundamental Keynesian conception that in the long run demand is autonomous of saving in the determination of output along a growth path. In Section 5.3 we construct an alternative growth model in which the utilisation of capacity is conceived to vary both in the short and in the long run. Based on this conception the long-run average utilisation of capacity is endogenously determined and is systematically different from normal capacity utilisation. A novel feature of our analytical framework is that it is based on historical periodization in which trend economic growth from one period to the next is conceived to be determined not only by the growth rate of autonomous demand but also by long-run changes in the value of the supermultiplier. A short conclusion is provided in Section 5.4.

5.2 The growth model with normal utilisation

The Keynesian demand-led growth model employed incorporates a 'supermultiplier' of induced expenditure originally developed by Hicks (1950) which links quantitatively autonomous demand to equilibrium output and income. This particular supermultiplier model has been articulated in the literature, notably by Serrano (1995) and, from a critical viewpoint, by Trezzini (1995, 1998). A feature of the model is that productive capacity is determined by long-run aggregate demand. In the model there are three basic components of aggregate demand (AD_t), consisting of autonomous demand (A_t), induced consumption expenditure ($c_t Y_t$) and induced investment (I_t^I) which also contributes to productive capacity:

$$AD_t = A_t + c_t Y_t + I_t^I \qquad [5.1]$$

where c_t is society's marginal propensity to consume with values $0 < c_t < 1$. The first component, autonomous demand, consists of those expenditures that are explained independent of changes in income and output occurring over the same time period. It is essentially that part of aggregate demand which, along with induced investment, is accommodated by saving (including taxation) that is endogenously generated by income. In a closed economy these expenditures consist of government expenditure, autonomous investment and autonomous consumption, while in an open economy they include exports which, though different from other components, are accommodated by foreign income and when this is less than imports the margin of difference is accommodated by foreign saving. For simplicity, we shall assume a closed economy. Again for simplicity, we shall also assume that autonomous demand does not create additional productive capacity – that is, it is non-capacity-generating expenditure. The second component, induced consumption, is that consumption which is a positive function of the current level of income and output. As is well known, its relationship to income is defined by the marginal propensity to consume whose value is conceived to depend on socio-institutional factors, most notably, the distribution of income and, connectedly, the taxation and welfare system.

The third component is induced investment, through which productive capacity is conceived to adjust to aggregate demand. Based on the accelerator principle, induced investment will depend on the amount of productive capacity that needs to be installed for a given technique of production to ensure the level of output accommodates expected demand. For simplicity, we will initially assume there is zero depreciation on fixed capital and employ a rigid accelerator to express induced investment in a familiar way as:

$$I_t^I = (K_{t+1} - K_t) = a_t(Y_{t+1}^e - Y_t) \qquad [5.2]$$

where K_{t+1} is the capital stock required in the future period to accommodate the expected level of demand, Y^{e}_{t+1}, and a is the capital–output ratio. In order to account explicitly for the role of capacity utilisation the capital–output ratio can be expressed as a/u, determined as follows:

$$K_t/Y_t = (K_t/Y^{*}_{t}) \;.\; (Y^{*}_{t}/Y_t) = a_t/u_t \qquad [5.3]$$

where a_t is the capital–output ratio, K_t/Y^{*}_{t}, when capacity is fully utilised (i.e. $u_t = 1$) and u_t is the degree of capacity utilisation, defined *as the ratio of actual output to full-capacity output,* Y_t/Y^{*}_{t}, for a given capital stock. Rearranging equation [5.3] we obtain an expression for capacity utilisation:

$$u_t = a_t Y_t/K_t \qquad [5.4]$$

On the plausible assumption that, typically, firms install productive capacity to produce with spare capacity to meet peak demand as well as to enable an expansion in output to capture greater sales revenue in the event of persistent higher demand, the degree of normal (or desired) utilisation, u^{n}_{t}, will have a value between zero and full capacity (i.e. $0 < u^{n}_{t} < 1$).[4] Given that u^{n}_{t} is the ratio of the desired level of output produced to the full capacity of installed capital, Y^{d}_{t}/Y^{*}_{t}, then the desired capital–output ratio, a_t/u^{n}_{t}, is determined as follows:

$$K_t/Y^{d}_{t} = (K_t/Y^{*}_{t}) \;.\; (Y^{*}_{t}/Y^{d}_{t}) = a_t/u^{n}_{t} \qquad [5.5]$$

And, therefore:

$$u^{n}_{t} = a_t Y^{d}_{t}/K_t \qquad [5.6]$$

In short, normal utilisation for an economic system is that which reflects the utilisation of capacity that firms determine will maximise their profit rates for a given technique, taking account of the possible fluctuations in actual demand and its impact on average costs over a period of time relevant to the installation of their existing capacity (see Ciccone, 1986, pp. 23–32). On this basis equation [5.2] for induced investment is re-written as:

$$I^{I}_{t} = (K^{d}_{t+1} - K_t) = (a_t/u^{n}_{t})(Y^{e}_{t+1} - Y_t) \qquad [5.7]$$

where K^{d}_{t+1} is the capital stock *desired* in the future period based on the expected level of demand, Y^{e}_{t+1}, and the *desired* capital–output ratio

a_t/u_t^n as based on the dominant techniques of production and the normal utilisation of productive capacity. This accelerator relationship supposes that through time net investment ensures the capital stock adjusts to produce output levels according to the desired capital–output ratio.

By substituting equation [5.7] for I_t^I in equation [5.1] we obtain the following aggregate demand function:

$$AD_t = A_t + c_t Y_t + (a_t/u_t^n)\,(Y_{t+1}^e - Y_t) \qquad [5.8]$$

Solving for equilibrium income:

$$Y_t = A_t + c_t Y_t + (a_t/u_t^n)\,(Y_{t+1}^e - Y_t) \qquad [5.9]$$

and, with expected growth in output, $g_t^e = (Y_{t+1}^e - Y_t)/Y_t$ and, re-arranging, we obtain:

$$Y_t = A_t / [1 - c_t - (a_t/u_t^n) g_t^e] \qquad [5.10]$$

If it is then assumed firms have perfect foresight, expected growth, g_t^e, will be equal to the growth rate of output (and income), g_t^y. And if we substitute the propensity to save, s_t, for $1 - c_t$, the following expression is obtained:

$$Y_t = A_t / [s_t - (a_t/u_t^n) g_t^y] \qquad [5.11]$$

A positive value of Y_t requires that given $0 < s_t \leq 1$, $s_t > (a_t/u_t^n) g_t^y$. The equilibrium income so determined may be called 'capacity income' because it corresponds to a level of output produced at the normal utilisation of capacity. By re-arranging equation [5.11] a familiar growth equation for a supermultiplier model is obtained:

$$g_t^y = [s_t - (A_t/Y_t)] / (a_t/u_t^n) \qquad [5.12]$$

The equilibrium growth rate, g_t^y, is determined by the ratio A_t/Y_t for a range of possible values up to a maximum value of $g_t^y = s_t/(a_t/u_t^n)$, when $A_t = 0$. This equilibrium growth is that necessary to ensure capacity saving, being that level of saving which is generated from income when output is produced at the normal utilisation of capacity is equal to autonomous expenditure plus induced investment.[5] It is also the growth rate at which the degree of utilisation conforms continuously to normal utilisation along a steady-state growth path in which the capital stock and output continuously grow at the same rate for a given production technique.

A major problem with this steady-state growth model is that it is not really compatible with the fundamental Keynesian notion that demand is autonomous in the determination of the trend growth of output (Trezzini 1995, pp. 48–56). This can be explained by reference to the growth equation [5.12]. According to this equation, for the given ratio A_t/Y_t which determines g_t^y to remain constant along the trend growth path, the growth rate of autonomous demand, g_t^A, must be equal to the steady-state growth rate: i.e. $g_t^A = g_t^y$. However, this means that the growth rate of autonomous demand is limited in the sense that $g_t^A < s_t/(a_t/u_t^n)$ consistent with $A_t > 0$, where $s_t/(a_t/u_t^n)$ is capacity saving as a ratio of the capital stock.[6] The reason for this limitation in the model is that growth in autonomous demand, which is equal to or greater than $s_t/(a_t/u_t^n)$, cannot be accommodated by the growth in capacity saving necessary for equilibrium along the steady-state growth path. If demand is truly autonomous there appears no logical reason why its growth should be so bound by capacity saving. Connected to this is the peculiarity in equation [5.12] of the inverse relationship between the ratio A_t/Y_t and g_t^A on the basis of $g_t^A = g_t^y$ and for given values of s_t, a_t and u_t^n. Again, if demand is truly autonomous there is no plausible basis for supposing that its growth should systematically increase as the ratio of autonomous demand to capacity income (i.e. A_t/Y_t) decreases and, *vice versa* (Trezzini 1995, pp. 52–3). In the steady-state growth model the logic for this inverse relationship is that as the magnitude of the latter ratio decreases (increases) an increasing (decreasing) proportion of capacity saving can be devoted toward induced investment and, thereby, toward augmenting (diminishing) the growth in capacity, its output and, causally, in the demand necessary to realise equilibrium growth. Hence, in this steady-state model the growth in autonomous demand ultimately depends on saving which is generated by the equilibrium growth in capacity income. This underlies the lack of autonomy of demand in the growth process when the trend rate of output growth and the saving which is generated by it is based on a *given* normal utilisation of capacity.

Our argument then is that under steady-state conditions in which it is supposed that there is a given normal utilisation of capacity along the trend growth path, aggregate demand is denied an autonomous role in the determination of economic growth. Nevertheless, as shown by Garegnani (1992), this theoretical problem can be surmounted by allowing the degree of capacity utilisation to vary both in the short and long run so that any level of autonomous demand (investment) can be accommodated by the generation of saving induced through changes in income and output facilitated by changes in capacity utilisation as well as in productive capacity. By allowing for persistent as well as temporary variations in the utilisation of capacity, long-run output has the elasticity to accommodate changes in aggregate demand beyond the steady-state for a given propensity to save (*ibid.*).[7] Importantly, this variability in capacity utilisation ensures that aggregate demand has an autonomous role in the growth process, which is crucial to the Keynesian approach (see Trezzini 1995, pp. 48–57; Palumbo and Trezzini 2003, pp. 110–14). It is clear, though, that this conception of the growth process cannot be reconciled with steady-state growth since the capital stock and output will be systematically growing at different rates.

5.3 Alternative growth model with endogenous utilisation

In an attempt to incorporate long-run elasticity of output into a super-multiplier growth model, Serrano (1995) proposed that consistent with variability in capacity utilisation the *average* utilisation of capacity which emerged over time would be the same as the normal utilisation of capacity. Unlike the steady-state model, the utilisation of capacity is not assumed to be constant but rather the *given* normal utilisation of capacity is proposed to correspond to an average of its fluctuations over time. This conception therefore brings in historical time with all the variables expressed as averages, including the expected growth in demand of firms, in the determination of an average rate of growth. To express this conception equation [5.12] can be re-written as:

$$g^y_t = [s_t - (A_t / Y_t)] / (a_t / u^a_t) \qquad [5.13]$$

where u^a_t is the average utilisation of capacity and $u^a_t = u^n_t$.

The problem with this conception is that there appears to be no compelling reason why the average utilisation of capacity which emerges over time should be equal to the normal utilisation of capacity. Indeed, as shown by Trezzini (1998, pp. 59–66), even on the assumption of perfect foresight, any deviation of actual growth from the equilibrium growth rate associated with a normal utilisation of productive capacity will require average utilisation to vary significantly from normal over a considerable period of time to restore the equilibrium rate of growth. Furthermore, the process of adjustment itself can cause the growth in capacity to change in relation to output growth since changes in utilisation which, in the long run, affect capacity, simultaneously affect demand. Moreover, in the long-run adjustment of capacity to demand, investment induced by deviations of average from normal utilisation will simultaneously affect demand as well as productive capacity. Hence, as Palumbo and Trezzini (2003, pp. 115–20) have argued, once it is acknowledged that the utilisation of capacity may vary such as to ensure long-run elasticity in output that can accommodate any feasible level of aggregate demand, then the adjustment process of capacity to demand is a path-dependent one that means average utilisation is, except by rare coincidence, unlikely to be equal to normal utilisation, notwithstanding investment decisions by firms to achieve it.

The question is: where do these analytical issues leave the demand-led growth theory presented in this chapter? The answer is a more modest theory which does not pretend *to account fully* for the growth process but nevertheless provides a framework for analysing the central causes of *trend growth and economic development* in a more concrete way, consistent with the notion that the growth in aggregate output is fundamentally determined by the growth in aggregate demand. A historical approach is proposed, suggested in Serrano (1995),[8] in which the growth model provides a demand-led framework for a concrete explanation of the average growth rate over historical time periods, termed here as 'epochs'. While these epochs must at a minimum be long enough for the adjustment of fixed productive capacity by firms to longstanding demand conditions, they can otherwise be defined arbitrarily according to their significance in explaining growth trends by reference to key historical events as well as to the character of the demand-led forces which are ascertained to determine economic development and growth performance. Thus, for example, an epoch could be defined by reference to the event of war, a change in the international economic regime, a fundamental change in policymaking, an unprecedented structural change in the economic system or a combination of these or other such historically related events. The theoretical counterpart to epoch in our model is 'period', in which the long-run average growth rate is conceived to be determined by the persistent forces of demand specified in our demand-led theory. In this approach the average growth rate in each period is conceived to be linked to that of the previous period so that growth in period t can only be properly explained by reference to the history of the growth process in period $t-1$ and, prior to this, period $t-2$ and so on back to period $t-n$. Hence, the long-run average growth rate in any period is determined by demand-led forces which have a historical context and, in concrete terms, are to be explained by reference to history. As further clarified below, in correspondence with epochs, the long run of the periods used are at a minimum long enough for fixed productive capacity to adjust to expected demand conditions consistent with long-run growth.

Firstly, in accordance with the foregoing concept of long-run elasticity in which changes in the degree of capacity utilisation play an active role in the adjustment of saving to autonomous demand (and induced investment) along with output adjusting to aggregate demand, the model shall suppose that the utilisation of capacity is endogenously determined in the growth process. This means that in any period the average utilisation of capacity is conceived to be endogenously determined.[9] Based on the reasoning given above, the average utilisation so determined is not conceived, except by coincidence, to equal the normal utilisation of capacity upon which firms base their investment decisions in adjusting their capacity to demand. This conception requires us to suppose that expectations of future demand by firms are not normally realised. Secondly, therefore, in the model the unrealistic assumption of perfect foresight is dispensed with, such that the expected growth of demand for a firm is not necessarily equal to its actual growth (i.e. $g^e_t \neq g^y_t$). While it is acknowledged that firms will continuously adjust their expectations of growth in demand to historical growth rates, unless the growth rate is stable for a very long period of time it is not plausible to assume that their expectations will be systematically correct. Once steady-state growth is abandoned and the growth rate is conceived to be determined by demand in a path-dependent way perfect foresight has little plausibility. Thirdly, as already anticipated the model will take a more general form and suppose the existence of fixed capital. This means we must account for the effect of the rate of depreciation of the capital stock on induced investment by rewriting equation [5.7] above to:

$$I^I_t = (a_t / u^n_t)(Y^e_{t+1} - Y_t) + (a_t u^n_t d_t) Y_t \qquad [5.14]$$

where I^I_t is induced investment and d_t is the average rate of depreciation of *utilised* capital in period *t*. The second term on the right-hand side of equation [5.14] clearly expresses the notion that the rate of depreciation of the capital stock increases with its utilisation. With respect to the effect of depreciation on induced investment, the equation shows that induced investment in our model is conceived to be based on the rate of depreciation expected by firms to occur *at the normal utilisation* of the capital stock. However, because average utilisation will be systematically different from normal, the depreciation of the capital stock which occurs will be systematically different from that expected. As is elaborated below, this unexpected depreciation of the capital stock can influence future induced investment by, in turn, contributing to the deviation of average from normal utilisation.

Fourthly, since the degree of utilisation of capacity that is realised will, except by chance, be different from normal, we need to account for the effect of this systematic deviation on induced investment. The model is able to account, however mechanically, for the manner in which capacity is conceived to adjust to aggregate demand in a demand-led growth theory. It is proposed that deviations of average utilisation realised in the previous historical period $t-1$ from normal utilisation in period t will tend to induce a change in investment in the current period, t, by firms endeavouring to adjust their capacity to demand so as to establish normal utilisation. This means that if normal utilisation in period t remains unchanged from that in period $t-1$, firms overall adjust their capacity to the historical deviation between average and normal utilisation occurring in period $t-1$. However, it is envisaged that in each period normal utilisation will be revised by firms when installing new capacity according to a complex of factors so that u_t^n can be different from u_{t-1}^n and so on. The degree of divergence of average from normal utilisation in period $t-1$ will clearly be a major factor in revising the normal degree of utilisation in period t.[10] On this basis it is supposed that the deviation of average from normal utilisation in one period will tend to induce a change in capacity-adjusting investment in the next period. This capacity adjustment mechanism supposes that period $t-1$ is sufficiently long that any deviation between average and normal utilisation can be considered systematic and firms can feasibly adjust their capacity in period t to expected demand conditions in the future period $t+1$. Nevertheless, this mechanism represents a simplification of the capacity-adjusting process, since the deviation between average and normal utilisation, especially in a period of stagnant economic growth, may only reflect a disparity between the actual and expected frequency of fluctuations in demand with peak demand well below full capacity. On the other hand, equality between average and normal utilisation may merely mask a significant increase in the amplitude of fluctuations in peak demand, requiring firms to make additional investment in capacity. In this respect, an underlying assumption of our model is that firms tend to adjust capacity at discrete intervals in each period when they install planned spare capacity which will on average be utilised over a long period of time according to the expected future growth in demand. Incorporating this conception with the stated qualifications into the determination of induced investment, equation [5.14] above is re-written as:

$$I_t^I = (a_t/u_t^n)\ (Y_{t+1}^e - Y_t) + a_t u_t^n\ d_t Y_t + (a_t/u_t^n - a_t/u_{t-1}^a)Y_t \qquad [5.15]$$

where u^a_{t-1} is the average degree of utilisation realised in period $t-1$ and the term $(a_t/u^n_t - a_t/u^a_{t-1})Y_t$ reflects the adjustment of capacity to demand to restore normal utilisation.[11] By re-expressing equation [5.15] it can be easily shown that *net* of the expected depreciation of capital, induced investment is the difference between the capital stock desired by firms to accommodate expected demand in period $t+1$ at normal capacity utilisation, K^d_{t+1}, and what the capital stock will otherwise be in period t at the existing average utilisation of capacity determined in period t–1, denoted as K^r_t:

$$I^I_t - (a_t u^n_t d_t)Y_t = (K^d_{t+1} - K^r_t) = (a_t/u^n_t)Y^e_{t+1} - (a_t/u^a_{t-1})Y_t \qquad [5.16]$$

where $I^I_t - (a_t u^n d_t)Y_t$ is induced investment *net* of expected depreciation of the capital stock. It will be convenient here to employ equation [5.15] rather than equation [5.16].[12] However, what this latter equation shows is that whereas in steady-state and other models discussed above induced investment changes at the same constant rate as output (and income), in our model it changes at a different rate from one period to the next according to changes in average utilisation brought about by unexpected changes in the growth rate of demand.[13] Accordingly, for this reason alone, the capital stock tends to grow at a different rate from one period to the next in our model.

These elements can be represented in our model by rewriting equation [5.10] above as follows:

$$Y_t = A_t/[1 - c_t - (a_t/u^n_t)g^e_t - a_t u^n_t d_t - (a_t/u^n_t - a_t/u^a_{t-1})] \qquad [5.17]$$

where all variables are expressed as 'averages' so that g^e_t refers to the expected average growth in demand in period t and the condition $1 > [c_t + (a_t/u^n_t)g^e_t + a_t u^n_t d_t + (a_t/u^n_t - a_t/u^a_{t-1})]$ is met. In absence of perfect foresight, expected average growth in demand (and hence, in output) will not be equal to the average growth in output in period t, g^y_t, such that $g^e_t \neq g^y_t$. Given the values of c_t (or s_t), a_t, d_t, u^n_t, u^a_{t-1} and g^e_t, which together determine the supermultiplier, and the level of autonomous demand, A_t, long-run average income and output is determined.[14] On the basis of this datum and the historically *given* capital stock (i.e. K_t) employed to produce output (i.e. Y_t) in period t, the average utilisation of capacity, u^a_t, will be endogenously determined as follows:

$$u^a_t = a_t Y_t / K_t \qquad [5.18]$$

and with $g^e_t \neq g^y_t$, then $u^a_t \neq u^n_t$. Hence, except when $g^e_t = g^y_t$, the average utilisation of capacity will be systematically different from normal and average utilisation will vary from one period to the next, such that $u^a_t \neq u^a_{t-1}$. The capital stock to determine average utilisation of capacity in period t in equation [5.18] is itself determined historically in the following way:[15]

$$K_t = K_{t-1} + I^I_{t-1} + (u^n_{t-1} - u^a_{t-1})a_{t-1}d_{t-1}Y_{t-1} \qquad [5.19]$$

The term $(u^n_{t-1} - u^a_{t-1})\ a_{t-1}d_{t-1}Y_{t-1}$ in equation [5.19] is the average depreciation of the capital stock in period $t-1$ which was *not expected* by firms when, through induced investment (i.e. I^I_{t-1}), they installed capacity to accommodate expected demand in period t (i.e. Y^e_t). Unexpected depreciation so affecting the capital stock is conceived to be systematic on account of the systematic difference between average and normal utilisation. Hence, for example, if $u^a_{t-1} > u^n_{t-1}$ because $g^e_{t-1} < g^y_{t-1}$, the depreciation of the capital stock will be greater than anticipated and, thereby, not compensated by induced investment, will tend to reduce the stock of capital available in period t. By affecting capacity in this way, unexpected depreciation will contribute to a higher average rate of utilisation determined in period t (i.e. u^a_t)[16] and, thereby, tend to contribute to its deviation from normal utilisation which, in the manner explained above, firms will endeavour to correct through induced investment in period $t+1$.[17]

On the basis of the analysis above the average growth rate in period t will be equal to:

$$g^y_t = Y_t - Y_{t-1} / Y_{t-1} \qquad [5.20]$$

where current average output, Y_t, is determined in equation [5.17] and output in the previous period is similarly determined according to the equation:

$$Y_{t-1} = A_{t-1} / [1 - c_{t-1} - (a_{t-1}/u^n_{t-1})g^e_{t-1} - a_{t-1}u^n_{t-1}d_{t-1} - (a_{t-1}/u^n_{t-1} - a_{t-1}/u^a_{t-2})] \qquad [5.21]$$

Now, for simplicity, we will denote the supermultipliers for period t and $t-1$ respectively as follows:

$$m_t = 1/[1 - c_t - (a_t/u^n_t)g^e_t - a_t u^n_t d_t - (a_t/u^n_t - a_t/u^a_{t-1})] \qquad [5.22]$$

$$m_{t-1} = 1/[1 - c_{t-1} - (a_{t-1}/u^n_{t-1})g^e_{t-1} - a_{t-1}u^n_{t-1}d_{t-1} - (a_{t-1}/u^n_{t-1} - a_{t-1}/u^a_{t-2})] \qquad [5.23]$$

The value of m_t will be different from m_{t-1} purely on the grounds that u^a_{t-1} is a different value to u^a_{t-2}. Thus, for example, even supposing $g^e_t = g^e_{t-1}$, $c_t=c_{t-1}$, $a_t=a_{t-1}$, $u^n_{t-1}=u^n_t$ and $d_t=d_{t-1}$, if $u^a_{t-1} > u^a_{t-2}$, then $m_t>m_{t-1}$. We can write the equations for the determination of long-run average output in period t and $t-1$ in the simple form:

$$Y_t=A_t \,.\, m_t \qquad [5.24]$$

$$Y_{t-1}=A_{t-1} \,.\, m_{t-1} \qquad [5.25]$$

Substituting equations [5.24] and [5.25] into [5.20] allows us to express the average growth rate of output in period t as:

$$g^y_t=A_t \,.\, m_t-A_{t-1} \,.\, m_{t-1}/A_{t-1} \,.\, m_{t-1} \qquad [5.26]$$

With re-arrangement and manipulation we can get the following demand-led growth equation for period t:

$$g^y_t=g^A_t+\Delta m_t(A_t/A_{t-1}) \qquad [5.27]$$

where g^A_t is the growth rate of autonomous demand and Δm_t is the change in the supermultiplier in period t as determined by $(m_t-m_{t-1})/m_{t-1}$.[18] This growth equation shows that the growth rate of output is determined by the growth rate of aggregate demand, as determined by two elements: (i) the growth rate of autonomous demand, g^A_t; and (ii) the *change* in the value of the supermultiplier, Δm_t. It is evident that if $m_t=m_{t-1}$ so that $\Delta m_t=0$, the growth of output will be determined wholly by the growth in autonomous demand; that is, $g^y_t = g^A_t$. While the growth in autonomous demand is conceived to be the main determinant, lasting changes in the supermultiplier can be a contributor to the determination of economic growth in this model.

The analytical limitations of our demand-led growth model should be mentioned here. In our model the *average* long-run growth of output and the *average* long-run growth in the capital stock over a period are the same, although it is supposed that output and the capital stock at any time within a period will be systematically growing at different rates associated with variations in the utilisation of capacity. This could not otherwise be the case in a 'growth equation' of the form of equation [5.27]. Our model is therefore not capable of accounting for the role of variations in the utilisation of capacity in the growth process *within any* period. Nor is it capable of accounting for interactions which occur between changes in capacity and demand along a growth path *within* a period. The complexity of these ongoing interactions in the growth process which, as discussed above, mean utilisation will be systematically different on average to normal (or desired), belongs to a separate analysis that accounts for cyclical changes in activity. Instead, our model accounts for variations between the average growth of output and capital accumulation associated with endogenous variations in average utilisation *between* different historical periods. It also endeavours to account for the process by which capacity adjusts to demand when the trend growth in demand deviates from that expected by firms and average utilisation systematically deviates from normal by reference to sequences of periods in which the trend growth rate changes from one period to the next. In this rather mechanical way the model represents interactions between demand and capacity along a path of changing *trend* rates of growth and accumulation. Notwithstanding these limitations, our model does provide an analytical framework of the fundamental Keynesian notion that the growth in output and capital accumulation is wholly determined by the growth in aggregate demand.

5.4 Conclusion

A central feature of the demand-led growth model is that unlike 'steady-state' models the growth rate depends not only on the growth rate of autonomous demand but also on the long-run change in the value of the supermultiplier. This stems from the historical periodisation incorporated into the model in which the supermultiplier will invariably be different from one period to next. Therefore, according to the model, trend growth in an historical period or epoch is explained not just by reference to factors determining the growth of autonomous demand but also by reference to factors which can cause long-run changes in the value of the supermultiplier such as changes in income distribution, technical change and the revision of expectations by firms about long-run demand. Moreover, in this model history is considered to play a central role in determining the value of the supermultiplier which, in any period, is dependent in part on events which have occurred in previous periods. The growth in autonomous demand nevertheless remains the main driving force of economic growth in the model. Hence, much of the explanation of growth will consist in identifying the key factors determining the growth in autonomous demand. These will include the longstanding fiscal policy of government, the long-running monetary policy of the central bank, developments in the financial system affecting the financing of private expenditure, market regulations and institutions affecting entrepreneurship and innovation and government policies on trade and industry, all of which influence in various ways the different components of autonomous demand. By reference to these factors, explaining growth will thus entail identifying the roles of the different components of autonomous demand: government spending, autonomous private investment, autonomous consumption and, for an open economy, exports.[19] Generated by a combination of these components it is evident that growth in autonomous demand can be driven by different sources in different historical periods. Hence, appealing to history, the role of government spending was a major driver in most advanced nations during the post-war recovery phase of the Second World War, whilst, in the last twenty years, autonomous consumption has played a more prominent role in generating demand than in previous historical periods. More intricately, the determination of the growth in autonomous demand as a whole will entail a causal interaction between its components and, therefore, should be part of any explanation of demand-led growth. In this respect, the sum will be greater than its parts. Hence, for example, large-scale government capital expenditure on transport and communications infrastructure, which improves productivity growth, is likely to augment the other components of autonomous demand by contributing to stronger growth in private investment in innovation and in autonomous consumption as well as contributing to a nation's better export performance. In this regard the causal relationship between these autonomous components of autonomous demand cannot be considered functional in the sense of the causal relationship supposed between income and demand, as defined by our supermultiplier. Instead, the causal relationship between them is conceived to be generally complex and contingent on a wide set of circumstances, such that they could only be properly explained in concrete terms in accordance with the historical approach proposed.

Notes

1. In our view this approach is consistent with the methodology employed by Adam Smith in the *Wealth of Nations* (1976 [1776]), in which a theoretical system informs an historical analysis of the major forces determining economic development. The important property of Adam Smith's theoretical system is that it is open to social and institutional factors playing a key explanatory role that can only be properly understood by reference to their history. On Smith's position, see Aspromourgos (2009, pp. 247–51).

2. Besides the Solow (1956) and Swan (1956) models, supply-side growth theory include the 'endogenous growth' models, of notably, Romer (1986; 1990) and Lucas (1988).
3. In classical analysis because quantities of commodities produced are *given* in the determination of normal prices and distribution the long-period method entails the gravitation of prices and distributive variables around their normal values but not quantities. Only in long-period marginalist economics in which normal prices and quantities are conceived to be simultaneously determined do quantities as well as prices gravitate around normal magnitudes. Thus, in classical analysis, the long-run equilibrium level of output and its composition upon which given, normal prices and distribution are determined along the growth path, can be conceived to be determined in a path-dependent manner so that it does not represent an equilibrium output in the sense that actual levels of output gravitate around it. Indeed, this methodological approach is adopted in the growth analysis which follows.
4. The notion that firms would permanently maintain excess productive capacity is originally attributable to Steindl (1952, pp. 4–14). Besides accommodating peak fluctuations in demand, Steindl argued that a 'more general reason' for maintaining excess capacity was that in competing with rivals, firms wanted to be in a position to expand their market share and establish their 'goodwill' in being able to reliably supply greater demand in the market.
5. This is simply given by the following equation:

$$s_t Y_t = A_t + (a_t/u_t^n)\,(Y_{t+1}^e - Y_t)$$

where it is assumed $Y_t > Y_t^m$, in which Y_t^m is the minimum level of income necessary given the existence of positive autonomous consumption, C^A, such that $s_t Y_t^m = C^A$. This condition is necessary in a global (or closed) economy for saving net of autonomous consumption to be positive (i.e. $s_t Y_t > C^A$).
6. By re-arrangement we can obtain a more comprehensible form for this term. Substituting K_t/Y_t^* for a_t into the term $s_t/(a_t/u_t^n)$ obtains $s_t/(K_t/Y_t^*/u_t^n)$ and then, by re-arrangement, to $s_t(Y_t^* u_t^n)/K_t$. This simplifies to $s_t Y_t/K_t$, where it is recalled that Y_t is capacity income.
7. This conception was largely proposed by Garegnani (1992) as a critique of the Cambridge conception that distribution was dependent on the rate of capital accumulation, which had been variously advanced by Kaldor (1955–56, pp. 94–100), Kahn (1959), Robinson (1962, pp. 11–13, 40–41) and Marglin (1984). In the Cambridge conception the limit to growth posed by existing capacity saving is essentially surmounted by generating growth in autonomous demand sufficient to cause a change in distribution which, in turn, induces a higher propensity to save so that additional saving is generated to accommodate the additional autonomous demand. This can be illustrated by reference to equation [5.12] above. Suppose, through government policy, g_t^A is increased to a rate which is higher than the existing g_t^y. This expansion in demand, through inflation, then brings about a redistribution of income from wages to profits, which, on the plausible assumption that capitalists have a higher propensity to save than wage earners, causes the value of s_t to increase, producing an increase in capacity saving necessary to facilitate

the expansion in autonomous expenditure. In this way the value of g_t^y will then adjust to a policy-determined g_t^A. It is supposed that in this process the resulting increase in the rate of capital accumulation will be associated with a higher rate of profit and, for a given technique, a lower real wage. In contrast, according to Garegnani's (1992) proposition, g_t^y can adjust to a higher g_t^A without any change in distribution by the degree of utilisation increasing in the long run. The increase in income (output) derived from a higher degree of utilisation of productive capacity is then able to generate the necessary saving to facilitate the additional autonomous demand. Importantly, this argument entails the rejection of a given normal utilisation and, thereby, of a steady-state growth model. For other related criticisms of the Cambridge approach to growth and distribution, see Vianello (1985), Ciccone (1986) and Garegnani and Palumbo (1998).

8. Also see Kaldor (1957, p. 601 n. 1).
9. This appears to be consistent with the conception briefly outlined by Ciccone (1987, pp. 103–6). Also see Amadeo (1986).
10. Hence, interpreted by reference to the frequency and amplitude of fluctuations in demand (and, hence, utilisation) which have occurred, the magnitude of divergence between u_{t-1}^n and u_{t-1}^a in period $t-1$ will provide important information to firms in the determination of u_t^n in period t (Amadeo, 1986, p. 155). Other major related factors which will influence the determination of normal utilisation are the technology embodied in newly installed capacity, the expected fluctuations in demand as based on historical experience and the degree of spare capacity planned for newly installed capacity (see Ciccone, 1986).
11. With respect to the third term on the right-hand side of equation [5.15], if we denote K_t^n as the capital stock with normal utilization and K_t^r the capital stock that would be realised in period t based on the average utilisation in period $t-1$, then $K_t^n - K_t^r = (a_t/u_t^n - a_t/u_{t-1}^a)Y_t$. Hence, for example, if $u_{t-1}^a > u_t^n$, this means for an existing level of demand and output, Y_t ,the capital stock that would be realised without any adjustment to induced investment in period t, K_t^r, is smaller than necessary for aggregate production to occur at a normal degree of utilisation: that is, $K_t^n > K_t^r$.
12. It should be noted that the aggregate level of investment in the model consists of induced investment determined according to equation [5.15] *plus* non-capacity creating autonomous investment which is part of and contributes to autonomous demand. On autonomous investment, see n. 19 below.
13. Note that $I_{t-1}^I - (a_{t-1}u_{t-1}^n d_{t-1})Yt_{-1} = (K_t^d - K_{t-1}^r) = (a_{t-1}/u_{t-1}^n)Y_t^e - (a_{t-1}/u_{t-2}^a)Y_{t-1}$, $I_{t-2}^I - (a_{t-2}\ u_{t-2}^n d_{t-2})Yt_{-2} = (K_{t-1}^d - K_{t-2}^r) = (a_{t-2}/u_{t-2}^n)Y_{t-1}^e - (a_{t-2}/u_{t-3}^a)Y_{t-2}, \ldots$ and so on.
14. This long-run position corresponds to equality between saving and autonomous expenditure *plus* induced investment, expressed as follows:

$$sY_t = A_t + a_t/u_t^n(Y_{t+1}^e - Y_t) + a_t\ u_t^n\ d_t Y_t + (a_t/u_t^n - a_t/u_{t-1}^a)Y_t$$

where the condition $Y_t > Y_t^m$ is met (see n. 5). Given the propensity to save, the level of saving adjusts, *via* the supermultiplier, to any given level of autonomous expenditure *plus* capacity-adjusting investment through changes in the long-run level of income (i.e. Y_t).

15. The capital stock can also be shown to be historically determined by reference to saving endogenously generated by income, as follows:

$$K_t = K_{t-1} + sY_{t-1} - A_{t-1} + (u^n_{t-1} - u^a_{t-1})\, a_{t-1} d_{t-1} Y_{t-1}$$

where A_t is the autonomous demand in period $t-1$ which absorbs part of saving but is assumed not to be adding to productive capacity.

16. Through this sequence of effects on average utilisation in period *t*, (i.e. u^a_t), the capital stock in period $t+1$ will, in turn, be affected since $K_{t+1} = K_t + I^I_t + (u^n_t - u^a_t)\, a_t d_t Y_t$.

17. In accord with equation [5.15], induced investment in period $t+1$ is:

$$I^I_{t+1} = (a_{t+1}/u^n_{t+1})\,(Y^e_{t+2} - Y_{t+1}) + a_{t+1} u^n_{t+1} d_{t+1} Y_{t+1} + (a_{t+1}/u^n_{t+1} - a_{t+1}/u^a_t) Y_{t+1}$$

Hence, while higher utilisation means less capital is required per unit of output (in period *t*), it leads to a faster rate of depreciation of capital per unit of output which tends to induce a greater level of investment in the future (i.e. $t+1$).

18. It follows from equation [5.18] above that the average rate of capital accumulation, g^k_t, will be equal to the average growth rate of output in period *t*; that is, $g^k_t = g^y_t$.

19. Autonomous private investment will mainly, though not exclusively, be connected to innovation, consisting of investment in research and development and in the installation of technically superior capital equipment or that necessary to produce new innovative products. Relaxing our assumption that all autonomous investment is non-capacity creating, then clearly capacity-creating private and public investment, which 'sets the pace' for demand growth, is included as part of, and contributes to, autonomous demand.

References

Amadeo, E. J. (1986) 'The role of capacity utilisation in long period analysis', *Political Economy: Studies in the Surplus Approach*, 2(2): 147–60.

Aspromourgos, T. (2009) *The Science of Wealth: Adam Smith and Framing of Political Economy* (London: Routledge).

Ciccone, R. (1986) 'Accumulation and capacity utilization: some critical considerations on Joan Robinson's Theory of Distribution', *Political Economy: Studies in the Surplus Approach*, 2(1): 17–36.

Ciccone, R. (1987) 'Accumulation, capacity utilisation and distribution: a reply', *Political Economy: Studies in the Surplus Approach*, 3(1): 97–111.

Garegnani, P. (1984) 'Value and Distribution in the Classical Economists and Marx', *Oxford Economic Papers, New Series*, 36, 295–325.

Garegnani, P. (1990) 'Sraffa: Classical versus Marginalist Analysis', in K. Bharadwaj and B. Schefold (eds), *Essays on Piero Sraffa* (London: Unwin Hyman): pp. 112–41.

Garegnani, P. (1992) 'Some Notes for an Analysis of Accumulation', in J. Halevi, D. Laibman and E. J. Nell (eds), *Beyond the Steady State* (London: Macmillan), 47–71.

Garegnani, P. and Palumbo, A. (1998) 'Accumulation of Capital', in H. D. Kurz and N. Salvadori (eds) *The Elgar Companion to Classical Economics*, Volume A–K (Cheltenham U.K: Edward Elgar): pp. 10–18.

Hicks, J. R. (1950) *A Contribution to the Theory of the Trade Cycle* (Oxford: Clarendon Press).

Kahn, R. F. (1959) 'Exercises in the Analysis of Growth', *Oxford Economic Papers*, New Series, 11(2): 143–56.

Kaldor, N. (1955–56). 'Alternative theories of distribution', *Review of Economic Studies*, 23(2): 83–100.

Kaldor, N. (1957) 'A model of economic growth', *Economic Journal*, 67(268): 591–624.

Lucas, R. E. (1988) 'On the mechanics of economic development', *Journal of Monetary Economics*, 22: 3–42.

Marglin, S. A. (1984) *Growth, Distribution and Prices* (Cambridge, MA: Harvard University Press).

Palumbo, A. and Trezzini, A. (2003) 'Growth without normal capacity utilization', *European Journal History of Economic Thought*, 10(1); 109–35.

Robinson, J. (1962) *Essays in the Theory of Economic Growth* (London: Macmillan).

Romer, P. M. (1986) 'Increasing returns and long-run growth', *Journal of Political Economy*, 94: 1002–37.

Romer, P. M. (1990) 'Endogenous technical change', *Journal of Political Economy*, 98: S71–S102.

Serrano, F. (1995) 'Long period effective demand and the Sraffian supermultiplier', *Contributions to Political Economy*, 14: 67–90.

Smith, A. (1976 [1776]) *An Inquiry into the Nature and Causes of the Wealth of Nations*, Glasgow edn, Vols. I–II (Indianapolis: Liberty Classics).

Solow, R. M. (1956) 'A contribution to the theory of economic growth', *Quarterly Journal of Economics*, 70: 65–94.

Sraffa, P. (1960) *Production of Commodities by Means of Commodities: Prelude to a Critique of Economic Theory* (Cambridge: Cambridge University Press).

Steindl, J. (1952) *Maturity and Stagnation in American Capitalism* (Oxford: Blackwell).

Swan, T. W. (1956) 'Economic growth and capital accumulation', *Economic Record*, 32: 334–61.

Trezzini, A. (1995) 'Capacity utilisation in the long run and the autonomous components of aggregate demand', *Contributions to Political Economy*, 14: 33–66.

Trezzini, A. (1998) 'Capacity utilisation in the long run: some further considerations', *Contributions to Political Economy*, 17: 53–67.

Trezzini, A. (2005) 'The irreversibility of Consumption and Economic Growth', *Quaderno di Ricerca del Centro Studi e Documentazione 'Piero Sraffa'*, no. 6.

Vianello, F. (1985) 'The Pace of Accumulation', *Political Economy*, 1(1): 69–87.

6
Normal Paths of Growth Shaped by the Supermultiplier

*Óscar Dejuán**

6.1 Introduction

The Keynesian principle of effective demand states that the equilibrium level of output in a given period is a multiple of the expected autonomous demand (Kalecki, 1971; Keynes, 1936). Can we extrapolate this principle to a long-run dynamic analysis and conclude that the rate of growth of output will eventually depend on the expected rate of growth of autonomous demand? A positive answer would be a significant step towards a long-period theory of output which, according to Eatwell and Milgate, provides the most solid ground for demand-led growth (Milgate, 1982; Eatwell, 1983; Eatwell and Milgate, 1983; Eatwell, 2012).

A theory of demand-led growth becomes more compelling and manageable when it is expressed by the multiplier and the accelerator mechanisms. As a matter of fact, they have been present from the first growth model (Harrod, 1939). Harrod showed that, for an economy defined by the capital/output ratio and the saving propensity, a rate of growth exists that warrants the absorption of the full-capacity output and full-capacity savings, year after year. The model, however, was so unstable that it was soon rejected as a 'knife edge':

> On either side of this line it is a 'field' in which centrifugal forces operate, the magnitude of which varies directly as the distance of any point in it from the warranted line. Departure from the warranted

* Ana González Martínez (University of Castilla – La Mancha) helped me with the search for the data contained in Section 6.2. Marc Lavoie (University of Ottawa) made useful comments on the first version of this chapter.

> line sets up an inducement to depart farther from it. The moving equilibrium of advance is thus a highly unstable one. (Harrod, 1939)

Keynes' most direct disciples tried to accommodate demand-led growth in a steady-state model where endogenous changes in distribution made possible the adjustment to a different path of growth (Robinson, 1956; Robinson, 1962; Kaldor, 1955–56; Kaldor, 1961). Kaleckian economists consider that capacity utilisation is the endogenous variable that makes possible the adjustment to the autonomous trend (Lavoie, 2010; Hein *et al.*, 2011). The same opinion is shared by most Sraffian economists (Garegnani, 1983; 1992; Ciccone, 1986; 1987; Trezzini, 1995; 1998; Palumbo and Trezzini, 2003). Apparently they have defected to the original project of building a normal theory of output parallel to the normal theory of prices.

Not all Sraffian economists share this view. John Eatwell in the Memorial of Pierangelo Garegnani (22 February 2012) defends the advisability of filling the classical-Keynesian core of economic theory with a theory of normal output with the typical gravitation properties. Cesaratto (2012) separates core-Sraffians, the group based in Rome under the leadership of Garegnani, from periphery-Sraffians which includes at least three models based on the supermultiplier (Serrano, 1995; Bortis, 1997; Dejuán, 2005). According to Cesaratto the last group shows a plausible way of reconciling demand-led growth with the independence of technology (the optimal capital/output ratio, which implies a normal degree of capacity utilisation) and the independence of distribution. This chapter aims to clarify these issues and to show the stabilising properties of the multiplier–accelerator interaction, when it is properly formulated.

Beginning (Section 6.2) with a glance at the empirical facts that the theoretical approach proposes to explain, is the Spanish case presented as an example. Section 6.3 develops an investment function based on a flexible accelerator which adjusts capacity to a growing demand and takes into account excesses of inventories and the capacity to remove them. Section 6.4 integrates the accelerator of investment with the multiplier of consumption. The resulting supermultiplier is able to explain macroeconomic equilibrium at a given moment and through time. The autonomous trend, if it lasts long enough, determines the actual growth of demand and output. It also shapes the warranted (normal capacity) rate of growth, which becomes an endogenous variable. In Section 6.5 it is concluded that the supermultiplier, if properly formulated, is a stabilising mechanism leading the economy towards the normal path of growth compatible with the autonomous trend (given technology, distribution and expenditure patterns).

6.2 A glance at a growing economy

For economists inspired by Schumpeter (1912), the railway provides a good metaphor to explain the dynamics of a capitalist economy. The trucks correspond to the productive capacity that serves induced consumption which represents 60–70 per cent of total output. The locomotive corresponds to the productive capacity that produces goods and services for autonomous demand in a broad sense. In the locomotive there are two clearly separated engine rooms. The front part, which houses the drivers, can be identified with the capacity serving autonomous demand in a strict sense. It includes: (a) autonomous consumption by households; (b) residential investment; (c) modernisation investment by firms that transforms the existing capacity, instead of expanding it; (d) real public expenditure; and (d) exports. The machines that will expand the productive capacity of the economy are produced in the rear part of the locomotive. This is identified as the 'expansionary investment' and is explained through the accelerator mechanism.

The railway metaphor conveys several lessons whose economic translation is obvious. (1) The train will run at the speed set by the locomotive. The trucks may go slower than the locomotive during the first moments after acceleration; slower, after a deceleration. But this is a transient phenomenon; eventually the rates of growth will converge. (2) Whatever the potential of this locomotive, the actual speed of the train is determined by the engine drivers. (3) When they decide to go faster, they accelerate the locomotive (the rate of capacity utilisation). Later, to avoid burning out the engine, they switch in additional power which implies a rise in the relative weight of the rear part of the locomotive. (The acceleration of investment brings about a rise in the share of the equipment sector).

With these Schumpeterian lenses, we are now going to look at the Spanish economy. Table 6.1 and Figure 6.1 show two long waves of prosperity (1964–74 and 1997–2008) and two long recessions (1975–96 and 2008 till the present day).[1] These movements can be related to the dynamics of autonomous demand in a strict sense. The car industry was the main engine of growth during the 1960s and 1970s. It produced for the home market (durable consumption goods paid for by consumer credit) and for foreign markets – cars have been the top Spanish export since the 1960s. From 1996 on, construction became the key engine of the Spanish

Table 6.1 Key economic indicators during long cycles, Spain 1964–2010

%	1964–74	1975–96	1997–07	2008–10
g(cars) (rate of growth of car production)	21.14	5.28	1.34	–7.47
g(h) (rate of growth of residential invest)	6.85	–0.43	7.45	–19.62
$\mathbf{g_i}$ (rate of growth of productive investment)	10.51	2.69	7.38	–5.66
$\mathbf{g_y}$ (rate of growth of GDP)	5.79	2.22	3.84	–3.66
i (share of productive investment in GDP)	23.96	19.23	23.22	21.05
u (rate of capacity utilization)	83.64	78.43	80.01	73.47

Source: INE, Ministerio Industria (Spain), ANFAC, AMECO.

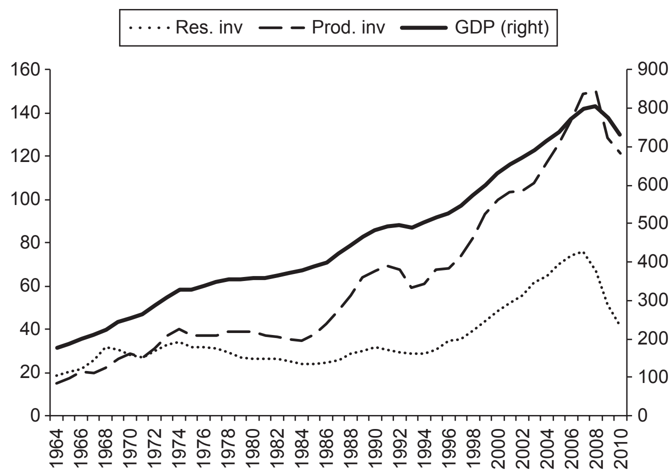


Figure 6.1 GDP, productive investment and residential investment, Spain 1964–2010 (billion Euros year 2000)
Source: See Table A (Appendix).

economy under the stimulus of easy mortgage loans. The saturation of the market was already clear in 2006. Its collapse after the financial crisis of 2007–08 plunged the economy into a serious recession.

During the long waves of prosperity accumulation accelerated, leading to a rise in the share of investment in aggregate demand (above 23 per cent) and the rate of capacity utilisation (above 80 per cent). During recessions both the investment share and the rate of capacity utilisation fell. The average rate of capacity utilisation over the 47 years under review is 80 per cent.

Table 6.2 shows the rate of capacity utilisation in the economies of the OECD for which the complete sets of data are available. We appreciate there are some level-differences that probably accord to the way each country defines normal capacity. The rate rises in boom periods but not in the steady way that would correspond with a demand expanding for several years. The low standard deviation proves that an inner force (investment) keeps the rate of capacity utilisation close to its normal position.

6.3 Expansionary investment and the flexible accelerator

The kernel of a growth model is the *accumulation of capital* that we shall identify with *productive investment of the expansionary type*. It can be called *induced investment,* provided we realise that it is different from the 'inducement' associated with consumption. Expansionary investment aims to create capacity in order to match efficiently the expected increases in demand.

According to the principle of acceleration, investment expenditure undertaken at the end of period t (I_t) can be computed as the difference between required capacity for the next and subsequent years ($K_{R,t+1}$) and installed capacity in t (K_{It}). It can be expressed by any of the following expressions which derive one from another, taking into account the definitions of the variables. D_t stands for aggregate demand in year t and g_d for its expected rate of growth. k is the optimal capital/output ratio.

Table 6.2 Rate of capacity utilisation

	Mean	**Standard deviation**
Australia	1.00	0.01
Belgium	0.78	0.03
Canada	0.76	0.11
Denmark	0.81	0.03
France	0.86	0.05
Germany	0.84	0.03
Italy	0.76	0.03
New Zealand	0.87	0.04
Norway	0.85	0.08
Spain	0.08	0.03
UK	0.79	0.07
USA	0.90	0.06

Source: Own computation with data from OECD and AMECO.

It corresponds to the normal rate of capacity utilisation that we shall normalise at unity ($u^*=1$). The actual rate is the ratio between actual and normal output: $(u_t = Y/Y^*)^2$. $E_{kt} = K_{It} - K_{Rt}$ is the excess of capacity.

$$I_t = K_{R,\,t+1} - K_{It} \qquad [6.1a]$$

$$I_t = k \cdot D_t \cdot (1 + g_d) - \frac{k}{u_t} \cdot Y_t \qquad [6.1b]$$

$$I_t = k \cdot g_d \cdot D_t + \left(k \cdot D_t - \frac{k}{u_t} \cdot Y_t \right) \qquad [6.1c]$$

$$I_t = k \cdot g_d \cdot D_t - E_{kt} \qquad [6.1d]$$

Most of the preceding formulae are divided into two parts. The first one ($k \cdot g_d \cdot D_t$) refer to *ex ante investment*: production of the capital goods required to respond to increasing demand. The second part (that subtracts excess of capacity) refers to *ex post investment*: actual investment expenditure at the end of the year. Such a subtraction brings about undesired inventories (E_{yt}) that will be discounted from next year's demand. Actual production in year $t+1$ will be: $Y_{t+1} = D_{t+1} - E_{yt}$.

When the economy is growing along its long-period path of growth we can write $Y_t = D_t$ and $u_t = 1$ and use the pure acceleration mechanism:

$$I_t = k \cdot g_d \cdot Y_t \qquad [6.2]$$

In the general case investment depends on the expected growth of demand and the deviations from normal capacity. Both concepts have been a matter of controversy and require further clarification.

(a) Normal capacity

Competition forces firms to invest in the best (available) technology and to use it in the most efficient way, i.e., at normal capacity. After a sudden and unexpected rise in demand, firms are supposed to raise capacity utilisation to gain as many customers as possible. The actual rate of profit may also increase for a while because installed capacity is being used more hours per day. But if firms continue over-utilising capacity they risk losing customers in the next peak of demand, impairing the risk-free profitability (profits free from the risk of losing customers in the peaks of demand). The maximisation of profits in the long run (or mere survival) compels firms to adjust capacity towards its normal rate.

Ciccone is right in assuming that firms plan investment on the basis of the regular fluctuations in demand which allows for the definition of the normal rate as the average of the cycle (Ciccone, 1986; 1987). There is no reason to accelerate investment in the peak of the cycle where $u_t>1$, since this possibility has already been taken into account in the investment decision. This conclusion does not hold, however, after permanent increases in demand like those which result from an acceleration in the autonomous trend.

It is important to draw a distinction between the normal rate of capacity and the normal rate of employment in long-period analysis. A normal rate of employment is not an equilibrium condition either in the short run or in the long run. The reason is clear: unemployment (unlike under-utilisation or over-utilisation of capacity) does not affect the profitability of firms.

(b) Expected growth of demand

Eatwell (1983) states that the proper independent variable of Keynes' *General Theory* is not investment, as such, but the long-term expectations of growth envisaged by investors (g_d). He cites the following passage:

> If we suppose a state of expectations to continue for a sufficient length of time for the effect on employment to have worked itself out so completely that there is, broadly speaking, no piece of employment going on which would not take place if the new state of expectation had always existed, the steady level of employment thus attained may be called the long period employment corresponding to that state of expectation. (Keynes, 1936, p. 48)

Keynes emphasised the importance of the 'animal spirits' of entrepreneurs to such an extent that most of his disciples have interpreted them as the prevalence of subjective factors in investment decisions. In the long-run perspective considered by investors, the 'animal spirits' of entrepreneurs are usually anchored to objective factors like the autonomous trend. In the model a persistent autonomous trend will forge expectations of demand growth. If one million houses are built in year one, 1,025 million in year two, 1,051 in year three and so on, an induced demand for machinery suggests an expected 2.5 per cent growth of demand at the factory. Even if in one year firms' production remains constant (because they were running excesses of inventories and/or capacity), they will eventually need to match the demand growth imposed by the autonomous trend. With technology constant, house-building cannot grow at a rate of 2.5 per cent if the equipment sector is not expanding production at at least the same rate.

Contrary to Palumbo and Trezzini's (2003) view, the flexible accelerator function (see equations [6.1a]–[6.1d]) does not require firms to have perfect foresight or a huge amount of information. Neither does it require the economy to be permanently in a fully adjusted path of growth. At the end of the year, firms decide investment expenditure having regard to the demand orders they receive and the ongoing imbalances.

Models that incorporate the accelerator mechanism will be stable provided demand expectations depend on objective factors (such as the autonomous trend) rather than the ups and downs of current income and capacity utilisation. The extreme instability of Harrod's (1939) model stems from the fact that the expected rate of growth depends only on psychological factors and is unduly influenced by the actual rate of capacity utilisation. Such behaviour refutes the autonomy of the autonomous trend.

6.4 Dynamic equilibrium through the supermultiplier

According to the Keynesian-Kaleckian principle of effective demand, the equilibrium level of output at a given moment does not depend on the productive capacity of the economy. Output and employment are supposed to adjust to expected aggregate demand which includes: autonomous demand (Z_t), expansionary investment by firms (I_t) and induced consumption by households:

$$Y_t = D_t = C_t + I_t + Z_t \qquad [6.3]$$

In our closed and private economy autonomous demand has been identified with residential investment. We take as given the expected demand at year t (Z_t) and its expected rate of growth (the autonomous trend, g_z). To simplify the exposition we shall assume that rentiers buy these houses with mortgage loans and let them to workers. The future stream of rents received by rentiers will be consumed.

It is well documented that induced consumption is a relatively high and stable proportion of disposable income. In a Kaleckian mood, the average propensity to consume can be computed as a weighted average of the propensities to consume of workers out of wages (c_w) and of capitalists out of profits (c_b). The proper weighting is the share of wages and profits in income (ω and β, respectively). To make the graphical representation of the model easier assume that all wages are consumed ($c_w = 1$) and all profits are saved ($s_b = 1$). In fact, workers are forced to save part of their wages in order to pay rent. But since rentiers consume all these rents we can write $c_w = 1$.

$$C_t = c \cdot Y_t = c_w \cdot \omega \cdot Y_t = \omega \cdot Y_t \qquad [6.4]$$

Let us consider, to begin with, an economy that is growing along its long-period path without excesses of capacity and inventories. The autonomous trend has prevailed long enough to anchor long-term expectations of demand, so $g_d = g_z$. Actually all variables are growing at the same rate, that we can identify with the warranted rate (g^*_t). In these circumstances we can compute investment by the pure acceleration formula given in [6.2] and determine the level of output by any of the following expressions:

$$Y_t = c \cdot Y_t + k \cdot g^*_t \cdot Y_t + Z_t \qquad [6.5]$$

$$Y_t = \frac{1}{1-c} \cdot (Z_t + I_t) = \mu \cdot (Z_t + I_t) \qquad [6.6]$$

$$Y_t = \frac{1}{1-c-k \cdot g_t^*} \cdot Z_t = \mu^* \cdot Z_t \qquad [6.7]$$

Expression [6.6] uses the simple multiplier. The equilibrium level of output is μ times the (broad) autonomous demand that includes residential investment (Z_t) and productive investment of the expansionary type (I_t). The last one can be computed *ad hoc* using any formula of the acceleration in [6.1]. Even simpler, it is g^* times the installed capacity or $(1+g^*)$ times the previous level of investment.

Expression [6.7] combines the multiplier and the acceleration in a single expression that Serrano (1995) (after Hicks, 1950) called the supermultiplier. Output in year t is μ^* times the (strict) autonomous demand for this year (Z_t).

Such an economy will grow steadily along a fully adjusted path. Harrod (1939) called it the 'warranted path'. If the rate of growth envisaged by entrepreneurs is g^* and they increase production at this rate, the entire supply will be absorbed by demand and all the savings generated in the process of production will be invested. Dividing [6.5] through income and clearing for g^* we get the following expression:

$$g_t^* = \frac{1-c-z_t}{k} = \frac{s-z_t}{k} = \frac{i_t}{k} = \frac{s_{it}}{k} \qquad [6.8]$$

As in Harrod (1939), the warranted rate depends on technology (the optimal capital/output ratio, k) and expenditure patterns (the aggregate propensity to save, $s=1-c$). But it is also affected by the structure of the economy, which changes with the autonomous trend. Such a structure is represented by the share of residential investment in income in a given period ($z_t=Z_t/Y_t$). Alternatively it can be represented by the share in income of productive investment (i_t) that will attract, through the credit channels, an equivalent amount of savings ($s_{i,t}$).

So far the workings of an economy have been described along its long path of growth. This is not a realistic description of how capitalist economies actually grow. If the autonomous trend is truly 'autonomous', it will change from time to time. In the example in Table 6.A (in the Appendix) we suppose that after year three g_z rises from 0.025 to 0.5. Firms adjust capacity to match the increasing demand and, at the end of the year, speed up investment over and above the level corresponding to the pure accelerator process. Consider [6.1d]: $I_t=k \cdot g_d \cdot D_t-E_{kt}$. A negative E_{kt} means a shortage of capacity that is filled by depleting inventories (negative E_{yt}). Next year production will be higher than the level justified by expected demand since firms have to replenish their stocks. The actual rate of growth of this economy can be replicated by the following formula where the two adjustment mechanisms are apparent: in the short run, the actual rate of capacity utilisation (u_t); in the long run, the structure of demand (z_t):

$$g_{y,t} = \frac{(1-c-z_t)\cdot u_t}{k} \qquad [6.9]$$

Figure 6.2, derived from the data in table A (Appendix), presents the traverse from one equilibrium path to another. For the given value of the parameters the adjustment is cyclical. The rate of capacity utilisation rises to recover cyclically its normal position ($u_t=1$).[3] Overutilisation of capacity speeds up the pace of accumulation raising the share of investment in income (i) until it enables a new steady growth path ($g_y=g_d=g^*=g_z$).

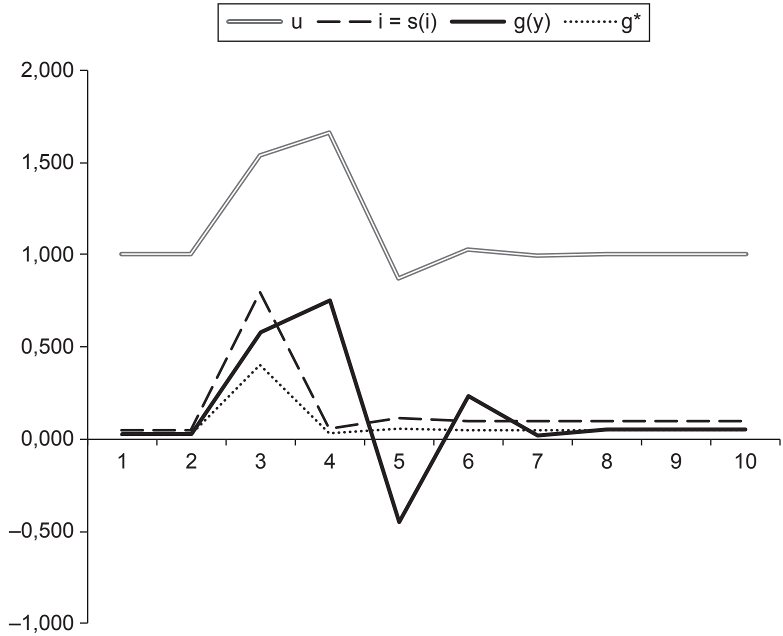


Figure 6.2 The traverse (after an acceleration of the autonomous trend)

Figure 6.3 shows the structural adjustments that make the warranted rate equal to the autonomous trend. Given technology (represented by the labour coefficient, *l*, and the optimal capital coefficient, *k*), we can draw the frontiers of distribution and growth. In the north-east quadrant we detect that the result of a rise in the autonomous trend is an increase in the share of expansionary investment in income ($i_t = I_t / Y_t$ rises). It occurs at the expense of per-worker consumption-type expenditure ($ç_t$); more specifically, at the expense of autonomous demand (z_t) because we have assumed a constant propensity to consume. The result looks paradoxical. The fall of *z* comes after the acceleration in autonomous demand (g_z). As a matter of fact the yearly production of houses has increased a lot but, since income has increased faster, the ratio Z_t/Y_t has fallen.

The south quadrants of Figure 6.3 look at the financial side of the economy; more precisely at the part related to business loans for expansionary investment. To simplify the exposition we shall suppose that all types of investment are financed by bank loans allocated on a demand basis.

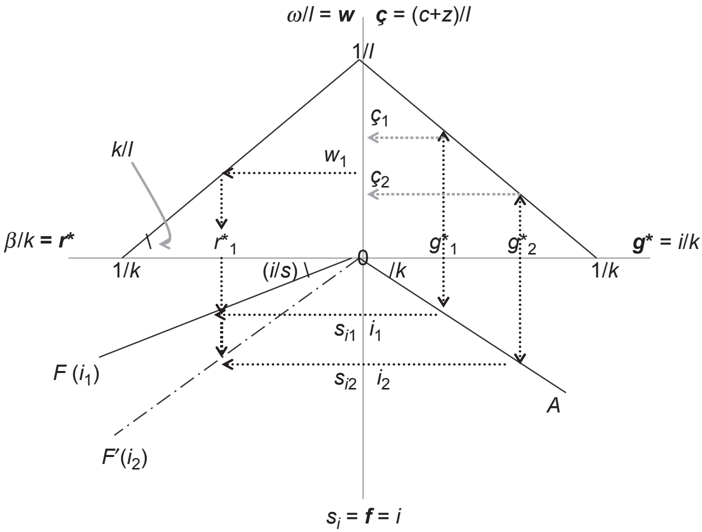


Figure 6.3 Structural adjustments towards the new autonomous trend

The demand for credit to finance accumulation (in the south-east) is a function of the rate of growth:

$$f_t^d = i_t = g_t^* \cdot k \qquad [6.10]$$

The supply of credit for expansionary investment is supposed to match this demand. It will absorb the savings generated in production that are not lent to rentiers. In relative terms it amounts to $s_{it} = 1 - c - z_t = s - z_t$. The equilibrium share of accumulated savings may also be obtained by multiplying the unit profit ($r^* \cdot k = \beta$), times the propensity to save out of profits ($s_b = 1$), times the share of expansionary investment in total savings:[4]

$$f_t^s = s_{it} = s - z_t = \beta \cdot s_b \cdot (i_t / s) \qquad [6.11]$$

The sub-index '*t*' borne by s_i, i and z remind us that these parameters adjust when the economy shifts to a different path of growth. The remaining ones (k, r^*, β, s, s_b) keep at their initial level. The term in brackets is the structural variable that makes the new path of growth possible and sustainable. Angle <r0F> opens whenever i increases.

Note that the change in the supply of savings for accumulation purposes has been driven neither by technology ($k_t = k$, $u_t = 1$), nor by distribution ($w = w_1$ and $r^* = r_1^*$),[5] nor by the saving propensity ($s = 1 - c$; $s_p = 1$). It is simply the result of the acceleration of expansionary investment in order to recover normal capacity after the acceleration in the autonomous trend.

A final word about the requirements of the supermultiplier model and the maximum autonomous trend compatible with it. The model computes the consumption and investment derived from the production of goods feeding autonomous demand. When autonomous demand is negligible we cannot compute and use the supermultiplier given in [6.7] (for the same reason that we cannot compute the simple multiplier when income is entirely devoted to consumption). We can still use the multiplier-accelerator model (our expression [6.6]) which would lead to the following equilibrium rate:

$$\hat{g}^* = \frac{1 - c}{k} = \frac{s}{k} \tag{6.12}$$

This is Harrod's warranted rate that now indicates the maximum rate of growth of autonomous demand that could be efficiently matched given technology (k) and expenditure patterns (c,s). A further increase in the autonomous trend cannot be attended by reducing the weight of autonomous demand. Firms will over-utilise capacity and u_t will be above 1, despite the effort of investors. Yet (as shown by Dejuán, 2005) the model continues to be stable and u_t converges to the following expression (it derives from [6.9], after equating $z_t = 0$):

$$u'' = \frac{k \cdot g_z}{1 - c} = \frac{k}{s} g_z = \frac{g_z}{g^*} \tag{6.13}$$

This recalls the thesis of the endogeneity of the utilisation rate (Palumbo and Trezzini, 2003; Hein *et al.*, 2011). Two provisos: (1) this applies only to the elementary model ($Y = C + I$) which is only found in text books; and (2) over-utilisation is restricted by technology.

According to Schumpeter (1912), capitalist economies overcome their natural tendency to stagnation thanks to the presence of innovative entrepreneurs who launch the production of new goods or traditional goods for new markets. The supermultiplier model we have studied is meant for this type of economy.

6.5 Conclusions

This paper extrapolates the Keynesian-Kaleckian principle of effective demand to the long run, in an attempt to build a demand-led growth theory coherent with the surplus approach. It has been shown that the autonomous trend determines not only the actual rate of growth of actual output, but also the warranted (normal capacity) rate. It 'rules the roost' of growth rates provided it is persistent enough. At least there is a tendency towards it, whenever firms try to be efficient and maximise profits in a long-run perspective.

The flexible supermultiplier nets two adjustment mechanisms. After the acceleration in the autonomous trend, firms will increase the utilisation rate to match demand. Afterwards they will speed up accumulation until the share of the expansionary investment in aggregate demand makes it possible to follow a new fully adjusted path of growth. The rate of capacity utilisation is endogenously determined in the short run. In the long run, the accumulation of capital re-establishes normal capacity utilisation and it is the warranted rate which becomes endogenous.

Contrary to the dominant opinion, it is contended that the multiplier-accelerator is a stabilising mechanism, leading the economy to fully adjusted positions (usually in a cyclical way). Three conditions are required to accomplish this task:

1. Autonomous demand and its expected rate of growth are truly autonomous.
2. Investment decisions are based on long-term expectations of demand growth that will be anchored on the autonomous trend provided it lasts long enough.
3. Investors take into account the excesses or shortages of capacity and inventories to get rid of them.

This conclusion stands as a critique to the Kaleckian and Sraffian models referred to in the introduction, which load the burden of adjustment (both in the short and in the long run) on the rate of capacity utilisation. It is felt that they do not take the restrictions imposed by technology and enforced by competition seriously enough. In a competitive economy firms are compelled to use the best technology in the best way. Firms can accommodate stationary fluctuations in demand and ordinary cycles by pushing up the rate of capacity utilisation. But if increases in demand are permanent (as happens when the autonomous trend rises), firms are bound to speed up investment to avoid the risk of losing customers in the next seasonal peak of demand.

The adjustment brought about by the supermultiplier does not require changes in distribution and/or in savings propensities, although they would make the adjustment towards a new autonomous trend easier. Accord is found with Palumbo and Trezzini's (2003) criticism of the first post-Keynesian economists referred to in the introduction: it is not sensible to assume a fall in wages during the boom, just to ensure the continuity of a mythical steady state. But it is felt that Palumbo and Trezzini are wrong when they conclude that an inbuilt tendency towards normal capacity implies perfect foresight in a steady-state pattern of growth. It has been shown that the flexible accelerator is able to work out of equilibrium. Nor should it be concluded that the endogeneity of the warranted rate implies that savings determine investment. After a change in the autonomous trend income will grow until the yearly generation of savings coincides with the yearly investments required by the new trend. The structural change occurs in the allocation of these savings between productive and residential investment (capacity-creating and non-capacity creating investment).

The opinion presented here is that the multiplier-accelerator model goes a long way in the determination of a theory of normal output. It is different but complementary and compatible with the Sraffian theory of normal prices. Both theories share the same set of exogenous data: technology and distribution (the real wage, as proposed). In addition, the theory of normal output includes certain expenditure patterns and the expected rate of growth of autonomous demand. The resulting theory of output shows (1) that the actual rate of growth of GDP is ultimately determined by the autonomous trend; and (2) that the structure of demand and productive capacity tends to accommodate to the requirements of the autonomous trend.

Palumbo and Trezzini rightly point out that the autonomous trend lacks the persistency of the traditional exogenous variables of the surplus approach (labour productivity, the real wage, workers' propensities to consume and save and so on). Certainly, in epochs without a well-established autonomous trend, the economy will suffer from macroeconomic instability. Yet the tendency towards the normal rate of growth and the proper structure of demand is always at work through the investment function. These tendencies will carve a path when the autonomous trend is persistent enough. China, an export-led economy with modern technology and low labour costs, can achieve a normal path of growth that trebles the path of the USA. So we can expect that the share of expansionary investment on income will be much higher in the Chinese economy. The rate of capacity utilisation, however, will be around its normal level in both countries. Within a given country, we can distinguish long waves of prosperity and depression. We saw in the comment about Table 6.2 that the capacity rate rises in boom periods but remains tied to the normal one.

Once the core of economic theory is well established, researchers are supposed to focus on the institutional framework that determines the exogenous data and the interactions between them. During the economic boom prior to the first great depression of the 21st century, for example, economists should have asked the following questions: For how long can residential investment act as a locomotive if the house prices are increasing three times faster than nominal wages? Can banks fill the gap with mortgage credits without leading households into a debt trap?

Notes

1. Inside these long waves we find cyclical movements of a shorter length. Within the long recession from 1975 to 2006 there was an economic boom between 1987 and 1992. In the last long wave of expansion there was a bust in the first two years of the 21st century. Car production is not included in Figure 6.1 because the data are in physical units.
2. Alternatively, it could be defined by the ratio between required and installed capacity $u_t'=K_{Rt}/K_{It}$. Both ratios coincide when there are no undesired stocks. They are accounted for in u_t, not in u_t'.
3. The table does not consider any restriction in the adjustment process. If we fix maximum and minimum levels for u_t and e_t the adjustment will be longer and smoother. If we consider that only a part of undesired inventories can be stored, the adjustment will be faster. There are many variants of the model that alter the shape of the process without changing the final results.
4. In our numerical example this ratio equals the share in profits of the savings financing expansionary investment: $S_I/B=I/S=(I/Y)/(S/Y)=i/s$.
5. The last column of Table 6.A (Appendix) shows that the actual rate of profit (r_t) is positively related to capacity utilisation.

References

Bortis, H. (1997) *Institutions, Behaviour and Economic Theory* (Cambridge, UK: Cambridge University Press).

Cesaratto, S. (2012) ‘NeoKaleckian and Sraffian controversies on accumulation theory’, unpublished paper.

Ciccone, R. (1986) ‘Accumulation and capacity utilization: some critical considerations on Joan Robinson’s Theory of Distribution’, *Political Economy, Studies in the Surplus Approach*, 2: 17–36.

Ciccone, R. (1987) 'Accumulation, capacity utilization and distribution: a reply', *Political Economy, Studies in the Surplus Approach*, 3: 97–111.

Dejuán, Ó. (2005) 'Paths of accumulation and growth: towards a Keynesian long-period theory of output', *Review of Political Economy*, 17: 231–52.

Eatwell, J. (1983) 'The long period theory of employment' *Cambridge Journal of Economics*, 7: 269–85.

Eatwell, J. (2012) 'The theory of value and the foundations of economic policy', *In memoriam of Pierangelo Garegnani*. Roma.

Eatwell, J. and Milgate, M. (eds) (1983) *Keynes's Economics and the Theory of Value and Distribution* (New York: Oxford University Press).

Garegnani, P. (1983) 'Two Routes to Effective Demand', in J. Kregel (ed.) *Distribution, Effective Demand and International Economic Relations* (London: Macmillan).

Garegnani, P. (1992) 'Some notes for an analysis of accumulation', in J. Halevi, D. Laibman and E. J. Nell (eds), *Beyond the Steady State. A Revival of Growth Theory* (London: Macmillan).

Harrod, R. F. (1939) 'An essay in dynamic theory', *Economic Journal*, 49: 14–33.

Hein, E., Lavoie, M. and Treeck, T. V. (2011) 'Some instability puzzles in Kaleckian models of growth and distribution: a critical survey', *Cambridge Journal of Economics*, 35: 587–612.

Hicks, J. R. (1950) *A Contribution to the Theory of the Trade Cycle* (Oxford: Clarendon).

Kaldor, N. (1955–56) 'Alternative theories of distribution', *Review of Economic Studies*, 23: 63–100.

Kaldor, N. (1961) 'Capital Accumulation and Economic Growth', in F. A. Lutz and D. C. Hague (eds), *The Theory of Capital* (New York: St Martins).

Kalecki, M. (1971) *Selected Essays in the Dynamics of the Capitalist Economy (1933–1970)* (Cambridge: Cambridge University Press).

Keynes, J. M. (1936) *The General Theory of Employment, Interest and Money* (London and New York: Harcourt Brace Javanovich).

Lavoie, M. (2010) 'Surveying Short-run and Long-run Stability Issues with the Kaleckian Model of Growth', in M. Setterfield (ed.), *Handbook of Alternative Theories of Economic Growth* (Cheltenham, UK: Edward Elgar).

Milgate, M. (1982) *Capital and Employment. A Study of Keynes's Economics* (London: Academic Press).

Palumbo, A. and Trezzini, A. (2003) 'Growth without normal capacity utilization', *European Journal of the History of Economic Thought*, 10: 109–35.

Robinson, J. (1956) *The Accumulation of Capital* (London: Macmillan).

Robinson, J. (1962) *Essays on the Theory of Economic Growth (*London: Macmillan).

Schumpeter, J. (1912) *Theorie der Wirtschaftlichen Entwicklung* [*The Theory of Economic Development*] (Leipzig: Verlag Dunker & Humbolt).

Serrano, F. (1995) 'Long period effective demand and the Sraffian supermultiplier', *Contributions to Political Economy*, 14: 67–90.

Trezzini, A. (1995) 'Capacity utilisation in the long run and the autonomous components of aggregate demand', *Contributions to Political Economy*, 17: 33–66.

Trezzini, A. (1998) 'Capacity utilisation in the long run: some further considerations', *Contributions to Political Economy*, 17: 53–67.

APPENDIX

Data

Technology: $\boldsymbol{l}=0.8$, labour coefficient; $\boldsymbol{k}=2$, optimal or normal capital/output ratio.

Distribution: $\boldsymbol{w}=1$, real wage. $\boldsymbol{r^*}=0.1$ (normal rate of profit corresponding to $k=2$ and $w=1$). The shares in income of wages and profits are also constant: $\omega=w\cdot l=0.8$; $\beta=1-\omega=0.2$.

Expenditure patterns: $\boldsymbol{c_w}=1$, propensity to consume out of wages; $\boldsymbol{s_b}=1$, propensity to save out of profits. The aggregate propensity to consume and save are also constant: $\boldsymbol{c}=c_w\cdot\omega=0.8$; $\boldsymbol{s}=s_b\cdot\beta=0.2$.

Autonomous trend: $\boldsymbol{g_z}$. It starts at 0.025. In period 3 it shifts to 0.05. (Remember to change the g incorporated in columns Z, D and I. Also in the supermultiplier, that rises from 6.66 to 10 when g_z doubles). The initial value of autonomous demand is $Z_t=1$.

Definitions of variables

$\boldsymbol{Z_t}$: Proper autonomous demand that we identify with residential investment.
$\boldsymbol{D_t}=\mu^*\cdot Z_t$, aggregate demand to which production adjusts. $\mu^*=1/(1-c-k\cdot g)$, supermultiplier.
$\boldsymbol{Y_t}=D_t-E_{y(t-1)}$, actual production (after discounting the inherited excess of inventories).
$\boldsymbol{K_{Rt}}=k\cdot D_t$, required capital.
$\boldsymbol{K_{It}}=K_{I(t-1)}+I_{t-1}=(k/u_t)Y_t$, installed capacity at the beginning of period t.
$\boldsymbol{E_{kt}}=K_{It}-K_{R,t}$ excess of capacity.
$\boldsymbol{L_t}=l\cdot Y_t$, labour employed.
$\boldsymbol{W_t}=w\cdot L_t$, mass of wages.
$\boldsymbol{B_t}=r\cdot K_{I,t}=Y-W$, mass of profits ('benefits').
$\boldsymbol{C_t}$: $c\cdot Y_t$, induced consumption. (It coincides with wages, although a part of C corresponds to the consumption of rentiers from the rents paid annually by working households).
$\boldsymbol{S_t}=s\cdot Y_t=s_b\cdot B_t$
$\boldsymbol{I_t}=K_{R,t+1}-K_{It}=k^*\cdot D_t\cdot(1+g_z)-K_{It}=k^*\cdot g_z\cdot D_t-E_{kt}$, *ex post* productive investment (of the expansionary type).
$\boldsymbol{E_{yt}}=Y_t-(C_t+I_t+Z_t)$, excess of inventories or undesired stocks. A negative number implies that inventories are below the desired level. It coincides with the excess of savings: $E_{yt}=S_t-(I_t+Z_t)$.
$\boldsymbol{e_t}=E_{yt}/Y_t$, share in income of undesired inventories.
$\boldsymbol{u_t}=Y_t/Y_t^*$ rate of capacity utilisation. (It accounts for unwanted inventories that are not anticipated when capacity utilisation is defined as $u_t=K_{Rt}/K_{It}$).
$\boldsymbol{z_t}=Z_t/Y_t$, share in income of autonomous demand.
$\boldsymbol{i_t}=I_t/Y_t$, share in income of expansionary investment. It coincides with the share of the savings financing expansionary investment: $s_{it}=(Y_t-C_t-Z_t)/Y_t=1-c-z_t$.
$\boldsymbol{g_{yt}}=(Y_t-Y_{t-1})/Y_{t-1}$, actual rate of growth of output.
$\boldsymbol{g_t^*}=i/k=(1-c-z_t)/k$, warranted rate of growth.
$\boldsymbol{r_t}=B_t/K_{It}$ actual rate of profit.

Table 6.A Adjustment towards a new autonomous trend

t	Z	D	Y	KR	KI	E(k)	L	W	B	C	S	I	E(y)	e	u	z	i=s(i)	g(y)	g*	r
1	1,000	6,667	6,667	13,333	13,333	0,000	5,333	5,333	1,333	5,333	1,333	0,333	0,000	0,000	1,000	0,150	0,050	0,025	0,025	0,100
2	1,025	6,833	6,833	13,667	13,667	0,000	5,467	5,467	1,367	5,467	1,367	0,683	–0,342	–0,050	1,000	0,150	0,100	0,025	0,050	0,100
3	1,076	10,763	11,104	21,525	14,350	–7,175	8,883	8,883	2,221	8,883	2,221	8,251	–7,107	–0,640	1,548	0,097	0,743	0,625	0,372	0,155
4	1,130	11,301	18,407	22,601	22,601	0,000	14,726	14,726	3,681	14,726	3,681	1,130	1,421	0,077	1,629	0,061	0,061	0,658	0,031	0,163
5	1,187	11,866	10,444	23,731	23,731	0,000	8,355	8,355	2,089	8,355	2,089	1,187	–0,284	–0,027	0,880	0,114	0,114	–0,433	0,057	0,088
6	1,246	12,459	12,743	24,918	24,918	0,000	10,195	10,195	2,549	10,195	2,549	1,246	0,057	0,004	1,023	0,098	0,098	0,220	0,049	0,102
7	1,308	13,082	13,025	26,164	26,164	0,000	10,420	10,420	2,605	10,420	2,605	1,308	–0,011	–0,001	0,996	0,100	0,100	0,022	0,050	0,100
8	1,374	13,736	13,747	27,472	27,472	0,000	10,998	10,998	2,749	10,998	2,749	1,374	0,002	0,000	1,001	0,100	0,100	0,055	0,050	0,100
9	1,442	14,423	14,421	28,846	28,846	0,000	11,536	11,536	2,884	11,536	2,884	1,442	0,000	0,000	1,000	0,100	0,100	0,049	0,050	0,100
10	1,514	15,144	15,144	30,288	30,288	0,000	12,115	12,115	3,029	12,115	3,029	1,514	0,000	0,000	1,000	0,100	0,100	0,050	0,050	0,100

7
The Pattern of Economic Growth of the Brazilian Economy 1970–2005: A Demand-Led Growth Perspective

Fabio N. P. de Freitas and Esther Dweck

7.1 Introduction

The present work analyses the pattern of economic growth prevailing in Brazil from 1970 to 2005. The analytical framework for this investigation is based on the classical supermultiplier demand-led growth model combined with the hypothesis that the balance of payments was the main potential (and often the most effective) constraint for the expansion of the Brazilian economy during the period in focus. We adopt a demand-led growth accounting methodology to analyse the Brazilian economy, based on a multi-source database compiled for this investigation.

The chapter is organised as follows. First, the Brazilian growth experience is put into context. Growth during the period 1970–2005 is compared to previous periods, starting from 1900. The Brazilian experience is also contrasted with the growth trajectory of other reference countries and the world economy. The following section presents our theoretical framework, as well as the empirical methodology adopted for this investigation, arguing that the labour constraint is not a prevalent characteristic of the Brazilian growth experience. In Section 7.4, the analytical framework is applied to investigating the pattern of economic growth prevailing in Brazil from 1970 to 2005. The last section provides the main results and our conclusions.

7.2 The Brazilian economic growth experience in perspective

Brazil's was one of the fastest-growing economies in the twentieth century. From the early 1900s until the end of the World War II (WWII), it grew at a relatively high average rate (3.8 per cent per year approximately). However, it also experienced huge instability, which was an important feature of the primary export pattern of development that characterised the period. The Brazilian economy performed particularly well from the beginning of the state-led industrialisation process after the end of WWII, until the Latin America crises in the 1980s. Indeed, as Figure 7.1 shows, from the late 1950s until the 1970s Brazil's economy grew at an average annual rate of 6.8 per cent.

The worst periods in terms of growth performance were the 1980s and the 1990s, when the country grew at an average rate of 2 per cent per year. Prolonged stagnation due to the Latin America crises started at the beginning of the 1980s and continued in the 1990s, when intensive liberal reforms were implemented. In the 2000s, the Brazilian economy recovered and grew at higher rates (3.3 per cent per year on average), in line with the better performance of the world economy (see Table 7.2 ahead). Even so, the country was not able to recover the expansion rate pace that prevailed in the primary export period at the beginning of the last century.

It is worth comparing the Brazilian performance with that of other reference countries. First, the levels of GDP per capita of Brazil, South Korea and the USA will be compared. The US economy is the benchmark country

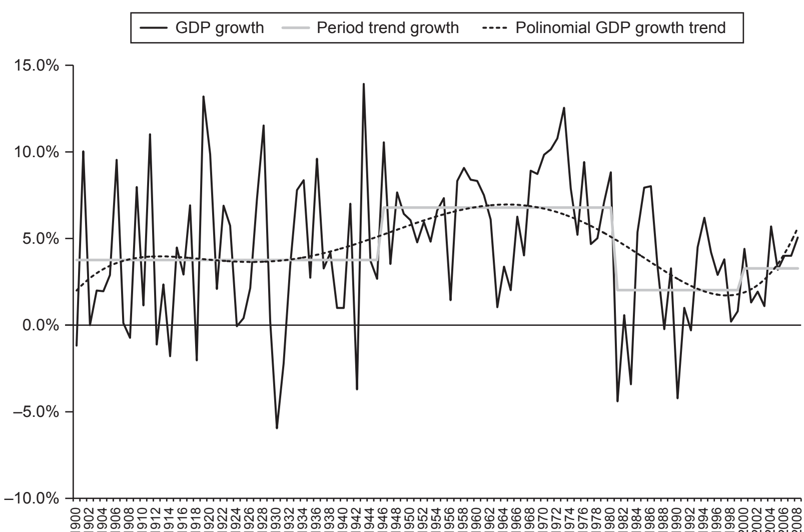


Figure 7.1 Brazilian annual real growth rates, 1900–2008
Source: Authors' elaboration based on Maddison (2010).

for the convergence analysis. South Korea has been included to allow for a contrast between Brazil and a successful economy in the Asian periphery.

Figure 7.2 and Table 7.1 show that both Brazil and South Korea engaged in a process of catching up with the US economy until the 1980s. Starting from relative shares of, respectively, 17.5 and 8.9 per cent of US per capita GDP in 1950, Brazil and South Korea achieved relative shares of 28.0 and 22.1 per cent in 1980, respectively. After that, the Latin America crises interrupted the Brazilian convergence process, and the country lagged behind until the beginning of the 2000s. Indeed, after two decades of stagnation Brazil reached a relative share of 19.4 per cent of US per capita GDP in the late 1990s. In the following decade, it managed to recover only modestly, achieving 20.6 per cent in 2008. In contrast, South Korea's convergence process continued in the period under review and, despite a brief interruption due to the Asian Crisis, the country reached a relative share of 62.9 per cent of US per capita GDP in 2008.

Another revealing comparison can be drawn between the economic growth experiences of Brazil, South Korea, the United States of America and the world economy. Indeed, in Table 7.2 one can observe the same pattern of convergence and divergence in terms of GDP per capita levels as discussed before. Additionally, one can also observe that Brazil only attained a higher growth rate than the world economy in the state-led

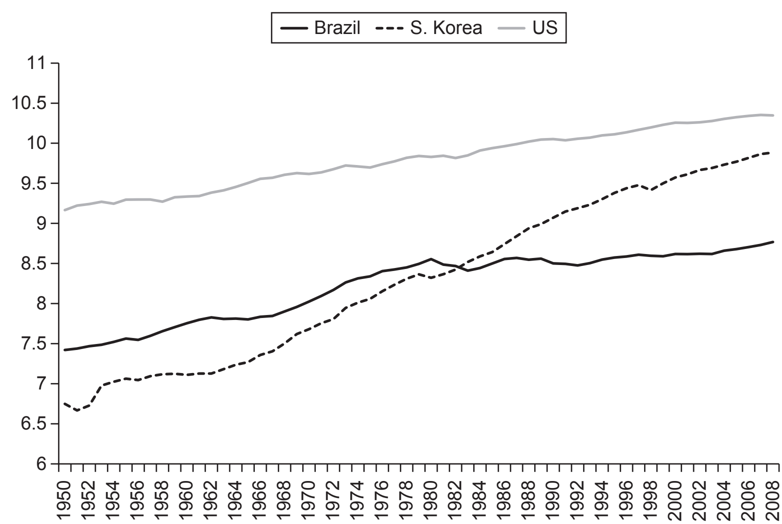


Figure 7.2 Per capita GDP levels 1950–2008, Brazil, South Korea and USA
Source: Authors' elaboration based on Maddison (2010).

Table 7.1 Relative per capita GDP (% of US per capita GDP)

Year	Brazil	South Korea
1950	17.5	8.9
1980	28.0	22.1
2000	19.4	50.5
2008	20.6	60.9

Source: Authors' elaboration based on Maddison (2010).

Table 7.2 Average real GDP growth rates (% per year)

Periods	Brazil	South Korea	US	World
1950–1980	6.8	7.5	3.6	4.5
1980–2000	2.0	7.5	3.2	3.0
2000–2008	3.3	4.4	2.1	4.2

Source: Authors' elaboration based on Maddison (2010).

development phase that lasted from 1950 to 1980, while South Korea managed to maintain an expansion rate above the world economy's for the whole period in focus (1950–2008).

The comparison drawn above provides the background to the analysis of our main subject. In the period concerned, Brazil presented its highest growth rates in the 1970s, experienced a long-lasting period of stagnation in the 1980s and 1990s, finally to present a modest recovery of GDP growth rates at the end of the 2000s. So, the present study aims to analyse the key features of the pattern of economic growth shown by the Brazilian economy in that period.[1]

7.3 The analytical framework

The Brazilian growth experience since the 1970s has been the subject of important debates. More recently, the literature[2] dedicated to interpreting and explaining the huge decline in Brazilian growth rates has been influenced by neoclassical economic theory, according to which the price system is supposed to convey information on the relative scarcity of resources, which is transmitted to consumers and producers and drives the choices they make towards the full utilisation of available resources. Hence, long-term economic growth should be characterised as a supply-constrained process, in which the rate of expansion depends on the growth of capital and labour inputs available to the economy, as well as on productivity growth. Therefore, according to that interpretative literature, the causes behind the trend break in Brazil's GDP growth rate would be: low domestic saving rate, low investment in human capital, and low (or even negative) growth rate of total factor productivity. These proximate causes have been subject to quantitative evaluation using a supply-side growth accounting empirical approach, inspired by the neoclassical growth theory. Furthermore, the literature referred to has also attempted to identify the more fundamental causes behind the decline in the Brazilian GDP growth trend. It has been suggested that 'market-unfriendly' institutions and pervasive market (and government) imperfections inherited from the post-WWII state – inducing an inward-oriented development strategy – are behind the poor growth performance of the Brazilian economy since the 1980s.

Nevertheless, contrary to the usual neoclassical viewpoint, one can argue that the Brazilian economy has not been normally constrained by the availability of resources in general and of labour in particular. As occurs in many developing economies, Brazil's economy shows a high degree of structural heterogeneity and a significant labour surplus. The great disparity in the levels of labour productivity observed across economic sectors is a quantitative expression of that kind of heterogeneity.

Table 7.3 shows significant levels of cross-sector divergence in labour productivity for the whole period. The low relative productivity of agriculture, forestry and fishing in the 1970s indicates the existence of a labour surplus in this sector. As Brazil grew and developed, labour productivity in this sector increased at a rate above the average productivity growth rate of the economy and, at the same time, the sector's employment share declined. Yet, in the same period, due to rapid urbanisation, the employment shares of wholesale and retail trade, hotels and restaurants, and community, social and personal services increased and their labour productivity grew at a lower rate than the average productivity rate of the economy. It is worth noting that, in contrast to the former's growth movement, this latter contributed to the increase in structural heterogeneity. So the Brazilian industrialisation-*cum*-urbanisation process was characterised by the reduction of labour surplus in primary activities, alongside an increase in mostly informal low-productive, low-paid services localised in urban centres. In sum, the industrialisation process was not able to produce a decline in the degree of structural heterogeneity or to eliminate the surplus labour existing in the Brazilian economy. Hence, it seems implausible that a generalised labour force

Table 7.3 Relative productivity, employment structure and labour productivity growth by sector

Variable	Sector	Agriculture, Forestry, and Fishing	Mining and Quarrying	Manufacturing	Public Utilities	Construction	Wholesale and Retail Trade, Hotels and Restaurants	Transport, Storage, and Communication	Finance, Insurance, and Real Estate	Community, Social and Personal Services (including Government Services)	Sectoral Sum
Relative Sectoral Labour Productivity levels	1970	19	269	180	133	142	83	99	460	164	100
	1980	17	205	190	250	138	69	141	357	110	100
	1990	28	372	143	470	135	41	132	377	95	100
	2000	37	646	166	1010	141	36	122	267	92	100
	2005	45	620	167	888	134	33	119	251	88	100
Sectoral Employment Structure	1970	47.9%	0.5%	14.2%	1.1%	6.1%	9.6%	3.3%	4.0%	13.2%	100%
	1980	37.2%	0.6%	13.6%	0.7%	8.5%	10.9%	3.2%	6.3%	19.0%	100%
	1990	25.5%	0.6%	15.5%	0.6%	6.7%	16.5%	3.9%	6.5%	24.3%	100%
	2000	20.7%	0.4%	13.0%	0.3%	6.2%	19.7%	4.3%	6.5%	28.9%	100%
	2005	18.7%	0.4%	13.1%	0.4%	5.6%	21.1%	4.5%	6.8%	29.4%	100%
Sectoral Labour Productivity growth	1970–1980	3.4%	2.0%	5.4%	11.7%	4.5%	2.9%	8.6%	2.2%	0.7%	4.8%
	1980–1990	3.0%	4.1%	–4.7%	4.4%	–2.1%	–6.9%	–2.6%	–1.4%	–3.3%	–1.9%
	1990–2000	4.0%	6.7%	2.5%	9.0%	1.4%	–0.4%	0.2%	–2.5%	0.6%	1.0%
	2000–2005	4.1%	–0.7%	0.2%	–2.4%	–0.9%	–1.8%	–0.5%	–1.0%	–0.6%	0.1%

Sources: Authors' elaboration based on Groningen Growth and Development Centre 10-sector Database (June, 2007), and Timmer and de Vries (2007).

constraint was a prevalent characteristic of the Brazilian economy's growth trajectory during the period under study.

The behaviour of labour productivity growth in relation to the GDP growth path that we observe supports the last conclusion. In fact, as shown in Figure 7.3, both series present a similar behaviour and move normally in the same direction. According to the neoclassical standpoint this phenomenon could be explained by assuming that the behaviour of the labour productivity growth rate drives that of GDP. Certainly, assuming that the economy has achieved almost continuous full employment, movements in the former are expected to produce similar movements in the latter. But once it is assumed that the Brazilian economy in the period under study was not characterised by the existence of continuous full employment, that kind of explanation completely loses its plausibility, and an alternative must be offered. A potential candidate is the hypothesis subjacent to the so-called 'Kaldor-Verdoorn Law' literature[3], according to which the latter empirical regularity suggests that the behaviour of GDP growth is the cause behind the observed labour productivity growth movements. This would be the case due to the existence of static and dynamic economies of scale associated with the process of economic growth.

Figure 7.3 Labour productivity and GDP growth rates, Brazil 1951–2009
Source: The Conference Board Total Economy Database (September 2010).

Consequently, conceiving economic expansion as a demand-led process, an increase in the demand-induced growth rate is thought to push labour productivity growth upward and to soften the impact of faster growth on labour requirements (and *vice versa*). Generally, the inducement we are analysing is not strong enough to prevent the GDP growth rate from being positively related to the employment growth rate. It only implies that the values of output-elasticity of employment range between zero and one. We conclude, therefore, that the availability of the labour force was not an effective constraint on the long-term expansion of the Brazilian economy during the period under review.

7.4 Theoretical background

The present chapter examines the hypothesis that the Brazilian economic expansion of 1970–2005 was a demand-led growth process subject to a balance of payments constraint. From the theoretical standpoint our analytical framework is based on the classical (or Sraffian) supermultiplier demand-led growth model proposed by Serrano (1995 and 1996).[4] A simplified, small open economy version of the model has been used as a reference. In this section we will briefly discuss the main hypothesis and the implications of the model adopted for this investigation.

First, the equilibrium between real aggregate supply and aggregate demand will be considered. By using the maximum number of components on the demand side that our database allows, the following equation is obtained:

$$Y + M = C_{HND} + C_{HD} + I_H + C_G + I_G + I_{SE} + I_{PE} + X \tag{7.1}$$

where Y is the gross domestic product; M is imports; C_{HND} is household non-durable consumption; C_{HD} is household durable consumption; I_H is household (residential) investment; C_G is government consumption; I_G is government investment; I_{SE} is state-owned enterprises investment; I_{PE} is private enterprises investment; and X refers to exports.

Now, it is assumed that total imports are related to total aggregate demand as expressed in the equation below:

$$M = (1 - \mu)(C_{HND} + C_{HD} + I_H + C_G + I_G + I_{SE} + I_{PE} + X) \tag{7.2}$$

where μ is the share of domestic content in total demand.

Additionally, it is also assumed that household non-durable consumption and private enterprise investment are induced expenditures:

$$C_{HND} = cY \tag{7.3}$$

$$I_{PE} = hY \tag{7.4}$$

where c is the propensity to consume non-durables, and h is the propensity of private enterprises to invest. Finally, Z designates the total autonomous expenditure, as follows:

$$Z = C_{HD} + I_H + C_G + I_G + I_{SE} + X \tag{7.5}$$

According to the supermultiplier model, aggregate consumption expenditures contain an induced component related to the purchasing power introduced into the economy by current production decisions and is usually associated with the wage bill generated by such decisions. So, one major determinant of the propensity to consume (c) is the wage share in aggregate income. As regards aggregate private enterprise investment, equation [7.4] is intended to capture how the level of economic activity can influence the investment decisions of that particular agent. Private enterprises' propensity to invest (h) is considered an endogenous variable in the classical supermultiplier model, since the behaviour of h can be explained by the deviations between the realised output capacity utilisation rate and the normal capacity utilisation rate. In this sense, the process of capitalist competition is expected to induce an increase of h insofar as a positive deviation occurs and *vice versa*. The variation in h will then induce an adjustment of productive capacity to match aggregate demand. As a result, the model predicts that the level of capacity utilisation tends to gravitate around its normal level. Furthermore, the model also predicts that the average value of h is thought to be positively related to the GDP growth rate trend.

Yet, substituting expressions [7.3], [7.4] and [7.5] in equations [7.1] and [7.2], and then solving the resulting system for the real GDP level will result in the following expression for the real GDP equilibrium level (Y^*):

$$Y^* = \left(\frac{\mu}{1 - \mu(c + h)} \right) Z = \alpha Z \tag{7.6}$$

where α is the supermultiplier, which, as shown in equation [7.6], depends on how the propensity to consume, the propensity to invest and the share of domestic content in demand behave.

Hence, according to the model, the behaviour of the supermultiplier (α) and the total autonomous expenditures (Z) can explain GDP behaviour. It is worth noting that the variables contained in the supermultiplier formula exert their influence on GDP growth rates exclusively by means of their rate of change – and not by their levels. This feature is significant for the analysis of the relation between functional income distribution and economic growth. According to the supermultiplier model, the main channel through which the share of wages (or the share of profits) affects the GDP growth rate is by influencing the propensity to consume and, hence, the supermultiplier itself. However, according to the model, the propensity to consume depends on the level of wage share. So, in order for this variable to exert its influence on the GDP expansion rate it must present a positive or negative growth rate. Therefore, in the classical supermultiplier model, functional income distribution can influence GDP growth insofar as it varies. In other words, a *once and for all* change in income distribution can only have a *temporary* effect on GDP growth rate (that is, it produces only a *level* effect on GDP).[5]

In our theoretical framework the balance of payments constraint is the main obstacle to economic expansion. It is a financial constraint on demand-led growth that is related to the availability of international currency, which in the period under study means basically the availability of dollars (Serrano, 2003; Medeiros and Serrano, 1999). Indeed, the accumulation of foreign reserves can keep the overall balance of payments in surplus. On the other hand, a deficit position cannot be sustained if the country in deficit is not an issuer of internationally accepted currency, such as the USA in the postwar period and, notably, after the end of the Bretton-Woods system. This is because a deficit position may set forces in train that are strong enough to limit the level of activity and the pace of economic expansion, so it could lead to an unsustainable loss of foreign reserves and/or pressure over nominal exchange rates and, therefore, over inflation rates. Eventually, and as a result of the action of those forces, authorities will be persuaded to limit the level of economic activity and the growth rate by using economic policy instruments to control aggregate demand.

It should be noted, however, that in our view the external restriction discussed above is not a permanent constraint. The literature[6] on balance of payments constrained growth, by contrast, deals with external constraints, and sees the balance of payments equilibrium as an attractor or an economic policy goal. Nevertheless, as we have seen, a surplus of the overall balance of payments is sustainable, and it is not presumed that governments of countries in that position will always push the economy to the limits defined by external constraints. Therefore, our theoretical framework conceives two possible growth regimes: a balance of payments constrained demand-led growth process, and a policy-constrained (or pure) demand-led growth process.

As Medeiros and Serrano (2001) observed, exports have a dual role in the process of economic growth.[7] First, they are a source of autonomous aggregate demand and, as such, export growth contributes directly to the expansion of GDP. Second, exports are a source of international currency and, therefore, indirectly enable the expansion of both domestic expenditures and GDP.[8] Consequently, in contrast with the literature on balance of payments constrained growth, our framework does not impose an export-led growth pattern on the interpretation of economic data. In fact, our theoretical framework is flexible enough to capture either a pattern of economic growth led by exports or a pattern of growth led by domestic expenditures. Such flexibility is particularly suitable for dealing with relatively closed economies like Brazil's, which, due to internal and/or external factors, has been alternating over time between the two regimes.

7.5 Empirical methodology

In order to examine the growth pattern of the Brazilian economy in the period under study, we adopted a demand-side growth accounting decomposition methodology rather than the supply-side growth accounting approach inspired by neoclassical growth theory.[9] It is worth noting that our empirical approach has been adapted from the structural decomposition analysis developed in the input-output literature,[10] which is largely based on Leontief's open input-output model. This model is compatible with our theoretical framework since it is a demand-oriented model.

The methodology has been adapted to capture the distinction between autonomous and induced variables that characterises the classical super-multiplier model. Indeed, it proved necessary to construct empirical proxies to distinguish between autonomous expenditures and induced expenditures. Government consumption was regarded as autonomous spending; household consumption was divided into autonomous and induced components. A distinction between durable and non-durable (including services) consumption was used as a proxy for distinguishing autonomous consumption from induced consumption, on the grounds that durable consumption is usually financed by credit, which is a source of purchasing power not directly related to current production decisions.[11]

Moreover, the amount of credit applied to durable consumption depends on institutional and social factors, such as the evolution of durable consumption patterns; income distribution patterns; institutionally determined base rates; the intensity of competition in the financial system that regulates banking; banking risk policies; tax structure; the legal framework behind the credit system; the public sector's role in the credit system; the relations between the financial sector and the central bank; and the monetary and credit policy framework adopted by the monetary authority, among others. It is acknowledged that households' debt-to-income ratio also influences the expansion of credit to durable consumption. Note, however, that such an influence is rather weak, particularly when longer periods of analysis are considered and, consequently, the institutional and social factors just mentioned are subject to significant changes, even when the analysis considers one country's experience only. On the other hand, non-durable consumption largely depends on the purchasing power introduced into the economy by current production decisions, namely payments of wages. Finally, as regards investment expenditures, since capitalist competition prompts a trend for the adjustment of productive capacity to match demand, private enterprise investment was considered induced spending. Other aggregate investment components (housing, government and state-owned enterprises investments) included in the analysis were considered autonomous, since they are neither driven by capitalist competition[12], nor directly related to the purchasing power introduced into the economy by production decisions.

To meet our methodology's requirements, a database was compiled from various sources: from the historical statistics of the system of national accounts (SNA);[13] complementary information from the work of Dos Santos and Pires (2007); and from some old input-output tables (specifically those for 1970, 1975, 1980 and 1985). The old input-output tables allowed for the division of household consumption into durable and non-durable components for the years before 1990.[14] Dos Santos and Pires' (2007) investment database allowed us to break down enterprise investment into private and state-owned components and, thus, to quantify private enterprise investment for the 2000s, which are not covered by official statistics.

The decomposition formula is derived from the national account identity between aggregate supply and aggregate demand.[15] The final formula is as follows:[16]

$$g = \alpha(1)\left[\frac{C_{HND}(0)}{Y(0)}\right]g_c + \alpha(1)\left[\frac{I_{PE}(0)}{Y(0)}\right]g_h + \alpha(1)\left[\frac{C_{HD}(0)}{Y(0)}\right]g_{C_{HD}}$$
$$+\alpha(1)\left[\frac{I_H(0)}{Y(0)}\right]g_{I_H} + \alpha(1)\left[\frac{C_G(0)}{Y(0)}\right]g_{C_G} + \alpha(1)\left[\frac{I_G(0)}{Y(0)}\right]g_{I_G}$$
$$+\alpha(1)\left[\frac{I_{SE}(0)}{Y(0)}\right]g_{I_{SE}} + \alpha(1)\left[\frac{X(0)}{Y(0)}\right]g_X + \frac{\alpha(1)}{\mu(1)}g_\mu + \alpha(1)\left[\frac{E(0)}{Y(0)}\right]g_E \quad [7.7]$$

where E is the value of inventory change; g is real GDP growth rate; and g_i refers to the real growth rate of variable i.

On the left-hand side (LHS) of the equation, the observed real GDP growth rate lies between a base period (denoted as period (0)) and a final period (denoted as period (1)).[17] The right-hand side (RHS) of the equation contains the factors that contribute to the expansion of the GDP. First, in order for the observed real GDP growth rate to match exactly the RHS of the equation, the last term referring to the contribution of inventory change growth rate was incorporated into the RHS. For a similar reason, all variables were deflated using the GDP price deflator, instead of using their own price deflators. This procedure has a significant implication: as for the changes in the variables' growth rates contained on the RHS of the equation, they will account changes expressed not only in volume, but also in relative prices (relative to the GDP price deflator). Second, on the RHS the contribution from each variable depends on its respective real growth rate multiplied by an accounting GDP elasticity specific to each variable being considered. All accounting elasticities depend on the value of the supermultiplier in the final period; all of them, except for the one related to the share of domestic content in demand, will also depend on the ratio of the respective variable value to the value of the GDP in the base period.

Finally, one should note that there is an important difference between the method of accounting for the contributions of the variables for autonomous expenditures, on the one hand, and the variables contained in the supermultiplier formula, on the other. The reason is that, in the latter case, the decomposition captures the contributions of the *propensity* to consume, the *propensity* to invest and the share of domestic content growth rates. That is, to be compatible with our theoretical framework, household non-durable (induced) consumption, private enterprise (induced) investment, and import contributions are related to the growth rate of those ratios. As for the autonomous expenditure variables, however, the decomposition captures their contributions in relation to their own growth rates.

The results produced by the growth accounting decomposition will be presented in two alternative forms. First, from the perspective of the sources of the contributing factors, the decomposition will quantify the contributions of both domestic and external sectors. The contribution of the domestic sector to GDP growth will be measured according to the following expression:

Domestic sector contribution

$$= \alpha(1)\left[\frac{C_{HND}(0)}{Y(0)}\right]g_c + \alpha(1)\left[\frac{I_{PE}(0)}{Y(0)}\right]g_h + \alpha(1)\left[\frac{C_{HD}(0)}{Y(0)}\right]g_{C_{HD}} + \alpha(1)\left[\frac{I_H(0)}{Y(0)}\right]g_{I_H}$$
$$+ \alpha(1)\left[\frac{C_G(0)}{Y(0)}\right]g_{C_G} + \alpha(1)\left[\frac{I_G(0)}{Y(0)}\right]g_{I_G} + \alpha(1)\left[\frac{I_{SE}(0)}{Y(0)}\right]g_{I_{SE}} \quad [7.8]$$

The contribution from domestic sector will then be divided into two components: private sector and public sector, which are represented, respectively, by the two following equations:

Public sector contribution

$$= \alpha(1)\left[\frac{C_G(0)}{Y(0)}\right]g_{C_G} + \alpha(1)\left[\frac{I_G(0)}{Y(0)}\right]g_{I_G} + \alpha(1)\left[\frac{I_{SE}(0)}{Y(0)}\right]g_{I_{SE}} \quad [7.9]$$

and

Private sector contribution

$$= \alpha(1)\left[\frac{C_{HND}(0)}{Y(0)}\right]g_c + \alpha(1)\left[\frac{I_{PE}(0)}{Y(0)}\right]g_h + \alpha(1)\left[\frac{C_{HD}(0)}{Y(0)}\right]g_{C_{HD}}$$
$$+ \alpha(1)\left[\frac{I_H(0)}{Y(0)}\right] \quad [7.10]$$

The external sector contribution, in turn, captures the influence of export growth and of changes in the share of domestic content in demand, and is measured according to the following equation:

$$\textit{External sector contribution} = \alpha(1)\left[\frac{X(0)}{Y(0)}\right]g_x + \frac{\alpha(1)}{\mu(1)}g_\mu \quad [7.11]^{18}$$

So the real GDP growth rate will be decomposed into its contributing factors according to the following equation:

$$\begin{aligned} g &= \textit{Domestic sector contribution} \\ &\quad + \textit{External sector contribution} + \alpha(1)\left[\frac{E(0)}{Y(0)}\right] g_E \\ &= \textit{Private sector contribution} + \textit{public sector contribution} \\ &\quad + \textit{External sector contribution} + \alpha(1)\left[\frac{E(0)}{Y(0)}\right] g_E \end{aligned} \qquad [7.12][19]$$

Second, from another perspective, the RHS terms of decomposition equation [7.7] are grouped according to whether contributions come from autonomous expenditures or from changes in the supermultiplier. In our model contributions of autonomous expenditures will be quantified as follows:

$$\begin{aligned} &\textit{Autonomous expenditures contribution} \\ &= \alpha(1)\left[\frac{C_{HD}(0)}{Y(0)}\right] g_{C_{HD}} + \alpha(1)\left[\frac{I_H(0)}{Y(0)}\right] g_{I_H} + \alpha(1)\left[\frac{C_G(0)}{Y(0)}\right] g_{C_G} \\ &\quad + \alpha(1)\left[\frac{I_G(0)}{Y(0)}\right] g_{I_G} + \alpha(1)\left[\frac{I_{SE}(0)}{Y(0)}\right] g_{I_{SE}} + \alpha(1)\left[\frac{x(0)}{Y(0)}\right] g_x \end{aligned} \qquad [7.13]$$

For its part, changes in the supermultiplier contribution will be measured quantitatively according to equation [7.14]:

$$\begin{aligned} &\textit{Supermultiplier contribution} \\ &= \alpha(1)\left[\frac{C_{HND}(0)}{Y(0)}\right] g_c + \alpha(1)\left[\frac{I_{PE}(0)}{Y(0)}\right] g_h + = \left[\frac{\alpha(1)}{\mu(1)}\right] = g_\mu \end{aligned} \qquad [7.14]$$

Hence, combining equations [7.13] and [7.14] and including the term related to the contribution of inventory change results in an alternative decomposition for the real GDP growth rate, which can be expressed as follows:

$$\begin{aligned} g &= \textit{Autonomous expenditures contribution} \\ &\quad + \textit{Supermultiplier contribution} + \alpha(1)\left(\frac{E(0)}{Y(0)}\right) g_E \end{aligned} \qquad [7.15]$$

Table 7.4 enables the decomposition quantitative results that are presented in Section 7.4 to be connected with the decomposition formulas discussed above. The first column ('Variables') lists the symbols for the variables used in the decomposition analysis. Column six ('Total')

Table 7.4 Decomposition equations in table form

Variables	Domestic Sector		External Sector	Inventory Changes	Total	Autonomous Expenditures	Supermultiplier	Inventory Changes
	Public	Private						
C_G	$\alpha(1)\left[\frac{C_G(0)}{Y(0)}\right]g_{C_G}$				$\alpha(1)\left[\frac{C_G(0)}{Y(0)}\right]g_{C_G}$	$\alpha(1)\left[\frac{C_G(0)}{Y(0)}\right]g_{C_G}$		
I_G	$\alpha(1)\left[\frac{I_G(0)}{Y(0)}\right]g_{I_G}$				$\alpha(1)\left[\frac{I_G(0)}{Y(0)}\right]g_{I_G}$	$\alpha(1)\left[\frac{I_G(0)}{Y(0)}\right]g_{I_G}$		
I_{SE}	$\alpha(1)\left[\frac{I_{SE}(0)}{Y(0)}\right]g_{I_{SE}}$				$\alpha(1)\left[\frac{I_{SE}(0)}{Y(0)}\right]g_{I_{SE}}$	$\alpha(1)\left[\frac{I_{SE}(0)}{Y(0)}\right]g_{I_{SE}}$		
I_H		$\alpha(1)\left[\frac{I_H(0)}{Y(0)}\right]g_{I_H}$			$\alpha(1)\left[\frac{I_H(0)}{Y(0)}\right]g_{I_H}$	$\alpha(1)\left[\frac{I_H(0)}{Y(0)}\right]g_{I_H}$		
C_{HD}		$\alpha(1)\left[\frac{C_{HD}(0)}{Y(0)}\right]g_{C_{HD}}$			$\alpha(1)\left[\frac{C_{HD}(0)}{Y(0)}\right]g_{C_{HD}}$	$\alpha(1)\left[\frac{C_{HD}(0)}{Y(0)}\right]g_{C_{HD}}$		
C_{HND}		$\alpha(1)\left[\frac{C_{HND}(0)}{Y(0)}\right]g_C$			$\alpha(1)\left[\frac{C_{HND}(0)}{Y(0)}\right]g_C$		$\alpha(1)\left[\frac{C_{HND}(0)}{Y(0)}\right]g_C$	
I_{PE}		$\alpha(1)\left[\frac{I_{PE}(0)}{Y(0)}\right]g_h$			$\alpha(1)\left[\frac{I_{PE}(0)}{Y(0)}\right]g_h$		$\alpha(1)\left[\frac{I_{PE}(0)}{Y(0)}\right]g_h$	
μ			$\frac{\alpha(1)}{\mu(1)}g_\mu$		$\frac{\alpha(1)}{\mu(1)}g_\mu$		$\frac{\alpha(1)}{\mu(1)}g_\mu$	
X			$\alpha(1)\left[\frac{X(0)}{Y(0)}\right]g_X$		$\alpha(1)\left[\frac{X(0)}{Y(0)}\right]g_X$	$\alpha(1)\left[\frac{X(0)}{Y(0)}\right]g_X$		
E				$\alpha(1)\left[\frac{E(0)}{Y(0)}\right]g_E$	$\alpha(1)\left[\frac{E(0)}{Y(0)}\right]g_E$			$\alpha(1)\left[\frac{E(0)}{Y(0)}\right]g_E$
Total	$\sum$	$\sum$	$\sum$	$\sum$	g	$\sum$	$\sum$	$\sum$

reproduces the decomposition equation [7.7] in tabular form, where the column total cell equals the observed real GDP growth rate. This column also divides the table into two parts, each containing one of the two alternative methods of presenting the decomposition results as previously described in equations [7.8] to [7.15]. On the left-hand side, columns two ('Domestic sector – Public'), three ('Domestic sector – Private') and four ('External sector') represent equations [7.9], [7.10] and [7.11], respectively. Column five ('Inventory changes') represents the contribution for inventory change. In its turn, the sum of the totals for columns two, three, four and five corresponds to equation [7.12]. The right-hand side of column six ('Totals') presents the decomposition in an alternative way. So, column seven ('Autonomous expenditures') gives equation [7.13] in tabular form, while column eight ('Supermultiplier') contains equation [7.14]. Finally, the sum of the totals contained in columns seven, eight and nine represents equation [7.15], where the total cell of column nine ('Inventory changes') represents the contribution from inventory change.

Before concluding this section, the analytical status of the empirical methodology previously described will be briefly addressed. Similarly to the neoclassical supply-side growth accounting decomposition, the demand-side decomposition method adopted here is not intended to provide an empirical test of the theoretical framework subjacent to it. In fact, it is assumed that such framework is valid for an interpretation of the pattern of economic growth prevailing in the Brazilian economy in the period under study. In this sense, the decomposition exercise is to be understood, at best, as an indirect way of evaluating our theoretical framework, depending on its ability to provide a plausible and consistent explanation for the observed facts.

7.6 Some results

This section analyses the growth performance of the Brazilian economy for the whole period between 1970 and 2005.[20] We begin with an overview of the external situation and how it connects with the process of economic growth in Brazil during that period.

As shown in Figure 7.4, in the 1970s the Brazilian economy experienced huge current account deficits, mostly financed by financial account surplus positions. However, a deficit situation in the overall balance of payments emerged as a consequence of the two oil shocks and their impact on the world economy. In particular, the second oil shock in 1979 and the consequent increase in US interest rates had a major negative effect on Latin America peripheral countries, leading to the external debt crises that

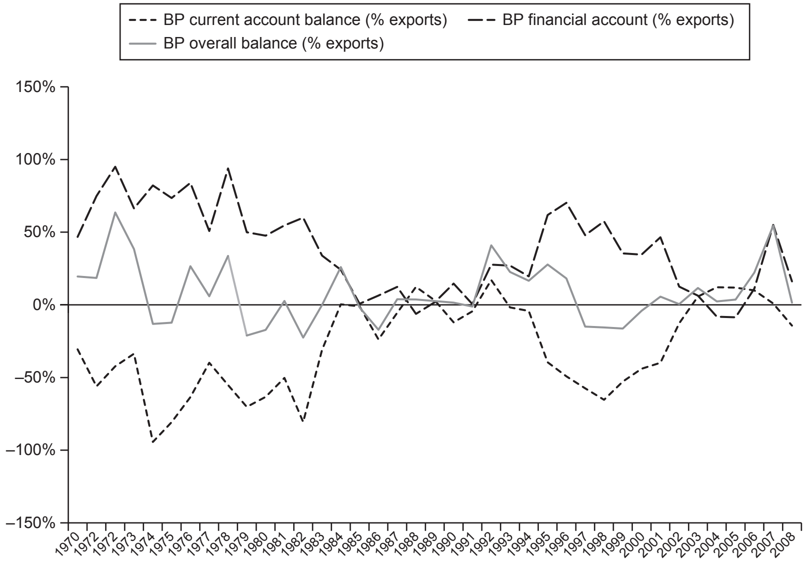


Figure 7.4 Selected balance of payments accounts, Brazil 1970–2008 (percentage of exports)
Source: Authors' elaboration based on IBGE, FGV, and Dos Santos and Pires (2007).

jeopardised the region's economic performance in the 1980s. Figure 7.4 also shows that in the 1980s, as international capital flows to Brazil were interrupted, the problem of external constraint became more acute.

Moreover, in order to keep the current account deficits under control, the Brazilian monetary authorities continued to promote nominal exchange rate devaluations. Combined with a high level of formal and informal indexation of nominal contracts and the escalation of social conflicts over income distribution, those measures also caused the inflationary process to escalate. Brazil entered a period of chronically high inflation that lasted until the early 1990s, when the return of international capital flows enabled the Real stabilisation plan to be successfully implemented in 1994, which in turn put an end to the policy of continuous nominal exchange rate devaluations and managed partially to eliminate inflation-indexed nominal contracts.

On the other hand, the return of the Brazilian economy to the international financial circuit in the early 1990s prompted a real exchange rate appreciation, which, along with the trade liberalisation policies that had been introduced in the late 1980s and intensified in the early 1990s, produced the high current account deficits that were characteristic of the late 1990s. So the balance of payments constraint continued to be a binding constraint on GDP growth. In fact, during this period the Brazilian economy faced a series of external shocks related to the Mexican, Asian and Russian crises. Finally, in 1998 Brazil experienced a balance of payments crisis and had to appeal to the IMF for help.[21]

The Brazilian crisis prompted an important change in the national macroeconomic policy regime. The new policy framework combined inflation targeting, large primary government budget surplus and floating (but still widely managed) exchange rates. In particular, the focus on price stabilisation and sound finance helped shape a conservative consensus on matters of economic policy. At least until 2003, such conservative consensus coexisted with a situation of considerable external pressure (particularly in the year 2002, due to the Brazilian presidential election). However, the balance of payments situation has been improving since 2003 mostly as a result of real exchange rate depreciation after 1999; favourable development of international commodity trading; and resumption of large gross private capital flows (Serrano and Summa, 2012). The current account deficit was converted to surplus and the overall balance of payments also showed a surplus, which enabled a policy of foreign reserves accumulation and payment of accumulated foreign debt (including IMF loans) to be implemented.[22] The continuous accumulation of foreign reserves, however, indicated that the economy was able to expand at higher rates and, therefore, that the binding constraint in the period was not the external one – a conclusion supported by the trend towards real exchange rate appreciation, observed from 2003 on, as a result of the *modus operandi* of the inflation targeting regime adopted in Brazil. The Brazilian central bank took advantage of the favourable external environment to adopt high positive interest rate differentials to control the nominal exchange rate in order to achieve the official inflation rate target (Serrano, 2010). Hence, from 2003 on, the Brazilian economy was characterised by a demand-led growth process chiefly influenced by internal factors, most of them related to the political dispute involving the maintenance or the modification of the macroeconomic policy regime inside and outside the government (Barbosa and Souza, 2010; Serrano and Summa, 2012).

Now, applying the decomposition methodology enables us to identify the pattern of economic growth prevailing in the Brazil during the period analysed. Table 7.5 shows the results.

As shown in Table 7.5, from 1970 to 2005 Brazil grew by 3.99 per cent per year. The domestic sector was the main source of GDP growth in the period, contributing 2.5 percentage points (hereafter denoted pp), while

Table 7.5 Average annual rate of growth decomposition, Brazil, 1970–2005 (constant 1980 prices)

Variables	Domestic Sector		External Sector	Inventory Change	Total	Autonomous Expenditures	Super-multiplier	Inventory Change
	Public	Private						
C_G	1.83%				1.83%	1.83%		
I_G	0.11%				0.11%	0.11%		
I_{SE}	–0.004%				–0.004%	–0.004%		
I_H		0.35%			0.35%	0.35%		
C_{HD}		0.42%			0.42%	0.42%		
C_{HND}		–0.15%			–0.15%		–0.15%	
I_{PE}		–0.04%			–0.04%		–0.04%	
μ			–0.11%		–0.11%		–0.11%	
X			1.43%		1.43%	1.43%		
E				0.16%	0.16%			0.16%
Total	1.93%	0.57%	1.32%	0.16%	3.99%	4.13%	–0.30%	0.16%

Source: Authors' elaboration based on IBGE, FGV and Dos Santos & Pires (2007).

the external sector had a contribution of 1.32 pp. Within the domestic sector, the public sector was the major source of GDP growth, with a contribution of 1.93 pp, while the private sector made a more modest contribution of 0.57 pp. These results suggest that the Brazilian economy was characterised by a domestic demand-led growth pattern or, more specifically, by a 'government demand-led' growth pattern, in which the expansion of the domestic market figured as the major source of economic growth.[23]

Next, looking at the other method of presenting the decomposition results, autonomous expenditure growth figures as the main source of economic growth, with a positive contribution of 4.13 pp, in contrast to the supermultiplier's change negative contribution of −0.30 pp. This result shows that, during the whole period under review, autonomous expenditure growth was the major determinant of the GDP growth rate trend. For some periods, nevertheless, changes in the supermultiplier made a significant contribution to GDP growth, as can be seen where the GDP growth rate and the autonomous expenditure growth rate diverge in Figure 7.5.

The average annual autonomous expenditure growth rate was 4.7 per cent for the whole period. Exports grew, on average, 6.3 per cent

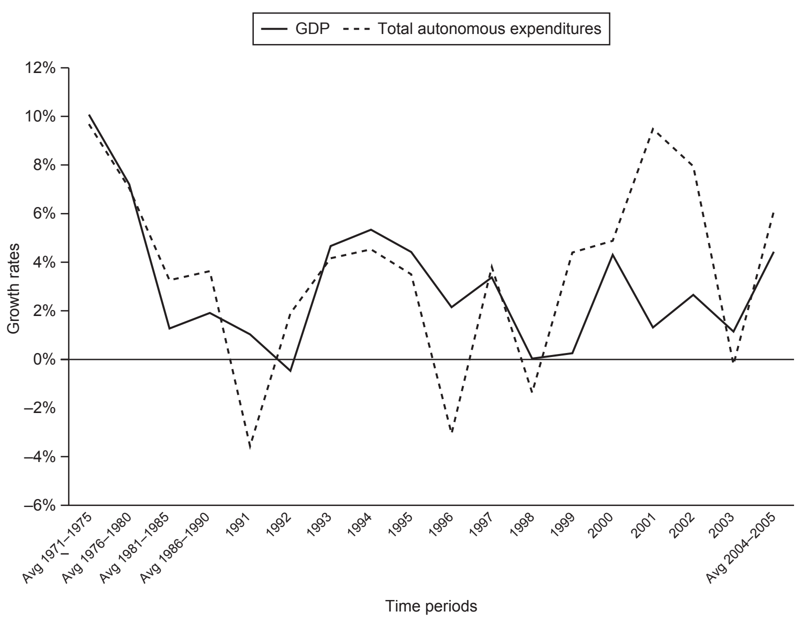


Figure 7.5 GDP growth rates and autonomous expenditure growth rates
Source: Authors' elaboration based on IBGE, FGV, and Dos Santos and Pires (2007).

annually, while autonomous private and public expenditures grew, respectively, 3.4 per cent and 4.7 per cent per year on average. This result contrasts with the one related to the growth decomposition presented above. Indeed, if we took growth rates as a reference, the Brazilian economy should be expected to follow an export-led growth pattern. Still, as the decomposition shows, this is not the case. The reason is that a higher export growth rate is more than compensated by the lower value of exports' share of GDP in the base year (that is, 7.0 per cent in 1970). On the other hand, public autonomous expenditures had a GDP share of 18.0 per cent in 1970. That share, combined with the growth rate of public autonomous expenditures in the period, explains why the public sector made the highest contribution to GDP expansion. Two explanations for the initial low share of exports in GDP are the inward-oriented state-led development process in the post-WWII period and the continental size of Brazil.[24] The shares of public and private autonomous expenditures, as well as the share of exports, are shown in Figure 7.6 below. It is worth mentioning, however, that the situation has changed since the 1970s. As shown in Figure 7.6, by the end of the period, exports' share of GDP had increased to around 15 per cent, whereas public autonomous expenditures were of the order of 22 per cent.[25] In the light of that, it can

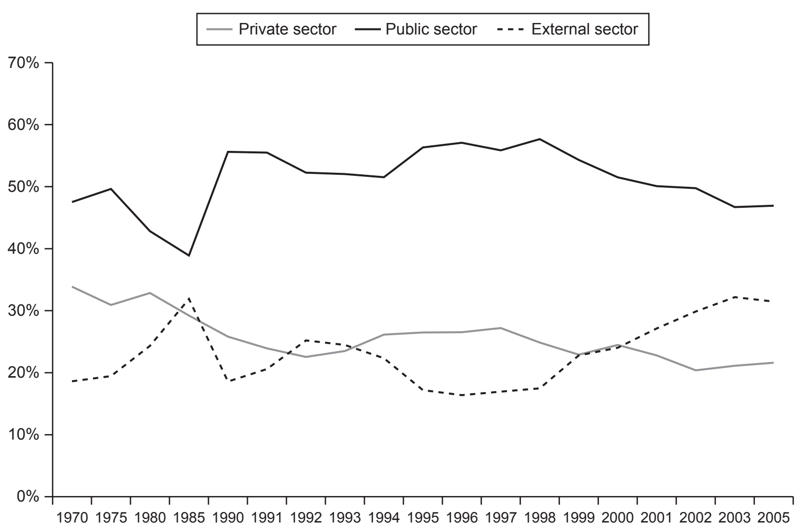


Figure 7.6 Autonomous expenditures share of GDP
Sources: Authors' elaboration based on IBGE, FGV, and Dos Santos and Pires (2007).

be suggested that the Brazilian economy has become more dependent on the dynamism of external markets.[26] Yet it is not clear whether this result means a structural change has occurred in the Brazilian economy's growth pattern.

Another important issue concerns how the components of autonomous expenditure growth behave over time. Figure 7.7 shows that exports were the most volatile component, followed by private and public autonomous expenditures. This suggests that the observed pattern of expansion of the Brazilian economy, which is mostly driven by public sector expenditures, would be more volatile if private and external sectors were dominant sources of economic growth. Therefore, should the Brazilian economy become more dependent on the dynamism of external markets (a possibility suggested above), one would also expect it to be more susceptible to fluctuations in exports.

Our analysis now turns to the behaviour of the supermultiplier. As illustrated in Figure 7.8, the supermultiplier showed a declining trend during the period with which we are concerned. In the 1970s

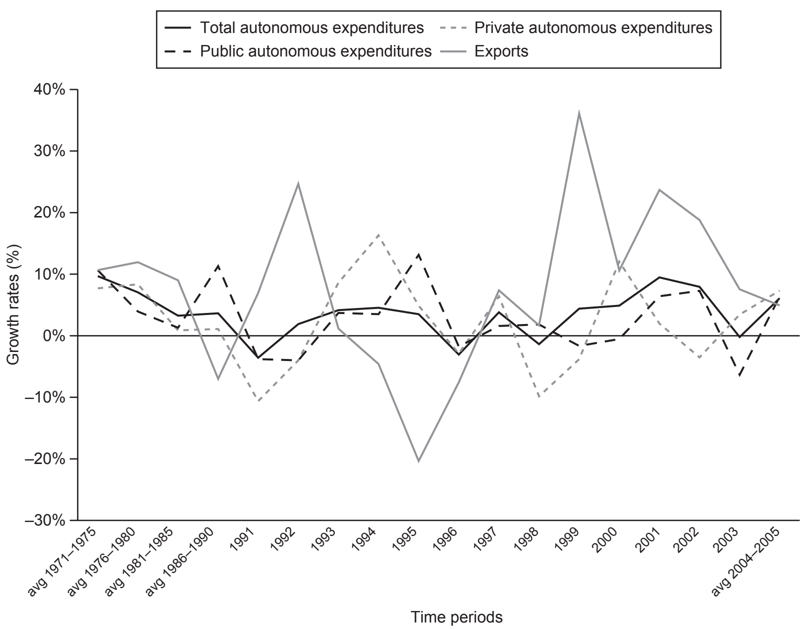


Figure 7.7 Annual average growth rates of autonomous expenditure components (percentage per annum)
Sources: Authors' elaboration based on IBGE, FGV, and Dos Santos and Pires (2007).

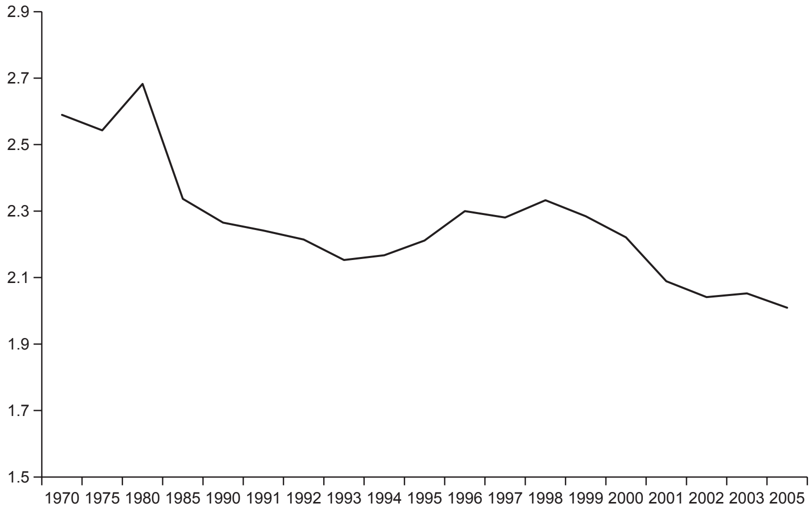


Figure 7.8 Supermultiplier
Sources: Authors' elaboration based on IBGE, FGV, and Dos Santos and Pires (2007).

its value ranged between 2.5 and 2.7, but by the end of the decade it had declined to approximately 2.0. Table 7.5 demonstrates that this decline was due to three factors: a reduction in propensity to consume, a decline in private enterprise propensity to invest, and a decrease in the share of domestic content in total demand.

As Table 7.5 shows, the growth in propensity to consume contributed –0.15 pp to the GDP growth rate (4.0 per cent), almost half the value of the total contribution of the supermultiplier to economic growth (–0.30 pp). As we have argued above, according to the supermultiplier model the wage share (of GDP) is one of the major determinants of propensity to consume. Figure 7.9 illustrates this point by showing that propensity to consume and wage share tend to move in the same direction. So the worsening of functional income distribution apparently contributed to the decline of propensity to consume.

Next, the contribution of the share of domestic content will be assessed. According to Table 7.5, that component scored a contribution of –0.11 pp to the observed GDP growth rate and explained one-third of the decline in the supermultiplier. As shown in Figure 7.10, the share of domestic content presented an upward trend until 1989, with a peak value of almost 95 per cent, followed by a significant downward

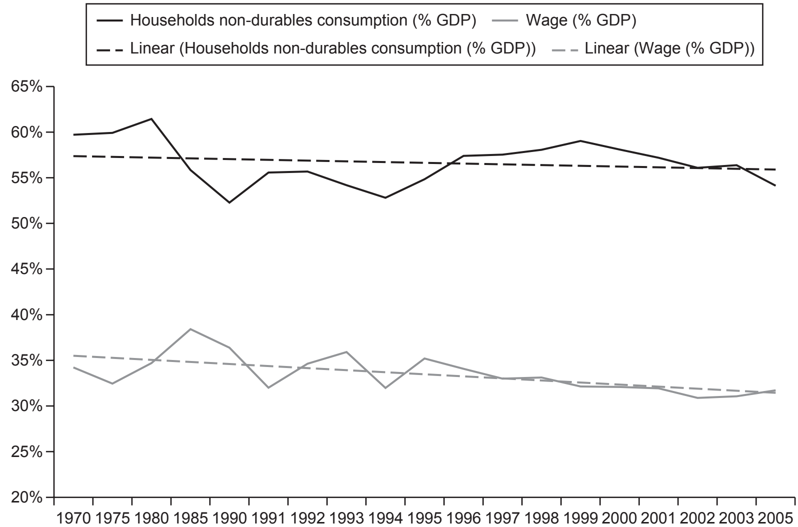


Figure 7.9 Household non-durables consumption and wages (as percentage of GDP)
Sources: Authors' elaboration based on IBGE, FGV, and Dos Santos and Pires (2007).

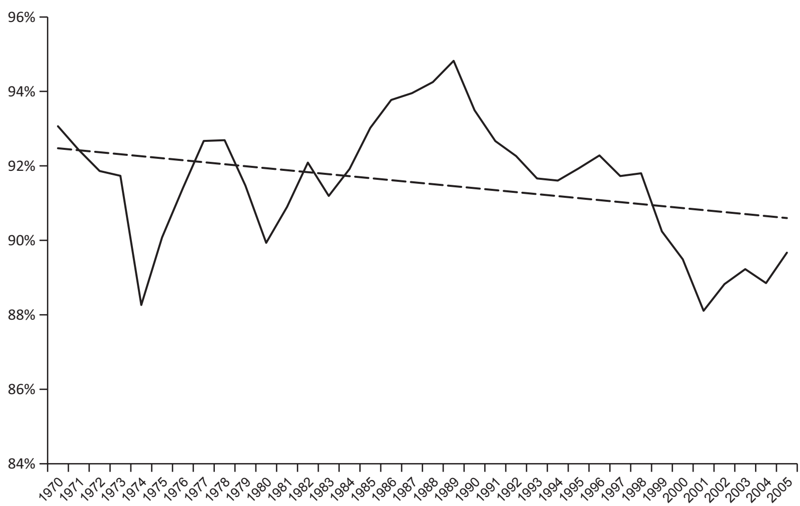


Figure 7.10 Share of domestic content in demand (percentage)
Sources: Authors' elaboration based on IBGE, FGV, and Dos Santos and Pires (2007).

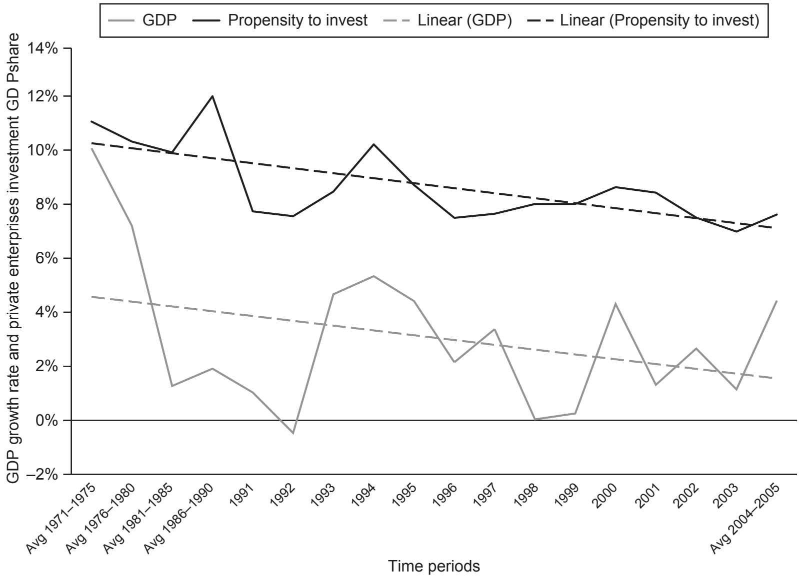


Figure 7.11 GDP growth rates and private enterprise investment share of GDP (percentage)

Source: Authors' elaboration based on IBGE, FGV, and Dos Santos and Pires (2007).

trend falling to approximately 90 per cent in 2005. This last trend resulted from a combination of trade liberalisation policies initiated in 1987 and the overvaluation of the real exchange rate.[27]

Table 7.5 shows that the contribution made by private enterprise's propensity to invest to the GDP growth rate scored –0.04 pp. Furthermore, as mentioned before, according to the classical supermultiplier model, the relation between the ratio of private enterprise investment to GDP, on the one hand, and the GDP growth rate trend, on the other, is expected to be positive. According to Figure 7.11, the observed behaviour of these two variables in our period is compatible with the relation proposed by the model. The downward trend of the GDP growth rate is followed by a decrease in private enterprise propensity to invest, as measured in this chapter.[28]

7.7 Concluding remarks

Between 1970 and 2005 Brazil experienced a huge decline in GDP growth rates, from an average rate of almost 8.5 per cent per year in the 1970s to approximately 2.3 per cent per year from the 1980s to 2005.

The neoclassical standpoint has dominated the more recent literature dedicated to studying that period. However, we are adopting a completely different perspective, based on a combination of the classical supermultiplier demand-led growth model and the hypothesis that the balance of payments was the main potential (and often the most effective) constraint on the expansion of the Brazilian economy during the period under study.

The main results of our investigation are the following. First, public sector expenditures were the major source of economic growth in the period analysed. The main influence of external conditions in general and exports in particular was the indirect effect they exerted on the growth of domestic demand. Although exports made a significant contribution, their growth path proved to be very unstable and, even when their contribution to economic growth was relatively high, they were unable to sustain alone the pace of GDP growth observed in the 1970s. This explains why in certain periods the Brazilian economy did experience, as suggested by Medeiros and Serrano (2001), a pattern of export-led 'stagnation'.

Second, our empirical investigation showed that the classical supermultiplier model is able to explain the main features of Brazilian growth in those decades. The contribution of autonomous expenditures to economic growth considered in the model is able to explain most of the observed GDP growth rate trend for the whole period. The share of GDP that wages represent seemed to be a significant determinant of propensity to consume. The propensity to invest by private enterprises was positively related to the GDP growth rate trend. Furthermore, our methodology proved flexible enough to enable a history-friendly approach to the analysis of growth experiences. In particular, the approach allowed for the consistent incorporation into the analysis of features pertaining to institutional and power relations, also proving compatible with the main theoretical and methodological aspects of the classical surplus approach to political economy.

Notes

1. Lack of space prevents the analysis here of the sub-periods between 1970 and 2005, which will be the subject of a future work.
2. See Abreu and Verner (1997), Ferreira and Malliagros (1998), Pinheiro *et al.* (2001), Gomes *et al.*, (2003), Bacha and Bonelli (2004) and Loayza *et al.* (2005).

3. See Kaldor (1989[1981]), McCombie *et al*. (2003), Thirlwall (1983) and, for evidence related to the Brazilian economy, see Marinho *et al*. (2002).
4. See also Cesaratto *et al*. (2003) and Serrano and Freitas (2007).
5. This specification must be contrasted with the results obtained in the neo-Kaleckian growth models, especially those connected with the contributions of Marglin and Bhaduri. As is known, those models allows for the existence of a regular relation between the wage share (or the profit share) and the trend GDP rate of growth.
6. See Kaldor (1978[1971]), Kaldor (1989[1983]), McCombie and Thirlwall (1994) and McCombie and Roberts (2002).
7. The balance of payments constrained growth literature does not properly acknowledge the dual role of exports, frequently ignoring the possibility or the likelihood of a pure demand-led growth regime in the sense defined above.
8. It must be noted that, as a source of international finance, exports stand out as a source of international currency without present or future international payment flows. The same is not true for other sources of external finance such as, for example, direct investment, portfolio investment, trade credits and international loans.
9. See Solow (1957) for the seminal contribution and Hulten (2010) for a recent survey of the literature.
10. For a survey of the literature see Miller and Blair (2009, chapter 13).
11. Following Serrano (1995), here we are using the Kaleckian classification of autonomous expenditures as related to the source of purchasing power introduced in the economy. For a very clear account of Kalecki's use of this classification see López and Assous (2010, chapter 2, appendix 2.1).
12. State-owned enterprises' investment was classified as an autonomous expenditure because, for most of this period, capitalist competition did not exert a major influence on its behaviour.
13. Data pertaining to periods until 1989 were provided by the Getulio Vargas Foundation (FGV in Portuguese). Data from 1990 on have been provided by the Brazilian Institute of Geography and Statistics (IBGE in Portuguese), which is now responsible for the SNA.
14. For data after (and including) 1990 the Brazilian National Accounts System adopted the United Nations 1993 revision of the SNA methodology, which requires the compilation of use tables to allow for the identification of durable consumption.
15. Note that for the decomposition to be exhaustive we have to include in the RHS of equation [7.1] an extra variable that captures the unplanned inventory accumulation. The inclusion of this variable transforms the equilibrium condition represented in equation [7.1] into an identity. For more details, see the Appendix.
16. The detailed derivation of the decomposition formula is shown in the Appendix.
17. The symbols (0) and (1) attached to a variable in level mean that the level of the variable in question is evaluated in the base period and in the final period respectively.
18. Note that this way of calculating the external sector contributions to economic growth contrasts with the usual method that uses net exports.

19. Again, the inclusion of the inventory change growth contribution is deemed necessary to guarantee equality between the RHS of the equation and the observed real GDP growth rate.
20. For a more detailed account of the Brazilian economy in this period see Serrano (1998), Serrano and Summa (2012) and Barbosa and Souza (2010)
21. Note that from 1997 to 2000 the deficit in the overall balance of payments is evidence of the severity of the external problems faced by the Brazilian economy in the late 1990s.
22. It should be noted, however, that this policy gained momentum from 2006 on and, therefore, in a period not covered by our analysis.
23. However, it does not mean that the external sector was not an important determinant of the growth process, since, as indicated above, its influence on growth can also be exerted indirectly by means of the balance of payments constraint to the domestic demand expansion.
24. One should note that the continental size of the Brazilian economy continues to contribute to the prevalence of a domestic demand-led growth pattern.
25. The figures for the shares of autonomous expenditures in GDP must be analysed with caution. One implication of the classical supermultiplier model is that a lower trend of GDP growth rate is expected to be related to a higher share of total autonomous expenditures in GDP. This occurs because lower GDP growth rates tend to lower the propensity to spend (that is, a lower ratio of private enterprise investment to GDP) and, consequently, to lower the supermultiplier value. Equation [7.6] shows that the share of total autonomous expenditures and the supermultiplier value are inversely related. Therefore, from the viewpoint of the model adopted here, the share of autonomous expenditures in GDP should not be interpreted as an indicator of the dynamism of this expenditure unless the behaviour of the variables contained in the supermultiplier formula are thoroughly analysed, particularly the behaviour of private enterprise investment.
26. More recent data (after 2005) reveal that as a share of GDP, exports have declined to values ranging between 11 per cent and 12 per cent. Nevertheless, these values are higher than the average export share observed from 1970 to 1999.
27. The Brazilian crisis of 1998 interrupted the over-valuation trend for a while and led to a depreciation trend that lasted until 2003. From then on, as mentioned before, the Brazilian central bank maintained relatively high international interest rate differentials to keep inflation under control. The management of nominal exchange rate led to the appreciation of real exchange rate or to the maintenance of a relatively overvalued real exchange rate from 2003 on.
28. It should be noted that the positive correlation between GDP growth rates and investment (in physical capital) to GDP ratio is one of the most robust empirical relations obtained in the economic growth applied literature (see, for example, Sala-i-Martin, 1997). In addition, many very influential models in the growth literature are not compatible with this stylised fact, such as the Solow model (Solow, 1956) and the neo-Kaleckian models (see Marglin and Bhaduri, 1990).

References

Abreu, M. and Verner, D. (1997) *Long-Term Brazilian Economic Growth: 1930–94* (Paris: OECD).

Bacha, E. L. and Bonelli, R. (2004) 'Accounting for Brazil's Long Term Growth', *Discussion Paper* 1018 (Rio de Janeiro: IPEA – Institute for Applied Economic Research).

Barbosa, N. H. and Souza, J. A. P. (2010) 'A inflexão do Governo Lula: política econômica, crescimento e distribuição de renda', in E. Sader and M. A. Garcia (eds), *Brasil entre o Passado e o Futuro* (São Paulo: Boitempo Editorial): pp. 57–110.

Cesaratto, S., Serrano, F. and Stirati, A. (2003) 'Technical change, effective demand and employment', *Review of Political Economy*, 15(1): 33–52.

Dos Santos, C. H. and Pires, M.C. (2007) 'Reestimativas do Investimento Privado Brasileiro: qual a sensibilidade do investimento privado "referência 1985" a aumentos da carga tributária?', *Discussion Paper* 1297 (Rio de Janeiro: IPEA – Institute for Applied Economic Research).

Elias, V. J. (1978) 'Sources of Economic Growth in Latin American Countries', *Review of Economics and Statistics*, LX(30): 362–70.

Elias, V. J. (1992) *Sources of growth – a study of seven Latin American economies* (San Francisco: ICS Press).

Ferreira, P. C. and Malliagros, T.G. (1998) 'Impactos produtivos da infra-estrutura no Brasil – 1950–1995', *Pesquisa e planejamento econômico*, vol. 28, no. 2: 315–38 (Rio de Janeiro: IPEA).

Gomes, V., de A. Pessôa, S. and Veloso, F. (2003) 'Evolução da Produtividade Total dos Fatores na Economia Brasileira: Uma Análise Comparativa', *Pesquisa e Planejamento Econômico*, 33(3), December (Rio de Janeiro: IPEA).

Groningen Growth and Development Centre 10-sector database, June 2007, http://www.ggdc.net/

Hulten, C. R. (2010) 'Growth Accounting', in B. H. Hall and N. Rosenberg (eds), *Handbook of the Economics of Innovation* (Amsterdam and Oxford: North-Holland): pp. 987–1031.

Kaldor, N. (1978 [1971]) 'Conflicts in National Economic Objectives', in N. Kaldor, *Further Essays on Economic Theory* (New York: Holmes & Meier).

Kaldor, N. (1989 [1981]) 'The Role of Increasing Returns, Technical Progress and Cumulative Causation in the Theory of International Trade and Economic Growth', in N. Kaldor, *Further Essays on Economic Theory and Policy* (New York: Holmes & Meier).

Kaldor, N. (1989 [1983]) 'The Role of Effective Demand in the Short Run and the Long Run', in N. Kaldor, *Further Essays on Economic Theory* (New York: Holmes & Meier).

Loayza, N., Fajnzylber, P. and Calderón, C. (2005) *Economic Growth in Latin America and the Caribbean: Stylized Facts, Explanations, and Forecasts* (Washington: World Bank).

López, J. and Assous, M. (2010) *Michal Kalecki* (Basingstoke and New York: Palgrave Macmillan).

Maddison, A. (2010) *Historical Statistics of the World Economy: 1–2008 AD* (Groningen: Growth and Development Centre).

Marglin, S. A. and Bhaduri, A. (1990) 'Profit Squeeze and Keynesian Theory', in S. A. Marglin and J. B. Schor (eds) *The Golden Age of Capitalism* (Oxford: Oxford University Press).

Marinho, E. L., Nogueira, C. G. and Rosa, A. L. T. (2002) 'Evidências Empíricas da Lei de Kaldor-Verdoorn para a Indústria de Transformacão no Brasil (1985–1997)', *Revista Brasileira de Economia*, 56(3): 457–82.

McCombie, J. S. L. and Thirlwall, A. P. (1994) *Economic Growth and the Balance of Payments Constraint* (London: Macmillan).

McCombie, J., Pugno, M. and Soro, B. (2003) *Productivity Growth and Economic Performance: Essays on the Verdoorn's Law* (Basingstoke and Hampshire: Palgrave Macmillan).

McCombie, J. S. L. and Roberts, M. (2002) 'The Role of the Balance of Payments in Economic Growth', in M. Setterfield (ed.), *Demand-Led Growth: Challenging the Supply Side Vision of the Long Run* (Aldershot: Edward Elgar).

Medeiros, C. A. and Serrano, F. (1999) 'Padrões Monetários Internacionais e Restrição Externa ao Crescimento', in J. L. Fiori (ed.), *Estados e Moedas no Desenvolvimento das Nações* (Petrópolis: Vozes).

Medeiros, C. A. and Serrano, F. (2001) 'Inserção externa, exportações e crescimento no Brasil', in J. Fiori and C. Medeiros (eds), *Polarização Mundial e Crescimento* (Petrópolis: Vozes).

Miller R. E. and Blair P. D. (2009) *Input–Output Analysis: Foundations and Extensions* (Cambridge: Cambridge University Press).

Pinheiro, A. C., Gill, I. S., Servén, L., Thomas, M. R. (2001) *Brazilian Economic Growth 1900–2000: Lessons and Policy Implications* (Washington, DC: World Bank).

Sala-i-Martin, X. (1997) 'I just ran two million regressions', *American Economic Review*, Papers and Proceedings of the Hundred and Fourth Annual Meeting of the American Economic Association, 87(2):178–83.

Serrano, F. (1995) 'Long period effective demand and the Sraffian supermultiplier', *Contributions to Political Economy*, 14(1): 67–90.

Serrano, F. (1996) 'The Sraffian supermultiplier', unpublished dissertation, Cambridge University, England.

Serrano, F. (2003) 'From "static" gold to the floating dollar', *Contributions to Political Economy*, 22(1): 87–102.

Serrano, F. (2010) 'Taxa de juros, taxa de câmbio e o sistema de metas de inflação no Brasil', *Revista de Economia Política*, 30: 63–72.

Serrano, F. and Freitas, F. N. P. (2007) El Supermultiplicador Sraffiano y el Papel de la Demanda Efectiva en los Modelos de Crecimiento, *Circus*, 1: 19–35.

Serrano, F. and Summa, R. (2012) 'Macroeconomic policy, growth and income distribution in the Brazilian economy in the 2000s', *Investigación Económica*, 71, October–December: 55–92.

Solow, R. M. (1956) 'A contribution to the theory of economic growth', *Quarterly Journal of Economics*, 70(1): 65–94.

Solow, R. M. (1957) 'Technical change and the aggregate production function', *Review of Economics and Statistics*, 39(3): 312–20.

Thirlwall, A. P. (1983) 'A plain man's guide to Kaldor's growth laws', *Journal of Post Keynesian Economics*, 5(3): 345–58.

Timmer, M. P. and de Vries, G. J. (2007) 'A Cross-Country Database for Sectoral Employment and Productivity in Asia and Latin America, 1950–2005', *Groningen Growth and Development Centre Research Memorandum GD-98*, (Groningen: University of Groningen).

Appendix

List of symbols

Y – Gross domestic product
M – Imports
C_{HND} – Households non-durable consumption
C_{HD} – Households durable consumption
I_H – Households (residential) investment
C_G – Government consumption
I_G – Government investment
I_{SE} – State-owned enterprises investment
I_{PE} – Private enterprises investment
X – Exports
E – Inventory change
μ– Share of domestic content in demand
c – Households' propensity to consume
h – Private enterprise's propensity to invest
Z – Total autonomous expenditures
α – The supermultiplier
g – GDP growth rate
gi – Growth rate of variable i

The decomposition methodology

We will start with the national account identity between aggregate supply and aggregate demand. Using the maximum number of components of the demand side that our database allows results in the following equation:

$$Y+M=C_{HND}+C_{HD}+I_H+C_G+I_G+I_{SE}+I_{PE}+X+E$$

Now, it is assumed that imports are related to total aggregate demand as expressed in:

$$M=(1-\mu)\ (C_{HND}+C_{HD}+I_H+C_G+I_G+I_{SE}+I_{PE}+X+E)$$

Next, it is assumed that:

$$C_{HND}=cY$$

$$IPE=hY$$

$$Z=C_{HD}+I_H+C_G+I_G+I_{SE}+X$$

Then, by substituting the above relations on the first two equations, the following is obtained:

$$Y=\mu cY+\mu hY+\mu(Z+E)$$

This equation will serve as a starting point for the subsequent GDP growth decomposition analysis. So, GDP change will be taken as described in the following equation:

$$\begin{aligned}Y(1)-Y(0)=\mu(1)c(1)Y(1)-\mu(0)c(0)Y(0)+\mu(1)h(1)Y(1)\\ -\mu(0)h(0)Y(0)+\mu(1)[Z(1)+E(1)]-\mu(0)[Z(0)+E(0)]\end{aligned}$$

Adding and subtracting the terms $\mu(1)c(1)Y(0)$ and $\mu(1)h(1)Y(0)$ to/from the RHS of the equation and assuming the fact that $\Delta Y=gY(0)$, the following is obtained:

$$\begin{aligned}gY(0)=\mu(1)c(1)gY(0)+\mu(1)h(1)gY(0)+[\mu(1)c(1)-\mu(0)c(0)]Y(0)\\ +[\mu(1)h(1)-\mu(0)h(0)]\ Y(0)+\mu(1)\ [\Delta Z+\Delta E]+\Delta\mu[Z(0)+E(0)]\end{aligned}$$

Dividing both sides of the equation by $Y(0)$ results in the following:

$$\begin{aligned}g=\mu(1)c(1)g+\mu(1)h(1)g+\mu(1)c(1)-\mu(0)c(0)+\mu(1)h(1)-\mu(0)h(0)\\ +\mu(1)\left[\frac{\Delta Z}{Y(0)}\right]+\mu(1)\left[\frac{\Delta E}{Y(0)}\right]+\Delta\mu\left[\frac{Z(0)+E(0)}{Y(0)}\right]\end{aligned}$$

By adding and subtracting $\mu(1)\,c(0)$ and $\mu(0)\,h(0)$ to/from the RHS, the following is obtained:

$$\begin{aligned}g=\mu(1)\ c(1)\ g+\mu(1)\ h(1)\ g+\mu(1)\ \Delta c+\Delta\mu\, c(0)+\mu(1)\ \Delta h+\Delta\mu h(0)\\ +\mu(1)\left[\frac{\Delta Z}{Y(0)}\right]+\mu(1)\left[\frac{\Delta E}{Y(0)}\right]+\Delta\mu\left[\frac{Z(0)+E(0)}{Y(0)}\right]\end{aligned}$$

Solving the above equation for the growth rate results in:

$$\begin{aligned}g=\alpha(1)\Delta c+\frac{\alpha(1)}{\mu(1)}\Delta\mu c(0)+\alpha(1)\Delta h+\frac{\alpha(1)}{\mu(1)}\Delta\mu h(0)+\alpha(1)\left[\frac{\Delta Z}{Y(0)}\right]\\ +\alpha(1)\left[\frac{\Delta E}{Y(0)}\right]+\frac{\alpha(1)}{\mu(1)}\Delta\mu\left[\frac{Z(0)+E(0)}{Y(0)}\right]\end{aligned}$$

First collect all the terms in which $\Delta\mu$ appears. Then, put $\frac{\alpha(1)}{\mu(1)}\Delta\mu$ in evidence and use the definition that $\Delta\mu=g_{\mu}\mu(0)$ to arrive at the fourth term on the RHS of the equation below. In addition, the fact that:

$$\Delta Z=\Delta C_{HD}+\Delta I_H+\Delta C_G+\Delta I_G+\Delta I_{SE}+\Delta X$$

is used to obtain the third term on the RHS of the following expression.

$$g = \alpha(1)\Delta c + \alpha(1)\Delta h + \alpha(1)\left[\frac{\Delta C_{HD} + \Delta I_H + \Delta C_G + \Delta I_G + \Delta I_{SE} + \Delta X}{Y(0)}\right]$$
$$+ \frac{\alpha(1)}{\mu(1)} g_\mu \left[\frac{\mu(0)\{c(0)Y(0) + h(0)Y(0) + Z(0) + E(0)\}}{Y(0)}\right]$$
$$+ \alpha(1)\left[\frac{\Delta E}{Y(0)}\right]$$

However, it is known that:

$$Y(0) = \mu(0)c(0)Y(0) + \mu(0)h(0)Y(0) + \mu(0)(Z(0) + E(0)).$$

So, the fourth term on the RHS is equal to $\frac{\alpha(1)}{\mu(1)} g_\mu$. On the other hand, the third term on the RHS can be dismembered to isolate the individual contributions of each type of expenditure involved. Further, all contributions can be expressed in terms of the real growth rate of the variables involved. As a consequence, the following equation (which appears in the text – see p. 170) is obtained:

$$g = \alpha(1)\left[\frac{C_{HND}(0)}{Y(0)}\right]g_c + \alpha(1)\left[\frac{I_{PE}(0)}{Y(0)}\right]g_h + \alpha(1)\left[\frac{C_{HD}(0)}{Y(0)}\right]g_{C_{HD}} + \alpha(1)\left[\frac{I_H(0)}{Y(0)}\right]g_{I_H}$$
$$+ \alpha(1)\left[\frac{C_G(0)}{Y(0)}\right]g_{C_G} + \alpha(1)\left(\frac{I_G(0)}{Y(0)}\right)g_{I_G} + \alpha(1)\left(\frac{I_{SE}(0)}{Y(0)}\right)g_{I_{SE}}$$
$$+ \alpha(1)\left[\frac{X(0)}{Y(0)}\right]g_X + \frac{\alpha(1)}{\mu(1)} g_\mu + \alpha(1)\left(\frac{E(0)}{Y(0)}\right)g_E$$

Part II
Understanding the International Economic Order

8

Continuity and Change in the International Economic Order: Towards a Sraffian Interpretation of the Changing Trend of Commodity Prices in the 2000s

*Franklin Serrano**

8.1 Introduction

In the first decade of the twenty-first century we can observe some elements of continuity and others of change in the international economic order. In terms of continuity, what is perhaps most striking is the resilience of the 'floating dollar standard', which was neither a cause of the major world crisis of 2008, nor was negatively affected by it.[1] In terms of change, there is the new tendency towards a greater relative autonomy of the (relatively fast) rates of growth of GDP of many developing economies from the (low) growth rates of the advanced capitalist countries, A second (and intimately connected) change is the increasing absolute (dollar) and relative prices for internationally traded 'commodities' (food and raw materials in general).

Both new trends are ultimately related to four interconnected features of the world economy after 2003. The first was the relatively low rates

*In addition to the presentation at the Sraffa 2010 Conference at Roma Tre University, the chapter draws on two later presentations on these topics, at CEFIDAR, Buenos Aires, in August 2011 and at the 'Congreso Internacional Crisis de la teoría económica y políticas alternativas ante la crisis global', 7–11 November 2011 at UNAM, Mexico City. The author is grateful to the organisers and participants of these three events. Special thanks are due to very useful discussions on the topic with my colleagues, Professors Carlos Medeiros, Eduardo Crespo and Eduardo Pinto and also to Fabian Amico and an anonymous referee for their comments on an earlier draft. None of them, of course, bears any responsibility for the remaining errors and mistakes.

of interest in the USA and other advanced countries and the accompanying massive gross capital flows towards the developing economies in general. The second was the shift towards policies of fast growth of internal demand in large developing countries and the acceleration of South–South trade. Third, the revival of natural resource nationalism in many developing and transition countries. Finally, there has been a marked improvement in policies for managing the balance of payments in many developing countries. These improved policies have variously included managed floating exchange rate regimes, a massive accumulation of foreign exchange reserves, the early repayment (or default in the case of Argentina) of official external debt, the setting up of sovereign funds and the selective taxation of some commodity exports (and at times subsidy of some commodity imports).[2]

All these elements taken together have led to the remarkable absence of balance of payments crises originating in the periphery and a quick recovery from the late 2008 world financial crisis.

These changes appear to be the combined result of (i) the favourable state of the world economy in terms of both growth of trade and access to capital flows; and (ii) an important reaction to economic policies adopted by developing countries to the various financial and balance of payments crises of the late 1990s up to 2002. In any case those changes in the international situation are behind the partial decoupling of the trend growth rates of developing countries as a whole (both for commodity and industrial-exporting countries) observed in the 2000s. Ultimately they were also very important for the change in relative commodity prices, which in this decade saw a reversal of the falling trend of the two previous ones.

The purpose of this chapter is to provide a preliminary attempt to understand the broad proximate causes of this marked, and to a large extent surprising, change in dollar and relative international 'commodity' prices in this new international context. For reasons of space references to the other new feature of the world economy, the tendency towards a decoupling of trend rates of growth of developing countries and of the political and policy changes in many developing countries that ultimately rendered these two trends possible, despite its importance, will be kept to the minimum required for an understanding of the question in focus, namely, the direct causes of increased dollar and relative international 'commodity' prices.[3] Again due to space limitations discussion is mainly on the analysis of commodity market behaviour and on the rapid increase of dollar commodity prices in the 2000s, which are generally not well understood instead of focusing on the much better-known causes of the low rate of increase of the dollar prices of manufactured exports during this period, which are connected with money wage and productivity trends in both the advanced and the newly industrialised Asian countries.

We consider that the modern classical surplus approach, revived by Sraffa and Garegnani since the early 1960s, provides a useful analytical framework on which to base this study. It is well known that the classical notion of price of production, revived by Sraffa, focuses on the objective material elements in the 'cost of production' of all produced goods. This should also apply to the relatively standardised foods and raw materials traded in international markets that are nowadays known as 'commodities'.[4] Moreover, the theory of prices of production foregrounds the necessary connections between costs of production' and distributive variables, and in particular the (social) rules that govern the distribution of income across between wages, profits and different types of rents (and also exchange rates), in a particular historical situation.

Most analyses of the recent increase in commodity prices have focused almost exclusively on the demand side, whether that of final users or for speculative purposes. There has been a relative neglect of the role of certain supply constraints and of more persistent cost-of-production elements. And there has been, if anything, an even greater neglect of the close connection between these changes in the costs of production and associated changes in the distribution of income.

We should perhaps also note that this neglect of cost-of-production and income distribution aspects in the recent literature contrasts sharply not only with the Sraffian classical surplus approach but also with the analyses of the pioneers of development economics, such as Singer, Prebisch and Lewis, who focused on precisely these aspects when studying the long-run trends of terms of trade.[5] In order to illustrate the close links between our own Sraffian standpoint and that of the pioneers of development economics, we shall discuss below the extent to which both the main changes in the international economic order, namely, the partial decoupling of the growth rates of developing countries and the new trend of relative commodity prices, were in part anticipated as possibilities by pioneer development economist Arthur Lewis.

The analysis will argue that changing international relative commodity prices in the 2000s reflect changes in their relative production costs. Nominal unit dollar costs of commodities have increased rapidly due to the deliberate restriction of the supply of oil (a result of both OPEC and the revival of 'natural resource nationalism'), rising costs of minerals (due in part to capacity constraints on more efficient mines) and the fast growth of real wages in major mineral and agricultural commodity-exporting countries (together with the real appreciation of many of their currencies). On the other hand, unit dollar costs of manufactured export products rose slowly due to the slow growth of real wages in advanced capitalist countries, combined with real wages which, though growing fast, are not keeping pace with growth of industrial productivity in the developing countries that export industrial goods (particularly in China and other Asian NICs).

The argument of this chapter will proceed as follows. Section 8.2 briefly describes the main characteristics of the recent boom in commodity prices. Section 8.3 addresses the *demand*-side elements of the boom in dollar commodity prices. Section 8.4 looks at some of the supply and *cost-of-production* aspects of the rise in dollar prices of specific broad groups of commodities. Section 8.5 first discusses the changes in the balance of payments position of commodity-exporting countries in the 2000s, including the partial decoupling of growth rates of developing countries in general, in order to show the important role for the cost and dollar prices of commodities in general of the trend towards a *nominal (dollar) and real appreciation of the currencies of major commodity-exporting countries*. Section 8.6 deals briefly with some better-known causes of the slow growth of non-commodity international dollar prices, which are needed to understand the question of the rise in the *relative* prices of commodities. Section 8.7 concludes by comparing and constrasting our own suggested interpretation of the recent change in relative prices with the old views of Singer, Prebisch and Lewis.

8.2 The recent boom in commodity prices

Apart from being extremely volatile, international commodity prices in general (agricultural and mineral) have increased substantially between 2000 and 2010, both in absolute (US dollar) terms and in relative terms, whether we compare them with international dollar prices of manufactured goods or in terms of the overall internal price indexes of most countries.

Dollar prices for oil started recovering, after reaching extremely low historical levels in 1999.[6] Nominal increases in metals and food prices came later. By 2003 oil prices started growing much faster and metals prices even more so, and by 2007 the rapid rise in food prices began. In the wake of the global financial crisis, prices of all types of commodities fell drastically in 2009, but quickly recovered in 2010. Prices reached a peak in mid-2011 with food and metals (but not oil) at higher nominal dollar levels than in 2008. Over the whole decade, it was dollar energy prices that increased most, followed by metals, while food prices increased by much less. According to IMF data, crude oil prices increased at an annual rate of 17.84 per cent a year from 1999 to 2002, and then by 18.14 per cent a year from 2003 to 2010. Prices of metals actually decreased by 0.2 per cent a year from 1999 to 2002, but then increased at a very fast rate of 20.36 per cent a year from 2003 to 2010. Food prices increased from 1999 to 2002 at a low rate of 0.28 per cent a year, but rose by 4.3 per cent a year from 2003 to 2010.

Over the same period, international industrial prices and world inflation did not keep pace with such fast commodity price increases, resulting in a large increase in the *relative* price of all types of commodities. Indeed, world inflation actually fell from an average of 4.44 per cent a year during 1999–2002 to 3.87 per cent in 2003–10. Consumer price inflation in the richest developed economies remained practically stable around 2 per cent over the whole decade. But inflation in developing countries fell from 8.19 per cent during 1999–2002 to 6.72 per cent in 2003–2010 in the case of Latin American countries. International dollar prices for manufactured products also increased at a much lower rate than commodity prices. WTO's index of manufacturing unit values (MUVs) actually decreased at a rate of 2 per cent a year from 1999 to 2002 and then increased at an annual rate of 4.85 per cent from 2003 to 2010. For commodities as a whole, 2003–08 was the longest phase of fast-rising dollar prices and had the largest overall through to peak price increase (131 per cent) since 1900 (World Bank, 2009, 2012).

8.3 Demand aspects

8.3.1 The 'China demand' effect for metals

Most analysts attribute the rapid increase in both dollar and relative commodity prices to the acceleration of the growth of the world economy in the 2000s, based on the very fast growth rates of some developing countries. According to this view, the increasing weight of developing countries in the world economy has been characterised by a process of heavy industrialisation and urbanisation, as well as by the spread of new Western consumption habits. This process has increased the demand for oil, metals and food in these countries and is the basis for a 'super cycle' of commodity prices that could last for a couple of decades.

Among these fast-growing developing countries, it is usually China's extremely rapid commodity import growth rate that is emphasised. Indeed, between 2002 and 2003, Chinese imports of commodities increased more than 40 per cent. In 2003 China was responsible for 26.5 per cent of the world demand for steel, 19.8 per cent for copper and 19 per cent for aluminum. This account, however, has a number of serious limitations as an explanation for rising commodity prices.

First of all, the world economy did not actually grow significantly faster in the 2000s compared with the 1990s. In fact, it grew much faster in the second half of the 1990s than in the 2000s, as the US 'dotcom boom' then coincided with China's very high growth rate. It is true that the growth rate of developing countries was, as a whole, faster in the 2000s than in the 1990s but this largely just made up for the reduction in the growth rate of the advanced economies.

The same growth pattern can also be seen in indices of world industrial production and the volume of world trade in merchandise.

Thus neither world GDP, nor world industrial production nor world volume of trade grew relatively faster in the 2000s than the 1990s, even if we exclude the big recession year of 2009 (as perhaps we should not, since periodic crises are, of course, a constituent part of the commodities and business cycles).

Another problem with explanations centred on the side of demand is that, as is well known, the income elasticity of world demand for most commodities is definitely below one, reflecting both technical change and the usual trend towards reduced intensity of commodities in GDP as income increases. Indeed, since the early 1970s, the decline of commodity intensities has been faster for food and energy commodities. In the case of metals we observe the same, though a little less pronounced, declining intensity, when we exclude China. But in this one case, the China effect on demand is so great that measures of the metals intensity of world GDP since 1995 increase, rather than decrease. Since then, indeed, there has been increasing metal intensity of world GDP. This means that since 1995 the income elasticity of the world demand for metals has definitely been substantially greater than one. In fact China's metal intensity relative to GDP was three times higher than that of the rest of the world in 1990 and became nine times higher than the world's average by 2008 (World Bank, 2009, 2012).

The third problem is that data from China confirm that the 'China demand effect' has only been highly relevant for the world economy as a whole as regards metals. While it is true that consumption and imports of all types of commodities in China have been growing rapidly, in most cases they started from a *very low base* level. Even for a few non-metal products where this was not the case and Chinese consumption has had a greater weight in world demand, China's role seems to have been mostly partially to compensate for the marked decline in the demand for commodities coming from the richer countries.

The big exception, where the Chinese demand has really had a big impact was the extraordinarily high increase in the use of metals in China. This has been associated with the very high growth rates in public and private investment, particularly in construction and infrastructure in general. In the period between 2002 and 2007, China's consumption of coffee increased 32 per cent, but as the world's total consumption actually decreased by 1.9 per cent over the same period, China's contribution to world demand growth in that period was just 0.1 per cent. Beef consumption grew 27 per cent in China over that period, during which world consumption of beef grew 7.2 per cent, to which China contributed only 2 per cent.

In the case of oil, China's consumption grew 48.7 per cent while the consumption in the rest of the word grew 6.6 per cent; the Chinese contribution to world demand growth was of 2.7 per cent over a five-year period (about 0.5 per cent a year).

The case of metals contrasts sharply with that general pattern. Chinese consumption of iron ore increased by 224.9 per cent between 2002 and 2007, while in the rest of the world it grew by 19.5 per cent, and China's contribution to world consumption growth was no less than 38.4 per cent. In the case of other metals such as aluminum, zinc and copper Chinese consumption grew at less spectacular rates over the period, between 70 and 125 per cent, but there was still a 'China effect' of between 10 and 20 per cent of the growth in world consumption of those metals over that five-year period (Jenkins, 2011).

The upshot of all this is that the China effect on the world demand for commodities, apart from metals, has been much smaller than is usually thought.

And even in the case of metals, where massive Chinese demand was of crucial importance for the fast growth of world demand, we find that the acceleration of the world demand for metals started around 1995, with a further acceleration around 2001, while, as we saw in Section 8.2, the prices of metals only start booming after 2003. Something seems to be missing in the demand-side explanation, even for metals.

8.3.2 Speculation

In part precisely because of the quite modest growth in 'requirements of use' (or final demand) for many commodities whose prices also increased sharply, many analysts have attributed the rise in commodity prices to increased speculation. Three elements have appeared, separately or together, as explanatory variables for the increase in speculation in commodity markets: falling interest rates in the USA; the devaluation of the dollar relative to other major currencies such as the euro; and financial deregulation and financial innovation in futures markets for commodities.

On this view, reduced short-term interest rates in the USA would have reduced the attractiveness of financial assets relative to commodities and cheapened the formation of speculative inventories. Moreover, the increase in dollar commodity prices would be a leading indicator that monetary policy is too expansionary. Low interest rates would later lead to a substantial increase in overall inflation and speculators would be anticipating this higher future inflation caused by monetary policy, using commodities as a hedge.

In reality, the idea that lower interest rates may to a certain extent increase the speculative demand for commodity prices is in itself reasonable. But the idea that the large increases in commodity prices that have occurred in the 2000s can be attributed to market expectations of a massive future acceleration of inflation in the USA does not make much sense and is not supported by independent evidence.. 'Core' and/or trend inflation in the USA and other developed countries has been kept low, in spite of drastic increases in dollar commodity prices. The ultimate cause of this regime of inflation moderation appears to be the weak bargaining power of American workers, reinforced by the relatively slow growth of employment in advanced countries and it seems clear that market participants understand that this trend is not likely to change any time soon.[7]

Other analysts argue that commodity speculators are hedging against the devaluation of the dollar against key currencies such as the euro or yen (i.e., currencies of rich countries that do not export commodities). It is pointed out that during some of the sharper short-term dollar commodity price increases the dollar was falling in value relative to these currencies. In fact this presumed correlation between dollar commodity prices and the value of the dollar relative to currencies of rich non-commodity-exporting countries is not at all robust and there are many periods, such as between 1984 and 1995, in which a large and almost continuous fall in the value of the dollar relative to these currencies has coincided with low and falling dollar commodity prices. It is more likely that in certain periods when this correlation can be observed, the falling dollar and rising commodity prices could be responding to a common third cause, such as low interest rate in the USA relative to those of other rich countries.[8]

It is unlikely, however, that lower American interest rates could, on their own, cause such large swings in dollar commodity prices. Indeed, while changes in the interest rate could be having some effect, it is interesting to note that this variable does not appear as a statistically significant driver of dollar commodity prices even in recent studies made by leading advocates of this view.[9] It is probable that this small effect has been totally overshadowed by the wild short-term fluctuations of expectations in the organised commodity markets.

Indeed, the process of financial deregulation and financial innovation, particularly in the American economy, brought an enormous amount of financial funds to commodities futures markets. By some estimates the value of funds directed to these markets by purely financial speculators grew from 13 billion dollars in 2003 to more than 260 billion dollars by March 2008. For many this was a major cause both of the high volatility of dollar commodity prices and, more unusually, also of their rising trend.[10]

On the other hand, those who want to deny that a speculative bubble could be amplifying the recent rise in dollar commodity prices, have two main arguments. First, the data do not show episodes of very large accumulation of physical inventories of commodities, which is considered a necessary 'signature' of speculative activity.[11] Moreover, it is argued that the volume of transactions in the futures markets merely reflects the compensating bets of different agents and does not affect the physical availability of commodities.[12] Therefore, activity in futures markets cannot affect the balance of supply and demand in the spot market, and thus has no effect on prices. Neither of these arguments, however, seems to be very solid.

A large accumulation of inventories is not really a necessary condition for speculation. Those who believe that it is argue that a large accumulation of inventories would be needed to sustain a speculative price increase, in order to compensate for the equilibrating role of the large fall in demand and large increase in supply that presumably occur when the spot price increases.

But perhaps one of the most important features of the products that we call commodities is precisely the difficulty of both decreasing significantly the quantity demanded and increasing significantly the quantity produced in the very short run. Very small imbalances between supply and demand can therefore lead to very large short-term primary fluctuations in spot prices, and divergent expectations between speculators allow the price movements to be magnified in both directions, as the same limited amount of inventories is sold and resold with speculators selling to other speculators who think the price will continue to rise, for instance.[13]

As for the argument of the irrelevance of activity in the futures markets, it fails to see the relevance of the potentially riskless gains of arbitrage between spot and future market prices of the same standardised commodity.

Both the spot and the forward prices of commodities are strongly influenced by the spot prices expected in subsequent periods. If there is a general expectation that the spot price will be higher at a subsequent date, both the spot and the forward prices will tend to increase right now. If speculators, expecting a higher price tomorrow, buy spot today, thinking of selling the commodity tomorrow, the spot price today will tend to increase. And such increase in the spot price is transmitted to the futures market, since now there is the option to reduce the supply allocated for future delivery and sell it today at the initially higher current spot price.

On the other hand, it is also true that if there is a large increase in purchases in the futures market today the spot price will rise today, because now investors have the option of selling more at the higher current forward price for future delivery. Recent financial deregulation and innovation, and in particular the extremely low margin requirements that allow extraordinarily high leverage ratios for financial speculators in commodity markets has greatly increased the availability of credit for speculative purchases in future commodities markets that are very quickly transmitted by arbitrage also to spot market prices.

There is also the counter-argument that volatility was also high for a number of commodities which have no organised futures market, such as rice and iron ore.[14] But this probably means that in these markets there were other forms of access to cheap and plentiful credit, not that no speculation occurred in these spot or the other futures markets.[15] Given all this, it seems reasonable to argue that, in spite of the controversy on the size of the effect of recent regulatory changes on future markets, overall speculation has indeed played an increasingly important role in world commodity markets over the 2000s.

But speculation usually works both ways, sometimes greatly intensifying the price increase when output is perhaps only a little lower than current final demand, and at other times causing dramatic price falls when output is greater than final demand. But then speculation cannot really explain the rising dollar and relative prices of commodities, for there seems to be no obvious reason why the massive short-term price increases have been, on average, so much higher than the, also very large, short-term decreases in commodity prices. Speculation, as we have seen, depends crucially on the expected spot market prices. Why, within such instability and volatility, were expected prices rising, rather than falling or being nearly random?

We thus see that to look only at the demand, whether final or speculative, is either misleading or, in the case of metals where world demand really boomed, incomplete. In order to understand the rise in commodity prices we must look at the amounts and specially the costs of the commodities brought to the market (supply). As Garegnani (1988, p. 254) put it: 'After all, we do explain and try to forecast the trend of the price of, say, copper, on the basis of the technical changes in its mining or the richeness of new mines, etc. – in spite of the fact that the prices of copper may fluctuate perhaps as much as 50% on either side of its trend value'.

8.4 Specific cost-of-production and supply aspects

8.4.1 Oil

Oil prices started to rise earlier and higher than other types of commodities. Here the China (and India) demand effect is quite widely known but particularly misleading. China's oil consumption amounted to less than 8 per cent of world consumption in 2008, and India's to around 3 per cent. And overall world oil consumption grew a little more than 2 per cent a year from 2000 to 2008.[16]

Clearly we must turn to the supply side for a satisfactory explanation of the rising trend of dollar oil prices. But we should be clear what exactly we mean by supply. One of the meanings of supply is the cost of production, and of particular interest is the cost of the oil production methods that have to be used to meet demand. Another meaning refers to the availability of physical quantities. And by availability we may mean either the existing inventories of already produced oil; the current productive capacity that would allow rapid expansion of output to meet demand; or the known deposits or reserves of oil under the ground (or ocean); or even the total physical or geological endowment of the resource left in the planet.

World supply of oil in the sense of existing inventories has not, on average, been scarce relative to demand over the years in which oil prices have been increasing, with world production closely matching the trend of world consumption.

Despite the popular view of increasing physical or geological scarcity and the Hubbert Curve, and the world 'peak oil' doctrine, the availability of oil in the sense of known reserves has not really been an issue (Radetzki, 2010). Not only were the world's oil reserves substantially higher in 2010 than in 2000 in absolute terms, but more importantly, the ratio of current proven reserves to current production reserves actually showed a modest increase – from 40 to 42 – between 2002 and 2007 (World Bank, 2009). In fact this fundamental overabundance of oil reserves together with the crucial role of oil as a basic good that is used directly and indirectly in the production of every other good is the key to understanding that royalties do not reflect a presumed looming physical scarcity of the resource, but mostly political power and strategic policies of states.[17]

Therefore, it is to governments' energy strategies and to productive capacity constraints and costs (including royalties charged) that we must mainly refer if we want to understand the supply side of the oil market.[18] These productive capacities and associated costs, together with American geopolitical and energy security policies, generate different types of monopoly, absolute and differential rents for public and private resource owners and producers.

A key feature of the world oil market is the fact that in the lowest cost and highest reserve regions, which are mostly OPEC countries, not only are the reserve deposits of oil vastly overabundant but the cartel also tends to produce at much less than potential full capacity.

Traditionally, Saudi Arabia has played the role of 'swing producer', keeping a sizeable planned degree of spare capacity to smooth the adjustment of current supply and demand in the world oil market and enforcing OPEC member coordination.

Saudi Arabia's ultimate objective seems to have been an attempt to restrict the supply of oil from OPEC countries. This policy has, over a long period of time, allowed market oil prices not to fall for too long below an informal and tacitly acknowledged floor: a price high enough to cover the production costs of the US and Canadian oil industries. This has been very important to ensure the long-run survival and profitability of the huge (and politically powerful) but relatively high-cost American oil industry. This fundamental aspect of American strategic energy security policy is based on the special geopolitical relationship between Saudi Arabia and the USA. The informal floor for oil supply prices creates a peculiar kind of classical *monopoly rent* in the OPEC countries which is then shared between OPEC members and the big multinational oil companies according to the (sometimes shifting) bargaining power of the two groups. Thus, OPEC's royalties are determined as a share of such specific monopoly rents.[19]

Note that the American floor price of production also includes, besides the usual elements of production costs, the royalties received by the (generally private and numerous) owners of American oil resources. These royalties are an *absolute rent* determined by the relative bargaining power of the owners in relation to the extraction industry, a bargain that is directly affected by various aspects of the American government's overall energy policy.[20]

Note that when market prices for oil are oscillating around this American floor price, production of oil in other countries and regions where the costs of production (including politically determined royalty rates, as in most cases the ownership of the subsoil resources is public) are lower than in the USA, albeit higher than those of the OPEC members, does generate classical *differential rents* for these countries.

But the American price of production is simply the oil production floor price. When world demand for oil increases sufficiently, beyond OPEC output and the productive capacity of American and other intermediate-cost regions, production becomes viable in regions of the world where costs (including absolute rent taken as state royalties) are much higher. And the actual price of oil production in these high demand conditions is given by the cost of the productive capacity that has to be activated to meet world demand.[21] This higher price of production generates further differential rents for all other lower-cost regions, even if their costs are above those that set the American floor price. Just to give a curious example, in 2008, when the demand for oil was growing fast and the market price for oil reached record levels, there were reports that in South Africa, extremely costly and highly polluting coal was being used to produce synthetic oil, something that seems to have been last done in Germany towards the end of World War II.

Equipped with this view of the structure of the world oil market, we can now turn to the increase in market prices during the last decade. The dollar market price for oil began to rise from the record low prices of 1999. Those market prices were initially substantially below the floor American price of production described above and seriously threatened both OPEC rents and royalties and the viability of the American oil industry. Then some members of OPEC, in particular Venezuela and Saudi Arabia, made an effort to coordinate all OPEC members and restrict both current output and investment in new capacity, in order to reduce the massive unplanned spare capacity both in OPEC and in other higher-cost producing countries.[22]

This tactic was quite successful and market prices began to recover rapidly. Later, as demand also began to increase more rapidly after 2003, with OPEC skillfully managing not to increase its production in line with it,[23] average market prices began to rise and most of the new production to meet rising world demand had to come from regions with higher costs for technological, geological, environmental or regulatory reasons, such as the tar sands of Canada, which nowadays seem to be setting the American floor price of production. Over time, exports to meet world demand even came from regions with very high production costs, such as offshore oil from Brazil.

This process of rising dollar prices of production and rents has been generally misunderstood as being either the consequence of the always predicted, and always missed, impending world peak of total oil production or, more plausibly, as an indication that 'all the cheap stuff has gone' with some analysts expressing doubts as to the true level of remaining reserves even in the OPEC countries.[24]

In many non-OPEC countries a very important element of the rising price of oil production seems to have been a substantial revival of 'natural resource nationalism' as a large number of developing countries, in most regions and even in Africa, seized the opportunity of rising market prices to renegotiate contracts with international private oil companies on more favourable terms. This process has increased state control of oil reserves and substantially raised the royalty rates and thus the absolute rent component of the price of production of oil. This movement stands in marked contrast to what had happened in the 1990s, where the subservient attitude of the state in oil producing countries was the international norm, as exemplified, for instance, by the contrast between contracts made by the Russian state under Presidents Yeltsin and Putin, or by the changing attitudes of Latin American and African governments in their relations with multinational oil corporations.[25]

8.4.2 Metals

Unlike oil but similarly to the case of coal, the ratio of current deposits to current production has decreased for many metals over the 2000s due to the rapid growth of world demand. Indeed, ratios of reserves to current production decreased between 2000 and 2007 for products such as bauxite, iron ore, nickel, tin and zinc, increasing somewhat for copper and lead (World Bank, 2009).

Note, however, that the mineral deposits described in these indices are those in existing mining areas, not the total fixed and finite physical availability of the metal in the earth's crust. In fact, there is not much interest in searching for new mining areas when existing ones have deposits that can last many decades at current rates of production. For instance, while the ratio of iron reserves to current production has fallen from 132 years' equivalent production to 79 since 2007, the amount of iron ore actually available on earth in the mid-2000s was estimated at about 120 million years' equivalent production, and about 2.5 billion years' for copper.[26] Thus, what really matters on the supply side is the existing productive capacity and the extraction costs. As in the case of oil, there is a long lead time for increases in mining productive capacity (some estimate more than 5 years on average). An unexpected change in demand may therefore leave the mining industry as a whole with large amounts of unplanned spare capacity, which can often lead to the complete shutdown of production in the higher-cost mines when demand reduces. Conversely, if demand accelerates unexpectedly, market prices rise to the point that it becomes viable to operate mines at high or rising extraction costs for long periods of time, generating differential rents for the producers with lower costs. Over longer periods of time this tendency towards increasing costs may be, and has historically been, checked by major improvements in extraction technology and technical progress in general. Also, similarly to oil, mining deposits often pay royalties as an institutionally and conventionally determined absolute rent for the private or more often state owners of subsoil rights.

Let us now turn to the rapid increase in metal dollar prices in the 2000s. In spite of the rapid rise in demand since the mid-1990s, driven by the acceleration of Chinese rates of infrastructure investment, and also of the increase in energy costs driven by the recovering oil price, dollar prices of metals only started to grow rapidly after 2003, and even more so than did energy prices after 2006.

The explanation for the initial period of very low dollar prices seems to be connected with the very high spare capacity for low-cost producers that came as the metals intensity of world demand fell from the mid-1980s to the mid-1990s. Over time, as demand continued to grow rapidly, spare capacity fell and at the same time new capacity has been growing slowly due to the long lead times and the further massive acceleration of demand after 2003. This process made the operation of higher-cost producers and increased differential rents viable. The market situation strengthened the bargaining power of the state in many developing countries, especially in relation to multinational private firms. This contributed to the marked revival of 'natural resource nationalism' that has been observed in mineral-producing countries in South America and Africa, and in Russia and parts of Asia. These governments have successfully increased royalties and thus the absolute rent component of the price of production of many metals.

Therefore, among the main causes of the boom in dollar metal prices after 2003 are rising energy costs, productive capacity constraints for lower-cost producers, rising extraction costs and 'natural resource nationalism' (Gopinath, 2011; Radetzki, 2012b).

8.4.3 Food prices

Dollar food prices increased much less and much later than oil or metal prices. Agricultural supply can respond to increases in demand much quicker than it is the case for minerals. On the other hand, due to its dependence on weather conditions, food output can change erratically over short periods. Thus, in the case of food, not only has demand grown relatively slowly but there has also been no clear sustained need to produce under increasing costs. Of course, the quality and quantity of specific types of land may be scarce and differential rents are generated on land with higher costs, but, contrary to what many believe, there is no inherent tendency towards decreasing returns in this sector.[27] Technical progress also appears in general to have been faster in food production than in mineral or oil extraction. For all these reasons agriculture price trends are much more clearly dominated by events on the supply side.

During the recent boom, current supply conditions as expressed by the stock-to-use ratios remained fairly stable (with the exception of wheat where the ratio fell), showing that production in general grew in line with demand.[28]

Agricultural dollar prices seem to have increased, first because of the effects of rapidly rising oil prices on energy and fertilizer costs, but also, for specific crops, because of the energy policies of the USA and the European Union.

The sudden very rapid increase in the demand for biofuels in these regions since 2006 seems to have strengthened the link between the oil price and some agricultural prices. It created an *opportunity cost* such that when the price of food that can be used as biofuel falls below a certain level determined in relation to the oil price, it is more profitable to divert the crop to meet the large and rising demand for biofuels rather than allow the market price to fall further.[29] Given this link, rising oil prices increase the floor below which the prices of these crops do not fall.[30]

8.5 General costs: the role of exchange rates

8.5.1 Decoupling growth under financial integration under the floating dollar standard

While the devaluation of the dollar relative to the key currencies of advanced economies does not seem to have played a major role in the rise of dollar commodity prices, its devaluation *relative to the currencies* *of the set of commodity-exporting countries* may have been an important element in the rise of dollar commodity prices. The revaluation of these currencies relative to the dollar increases the unit costs of production of all types of commodities, measured in dollars[31] (particularly where labour costs are more significant).[32] But in order to understand this effect we must discuss some key aspects of the balance of payments position of commodity-exporting countries during the 2000s.

After the series of balance of payments crises of the late 1990s and early 2000s many developing (or 'emerging') economies, including both industrial exporters and a large number of commodity-exporting countries, adopted a deliberate policy of reducing their external vulnerability. These countries made a huge effort to pay back their foreign debt (both private and official) and accumulate foreign exchange reserves, and many set up large sovereign stabilisation funds. Most also adopted heavily-managed floating exchange rate regimes to mitigate speculative pressures. When the world economy, international capital flows and commodity export volumes and dollar prices started growing rapidly after 2003, with international trade expansion fuelled largely by the rapid growth of both international trade and internal markets in the major industrial-exporting developing economies in Asia, these changes in macroeconomic policies allowed many of the commodity-exporting countries to grow without incurring major current account deficits and external debts. This, together with improved management of exchange rates and short-term foreign capital inflows, led to a drastic reduction in interest rate spreads for commodity-exporting countries and thus a major, and unprecedented, improvement in their balance of payments position. With a marked decrease in external vulnerability of these economies, no serious currency crises have since arisen in the commodity periphery.

This gave many developing countries (including in Africa and Latin America) scope to implement anti-cyclical macroeconomic policies and, more importantly, led to the growth of both external and internal markets, investment and productive capacities.

Led by China, the so-called South–South trade grew very rapidly in most regions. This led to a partial decoupling of the *trend* rates of growth of this large group of developing countries in the 2000s from the (slow) rate of growth of advanced countries, though trade and financial integration has made the cyclical fluctuations around these higher trends correlate very strongly with the fate of the world economy as a whole and hence with the cyclical fluctuations of advanced capitalist countries.[33] This decoupling of growth rates was anticipated in part by Lewis. In his Nobel lecture (published as Lewis, 1980) Lewis argued that it would be possible for developing countries to grow fast, even in the light of a slowdown in the growth of advanced capitalist countries, if the South–South trade among developing countries grew fast enough. This required that larger countries should rapidly expand both their internal markets and imports, acting as growth locomotives.

However, as correctly pointed out by Akyüz (2012), this decoupling of trend is only partial, because in spite of the improved policies and management of capital inflows the growth of developing countries in the 2000s still depended very much on these massive capital flows from the advanced capitalist economies. There has been a decoupling of the trend of growth[34] but certainly not a financial decoupling (see also Yeyati and Williams, 2012).

Indeed, this marked change in the international economic order has been very dependent on one particular element of continuity within it: the floating dollar standard. The fact that the US dollar remains the dominant international currency means that it is still the American central bank that sets the basic international interest rate and that American monetary policy and developments in the US (and to a much smaller extent European) financial markets play a key role in the dynamics of the pricing and quantities of international capital flows towards developing countries.[35]

Notwithstanding the much improved management of these international capital inflows by developing countries in general and by many important commodity-exporting countries in particular,[36] the relatively low public and private rates of interest in the USA were absolutely crucial in creating favourable conditions for the large amounts of capital flows of all kinds directed towards the developing world, and the experience of the 2000s cannot be understood without taking into account these elements that are, of course, completely beyond the control of the developing countries.

8.5.2 The devaluation of the dollar relative to commodity currencies

While the macroeconomic policies of commodity-exporting countries have (with a few exceptions) been successful in spurring growth while preventing large current account deficits, and the massive capital inflows have allowed the continuous build-up of foreign exchange reserves, what really matters to us here is that a tendency towards the appreciation of the real exchange rate of these countries as a whole gradually set in. In some countries it came as a gradual trend towards nominal appreciation which was accommodated in order to control domestic inflation. In a few money wages tended to grow rapidly relative to productivity due to much better conditions in the labour market and more nationalist or progressive governments (such as in Russia, Argentina and Venezuela). Some other countries had some combination of both tendencies. The consequence of this general tendency towards a real revaluation of the national currencies of commodity exporters relative to the US dollar has been, since 2003, the relatively rapid increase in the labour unit costs in dollars of most commodity-exporting countries. Thus, the revaluation of the 'commodity currencies' relative to the dollar has been another of the important neglected elements affecting the dollar cost of production and dollar prices of most commodities (particularly food but also minerals) during the recent boom.[37]

The importance of the balance of payments position (and hence of the exchange rates) of commodity-exporting countries in general as a main determinant of the terms of trade of the 'periphery' in earlier cycles has been explicitly emphasised by Patnaik (2002)[38] and by Ginzburg and Simonazzi (2004), where the connection between exchange rate devaluations and falling real wages in commodity-exporting countries is made clear. It is also most likely that the strong correlation between the growth of world demand and relative non-oil commodity prices over long periods found by Erten and Ocampo (2012) may be reflecting the fact (stressed by Ginzburg and Simonazzi) that in periods of rapid world growth, because of both trade and capital flows, the nominal and real exchange rate of many of the periphery countries tends to appreciate and periods of slow growth and crisis are marked by a series of devaluations in the commodity-exporting periphery. It is the peculiarity of the cycle of the 2000s that during the global crisis such devaluations were sharp but quickly reversed since the balance of payments position of developing commodity-exporting countries has improved drastically relative to earlier commodity price cycles.

Note that in this interpretation, although financial speculation is not seen as a direct determinant of the trend of commodity prices, international financial developments matter a lot, since the international capital inflows and their management by commodity-exporting countries are both seen as crucial aspects of the improved balance of payments position of these countries, which ultimately explains the course of their exchange rate.

8.6 The change in relative commodity prices: the 'China cost effect'

The large increases in nominal dollar prices we have discussed eventually transformed themselves into large increases in relative prices, because the international prices in dollars of non-commodity tradable industrial goods grew at modest rates and domestic inflation rates (which depend on prices of non-tradable services) were kept low in most countries.

The fact that in spite of the commodity price boom inflation did not accelerate in advanced countries allowed the boom to last much longer than had been the case in earlier cycles. On the demand side, the moderate inflation prevented governments from thinking they had to pursue restrictive policies, which if adopted would have cut back world demand for commodities and brought commodity prices down. On the supply side, the absence of persistent inflation prevented drastic increases in the prices of other goods and services as a reaction to rapidly rising dollar commodity prices, which if it had occurred sufficiently would have cancelled out the increase in *relative* commodity prices.

The ultimate reason for this low inflationary impact of the commodity boom seems to be the unusually low bargaining power of industrial workers in most advanced economies. The strong competition from low dollar labour cost industrial-exporting developing countries such as China appears to have been one of many key elements that have weakened the bargaining power of the workers in advanced economies.

We can see, then, that although the 'China effect' on demand has been greatly overestimated, another type of 'China effect', this time in terms of costs, has played an important role in the increase in the relative price of commodities that has not generally been recognised.

The low level and slow growth of dollar unit labour costs for more sophisticated manufactured goods exported from China, Mexico and many other industrialising developing countries seem to have ultimately played an important role in sustaining the great change in the relative prices of commodities. This has been due to a peculiar combination of the relatively low growth of money wages, very rapid growth in labour productivity and a deliberate attempt by governments to prevent or at least minimise the appreciation of the currencies of industrial-exporting developing countries relative to the US dollar (which in fact generally revalued much less than the currencies of most commodity-exporting countries).

8.7 Concluding remarks: Singer-Prebisch, Singer 2 and Lewis 1 and 2

It is well known that Singer (1950) and Prebisch (1949) attributed the long-run decline in relative commodity prices to the fact that real unit labour costs in the central industrial countries tended to remain fairly stable, even if the productivity rate of the manufactured goods exported by the centre was growing much faster than that of the commodities produced by the periphery. This, they stated, reflected the strong bargaining position of the workers in advanced economies. This bargaining power made both money and real wages grow more or less in line with productivity. In the periphery with its low productivity in food and virtually unlimited supply of labour, real unit labour costs of the exported commodities tended to fall as real wages in the periphery did not grow in line with the growth of productivity.

Lewis (1978) also argued that the terms of trade between the centre and the periphery depended on the levels of the relative real wage in these two regions and not really on the type of product that was exported.

As in the 1960s and 1970s, many developing countries began to industrialise and to produce and even export simpler industrial goods. Singer (1998) noted that this had not changed the terms of trade in favour of the periphery. He then formulated what he called the Singer 2 hypothesis, according to which the key point was that the periphery continued to export unsophisticated goods of low unit value and low technological content, implying a 'commodification' of those industrial goods.

However, since the 1990s a number of low-wage developing countries, especially in Asia, have started to export far more sophisticated industrial goods and even some services (as shown by the information technology services exports of India). Contrary to the models of Singer-Prebisch, Lewis and Singer 2, these developing countries now export many (but not all) of the same products as the advanced central countries. Thus, neither the Singer-Prebisch nor the Lewis nor the Singer 2 hypotheses appear to be an entirely valid explanation of the recent trend of relative commodity prices. In fact Lewis himself anticipated in 1976 that if low-wage countries, in addition to commodities, started also to export large quantities of sophisticated industrial goods, then the level of terms of trade would turn against the latter. This possible Lewis 2 hypothesis seems to have become relevant in recent years; since productivity has tended to grow much faster in industrial goods (especially in electronics) than in commodities and nowadays neither in the old countries of the centre nor in the new industrial-exporting periphery are real wages growing in line with productivity, the real unit labour costs of industrial goods would tend to fall relative to those of commodities.

In our view, what he have called the China cost effect shows that Lewis's intuition was right and that for the moment that secular declining trend of terms of trade has been reversed. For now real wages are not growing in line with productivity in the central countries, partly because of the fierce competition from low-wage workers in the periphery that now exports sophisticated industrial goods. Moreover, as we saw above, this movement towards higher relative prices for commodities is intensified by the tendency for the exchange rates of commodity-exporting countries to appreciate relative both to the dollar and to the currencies of the industrial-exporting developing countries. This seems to be leading to more rapid increases in real and dollar wages in the commodity-exporting countries and stagnant real wages in the advanced industrial economies, the latter constrained not only by the lower wages but also by the much more rapidly growing productivity of developing industrial-exporting countries such as China (and others mainly in East Asia).[39]

Notes

1. For reasons of space and scope very little will be said in this chapter about this crucial aspect of the international economic order (which in practice means that the US economy is not subjected to balance of payments constraints) apart from brief remarks in Section 8.5 below. For a more detailed discussion of the 'floating dollar standard' see Serrano (2003, 2004, 2008) and Feldman (2009).
2. On these changes see Serrano (2008), Akyüz (2012) and Frenkel and Rapetti (2011). Note that for many authors such as Frenkel and Rapetti (20011) the improved balance of payments situation is considered good because it presumably helps each developing country seen in isolation to avoid a tendency towards the overvaluation of the exchange rate of their currencies. Here (see section 8.5 below) the issue is that it allows commodity-exporting countries as a whole to avoid the opposite tendency towards competitive uncoordinated exchange rate devaluations that depress commodity prices and terms of trade.
3. A more detailed analysis of these new aspects of the international economic order can be found in Freitas, Medeiros and Serrano (2013).
4. Having clarified its precise meaning in this context the word commodities will not be written in quotes in the rest of the chapter.
5. For a classical surplus 'cost of production' interpretation of the previous period of rising commodity prices in the 1970s see Sylos-Labini (1982).
6. Most of the data in this and in the next section come from Ferreira (2012).
7. For data that confirm the limited impact of recent commodity price increases in the trend of US inflation (in contrast with what happened in the 1970s) see Tootell (2011). On the causes of the low bargaining power of workers in advanced capitalist countries see Pivetti (2011).
8. In fact, perhaps the only relevant, though indirect, effect of the nominal devaluation of the dollar relative to the currencies of the other advanced economies is that it moderates the nominal and real increase of oil and other commodities in terms of the domestic currencies of these countries.

This reduces the magnitude of the initial domestic inflationary shock of the dollar commodity price boom and perhaps helps the boom to continue, by preventing more drastic action of the central banks to fight domestic inflation. Such measures could end up slowing down the growth of aggregate demand in these rich economies, thereby cutting the world demand for commodities, which would lower current dollar commodity prices. This possibility is raised in UNCTAD (2008). This latter effect does not seem to have been decisive, for even in the USA, due to the weak bargaining position of workers, the fact that the economy faced the full impact of the rising dollar commodity and specially oil prices did not generate a price-wage-price spiral that could increase core inflation and lead to contractionary measures by the central bank (see previous footnote).

9. Jeffrey Frankel has been the most emphatic advocate of this view and has incorrectly been predicting high inflation in rich countries for many years (Frankel, 2006). In Frankel and Rose (2009) we find the admission that the lack of a reliable significant econometric effect of low interest rates of commodity prices is a 'disappointment'.
10. For the former view see Pollin and Heintz (2011) and the latter see Wray (2008).
11. This point was raised by Krugman (2008).
12. See Irwin and Sanders (2011).
13. Note that classical economists from Smith to modern Sraffians were quite aware of the importance of all this and other causes of irregular oscillations of market prices in the very short run and thus did not think it was either possible or useful to postulate either demand or supply as regular functions of prices, using instead only the twin concepts of effectual demand and natural price to analyse the long-run trends of prices and quantities (Garegnani, 1983).
14. Note that, on the other hand, for obvious material reasons there is really no organised market for *immediate* physical delivery of oil. What is called the spot market for oil is actually a 'relatively near future' market.
15. As due to the possibility of arbitrage expected spot prices affect strongly both the current spot and futures market prices, persistent diferences between current spot and future market prices can be traced to variable margins connected to perceived financial costs and risks and convenience yields attributed to assured delivery by market participants. See Kaldor (1939).
16. See Serrano (2008).
17. Note that, contrary to appearances, the fact that oil is not scarce should make environmental concerns more rather than less serious. Given that the use of oil causes many forms of undesirable pollution, its high energy efficiency and lack of scarcity mean that there is no reason why the serious negative externalities coming from its excessive use would be somehow reflected in its market price.
18. Ravagnani (2008) provides a detailed critique of attempts to use the notion of a given, known and scarce finite stock of resources and the associated 'Hotelling rule' for the determination of the price of production of oil and other non-renewable natural resources.
19. See Serrano (2004). Adelman has for decades been making the crucial point that Middle East oil is not scarce (Adelman, 2004; Adelman and Watkins, 2008). Roncaglia (2003) provides a very interesting analysis of the

relationship between OPEC members, big oil companies and the USA but unfortunately his argument that this could be seen as a case of a 'trilateral oligopoly' obscures certain issues. First, there is no such thing as oligopoly in classical competition theory, where it is the degree of capital mobility and not the number of firms that is the relevant measure of the degree of competition. Second, and more importantly, the idea of a trilateral oligopoly leads to a misunderstanding of the superficially puzzling behaviour of Saudi Arabia which reflects their special political relationship with the USA and is therefore not the result of an autonomous notion of Saudi national interest. Routledge (2003) drew attention to the crucial issue of the coordination between OPEC and the USA in order to prevent the price of oil from falling below the American 'floor supply price'.

20. As shown by Piccioni and Ravagnani (2002) and Ravagnani (2008). Note that Fratini (2009) draws attention to the fact that absolute rent, which should be taken as given as a share of the gross output or the value of the product is not the same thing as a monopoly rent (which would appear to be the difference between an exogenously given price of product and its cost of production). For an excellent discussion of the different types of rent in both the old and the modern classical surplus approach see Fratini (2008).
21. Both Schefold (2001) and Kurz and Salvadori (2009) draw attention to the importance of productive capacity constraints on the extraction of minerals being one important precondition to use differential rent theory in the case of non-renewable natural resources.
22. See Routledge (2003).
23. See Davidson (2008).
24. See Lynch (2006) and Davidson (2008) for a critique of these views.
25. On the fall and rise of natural resource nationalism, see Medeiros (2011). On the important role of the increase in the nationalisation of most of the world's oil reserves during the 2000s for the oil price increase, see Radetzki (2012a).
26. See Tilton (2009). Tilton adds for the sake of comparison that our solar system is about 5 billion years old and will not last more than 10 billion years. Again, as in the case of oil, this abundance of the resource only makes it even more unlikely that the negative environmental externalities associated with mining, which can be substantial, would ever be reflected in the market prices of minerals.
27. Pasinetti (1999) explains very well how, even in Ricardo, the assumption of increasing costs in corn production was an historically contingent empirical assumption instead of the result of a general or natural law.
28. See Baffes and Haniotis (2010).
29. Fratini (2008) mentions that such cases should be taken as further evidence of the existence of absolute rents. In the end he may well be right but we think it seems more appropriate to take these opportunity costs as yielding a type of differential rent, reserving absolute rent only to those rents that are determined directly by bargaining and political and institutional forces (and not by differences in cost or prices).
30. For empirical evidence on the boom in agricultural prices, see Baffes and Haniotis (2010) and World Bank (2009, 2012).
31. Radetzki (2008) argues that the direct effect of a devaluation of the dollar in general would be to increase the demand from commodity consumers (by

lowering the price of commodities in local currency) and increase the dollar cost for commodity producers. This observation may help us to explain why the devaluation of the dollar relative to other advanced countries' currencies had little effect on commodity prices as their weight on world demand for commodities is not that large and also why the devaluation of the dollar relative to the currencies of the commodity-producing countries has had important and persistent effects on the dollar cost of production and prices.

32. For econometric evidence of this impact of exchange rates on food prices see World Bank (2012).
33. For data on these trend and cyclical relations see Yeyati and Williams (2012) and Akyüz (2012).
34. Akyüz (2010, 2012) argues that China's growth is still very heavily dependent on exports. For a different view, which puts more emphasis on the Chinese internal market than Akyüz does, see Medeiros (2006), Anderson (2007) and Kotz and Zhu (2010).
35. See Rozada and Yeyati (2011) in particular for evidence on the overwhelming importance of American long interest rates on public debt and the yield of American high-risk domestic assets on the spreads charged on 'emerging' markets.
36. As shown by Frenkel and Rapetti (2011), due to their improved balance of payments situation emerging market spreads have fallen for the first time ever below the spread of highrisk American assets around 2004 and have remained so ever since, including in the turbulent period after the global crisis of 2008.
37. In the recent boom strong evidence that the nominal revaluation of commodity currencies (currencies of a few key commodity exporters) actually preceded increases in dollar price of commodities can be found in Chen, Rogoff and Rossi (2010) (though interpreted not as cause of the boom but as a rational expectation of future commodity prices). In fact, as it was pointed out to me by Fabian Amico it is not really important if the dollar commodity price rises begin to happen before or after the appreciation of the currencies. What matters for explaining the trend of prices is that the appreciation of the commodity currencies persists and in particular is not more than cancelled out by the usual series of large subsequent devaluations.
38. Patnaik (2002) however interprets and criticises Prebisch's views on the terms of trade as if for him those were determined in the long run by demand. As will be seen in the next section, Prebisch's theory is interpreted differently here.
39. See ILO (2013), for data on trends of real wage growth in the 2000s that seem to be consistent with the above interpretation. Real wages grew fast in high-productivity growth developing industrial exporting countries, very fast in some commodity-exporting countries and rather slowly in the advanced industrial economies.

References

Adelman, M. (2004) 'The real oil problem', *Regulation*, 27(1) (spring): 16–21.

Adelman, M. A. and Watkins, G. C. (2008) 'Reserve prices and mineral resource theory', *Energy Journal*, 29, Special Issue: 1–16.

Akyüz, Y. (2010) 'Export Dependence and Sustainability of Growth in China and The East Asian Production Network', *South Centre Research Paper*, RP 27, April.
Akyüz, Y. (2012) 'The Staggering Rise of the South?', *South Centre Research Paper* 44, March.
Anderson, J. (2007) 'Is China Export-Led?' *UBS Investment Research Asian Focus*, UBS, 27 September.
Baffes, J. and Haniotis, T. (2010) 'Placing the 2006/08 Commodity Price Boom into Perspective', *World Bank Policy Research Working Paper* 5371.
Chen, Y., Rogoff, K. and Rossi, B. (2010) 'Can exchange rates forecast commodity prices?', *Quarterly Journal of Economics*, 125(3), (August): 1145–94.
Davidson, P. (2008) 'Crude Oil Prices: "Market Fundamentals" or Speculation?', *Challenge*, 57(4), (July): 110–18.
Erten, B. and Ocampo, J. (2012) 'Super-cycles of Commodity Prices since the Mid-nineteenth Century', *UN DESA Working Paper* 109, February.
Feldman, G. (2009) 'Global imbalances and the world economic crisis: a critical appraisal', http://www.boeckler.de/pdf/v_2009_10_30_feldman.pdf, accessed 25 January 2013.
Ferreira, S. (2012) 'Condições Externas e a Dinâmica da Inflação no Brasil 1994–2010: Uma Interpretação Estrutural', PhD thesis, Instituto de Economia, UFRJ.
Frankel, J. (2006) 'The Effect of Monetary Policy on Real Commodity Prices', in J. Campbell (ed.), *Asset Prices and Monetary Policy* (Chicago: University of Chicago Press).
Frankel, J. and Rose, A. (2010) 'Determinants of Agricultural and Mineral Commodity Prices', HKS faculty research working papers, RWP10-038, Harvard.
Fratini, S. M. (2008) 'Alcune osservazioni sulla teoria classica della rendita', *QA – Rivista dell'Associazione Rossi-Doria*, March.
Fratini, S. M. (2009) 'La rendita assoluta di Marx e le equazioni di prezzo di Sraffa', Department of Economics Working Paper 0105, Roma Tre University.
Freitas, F., Medeiros, C. and Serrano, F. (2013) 'The Decoupling of the Trend of Economic Growth of the Developing Countries in the 2000s', Paper for the First World Keynes Conference, Izmir University of Economics, Izmir/Turkey 26–29 June 2013.
Frenkel, R. and Rapetti, M. (2011) 'Fragilidad externa o desindustrialización: Cuál es la principal amenaza para América Latina en la próxima década?', *Serie macroeconomía del desarrollo,* 116, CEPAL, November.
Garegnani, P. (1983) 'The classical theory of wages and the role of demand schedules in the determination of relative prices', *American Economic Review*, 73(2), (May): 309–13.
Garegnani, P. (1988) 'Actual and normal magnitudes: a comment on Asimakopulos', *Political Economy*, 4(2): 251–58.
Ginzburg, A. and Simonazzi, A. (2004) 'Disinflation in industrialized countries, foreign debt cycles, and the costs of stability', *Quaderni di Ricerca Centro Sraffa*, 4, Uniroma 3.
Gopinath, D. (2011) 'Supply responses will be delayed for hard commodities', *Trusted Sources Consulting*, September.
ILO (2013) *Global Wage Report 2012/13: Wages and Equitable Growth* (ILO).
Irwin, S. and Sanders, D. (2011) 'Index funds, financialization, and commodity futures markets', *Applied Economic Perspectives and Policy*, 33: 1–31.

Jenkins, R. (2011) 'El "efecto China" en los precios de los productos básicos y en el valor de las exportaciones de América Latina', *Revista CEPAL*, (April), 103: 77–93.

Kaldor, N. (1939) 'Speculation and Economic Stability', *Review of Economic Studies*, October.

Kotz, D. and Zhu, A. (2010) 'The Dependence of China's Economic Growth on Exports and Investment', *Review of Radical Political Economics*, XX: n.10.

Krugman, P. (2008) 'More on oil and speculation', http://krugman.blogs.nytimes.com/2008/05/13/more-on-oil-and-speculation/

Kurz, H. D. and Salvadori, N. (2009) 'On the Theory of Exhaustible Resources: Ricardo vs. Hotelling', ISER Discussion Paper, Institute of Social and Economic Research, Osaka University, 0756.

Lewis, A. (1976) 'The Diffusion of Development', in T. Wilson (ed.), *The Market and the State* (Oxford: Oxford University Press).

Lewis, A. (1978) *Evolution of the International Economic Order* (Princeton: Princeton University Press).

Lewis, A. (1980) 'The slowing down of the engine of growth', *American Economic Review*, 70(4), (September): 555–64.

Lynch, M. C. (2006) 'Crop circles in the desert: the strange controversy over Saudi oil production', http://masterresource.org/wp-content/uploads/2008/12/crop-circles.pdf

Medeiros, C. (2006) 'A China como um duplo pólo na economia mundial e a recentralização Asiática', *Revista de Economia Politica*, v. 26.

Medeiros, C. (2011) 'The political economy of the rise and decline of developmental states', *Panoeconomicus*, 58(1): 43–56.

Pasinetti, L. L. (1999) 'Economic Theory and Technical Progress', Lecture at the Royal Economic Society Annual Conference, Nottingham.

Patnaik, P. (2002) 'Globalization of capital and terms of trade movements', http://www.networkideas.org/featart/mar2002/globalisation_1.pdf

Piccioni, M. and Ravagnani, F. (2002) 'Absolute rent and the "normal price" of exhaustible resources', Centro di ricerche e documentazione 'Piero Sraffa', Quaderno di Ricerca 2, Roma Tre University.

Pivetti, M. (2011) 'On Advanced Capitalism and the Determinants of the Change in Income Distribution', Paper presented at the *Congreso Internacional Crisis de la teoria economica y politicas alternativas ante la crisis global*, 7–11 November, (Mexico City: UNAM).

Pollin, R. and Heintz, J. (2011) 'How Wall Street speculation is driving up gasoline prices today', *PERI Research Brief*, June.

Prebisch, R. (1949) 'O desenvolvimento econômico da América Latina e seus principais problemas', *Revista Brasileira de Economia*, 3(3), (July): 47–111.

Radetzki, M. (2008) *A Handbook of Primary Commodities in the Global Economy* (Cambridge: Cambridge University Press).

Radetzki, M. (2010) 'Peak Oil and other threatening peaks – Chimeras without substance', *Energy Policy*, 38(11), (November): 6566–9.

Radetzki, M. (2012a) 'Politics – not OPEC interventions – explain oil's extraordinary price history', *Energy Policy*, 46, (April): 382–5.

Radetzki, M. (2012b) 'The perseverance of the ongoing metal and mineral boom', M*ineral Economics*, 46, (August): 382–5.

Ravagnani, F. (2008) 'Classical theory and exhaustible natural resources: notes on the current debate', *Review of Political Economy*, 20(1): 79–93.

Roncaglia, A. (2003) 'Energy and market power: an alternative approach to the economics of oil', *Journal of Post Keynesian Economics*, 25(4), (Summer): 641–59.
Rozada, M. and Yeyati, E. (2011) 'Risk Appetite and Emerging Market Spreads', *Business School Working Papers*, 2011-03, Universidad Torcuato Di Tella.
Rutledge, I. (2003) 'Profitability and supply price in the US domestic oil industry: implications for the political economy of oil in the twenty-first century', *Cambridge Journal of Economics*, 27(1): 1–23.
Schefold, B. (2001) 'Critique of the corn-*guano* model', *Metroeconomica*, 52(3), (August): 316–28.
Serrano, F. (2003) 'From "static" gold to the floating dollar', *Contributions to Political Economy*, 22, (November): 87–102.
Serrano, F. (2004) 'Power relations and American macroeconomic policy, from Bretton Woods to the floating dollar standard', *http://www.networkideas.org/feathm/sep2006/pdf/power_relations.pdf*
Serrano, F. (2008) 'A economia americana, o padrão dollar flexível e a expansão mundial nos anos 2000', in J. Fiori, C. Medeiros and F. Serrano (2008), *O Mito do Colapso do Poder Americano* (Rio de Janeiro e São Paulo: Editora Record).
Singer, H. (1950) 'The Distribution of Gains between Investing and Borrowing Countries', *American Economic Review*. Papers and Proceedings, 40, May.
Singer, H. (1998) 'Beyond terms of trade: convergence/divergence and creative/uncreative destruction', *Zagreb International Review of Economics and Business*, 1(1), (May): 13ff.
Sylos-Labini, P. (1982) 'Rigid prices, flexible prices and inflation', *Banca Nazionale del Lavoro Quarterly Review*, 140: 37–68.
Tilton, J. (2009) 'Is mineral depletion a threat to sustainable mining?', http://inside.mines.edu/UserFiles/File/economicsBusiness/Tilton/Sustainable_Mining_Paper.pdf
Tootell, G. (2011) 'Do Commodity Price Spikes Cause Long-Term Inflation?' *Public Policy Briefs*, Federal Reserve Bank of Boston, n. 11-1.
UNCTAD (2008) 'Trade and Development Report, 2008', UNCTAD.
World Bank (2009) *Global Economic Prospects 2009: Commodities at the Crossroads*, World Bank.
World Bank (2012) 'Global Commodity Market Outlook', *Global Economic Prospects January 2012*, World Bank.
Wray, L. R. (2008) 'The Commodities Market Bubble: Money Manager Capitalism and the Financialization of Commodities', *Public Policy Brief*, Levy Economics Institute, 96.
Yeyati, E. and Williams, T. (2012) 'Emerging economies in the 2000s: real decoupling and financial recoupling', *Policy Research Working Paper Series* 5961, World Bank.

9

The Political Economy of the Rise and Decline of Developmental States

Carlos Aguiar de Medeiros[*†]

9.1 Introduction

The spread of industries in several peripheral countries after the Second World War and the great divergence that has opened up between them since the 1980s has sparked widespread debate on economic development. Interpretations based on neoclassical and on institutional economics[1] are the major fields of historical explanations. Despite their wide differences on the determinants of economic growth, they share a common perspective on three basic aspects: first, the supposition that strategies of development are built on a set of government policies and on institutions that model private behaviours (of course, they disagree on which policies and institutions promote economic development); second, a 'methodological nationalism'[2] in which individual countries' performances are essentially explained by domestic factors. The third is a corollary of the two above perceptions and says that the state as a major initiator of positive change (in resource allocation, as in the heterodox reasoning, or in the creation of market institutions, as in orthodox thought) is responsible for the success or failure of growth strategies. For the mainstream school, the wrong policies of populist states play the dominant role, for the heterodox, they are those of liberal or neo-liberal states. For both, a meritocratic state is central to successful development

*Another version of this paper was published in *Panoeconomicus*, 2011, 1, 43–56. The author thanks CNPQ for financial support, and Franklin Serrano and Andrea Ginzburg for the comments on the original paper presented at the Sraffa Conference, Roma 2–4 December 2010, and an anonymous referee. Any errors are mine.
†Associate Professor at IE, Federal University of Rio de Janeiro.

strategies (to avoid rent-seeking cases according to neoclassical authors, to discipline large firms according to the institutionalists).[3]

Stemming from a methodological perspective based on classical political economy, on Latin American structuralism and on a Gramscian view of the state's formation,[4] this chapter takes a critical stance on these three basic aspects. To begin with it considers that modern explanations about developmental states and the role of institutions neglect the different challenges and circumstances created by initial conditions and how different economic and social structures influence institutions (a bias opposed to the classical structuralism of the United Nations Economic Commission for Latin America and the Caribbean (ECLAC), which neglected the autonomous role of institutions). The major challenge in explaining development strategies is to articulate the two dimensions. Second, it proposes that the international environment is not identical for all developing nations and is influenced by the hegemonic state's economic and political actions, creating different development opportunities for their industrialisation. Finally, it considers that the state (and its developmental historical variant)[5] may not be viewed as an agent in isolation from social classes and relations with other states, but rather, as a central institution where the dominant class or some of its sectors lead a coalition of power, building a hegemonic project compatible with a particular accumulation strategy.[6]

In addition to this introduction, the chapter has three main sections. In the next two sections there is an attempt to illustrate some of these issues using the analysis of national patterns of industrialisation and development strategies experienced by peripheral countries during both the highest convergence period and the greatest divergence phase. The third and final section explores how new challenges have been answered by developing nations and some attempts to rebuild a 'neo-developmental state'.

9.2 National development strategies in the golden age of the post-war and developmental states (DS)

The partial and limited spread of industrialisation in the post-war period (especially among industries and activities closer to the innovative processes, such as the capital goods sector) was a consequence of national strategies led by development-oriented states specifically geared to creating modern industry and its infrastructure as the main engines of economic growth.

Under US hegemony national development was basically an accumulation strategy and a hegemonic project of industrial national capital coordinated by national states favouring the formation of large industrial companies and their markets. These developmental strategies were followed in several countries and took different routes according to the size of the economy, the natural resource base, income distribution, geopolitical position and the political power underlying that strategy.

With distinct levels of success and with a much lower income per person, a few countries (in Latin America, especially Brazil and Mexico, in Asia, the Asian Tigers, especially Korea and Taiwan, the Southeast Asian countries, India and China) followed a path of industrialisation somewhat similar to what European countries had experienced in restructuring during the post-war period. The strategy was to transplant the key industries typical of the American manufacturing pattern – metal mechanics, automobiles and chemicals – and their consumption patterns, centred on durable consumption goods financed by credit. As Prebisch (2011 [1949]) remarked, the typical Keynesian post-war policies on the periphery would require structural change to offset the external constraints, and industrialisation would be the basis for a policy directed towards high growth and unemployment reduction.

Between 1950 and 1980, the steady increase in per capita income in those countries resulted from the increased pace of industrial output growth and the transfer of surplus labour in agriculture to urban activities led by industry and services. In countries where this shift was greater, as in Korea or Brazil, the growth rate was higher, and in countries where it was less intense, as in India, the rate of growth of income per capita and per employed person was lower.

Despite the diversity of initial conditions within those countries, they faced challenges caused by the large technological gap in relation to industrialised countries, the narrowness of domestic markets, the problems of coordination and financing of complementary blocks of investments in new sectors and the restrictions on balance of payments.[7]

The influence of industrial success in the Soviet Union gave major political legitimacy to long-term planning. But national industrialization strategies were not only distinguished by planning. In those most successful countries such as Brazil and Korea (and later in China, since Deng Xiaoping's reforms), strategies were the result of industrial policy and public investment. In some countries, like Korea, Taiwan or Mexico, the state has directly controlled the 'commanding heights' of the financial sector (Haggard *et al.*, 1993). Besides finance, in many countries industrial incentives were directly applied through state enterprises operating in strategic industrial activities and infrastructure.

Thus, regardless of the higher or lower share of exports in the composition of industries' final demand, the late industrialisation of the last century was led by the states. The conventional distinction between a strategy based on import substitution industrialisation (ISI) associated with the state leadership in countries like Brazil and Mexico, and an export-oriented industrialisation (EOI) associated with a pro-market strategy that would have prevailed in Korea, Taiwan or Thailand, does not, in fact, coincide with the historical evidence on industrialisation pursued in those countries.[8] All strategies originally included import substitution processes and selective opening and put greater or lesser emphasis on industrial exports according to different factors.

An essential part of national development strategies was the macroeconomic regime in which the exchange rate and the fiscal and monetary policy were subordinated to the objectives of industrial development.[9] Until the 1970s, external financing was scarce and the constraint on foreign currency imposed strict control on foreign exchange.

However, in spite of common strategies, these countries followed different patterns. Two aspects stand out: the levels of income distribution associated with the industrialisation process (inequality in Latin America was much higher than in Asia), and the share of industrial exports in countries' total exports (much higher in Asia). Besides these structural dimensions, one important difference was the role played by foreign capital, which was larger in Argentina, Brazil and Mexico than in Korea, Taiwan or India. As regards the state's power to influence the economy, one may observe that although the common base was a coalition between the military, technocratic planners, and the private industrial sector, this coalition was stronger in Korea or Taiwan than in other countries. Both countries built a 'cohesive capitalist' state (Kohli, 2004) by dislodging the landowners and enjoyed the strong support of the USA in their strategy of communist contention in Asia. It was also influenced by Japanese institutions. Except for Korea and Taiwan, the developmental state was more 'fragmented and multi-class' and the industrial sector had to establish political alliances with powerful landowners.

Different patterns

Unlike Western Europe, industrialisation in peripheral countries led by developmental states was not accompanied by social democratic coalitions aiming at the distribution of income and full employment.[10] The goal of greater equity was subordinated to growth and industrialisation goals. Income distribution was quite uneven according to the different social coalitions supported in the state. The pattern of income distribution was essentially influenced by the way land and the modernisation of agricultural production had evolved. In countries where the productivity in food production[11] was lower and internal structural heterogeneity[12] higher, peasant income was lower and the heavier was the weight of the traditional oligarchy on political power – such as in Brazil, India, the Philippines or Indonesia where the state was more 'fragmented and multi-class'. In these cases, industrialisation took place accompanied by social marginalisation and the exclusion of rural masses and suburban areas from modern consumption, leading to large income concentration. In countries where land reforms and simultaneous modernization of agriculture took place (as in Korea and Taiwan) internal structural heterogeneity and social polarization was less intense, and the state was more cohesive.

Similarly, export performance followed a distinct path.

In East Asia, import substitution was quickly followed (as early as the 1960s) by industrial exports (textiles and clothing in the early stages), generating a greater trade diversification and a consequent positive effect on the balance of payments. In Latin America, the diversification of exports happened much more slowly and less intensively. Several hypotheses explain this discrepancy. Excessive protectionism[13] or pessimism about Latin America's ability to export industrial products[14] would have prevailed in the region, in contrast to the clearly export-oriented strategies of Asian countries.

What distinguishes these approaches is the lack of connection between strategies and economic structure. Following Diamand (1978) in the case of Argentina, Mahon (1992) and more recently Bresser Pereira (2010), here we argue that among countries with highly competitive export industries based on natural resources, such as those of Latin America, there came to be an external heterogeneity or an 'imbalance in the production structure' between the productivity of the primary export sector and that of the industrial sector. This imbalance led to the formation of an uncompetitive exchange rate for industry, which contributed to the specialisation of the export sector. Industrial policies favoured domestic industry through preferential exchange rates and tariffs, but these policies achieved poor results in promoting industrial exports.[15]

In Latin America, especially in Argentina, the high share of exports-based natural resources exacerbated a distributive exchange-rate conflict between the primary exporters, the industrial sector, non-tradable support activities and the working class.

In Asia, for a small group of countries like Korea, Taiwan and certainly city-states like Hong Kong and Singapore, the scarcity of natural resources made the export of industrial products the obligatory route to industrialisation. Poor natural resource bases were favourable to a more balanced economic structure, making it possible to establish a real exchange rate more favourable to industry. In these 'invited countries' (Medeiros, 1997) industrial exports were directed to the USA, which gave them preferential access. In addition the export companies of these countries benefited from the road previously travelled by Japan which provided these latecomers with some trade networks and investment (Medeiros, 1997; Ginzburg and Simonazzi, 2003).

Thus, due to dissimilar structural and geopolitical circumstances, these strategies have evolved[16] to yield different macroeconomic and social results.

This fact had important consequence for the ease with which Asian countries reacted to the debt crisis of 1980, but it had already manifested itself in the pattern of indebted growth followed by Brazil, Mexico and Korea in the 1970s.[17]

9.3 Neo-liberalism and the crisis of the developmental nation states

The Reagan-Thatcher offensive against organised labour, the external debt crisis in the peripheral countries and the collapse of the USSR in 1991, at a time when a new technological revolution based on information and telecommunications was rising, led to significant changes in the international division of labour. Under unrestrained competition, finance and productive internationalisation greatly enlarged.

The 'Washington–Wall Street complex'[18] (and its leadership through the World Bank and the International Monetary Fund) established itself as the centre of political power and of the ideology. This brought about a new strategy of accumulation and a new hegemonic project increasing the dominance of capital, in general, and finance capital, in particular, over other factions and interests.[19] It also coincided with a new US trade offensive to open the hitherto regulated internal markets of the newly industrialised exporting nations.[20]

Among the industrialised nations, the large corporations, exposed to intense international competition, sought greater autonomy from the state, the workers and the chain of domestic suppliers demanding greater state support for the globalisation processes of production and finance in new spatial and regional arrangements. Transplanting labour-intensive activities to peripheral countries was intense, remaking the international division of labour. The firm's strategy of going global introduced a fracture between national capitalism and national capital with important repercussions for macroeconomic policies and political coalitions.

The crisis of the developmental state was similar to the end of the Keynesian national welfare state in industrialised countries. If the latter was associated with the end of the subordination of monetary and fiscal policy to full employment, the former crisis was connected with the end of the subordination of fiscal and monetary policy to industrial development.

In fact, as has happened with national Keynesianism, development strategies based on industry and on internal markets were abandoned in many countries and a new hegemonic project led by cosmopolitan capital was established. This brought about in many 'fragmented' states what can be considered (in a Gramscian expression) a capitalist passive revolution aimed at rebuilding a more cohesive state around big business and financial interests.

Despite the differences observed in time and space, the discontinuity in development strategy involved two major forces: financial openness and a big business revulsion against the developmental state.

Financial openness played an important role in the crisis and discontinuity of national developmental strategies (in both industrialised and, mainly, semi-industrialised nations) insofar as it exposed the economies to volatile capital inflows and dissolved the role of domestic credit as a mechanism for coordinating investments (Haggard and Maxfield, 1993). It was in the wake of the exchange-rate crises that Washington Consensus structural reforms were introduced on a huge scale. The institutional position of the central bank (with an exclusive focus on price stability) was strengthened due to international constraints that followed the external crisis, and the role of manufacturing interests and their institutions was diminished.

As a corollary to this macroeconomic and institutional change, there was a split – to the extent to which the nations opened their economies – between the interests of the large corporations and the national industrial strategies that had hitherto been the basis of national development. Cultivated and promoted by their developmental state, domestic businesses (challenged in their own markets) began to seek new opportunities and strategies for accumulation, especially through the formation of joint ventures with multinational corporations and through majority interests or participation in the business of privatisations. Such opportunities of going global demanded new functions and policies from the state (Medeiros, 2009).

The rebellion by large corporations against developmental states occurred everywhere.[21] It was generally accompanied by public opinion that identified developmentalism and industrial policies – such as those implemented by countries like Brazil, Korea or Indonesia – with political authoritarianism, with 'crony capitalism' and, in the case of Brazil, with income concentration. The political legitimacy of industry-based accumulation strategies and, consequently, the hegemony of this project were profoundly shaken.[22]

But a great divergence took place.

In Latin America the intense debt crisis of the 1980s brought about high inflation rates and a deep recession. This caused a structural crisis in prevailing state-led growth and created new coalitions of internal and external interests around the agenda of the Washington Consensus reforms that spread all over the region in the 1990s (Medeiros, 2008a).

In Asia, the external shock of the 1980s was not so disruptive. In large countries such as China or India, the debt ratio was too low to make any substantial negative impact (Hughes and Singh, 1991). In East and South-East Asian countries, thanks to better solvency ratios, the surge of Japanese investments and the clustering of production chains in the region (Medeiros, 1997), the majority of the economies (more or less open) experienced high growth, thereby preserving the bulk of institutions which had developed earlier. This clustering of success and collapses in space and time highlights the limitations of 'methodological nationalism' (Ocampo *et al.*, 2007; Medeiros, 1997). Only in the 1990s, but mainly after the 1997 crisis, did a strong offensive against developmental institutions occur.

Thus, there were different national answers to the liberalisation process. In Latin American countries, liberalisation took the form of a radical U-turn (Palma, 2010), from a very weak national position; in Korea and other Asian countries, liberalisation was adopted later and tied to a trajectory of high growth. Other Asian countries like China, Taiwan or India did not dismantle the main developmental institutions.

The degree and impact of these changes on national developmental strategies essentially depended on the extension and circumstances of the external crisis but the different ways in which they occurred in Asian and Latin American countries were conditioned by the production structure, the regional dynamics and the power and political cohesion of the nation states.

Different paths

Throughout the 1990s, it was possible to identify various reactions to the liberalisation and technological pressures. One common response to the new challenges was the pursuit of an 'integrationist' strategy (Amsden, 2001), or as Lall (2000) puts it, 'a passive strategy dependent on foreign direct investment (FDI)'. This was based on two pillars: on the micro side, this strategy was built by the formation of new private alliances and re-specialisation in activities with absolute cost advantages (whether in industrial commodity chains, as in Mexico, or in natural resources, as in most South American countries and in Russia throughout the 1990s).[23] On the macro side, the strategy centred on exports and on external financing and investment as the main growth machine.

In Mexico, the liberalisation process, initiated after the 1982 default in external debt and bank nationalisation, accelerated at the beginning of the nineties, moving towards the North American Free Trade Agreement (NAFTA) established in 1994. Led by small group of technopols,[24] a victorious coalition formed by large Mexican groups, mainly in the non-tradable sector, and American multinational companies, inaugurated a growth strategy based on exports of labour-intensive industrial activities in a 'shallow' trade specialisation. This export model increased Mexican dependency on US markets and investments and promoted the rise and internationalisation of domestic conglomerates. The latter endeavour fractured the early connections between Mexican capitalism and Mexican big business, and the state played the role of protagonist in privatisation deals and in providing massive finance support.

In South America, the rebellion of cosmopolitan big business against the developmental state started during the late 1970s and the 1980s as a consequence of external crisis and hyperinflation in many countries, occurring alongside the expansion of the power and influence of the holders of dollarised assets, such as the traditional exporters, banks and non-tradable activities in association with foreign capital. Argentina's external debt resulted during the 1980s in huge wealth and debt transfers from state to big business. Starting in 1989, the Structural Adjustment Programme and massive privatisation supported by Washington institutions and the Argentinean elite generated a premature deindustrialisation and denationalisation, but simultaneously, a large centralisation of capital took place headed by commodity exporters and finance.

In Brazil, it was the external crisis of the 1980s, high inflation and the eruption of an autonomous labour struggle that undermined the development coalition. With more diversified industry that partially resisted the process of trade and financial opening, some important public enterprises (including a big development bank) were preserved from the massive privatisation and denationalisation of the mid-1990s.[25] The *desenvolvimentistas* – the technocratic, intellectual and industrial leaders who had led the old economic strategy – were not completely dislodged from the state as had happened in Mexico or Argentina. But the winners from these liberal transformations were the foreign investors and big business in the finance sector and in the production of commodities.[26]

Led by the bureaucracies close to Washington's institutions, the reforms removed industry and its bureaucracies (planning and labour ministries, intermediary government agencies, etc.) from the 'commanding heights' of the economy.

An essential feature of this strategy was a macroeconomic regime based on monetary stability and financial openness. The power of the financial sector in these countries increased not only because its assets grew faster, but because of the predominance in its economic policy of its main interests – higher interest rates and low inflation. Due to high levels of external debt and the growing influence of the International Monetary Fund (IMF) on domestic policies, this financial domination found its expression through orthodox central banks that assumed in these countries the 'commanding heights' of the economy.

Some of these changes and the demise of the developmental state also occurred in Korea at the beginning of the 1990s[27] and in many Asian countries following the 1997–98 external crisis. Throughout the 1990s, several Asian nations followed a mixed strategy based on industrial incentives and on foreign direct investment and exports integrated in commodity chains in a 'flying geese' model. Korea, under pressure from America, opened up its financial system, eliminating the influence hitherto exerted by government on credit and investment. Big *chaebols* decided that government intervention was a hindrance to new economic opportunities.

The national development strategy was not completely changed in China and India (both with military power and autonomous geopolitical presence) nor in Taiwan and Singapore, who followed a path of greater autonomy (or of greater resistance), preserving the national developmental strategy and its hegemonic project in a new context. Although it relinquished some previous economic regulatory mechanisms, the developmental state in dynamic East Asian countries survived.

9.4 National development strategies at the beginning of the new millennium

At the beginning of the new millennium great changes occurred in the world economy. Higher international growth, a substantial rise in commodity prices, lower rates of interest and a continuous expansion of industrial commodities chains (mainly) located in Asia were the principal factors. The rise of China as a great trading power was at the centre of these changes. These circumstances brought about better and more widespread economic opportunities for many peripheral countries. Even for less competitive Latin American countries, the rise in commodity prices allowed the rare simultaneous occurrence of economic growth with positive current account balance and a sharp contraction of external debt. The 2008 financial crisis brought about a great recession in industrialised countries (an effect that still continues) but did not change some of these new and structural circumstances for less industrialised countries. In this context, many countries introduced Keynesian expansionist measures against the hitherto predominant orthodox opinion. Politically, the once strong IMF, World Bank and World Trade Organization (WTO) lost influence in the face of the waves of crisis hitting the countries that had followed their main prescription. Nationalism gained more legitimacy.

Faced with these circumstances, developing countries adopted three different strategies, two of which were not very different from the route taken in the 1990s. The first, a 'passive and integrationist' strategy, like the one followed by Mexico and some East European countries in the 1990s, was merely a continuation of the neo-liberal strategy and gained more support despite its weak results. The second, a 'neo-developmental' strategy adopted by China, Taiwan, Singapore, or even India, explored new opportunities to upgrade their industrial structures including new policies without radically changing their previous mechanisms of industrial and financial regulation. Finally the third, a neo-Keynesian, took place in many primary exporter countries.

As has been argued here, specific institutions and mechanisms for coordinating production were created to solve the problems of industrialisation according to the particular production sectors and technological stages. Although the institutions required for coordinating industrial sectors in agrarian economies (the post-war challenge) are different from those required for industrial upgrading (the present challenge), they continue to be necessary, as the Asian experience indicates. As Lall (2000) observed, commercial policies, credit and subsidy policies, infrastructural development, skill development, technological incentive, and the attraction and delimitation of foreign direct investment (FDI) continue to be the instruments of industrial policy-making. In fact, industrial policies (not only horizontal but also vertical ones) are necessary both for the creation of the incentives of the innovation process in activities involving rapid transformations in the international economy, and for the construction of a new infrastructure.

Thus, the emergence of new challenges to the strategies of industrialisation, resulting from new information and telecommunications technologies (ITT) and from the formation of global and regional production chains, placed new demands on national industrial policies. The construction of a new transportation and communications infrastructure, the dissemination of new technologies and the pursuit of specialisation in specific production segments became part of the ordinary agenda of national projects of industrial upgrading. This 'neo-developmental' strategy is less centred on the internal market as the prevailing dimension of accumulation.[28] The processes of productive regionalisation and internationalisation of 'national champions' strongly expanded firms' investment horizons. The strategies of buying established technology and of adaptation based on process innovations, which typically distinguished Japan and Korea, were challenged by modularised production and new strategies based on greater proximity to proprietary activities and activities related to product innovation. In China, as well as Korea and Taiwan, a 'second phase of catching up' (Chang, 2006), based on innovation and the construction of proprietary national technologies, would be the basic challenge of industrial upgrading.

Based on these new challenges, a 'neo-development strategy' aimed at continuing the 'catch-up' strategy was developed by countries that knew how (or were able) to resist external and internal pressures.

In Asia, this new strategy was mainly followed by China which strongly combined public investment in infrastructure – the main driver of overall growth productivity – with a selective industrial policy in ITT in expansionist macroeconomics, including low interest rates, an anti-cyclical fiscal policy and the maintenance of a competitive rate of exchange. The subordination of finance and enterprise to the development goals was achieved through the maintenance of political centralisation and by the leadership of state enterprises in the 'commanding heights' of the economy. In India, with a much more fragmented society, some of the old regulation instruments were also preserved, favouring a less ambitious, but nevertheless active industrial policy.

As we have been observing throughout this chapter, a strategy is not only the outcome of a decision by a state; its coherence is socially and structurally conditioned. In Korea, after the liberal reforms implemented in the midst of an ample IMF financing adjustment, the previous industrial policies and institutions were dismantled. But the extraordinary expansion of exports (partially induced by Chinese expansion) that followed the 1997 crisis permitted a substantial reduction in sovereign debt and less interference from the IMF. In this new situation, some big *chaebols* rebuilt, with the Korean state, a new coalition for industrial modernisation based on innovation. Thus, due to a developed and homogeneous structure and pragmatic economic policy, institutional change in Korea did not interrupt its high road to investment and productivity. In a more interventionist Malaysia, something similar happened. A neo-developmental strategy is being followed without or 'beyond the developmental state' (Fine, 2005).

A neo-developmental strategy, as we have described it, is today less focused on the productive sector than was the case in the past and more centred on innovation processes in new technologies through several policies and instruments.[29]

As observed in South America between 2002 and 2008, a spectacular rise in the price of commodities permitted those countries to obtain higher growth rates, a sharp contraction in external debt and an accumulation of reserves. After the evident failure of neo-liberal strategies based on Washington Consensus reforms, a more pragmatic macroeconomic policy became established. At the same time, various nationalist movements spread from Patagonia to the Andes countries (Venezuela and Bolivia, also highly critical of the market and liberal institutions supported by the USA) as a backlash to the radical liberal experiments of the 1990s. These movements created regional policies and agreements such as ALBA (Bolivarian Alliance of America) and UNASUL (South American Nations Union) with alternative goals to the free trade initiatives led by the USA.

In Russia, a similar situation occurred, turning around the tragic decade of the 1990s and enabling the new government to construct significant sovereign reserves.[30] Countries as diverse as Argentina, Brazil and Russia could achieve higher growth prompted by internal markets that were now released from the external constraints that had blocked them throughout the 1990s. Higher minimum wages, higher social transfers and employment were achieved. And this occurred, essentially, without changing the pattern of economic growth.

Thus, various countries began constructing new development strategies situated somewhere between the neo-developmental strategy based on 'a second catching-up phase', and a passive and integrationist strategy. The new strategy, a 'neo-Keynesian' one, tries to distance itself, on the one hand from the previous strategy of national development, and on the other from the pro-finance and liberal macroeconomic policy advocated by the Bretton Woods institutions.

Without the particular conditions that support a 'high road' which we observed in some Asian countries, the state, in this third way, has less power to induce structural change. The economic and social cohesiveness required for this is missing.

As we have seen, for different structural and political reasons in major Latin American countries and Eastern Europe, the major private economic groups that in the past were the main beneficiaries of industrial policy are, nowadays, much more closely associated with international commodity chains in asymmetrical regional agreements and in non-tradable activities (in the case of Mexico or Eastern European countries that adopted the integrationist strategy); or became fragmented and failed to survive the radical process of liberalisation (as was the case in Argentina); or relocated to sectors based on natural resources and related support activities in services and construction (Brazil and Russia). Of course, in these countries, there are large segments of national manufacturing industry not connected to global chains that have resisted and survived. Nowadays, they are exposed to strong competition from China and need a more active industrial policy but these interests are diffuse and have less power to exert leadership in economic policy or to build political support for a comprehensive industrial policy.[31] On the other hand, opportunities to expand investment in natural resources have greatly enlarged.

But despite those new economic coalitions that explain why only a few countries are building solid developmental strategies based on technological catch-up, the re-birth of Keynesian (or, at least, more pragmatic) economic policy in many peripheral countries since 2008 has increased the possibilities for higher rates of growth and new perspectives on structural change. This has opened up more space for social and economic varieties of accumulation and growth strategies for the days ahead.

Notes

1. For a recent classification of the institutionalist approach, see Fine (2005).
2. For an original reference to this expression, see Gore (1996); see also Medeiros (1997).
3. Skocpol (1985) is an essential reference for this Weberian approach.
4. This tradition is very influential in some contemporary analyses of globalisation, such as Jessop (2002) and Morton (2007), and broad perspectives on capitalist state and institutions like the social structure of accumulation theory (SSA), McDonough, Reich and Kotz (2010).
5. The internationally consecrated formulation about the Developmental State is Johnson's (1982). Along these lines Amsden (2001) considered that the developmental state was predicated to perform four functions: development banking; local-content management; selective seclusion; and national firm

formation. In accordance with this description a state dedicated to promoting industry through indirect and discretionary interventions on prices and investment is considered to be developmental. However, the success of this strategy depends less on design and more on implementation and this depends on internal interests, conditioned by economic structures and the action of the hegemonic state.

6. Jessop (2002) in his definition of economic domination considers two dimensions. 'The first is internal to the economy and concerns the power of one or another fraction of capital …to impose its immediate interests on other fractions… . The second dimension of economic domination … refers to the capacity of capital in general, a given fraction of capital, or particular capitals to steer the evolution of other institutional orders in line with the demands of capital accumulation' (*op. cit.*, p. 29). In his critical observations on the present chapter Andrea Ginzburg considered that Jessop solved *ex ante* the problem of 'embeddedness'. I consider that this critique is somehow exaggerated when applied to this second dimension. Jessop follows here a Gramscian perspective that concedes to the state a relative autonomy.
7. The analysis of questions regarding the implementation of modern industries in agrarian societies was the main focus of study by the 'pioneers of development economics' such as Rosenstein-Rodan, Nurkse, Lewis, Hirschman, Furtado and Prebisch, inducing a new meaning to development economics.
8. State intervention is a phenomenon that has been common across the development experience, in the successful cases as well as the failures.
9. In his critical analysis of this chapter, Andrea Ginzburg noted that no effective demand analysis was provided here to explain the developmental strategies taken in this period. To clarify the point, the perspective adopted here follows the broad idea that during this period, given the expansionist macroeconomic policies adopted by the most successful developmental states, the main constraint on growth was exerted by the balance of payments. For a discussion see Medeiros and Serrano (2001).
10. During this same period and among industrialised countries a 'Keynesian National Welfare State' (Jessop, 2002) was built.
11. The distinction developed by Lewis (1978) and Furtado (1969) between tropical agriculture and that of a temperate climate is essential for the understanding of distinct starting points of the 'primary exporter' model and for the different levels of heterogeneity of social and economic structures.
12. See Anibal Pinto (1973). Here by internal structural heterogeneity is meant the productivity gap between food production and industrial goods and by external productivity or unbalanced economic structure the productivity gap between primary export sector and industry.
13. Brazil is an exception to this pattern. See Fishlow (1991).
14. For discussion see Pazos (1984).
15. In his comments on this chapter Andrea Ginzburg considered that Mahon and Diamand and the author accepted a strong price elasticity optimism and adopted in this analysis a supply-side static approach, as in 'Dutch Disease' hypothesis. This is a misunderstanding. The point under consideration is that although many industrial policies can induce export diversification, the existence of a competitive exchange rate for industry is a necessary (but

surely not sufficient) condition for this achievement. The fact that Korea or Taiwan had a competitive exchange rate during this period was the outcome not only of an industrial strategy (that was based in several policies) but of the fact that, unlike Latin American countries, they had no other way of obtaining currency. Structure matters as well as institutions.

16. Due to a higher balance-of-payments constraint the stronger incentives for substitution of imports in LAC induced a higher horizontal diversification than in Asia where it could be balanced by a higher priority in productivity. One may consider that the creation of *maquiladora* (in-bond industry) in Mexico occurred in 1965 as in Taiwan.
17. Although the ratio of debt to GDP was not very different in Korea and Brazil, the ratio of debt service to export that better reflects the country's external solvency was much lower in Korea. For an analysis of the 1970s external indebtedness in LAC, see Medeiros (2008a).
18. Wade and Veneroso (1998).
19. For discussions on these formulations see Serrano (2004) and Glyn (2006).
20. The liberating pressure was particularly strengthened in the environment of the WTO in the Uruguay Round that started in the 1980s.
21. 'In Korea, the giant conglomerates (the *chaebols*) have aggressively campaigned during the 1990s to convince the population that the government should abandon its industrial policy and financial regulation' (Chang, 2006, p. 253).
22. Fine (2005) examines different interpretations of the crisis of the developmental state, from what he calls the 'political school' to the 'economic school'. For the former, the developmental state was a singular historical construction that lost its functionality with the success of development. For the latter authors, the State crisis derived from financial opening and the lack of adaptation of institutions to the new environment; for others, such as Chang (2006), the financial opening destroyed the basic mechanism of investment coordination.
23. 'Within the context of the above-mentioned structural heterogeneity, LA has developed two types of successful "modern-sector" regional oligopolies: those involved in large scale capital intensive commodity production for exports, and those that have mastered the technique of organizing low-value-added labour intensive production chains – sometimes for exports (most agricultural products) and sometimes in services (eg. retail)' (Palma, 2010, p. 33).
24. This expression captures the political, ideological and economic work of a small group of academics who acted as organic intellectuals, in a Gramscian meaning, of the new socioeconomic bloc. Unlike the old technocrats, they had to be good politicians.
25. This resistance was only partial. Through the 1990s there was a disarticulation of some productive chains and an abandonment of some innovation-intensive activities, creating a kind of 'regressive specialization' (Coutinho, 1997), although some modern capital goods manufactures more integrated with mineral and raw material resources were preserved.
26. 'During the 1990s, Brazil's industrial elite withdrew without much conflict from areas attractive to international investors: industrialized food production and distribution, supermarket chains and automobile spare parts in the

beginning of the decade and, later, telecommunications, advanced services and financial institutions. The industrial elite has either migrated to the tertiary sector or retreated from business altogether, investing their capital in the financial market, pension funds or real estate. Other possibilities were "support activities" such as the building sector, packaging industry, car sales concessionaires, activities related to business and law consulting, business promotion, educational and cultural events, administrative and honorific positions in the "third sector" (NGOs) and the administration of real estate' (Rocco, 2007, p. 208).

27. The dissolution of the powerful Economic Planning Board into the Ministry of Finance and Economy was a milestone for the new state.
28. This is not to say that this strategy is led by exports: in case of China and India internal markets are still the main demand source for capital accumulation, but in both countries exports are much larger than they were in the past.
29. And this, as we have argued throughout this chapter, is conditioned by economic structure, the geopolitical position of the country and political coalitions between states and social classes.
30. With the end of the socialism and the radical liberalism of the 1990s, Russia resumed a national developmentalism strategy based, however, on natural resources. Thanks to its geopolitical position and greater state control over oil and natural gas exports, transfers to other sectors of the economy increased substantially; however, they did not result in greater export diversification.
31. The industrialisation of natural resources as an alternative industrial strategy has been under discussion in many small Latin American countries.

References

Amsden, A. (2001) *The rise of "the rest": challenges to the west from late-industrializing economies* (Oxford: Oxford University Press).

Bresser Pereira, L. C. (2010) *Globalização e Competição*. (Sao Paulo: Edit Campus).

Chang, H. J. (2006) *The East Asian Development Experience, The Miracle, the Crisis and the Future* (New York: Zed Books).

Coutinho, L. (1997) 'A Especialização Regressiva: um balanço do desempenho industrial pós-estabilização', in J. P. R. Veloso (ed.) *Brasil: Desafio em um país em Transformação* (Rio de Janeiro: Jose Olympio Edit), 81–106.

Diamand, M. (1978) 'Towards a Change in the Economic Paradigm Through the Experience of Developing Countries', *Journal of Development Economics*, 5, 19–53.

Fine, B. (2005) 'Beyond the Developmental State. Towards a political economy of Development', in C. Lapavitsas and M. Noguchi (eds) *Beyond Market Driven Development* (New York: Routledge).

Fishlow, A. (1991) 'Some Reflections on Comparative Latin American Economic Performance and Policy', in T. Banuri (ed.) *Economic Liberalization: No Panacea The Experiences of Latin America and Asia* (New York: Oxford University Press), 149–71.

Furtado, C. (1969) *Teoria e Política do Desenvolvimento Econômico* (São Paulo: Companhia Editora Nacional).

Ginzburg, A. and Simonazzi, A. (2003) 'Patterns of Industrialization and the Flying Geese Model: The Case of Electronics in East Asia', paper presented at the Conference on Clusters, Industrial Districts and Firms: the Challenge of Globalization, Modena, Italy 12–13 September.

Glyn, A. (2006) *Capitalism Unleashed* (London and Oxford: Oxford University Press).

Gore, C. (1996) 'Methodological Nationalism and the Misunderstanding of east Asian Industrialization', *European Journal of Development Research*, vol. 8, no. 1.

Haggard, S., Lee, C. H. and Maxfield, S. (1993) *The Politics of Finance in Developing Countries* (Ithaca and London: Cornell University Press).

Hughes, A. and Singh, A. (1991) 'The World Economic Slowdown and the Asian and Latin American Economies: A Comparative Analysis of Economic Structure, Policy and Performance', in T. Banuri (ed.) *Economic Liberalization: No Panacea The Experiences of Latin America and Asia* (New York: Oxford University Press), 57–100.

Jessop, B. (2002) *The Future of The Capitalist State* (Cambridge: Polity).

Johnson, C. (1982) *MITI and the Japanese Miracle* (New York: Stanford University Press).

Jilberto, A. and Hogenboom, B. (2007) 'Latin America Conglomerates in a Neo-Liberal Era', in A. E. F. Jilberto and B. Hogenboom (eds) *Big Business and Economic Development* (London: Routledge), 156–87.

Kohli, A. (2004) *State Directed Development, Political Power and Industrialization in the Global Periphery* (New York: Cambridge University Press).

Lall, S. (2000) 'The Technological Structure and Performance of Developing Country Manufactured Exports, 1985–98'. *Oxford Development Studies*, vol. 28, no. 3, 337–69.

Lewis, A. (1978) *Evolution of International Economic Order* (Princeton: Princeton University Press).

Mahon, J. (1992) 'Was Latin America too Rich to Prosper? Structural and Political Obstacles to Export Led Industrial Growth', *Journal of Development Studies,* vol. 28, no. 2, 241–63.

McDonough, T., Reich, M. and Kotz, D. (2010) *Contemporary Capitalism and Its Crises, Social Structure of Accumulation Theory for the 21st Century* (New York: Cambridge University Press).

Medeiros, C. A. (1997) 'Globalização e a inserção internacional diferenciada da Ásia e da América Latina', in M. C. Tavares and J. L. Fiori (eds) *Poder e Dinheiro* (Rio de Janeiro: Vozes), 279–347.

Medeiros, C. A. (2008a) 'Financial Dependency and Growth Cycles in Latin American Countries', *JPKE*, vol. 31, no. 1, 79–101.

Medeiros, C. A. (2008b) 'Desenvolvimento Econômico e Ascensão Nacional: rupturas e transições na China e na Rússia', in J. L. Fiori, C. Medeiros, and F. Serrano, *O Mito do Colapso do Poder Americano* (Rio de Janeiro: Vozes), 173–273.

Medeiros, C. A. (2009) 'Latin American Auction', *New Left Review*, no. 55, 109–32.

Medeiros, C. A. and Serrano, F. (2001) 'Inserção Externa, exportações e crescimento no Brasil', in J. L. Fiori and C. A. Medeiros (eds) *Polarização Mundial e Crescimento* (Petrópolis: Vozes), 105–35.

Morton, A. D. (2007) *Unravelling Gramsci, Hegemony and Passive Revolution in the Global Economy* (London: Pluto Press).

Ocampo, J. A., Jomo, K. S. and Vos, R. (2007) 'Explaining Growth Divergence', in J. A. Ocampo, K. S. Jomo and R. Vos (eds) *Growth Divergences* (New York: Zed Books), 1–25.

Palma, G. (2010) 'Why has productivity growth stagnated in most Latin American countries since the neo-liberal reforms?' *Cambridge Working papers in Economics*, 1030 (Available at http://www.econ.cam.ca.uk/dae/repec/cam/pdf/cwpe1030.pdf).

Pazos, F. (1984) 'Have Import Substitution Policies Either Precipitated or Aggravated the Debt Crisis?' *Journal of Interamerican Studies and World Affairs*, 27 (Winter), 57–73.

Pinto, A. (1973) *Distribuição de renda na América latina e desenvolvimento (*Rio de Janeiro: Zahar).

Prebisch, R. (2011 [1949]) 'O Desenvolvimento Econômico da América Latina e alguns de seus principais problemas', in Adolfo Gurrieri (ed.) *O Manifesto Latino-Americano e outros Ensaios* (Rio de Janeiro: Contraponto) 95–153.

Rocco, R. (2007) 'Economic Flexibilization and Denationalization in Brazil' in Alex E. F. Jilberto and Barbara Hogenboom (eds) *Big Business and Economic Development* (London: Routledge) 191–211.

Serrano, F. (2004) 'Relações de Poder e a Política Macroeconômica Americana, de Bretton Woods ao Padrão Flexível' in José Luis Fiori (ed.) *O Poder Americano* (Rio de Janeiro: Edit Vozes), 179–225.

Skocpol, T. (1985) 'Bringing the State Back In: Strategies of Analysis in Current Research' in Peter Evans, D. Rueschemeyer, T. Skocpol, *Bringing the State Back In* (Cambridge: Cambridge University Press).

Wade, R. and Veneroso, F. (1998) 'The Asian Crisis: the high debt model versus the Wall Street Treasury IMF Complex', *New Left Review*, no. 228, 3–25.

10

Harmonic and Conflict Views in International Economic Relations: a Sraffian View

*Sergio Cesaratto**

In dealing with the European crisis and the frequent accusation of German 'mercantilist behaviour' (Cesaratto, 2011, 2012b, 2013; Cesaratto and Stirati, 2011), it seemed natural to look for analytical approaches to international economic relations (IER) that went beyond naive pro-European rhetoric and mainstream economic beliefs in the harmonic virtues of international laissez-faire. Some Sraffian contributions to the demolition of these beliefs will be recalled below. The pre-laissez-faire, mercantilist tradition was another natural candidate for attention. Mercantilism, the world of non-harmony, may be envisaged as an underground tradition, which a group of northern European economists called 'the other canon' (www.othercanon.org), parallel to the laissez-faire tradition. Kalecki's view of net exports as a way of realising profits also buttresses a conflict view of IER. A cynical view also springs from political realism, a major tradition in political science. An intellectual father of political realism was Thomas Hobbes, contemporary of many British mercantilists. In this tradition, a social contract is enforceable at domestic level by attributing authority to the Prince, but not at international level where sovereign states do not submissively recognise any higher authority. Mercantilism and political realism converge in international political economy (IPE), a field that arose in the early 1970s as an attempt to bridge the gap between the disciplines of international economics and international relations (Strange, 1970). Political realism is commonly juxtaposed with a liberal tradition that holds a more harmonic view of IER. IPE has recently been colonised by neoclassical political scientists. The Sraffian criticism of neoclassical economics therefore appears remarkably

*I thank an anonymous referee for useful comments.

precious as a response to neoclassical imperialism. The nation-state is at the centre of mercantilism and IPE. The classical and Marxist approaches are not on easy terms with the notion of nation-state, so that it was also impossible to avoid this topic.

This chapter is a preliminary exploration of the complementarity of the classical conflict view of income distribution and the disharmonic traditions of IER in opposition to the harmonic beliefs of economic and political liberalism.[1]

10.1 Mercantilism and laissez-faire

In classic treatises on mercantilism, Furniss (1920), Suviranta (1923) and Heckscher (1955) regard it as a first *systematic* attempt to understand economic phenomena.[2] All trained in neoclassical economics, these economists found mercantilism profoundly different from laissez-faire doctrines, both classical and marginalist.

Assessing Adam Smith's famous criticism of mercantilism, to which Smith devoted a quarter of *The Wealth of Nations*, Suviranta (1923, p. 160) notes that '[t]he difference between the mercantile and the liberal point of view was not accidental, arising merely from confusion in thought, but it was deep-rooted in the different character of these economic systems'. In the latter point of view, the purpose of economic activity is '[p]roducing wealth for satisfying human wants, i.e. the ultimate end is consumption'. On the other hand, according to mercantilism '[t]he logical consequence of the fact that the people were primarily thought of as a capital material, was that consumption also came primarily to be servant of production, and not a means of satisfying human wants' (*ibid.*, p. 162).

In a similar vein, Heckscher (1955 [II], p. 285) considered mercantilism to be amoral, because it broke with the ethical foundations of the Middle Ages political views, and characteristically put the raison d'état, not individualism, centre-stage: 'the welfare of society or, in actual fact, the welfare of the state was substituted in place of the amelioration of the individual. This was a perfectly simple corollary of the raison d'état, or pure Machiavellism. ... In addition the raison d'état was conceived emphatically as materialistic or economic' (*ibid.*, p. 286). Mercantilism was also amoral with regard to its means. In particular, the pursuit of individual interest was seen as a function of the welfare of the state, and had to be regulated for that purpose (*ibid.*, p. 293 and *passim*). Heckscher regarded this view as being opposite to that of laissez-faire economists influenced by the utilitarian goal of improving social happiness, seen as the sum of individual welfare (*ibid.*): both the individual and the state served the 'community'. The mercantilists held a more cynical view in which both individuals and the community served the state (*ibid.*, pp. 328–9). According to Heckscher mercantilists thought in terms of nation-states, and did not regard the individuals composing a nation as equal.

Rational thought pervaded mercantilism: 'Rationalism characterized mercantilism to so high a degree. There was little mysticism in the arguments … this rationalism expressed itself in references to *nature*. Nature was conceived as a factor which also influenced the social sphere, social life being placed parallel to physical life of the individual; and society was regarded as a body with functions similar to those of the physical body' (*ibid.*, p. 308). Mercantilists, as well as laissez-faire theorists, therefore believed in the possibility of discovering natural laws. The difference with laissez-faire theorists was that mercantilists did not regard public intervention as interference with the benign working of those laws, but as their natural complement. Mercantilists did not believe in a 'immanent social rationality' (*ibid.*, p. 321) or objective economic harmony, whereas laissez-faire 'went so far in its belief in the domination of natural laws in society that it believed in an immanent reason in the free play of forces' (*ibid.*, p. 323).[3]

Heckscher regarded the victory of laissez-faire theories over mercantilism as temporary. During the nineteenth century they were submerged by historicist and nationalistic doctrines: 'Society was regarded as a growth in the highest degree naturally determined, to be changed only by slow and gently progressive treatment, bound to tradition, each individual nation containing inherent and more or less ineradicable peculiarities' (*ibid.*, p. 334). According to the Swedish economist, laissez-faire doctrines also failed in their alleged humanitarian aims. Indeed, what they did, Heckscher argues, was to support measures that protected the individual against the state, but they failed to protect him against the market, 'against the pressure of social conditions, which did not have their origins in definite measures of the state but which, on the contrary, demanded such measures if they were to be abolished. On this point, *laissez-faire* was obstructed by its belief in natural rights, i.e., its belief in a predetermined harmony, to which was added in practical policy the influence of employer and capitalist interests' (*ibid.*, p. 337). Heckscher observes that paradoxically but not incidentally, social reforms were supported and adopted by conservatives: 'economic policy being bound up with the duty of the patriarchal state to care for the welfare of its subjects', while 'the growing importance of socialism also goaded politicians into finding remedies' (*ibid.*, p. 337).

10.2 Domestic social surplus and foreign trade surplus

In *Theories of Surplus Value* Marx argued that although James Steuart, the last great mercantilist, represented the clearest expression of the theory of 'profits upon alienation', he 'does not share the illusion that the surplus-value which accrues to the individual capitalist from selling the commodity above its value is a creation of new wealth' (Marx, (1969 [1861–3], p. 41). What Marx seems to appreciate is the idea that the surplus value is the result of one side taking advantage of the other, although the 'unequal exchange' that generates a surplus for one trading side takes place, for Marx, in the labour, not the goods market. The classic treatises on mercantilism also underline the clash between the mercantilists' view of distribution and that of 'modern' marginal theory.

Furniss (1923, pp. 198–203) regards wage determination in mercantilism as affected by the interests of the dominant classes, something far removed from the marginalist notion of wages as the natural reward of labour linked to its (marginal) contribution to production. The contrast between the mercantilist view of labour as 'the source of national wealth' and the recommendation of low wages led Furniss to perceive 'the germs of the socialist doctrine' (*ibid.*, p. 25), the term he uses to denote the classical surplus approach, in mercantilism. Indeed, very few mercantilists clearly anticipated the classical concept of surplus, and almost none came close to seeing the origin of a foreign surplus in a domestic surplus of production over consumption.[4] However, the concept was there, for instance, when they regarded the unemployed as a burden on the nation, implying that they lived on a surplus produced by the employed population, and when they argued that productive utilisation of the unemployed would increase the riches of the nation (e.g. *ibid.*, pp. 41–7 and 89–95).[5] In a famous example, Petty alludes to the ability of one section of the population to sustain the rest, including those that produce export goods (e.g. Aspromourgos 1996, p. 23). Davenant is worth quoting: 'If all hands in this Kingdom that are able were employed in useful labour our manufactures would be so increased that the commonwealth could be thereby greatly enriched and the poor, instead of being a charge, would be a benefit to the Kingdom' (quoted by Furniss, p. 91).

Suppose that the social product P just consists of *necessities*: $P=N+N'$, where N and N' are the necessities of L workers and U unemployed, respectively, both receiving a real wage equal to w, so that $P=Lw+Uw$. Defining per capita output as $\pi=P/L$, we get: $L(\pi-w)=Uw$. In other words, the unemployed survive on the surplus produced by workers above their subsistence ($\pi > w$, otherwise the unemployed would already have died). If the unemployed are put to work, the social product becomes: $P' = (L+U)\pi = L'\pi$, and the social surplus: $S = P' - L'w = L'(\pi - w)$. Now a larger surplus can be used to sustain an unproductive class (not consisting of unemployed in this case) or exported.

However, only in a late mercantilist quoted by Furniss, William Hay, an author who was presumably influenced by Petty, do we find clear coordination between domestic social surplus and foreign trade surplus:[6] 'The source of wealth is from the number of its inhabitants; ... the more populous a country is, the richer it is or may be ... For the earth is grateful and repays their labour not only with enough but with an abundance ... Now whatever they have more than they consume, the surplus is the *riches* of the nation. This surplus is sent to other nations and is there exchanged or sold, and this is the trade of the nation. If the nation to which it is sent cannot give goods in exchange to the same value they must pay for the remainder in money; which is the balance of trade; and the nation that hath that balance in her favour must increase in wealth' (1751, quoted by Furniss, pp. 19–20, italics in Furniss).

Using the same symbols as above, the coordination between the two surpluses can be summarised as follows. In a surplus-producing economy we have: $S = P - N$. The social surplus can have a number of destinations: capitalists' consumption C_c, capital accumulation I and net exports $X - M$, that is: $S = C_c + I + X - M$. If, for simplicity C_c and I are zero, we obtain: $S = X - M$.

10.3 The centrality of national output, employment and trade surplus

As we have seen, mercantilism appears to have accorded primacy to production rather than to consumption, as in later laissez-faire theories.[7] The goal of maximising domestic production and employment, while minimising domestic consumption and imports of superfluous goods, aimed at obtaining the largest possible foreign trade surplus, which was seen by many mercantilists as the origin of net wealth for the nation. As suggested by Kalecki's lesson, we may now interpret foreign trade surplus as a way for capitalists to realise the domestic social surplus they do not consume or invest.[8] Although full coordination of internal and external surpluses was to some extent approached, it may not be said that the mercantilists were successful in this regard. So on one hand we are left with hints, by later mercantilists in particular, that the social surplus is the origin of net wealth (intended as that part of the social surplus that can be consumed or accumulated without endangering reproduction of the system on at least the same scale) and on the other hand with the idea that the origin of net wealth lies in the foreign trade surplus. How did they justify this second origin of net wealth or the importance they attributed to the trade surplus?

Profits upon alienation

Supposing, like Marx, that most mercantilists held a theory of 'profit upon alienation', it follows naturally that net gains for the nation as a whole can only be obtained by foreign trade (Heckscher [II], 1955, p. 193). One of the clearest expressions of this view is in widely quoted passages by Charles Davenant:

> It is the Interest of all Trading Nations, whatsoever, that their Home Consumption should be little, of a Cheap and Foreign Growth and that their own manufactures should be Sold at the highest Markets, and spent Abroad; Since by what is Consumed at Home, one loseth only what another gets, and the Nation in General is not at all the Richer; but all Foreign Consumption is a Clear and Certain Profit. (1697, quoted by Heckscher [II], p. 115)

So there are no 'profits' if the product is sold at home, but positive profits if consumption is kept at a minimum, cheap foreign commodities are imported, and net output sold abroad. Mercantilists had no clear notion of profits as the net income accruing to capital (Ehrlich, 1955) and as we have seen, almost without exception they did not clearly perceive the existence of a domestic surplus as the basis for a foreign surplus. They seemed to have even less idea that, given the real wage, the larger the domestic surplus, the more difficult its realisation in the domestic market, and the greater the need for 'external markets'. In any case, Kalecki's theory upholds their point of view as to the importance of net exports.

Fear of goods

The importance attributed by mercantilists to the trade surplus may be linked to the importance they seemed to attribute to the maximisation of domestic output and employment by minimising imports and maximising exports, and vice versa, to the maximisation of domestic employment to reduce imports and increase exports.[9] In this regard Johnson (1937, p. 302) describes a 'balance of work' as the difference between the labour content of exports and imports (see also Furniss, 1920, pp. 13–14; Suviranta, 1923, p. 142; Heckscher, 1955 [II], p. 366). Imports were seen as reducing domestic employment, and exports as labour 'paid by foreigners'. Here one may perhaps identify the 'fear of goods', which Heckscher saw as a leitmotif of mercantilism, as fear that limitation of the domestic market was incompatible with full employment, not such a strange concern in a low-wage economy (Heckscher, 1955 [II], pp. 121, 365).

Low wages were seen by mercantilists, though not unanimously, as a way to keep domestic consumption and imports at bay (*ibid.*, p. 364) and production costs low (*ibid.*, pp. 152–3), while encouraging hard work (Furniss, 1920, Ch. 4). A low-wage economy was therefore a central objective for a typical mercantilist (Furniss, 1920, pp. 8, 40; Heckscher, 1955 [II], pp. 46, 163–5, 153–4, 364–5). This calls for reflection.

We may distinguish a policy of *import substitution* from an *export-led* economy. The mercantilist position is evocative of both. In Keynesian terms, a policy of *import substitution* that, say, decreases the marginal propensity to import, has (*ceteris paribus*) a positive effect on domestic output and employment and – for a given level of exports – on the trade balance. In addition to the initial benefit for output and employment, by relaxing the foreign balance of trade constraint an import substitution policy also leaves more scope for growth policies based *on the growth of the domestic market*. So, in this case, the improvement in foreign trade is necessary to development of the domestic market, not an objective *per se*.

A low-wage economy would also keep imports at bay – assuming that imports of foreign goods are linked to the level of real wages. This policy, however, is hardly relevant to the development of a domestic market, but conducive to an export-led model. Once a decision to depress the domestic market is taken through a low-wage policy, then an *export-led model* is the only game in town, sustained in turn by the low labour costs. This description of a low-wage, export-led economy suits the mercantilist idea of imports as paying foreign labour, and exports as domestic labour paid by foreigners. According to the Kaleckian surplus approach, in such an economy capitalists maximise the domestic surplus they extract from workers, and get rid of it by net exports.[10]

Influx of money capital

Although a popular interpretation of mercantilism regards mere accumulation of specie as the aim of a trade surplus, mercantilists seem to have maintained that a net influx of precious metals was functional to growth of the domestic money supply, lowered the interest rate and favoured economic activity (Heckscher, 1955 [II], pp. 204, 208–9, 217–18 and *passim*). Heckscher notes that the idea that savings finance investment was absent in this literature, and capital was identified with money (*ibid.*, pp. 198–9), something that would not sound strange to modern unconventional monetary economists. This was Keynes's favourite explanation of mercantilism, expressed in Chapter 23 of the *General Theory*.

Power or plenty

Finally, the goal of a foreign trade surplus can be interpreted as a way to generate relative international political power.[11] The relative importance of power and wealth was the subject of controversy between Heckscher (1955 (I), p. 24 and *passim*) and Viner (1948); both acknowledged that the two objectives mutually sustained each other, but Heckscher attributed primacy to power and Viner to wealth. The mercantilist tradition has indeed been perceived as putting the nation-state at the centre of analysis, contrary to the cosmopolitan views of laissez-faire and Marx's theories: 'the state stood at the centre of mercantilist endeavours as they developed historically: the state was both the subject and the object of mercantilistic economic policy' (Heckscher, 1955 [I], p. 21). In short, mercantilism is 'the economic system of nationalism' (*ibid.*, [II], p. 13).[12]

In discussing presumed German mercantilism (Cesaratto, 2011, 2012b; Cesaratto and Stirati, 2011), I regarded export-led growth as a growth policy choice alternative to domestic-demand-led growth. In Germany, the Keynesian perspective was rejected by the influential Ordoliberal school as conducive to social indiscipline and inflation, regarded in turn as disturbing market-led resource allocation. Although not explicitly acknowledged, in this context, export-led growth remained the only game in town – a model that is conducive to, and simultaneously supported by, social discipline and wage restraint (Crouch, 2008, p. 479). As Ludwig Erhard (1897–1977) stated in 1953: 'foreign trade is not a specialized activity for a few who might engage in it, but the very core and even precondition of our economic and social order' (quoted by Cronin, 1996). Germany is indeed a perfect mercantile economy. At the micro level it has an excellent training, educational and R&D system; at the meso level the yardstick of a trade surplus creates an ideological climate that induces cooperation and discipline; at the macro level the systems keep wage growth in line with productivity growth and fiscal policy moderates domestic demand. Foreign policy has the promotion of German exports as priority. Paternalism is a traditional attitude of the German government; a sense of national community, traditions and nature is the main component of 'German ideology'. This perfectly suits the mercantilist tradition, particularly in its German version (Cameralism, Historical School, cf. Riha, 1985, chs 4 and 5). Codetermination is a further institutional pillar, but as Voltaire said (and as the late Vianello liked to quote): 'Incantations will destroy a flock of sheep if administered with a certain quantity of arsenic'. Just in case, a watchdog role was taken on by the Bundesbank in a unique wage-bargaining process directly involving the central bank and the leading trade union IG-Metall (Franzese and Hall, 2000, pp. 182–3). As we have seen, this model perfectly fits the Kaleckian view that net exports are a way of realising a domestic surplus, and that financial capital flows from core-surplus countries are a way to finance expenditure by trade peripheral-deficit countries, as confirmed by recent European experience (Cesaratto, 2012b).

10.4 The mercantilist tradition, liberalism and international political economy

Mercantilism was deemed dead in the late eighteenth century, when Adam Smith was self-confident enough to pass silently over the contribution of the last great quasi-contemporary mercantilist, James Steuart. Nonetheless mercantilist wisdom survived as an underground stream of thought parallel to mainstream laissez-faire economics, for instance in the work of protectionists Alexander Hamilton and Friedrich List in the German Historical School, up to modern developmentalism and in some versions of IPE.[13]

International political economy arose in the early 1970s in English-speaking countries (see Cohen, 2008). Economists Albert Hirschman and Charles Kindleberger are regarded as its forerunners. Simplifying, there are two competing political-philosophical inspirations of IPE, liberalism and political realism, which hold harmonic and conflict views of IER, respectively. The intellectual fathers of political realism are Thucydides, Machiavelli and Hobbes (see, e.g., Donnelly, 2000). All held a pessimistic view of human nature as being motivated by greed, suspicion and ambition. As a result, there is a conservative element in political realism – the human soul cannot change – but also a healthy reaction against facile utopias, which are seen as an obstacle to real change.[14] The international arena is regarded as an anarchic field governed by the rules of power: no Hobbesian social contract that delegates power to a super-national authority is acceptable to sovereign states, if not for contingent convenience.[15] The modern recovery and application of this approach to international relations is due to Edward Carr (notably a Marxist), Hans Morgenthau and Kenneth Waltz. On the other side, the liberal tradition maintains that affirmation of the market economy is a solution to domestic and international conflicts through the mutual advantages of free trade (on this, cf. the critical analysis by Albert Hirschman, 1977).

In this regard, an influential exponent of IPE, Robert Gilpin, distinguishes liberal, nationalist and Marxist traditions. 'Economic liberals', he writes, 'believe that the benefits of an international division of labour based on the principle of comparative advantage cause markets to arise spontaneously and foster harmony among the states; they also believe that expanding webs of economic interdependence create the basis for peace and cooperation in the competitive and anarchic state system' (Gilpin, 1987, pp. 12–13). On the other hand, 'Economic nationalists … stress the role of power in the rise of a market and the conflictual nature of IER'; they argue that 'economic interdependence must have a political foundation and that it creates yet another arena of interstate conflict, increases national vulnerability, and constitutes a mechanism that one society can employ to dominate another' (*ibid.*, p. 13). Gilpin identifies this approach with the mercantilist tradition. Finally, the Marxist tradition regards international relations as a field of imperial conflict and exploitation of peripheral countries.

Comparing the three approaches, Gilpin points out that, like liberals, Marxists tend to regard international trade as a modernisation force against the scepticism of nationalists (*ibid.*, p. 14).[16] Nationalists support the primacy of politics over economics, Marxists the opposite, and liberals maintain that the two spheres should remain relatively autonomous (*ibid.*, p. 26). Finally liberals and Marxists share an optimistic view of the human fate, the opposite of nationalists, who at international level base their stance on a conflict view of international relations (*ibid.*, p. 43).

The realist Gilpin is sympathetic to nationalists and Marxists, despite his personal liberal beliefs (*ibid.*, p. 25). He is also critical of the neoclassical foundations of the liberal view, accusing the dominant theory of being based on unrealistic assumptions, e.g. perfect competition, rationality, perfect information and the like. Any theory must indeed make simplifications, as long as they do not alter the substance, and this is what mainstream economists have largely done by making the alleged unrealistic assumptions. By virtue of its frail criticism of mainstream international economics, the realist tradition of IPE represented by Gilpin was exposed to liberal and neoclassical counter-criticism.[17] Not surprisingly, the latest generation of American IPE students has increasingly returned to neoclassical propositions and mainstream research methods. This new trend is called open economic policy (OEP) (Lake, 2009, pp. 50, 52). Benjamin Cohen's intellectual history of IPE (Cohen, 2008) sparked off a fierce debate on the evolution of IPE. Cohen notes divergence between American and British IPE, observing that the American school has become increasingly standardised, coming to resemble nothing so much as the methodology of neoclassical economics, featuring the same penchant for positivist analysis, formal modelling, and where possible, the systematic collection and evaluation of empirical data (*ibid.*, pp. 41–2). He also notes that political scientists 'have an inferiority complex when it comes to economics' (*ibid.*, p. 42).

Like Gilpin, defenders of traditional IPE point to the *limitations* of starting with the economic choices of rational individuals (the state was the main unit of analysis in traditional IPE). Another leading old-guard international political economist points out that a high price is paid by 'making preferences and interests exogenous, assuming that interests can be derived only from a rationalist model of human behaviour, excluding from analysis the constitutive aspects of institutional life, committing to an exclusively materialist conception of preferences and interests, and importing reductionist economic theories of politics' (Katzenstein, 2009, p. 127; see also Keohane, 2009, pp. 37–8). I do not regard this criticism of the assumption of rational choices by selfish individual agents as particularly illuminating. A non-individualistic theory of society calls rather for an alternative view of the economic fabric of society.

We cannot but refer to Marx's criticism of methodological individualism based on the classical 'surplus approach'. The methodological individualism of the early classical economists did not surprise Marx: after all, it was the ideology of a new form of society in which individuals broke previous institutional ties – feudal, religious etc. – with other individuals. This of course does not imply that ties have disappeared: they have just been superseded by more anonymous, market-dominated relations, creating the *illusion* that analysis can start with isolated individuals (e.g. Marx, 1957 [1857], pp. 82–3). The production and reproduction of social life is a collective fact for Marx, although in history the manner in which social surplus is produced and distributed has changed profoundly.[18] Individuals' interests and choices are moulded by their positions in production modes (cf. Marx, 1859). Note the degree to which the forgivable 'robinsonades' of Smith and Ricardo became the very foundations of economic and social theory with marginalism. In this theory, production is the *ex post* result of the (marginal) contributions of individual endowments of production factors, an unhistorical view in which socioeconomic relations of production are not the result of evolution of the modes by which humans produce and distribute social output and surplus.[19]

The fact that traditional IPE attributes centrality to the state as the basic unit of analysis is also problematical. On the one hand, the state is an indisputable guarantee of economic activity: 'as Carr has argued, every economic system must rest on a secure political base' (Gilpin, 1987, p. 47).[20] On the other hand, most IPE less convincingly assumes that 'society and the State form a unitary identity and that foreign policy is determined by objective national interest' when 'foreign policy (including foreign economic policy) is in large measure the outcome of conflicts between dominant groups in society' (*ibid.*, p. 48; see also Cohen, 2008, p. 125; Katzenstein 1977, p. 604; and Section 10.6 below).[21]

10.5 Comparative disadvantages?

Adam Smith accused the mercantile doctrine of looking after the interests of merchants and producers, while sacrificing those of consumers (1979 [1776], pp. 661–2). He upheld the advantages of international trade for all participating nations through exchange of surplus products, market expansion and thereby extension of the division of labour (1979 [1776], pp. 446–7). Smith held a theory of *absolute* advantages from trade which is very different from the theory of *comparative* advantages attributed to David Ricardo. A theory of absolute advantages is theoretically consistent with the pursuit of mercantilist policies, e.g. trade policies aimed at developing and safeguarding national absolute advantages. Despite his attack on mercantilism, Smith's theory of international trade is therefore not inconsistent with a disharmonic view of IER. Ricardo's theory of comparative advantages is commonly regarded as the final challenge to mercantilism: a harmonic view of international relations was seen to prevail over a conflict one, although Ricardo was very clear in limiting the validity of his celebrated theorem to the case of absence of capital mobility.

The Ricardian theory identified the origin of comparative advantages in technological differences, whereas the basic marginalist Heckscher-Ohlin-Samuelson (HOS) explanation of international trade explained specialisation on the basis of countries' different factor endowments. The international specialisation of free-trading countries with full employment thus depended on relative scarcity of factors. Sraffian authors, such as Parrinello and Steedman, have taken two critical directions. On one hand, the results of the capital theory controversy have been used to show the limited validity of the HOS theory for cases in which only land and labour are used as inputs. On the other hand, modern extensions of Ricardo's analysis have confirmed the limitations of comparative advantages that Ricardo himself pointed out.

Beginning with the first aspect, the HOS theorem may be expressed rigorously in terms of endowment of non-produced production factors, such as land and labour. On this basis the theory predicts that the country with the highest land-to-labour ratio exports land-intensive commodities. The inclusion of 'capital', however, undermined the prediction that the country with the largest 'capital supply' and the lowest interest rate exports the most capital-intensive commodity. To begin with, there is the standard problem of measuring the 'given amount of capital' irrespective of its distribution. Second, results in capital theory (assuming two sectors) show that ordering of sectors by capital intensity may change with the fall in interest rate and that the price of a more capital-intensive commodity will not fall monotonically with a fall in interest rate (Steedman, 1979a, pp. 4–5).

The international mobility of 'production factors' is seen by conventional theory as alternative to international trade: it is the same whether a relatively 'capital-rich' country exports capital-intensive commodities or 'capital' directly. The idea that capital flows arise from capital-rich countries lending *savings* to capital-poor countries is subject to capital theory criticism as much as to the domestic saving–investment nexus (Garegnani, 1983; Dalziel and Harcourt, 1997). A Kaleckian view would lead us to regard financial flows to trade deficit countries as part of a mercantilist strategy whereby surplus countries lend to deficit countries. In this view, loans precede import spending by peripheral countries, and foreign saving in core countries emerges as the result of their net exports to the periphery. The recent European crisis can be interpreted along these lines (Cesaratto, 2012b).

Marginal theory focuses on countries' different factor endowments, whereas Ricardo suggested that countries may differ because of their respective technology levels and, in the absence of capital mobility, specialise in producing the commodity for which they have the greatest comparative advantage, or the smallest comparative disadvantage. Brewer (1985) and Parrinello (2009) show that once the real wages in two potentially trading countries are, respectively, taken as given[22] and there is capital mobility, *absolute* and not comparative advantages determine the location of production. Too high a wage rate, or too low a productivity level may make a country uncompetitive, inducing capital to flow to the other country: '*We would say that a whole capitalistic economy is* *not competitive if all its capital-using techniques are unprofitable at the international equilibrium prices. This result overrules the claim that "a country must always possess a comparative advantage in something"* [as argued by Krugman]' (Parrinello, 2009, p. 52, italics in the original), which vindicates 'the intuitive idea that national competitiveness can be a source of possible economic conflict among the national economies of a global economy' (*ibid*., p. 50).[23]

10.6 Mercantilism, classical economists and Marx on the nation-state

As we have seen, mercantilism and Marxism both reject the approach, typical of laissez-faire economics, of considering the individual as the basic analytical unit, what Marx called 'robinsonades'. The social relations of production are Marx's analytical anchor, regarded as the way a society historically organises the production and distribution of social output and surplus. The mercantilist tradition does not attain the sophistication of Marx's historical materialism, but both reject the market as the place where free choices of a variety of individuals are reconciled in a harmonic way. National communities and states, rather than Marx's social relations of production and social classes, are the reference categories of the mercantilist tradition. In an important unpublished work, Marx (1975 [1845]) rejected this approach by fiercely criticising Friedrich List (1991 [1841]), an author we can classify in the mercantilist tradition of national political economy and developmental state. Marx's stance is not surprising, as Szporluk explains:

> Marx claimed that his theory, while the result of his own intellectual endeavour, was also the reflection of objectively working historical forces and would therefore be carried out as a predestined outcome of historical development. Marx further thought that the proletariat was that 'material force' whose historical task was to realise his philosophy. When one bears all of this in mind, it is easy to see why Marx found the theories of List, particularly his view of history and his program for the future, not only objectionable but aberrant … It was axiomatic to Marx that industrial progress intensified and sharpened the antagonism between the bourgeoisie and the proletariat, an antagonism that would in the immediate future explode in a violent revolution. List, in the meantime, preached class cooperation and solidarity in the building of a nation's power. Marx thought that the Industrial Revolution, and the concomitant rule of the bourgeoisie,

> promoted the unification of the world and obliterated national differences. (Communism, he thought, would abolish nations themselves.) List claimed that the same phenomenon, the Industrial Revolution, intensified national differences and exacerbated conflicts among nations. While Marx saw the necessity of workers uniting across nations against the bourgeoisie, List called for the unification of all segments of a nation against other nations. (1988, p. 4)

The belief in free trade and comparative advantages may have led classical economists to overlook the role of the nation-state (there are, of course, exceptions, for example in *The Wealth of Nations*, and also Ricardo's belief in capital immobility can be taken as a reference to a nationalist element, but not such as to refute the general attitude). According to List, classical economists' defence of free trade served the commercial interests of Britain in having open access to foreign markets (paraphrasing Joan Robinson (1966) and Carl Schmitt, it may be argued that List regarded free trade 'as the continuation of mercantilism in other forms'). According to Marx, however, classical economists were actually decoding the secular and cosmopolitan characteristics of capitalism (Marx, 1975 [1845]; Szporluk, 1988, p. 66 and *passim*), in particular the conflicting interests of capital and labour that, in his opinion, went beyond the provincial boundaries of national states. Marx's criticism of List also reveals the German national bourgeoisie's interests behind List's vivid description of national identities (there is a similarity with Adam Smith's criticism of mercantilist writers as prejudiced defenders of merchants' interests). According to Desai (2012), Marx accuses List of hypocrisy: 'As a spokesperson for a capitalist developmental state, List was not concerned with class exploitation, only with national exploitation: "However much the individual bourgeois fights against the others, as a class the bourgeois has a common interest, and this community of interest, which is directed against the proletariat inside the country, is directed against the bourgeois of other nations outside the country. This the bourgeois calls his nationality"' (Desai, 2012, p. 62, quotation from Marx (1975 [1845], p. 281).

The expectation of a forthcoming revolution in Britain and its generalisation elsewhere presumably led Marx to dismiss the importance of the development of backward nations and to regard nationalism as an impediment to revolution rather than as a necessary historical passage to developed capitalism. For Marx, the notions of political community and national identity are illusions and false consciousness, like religion (Szporluk, 1988, p. 58). Without discarding Marx's criticism of the reactionary aspects of nationalism and the global nature of capitalism (the first part of *The Communist Manifesto* is a tribute to global capitalism), we should not forget that history vindicated the factual relevance of List's arguments about nation-states in the economic and political fields.[24] For instance, most socialist revolutions overlapped with struggles for national independence, and practical examples of international labour solidarity are rare, to say the least.

In Germany, the most representative exponent of the Young Historical School, Gustav Schmoller, also distanced himself from methodological individualism in his famous 'The Mercantile System' (1897): 'The idea that economic life has ever been a process mainly dependent on individual action, an idea based on the impression that it is merely concerned with methods of satisfying individual needs, is mistaken with regard to all stages of human civilisation, and in some respects it is more mistaken the further we go back' (*ibid.*, p. 4). Schmoller's perspective is, of course, very different and somehow opposite to Marx's. Having its roots in Cameralism, German Historicism and the Romantic movements, Schmoller's Historical School views the nation-state as the supreme expression of human belonging to superior organisms (cf. Riha, 1985, chs 4 and 5). Schmoller talks of 'real *political* economies as unified organisms, the centre of which should be, not merely a state policy reaching out in all directions, but rather the living heart-beat of a united sentiment' (p. 50, italics in the original). And here we find the famous definition of mercantilism, later adopted by Heckscher:

> in its innermost kernel [mercantilism] is nothing but state making – not state making in a narrow sense, but state making and national-economy making at the same time; state making in the modern sense ... The essence of the system lies not in tariff barriers, protective duties, or navigation laws; but in something far greater: the total transformation of society and its organisation, the state and its institutions, the replacement of a local territorial economic policy with that of the national state. (*Ibid.*, pp. 50–1)

This is ideology, of course, but development economists, particularly Gerschenkron (1962, p. 24), later underlined the importance of nationalist ideologies 'igniting the imaginations' of people for the mobilisation of national resources in the early stages of industrialisation. The reason why developmental bourgeoisies emerge from previous social relations of production in some nation-states, i.e., how certain pre-industrial social relations of production generate progressive pro-growth dominant classes, interested in a developmental state, are questions that have not yet been much explored. Neoclassical institutionalists seem unable to go beyond the mantra of the central role played by the protection of property rights in igniting growth (e.g. Acemoglu and Robinson, 2012). This may just be part of a story, the best of which has still to be written.[25]

10.7 Conclusions

Research on the recent European financial crisis has prompted exploration of the harmonic and disharmonic views of international economic relations. The former, more liberal view is based on the Ricardian and neoclassical trade theories. The latter is derived from pre-Smithian mercantilist conflict views of international trade. Here we have investigated the contribution that Sraffian theory can offer the latter stream of thought. This contribution cannot be underestimated since it provides a rigorous analytical rebuttal of the neoclassical theory of international trade and capital flows and supports the existence of absolute advantages, which are a source of potential trade conflict between nations. Kaleckian theory may also vindicate mercantilist attention to trade surplus. Further research is needed on the topics surveyed in this chapter. What is at stake is the space in the economics profession left by the dominant theory to these promising fields of research.

Notes

1. Conflict views do not exclude harmonic domestic and international arrangements. In fact, the social-democratic Scandinavian compromise is based on a conflict view of distribution. In the conflict approaches harmony is a subjective political result. In marginalism, harmony is an objective natural outcome of laissez-faire.
2. I regard these treatises, written in an age that saw the affirmation of nationalism, as representative of the later *reception* of mercantilist ideas. I am more interested in this reception that in the controversies on mercantilism. For example, among historians of economic thought, Coleman (1957) did not regard it as a systematic body of ideas while Bob Coats (1958) defended this view.
3. Importantly, Heckscher points out that Smith, Ricardo and Malthus perceived objective disharmony, although they did not believe that public interference would improve things much (*ibid.*, pp. 328–9).
4. See the authors quoted by Furniss (1923, pp. 25–6), for instance Chamberlen (1649): 'This may be a note to all men, especially to statesmen to look no more upon the poor as a burden but as the richest treasure of a nation, if orderly and well-employed. Which is the more manifest if we consider first,

that though they multiply more than the rich they do not only feed and clothe themselves but the rich men are fed and clothed and grow rich *by what they get out of the poor's labor over and above their maintenance*. Secondly, that the poor bear a greater burden of taxes in the city and elsewhere. For the rich either abate what they get out of the poor's labor or (which is worse) permit them to starve for want of employment' (quoted by Furniss, p. 25, italics added). And Bellers: 'Regularly laboring people are the kingdom's greatest treasure and strength, for without laborers there can be no lords; and *if the poor laborers did not raise much food and manufacture than what did subsist themselves*, every gentleman must be a laborer and every idle man must starve' (quoted by Furniss, p. 25, italics added). Johnson (1937, p. 240) quotes Dudley North (1691) who also advances a clear idea of social surplus: '[Some labourers] are more provident, other more profuse…[some] raise more fruits from the earth, than they consume in supplying their own occasions; and a surplus remains with them and is property of the riches'.

5. Mercantilists oscillate between conceiving unemployment as involuntary, e.g. Davenant: 'a defect in our constitution that many continue in wretched poverty for want of employment, though willing enough to undertake it' (quoted by Furniss 1920, p. 82); and the idea of unemployment as a sin (*ibid*., ch. 4).
6. As noted above, according to Petty the necessities produced by one section of the population also sustain those employed in the export sector. We find here a clear coordination between internal and external surpluses.
7. Suviranta, 1923, pp. 122–3, 161; Heckscher (1955 [II], p. 124): 'the power of creating wealth is more important than the wealth itself'.
8. On Kalecki and the 'Sraffian supermultiplier' approach to accumulation theory, see Cesaratto (2012a).
9. 'When people had once arrived at the view that a surplus of goods was something undesirable, the connection between this and the amount of employment followed inevitably' (Heckscher, 1955 [II], p.122. Notoriously, in Heckscher a foreign trade surplus was necessary to get rid of what he named 'fear of goods' or 'fear of redundant stocks' (*ibid*., p. 59, fn. 3).
10. Serrano (2008, p. 14) criticises Kalecki for this partially unfortunate passage: 'If exports increase and at the same time there is an equal increase in imports, overall profits remain unchanged; international trade is boosted, *but production in the country does not increase*, nor will there be any inducement for expansion of investment activity' (Kalecki, 1934, p. 16, my italics; see also 1967, p. 152). No doubt there is an ultra-mercantilist element in this sentence: international trade is a zero-sum game. Indeed, although Kalecki correctly points out that only a trade surplus generates profits, he misses the point that an increase in exports, even if accompanied by a corresponding rise in imports, nonetheless provokes an equivalent increase in domestic output: a sort of 'balanced foreign trade-budget theorem'. However, Kalecki is not wrong when he argues that for a given output, the larger the profit share and the lower the share of profits that capitalists consume or invest, the larger the trade surplus necessary to realise profits.
11. Heckscher (1955 [II], p. 317) argues that the 'obsession with power also had this result, that the interest was taken not in the absolute total of commerce nor in the utility which it represented to the inhabitants of a particular

country, but only in the superiority gained over other countries, irrespective of whether there was no absolute increase at all or perhaps even an absolute decline'. The best quotation he provides is from an important German mercantilist: 'Whether a nation be to-day mighty and rich or not depends not on the abundance or scarcity of its powers or riches, but principally on whether its neighbours possess more or less than it. For power and riches have become a relative matter, dependent on being weaker and poorer than others' (*ibid.*, p. 22). Locke would hold a similar thesis (*ibid.*, pp. 22–3).

12. Heckscher distinguishes between the Romanticist notion of nationalism concerned with traditions, ethnicity etc., and the secular mercantilist identification of it with the state interest.
13. Mercantilism has also not disappeared in lay(wo)men's and politicians' preoccupation with foreign competition. Krugman (1997) devoted a book to disproving these preoccupations. Here we ignore other important traditions in international economic relations, for instance that of Wallerstein.
14. Cf. the symposium 'American Realism and the Real World' (2003) *Review of International Studies*: 29.
15. An inspiring figure of modern political realism, Reinhold Niebuhr (1892–1971), an American Protestant priest, wrote: 'Power sacrifices justice to peace within the community and destroys peace between communities' (quoted by Donnelly, 2000, p. 27).
16. '[L]iberals believe that trade and economic intercourse are a source of peaceful relations among nations because their mutual benefits of trade and expanding interdependence among national economies will tend to foster cooperative relations. Whereas politics tends to divide, economics tends to unite people' (Gilpin, 1987, p. 31).
17. This superficial criticism is also shared by many 'post-Keynesian' economists.
18. In a famous book, polymath Jared Diamond (2005 [1997]) regards the production of an economic surplus as the trigger of human civilisation. He presents a theory of human evolution that recalls the 'four stages' found in classical authors, such as Turgot and Smith (Meek, 1971). See also below n. 25.
19. Adam Smith was quite aware of the social nature of individuals (see Cesaratto, 1996, for a comparison with Schumpeter).
20. Carr (1981 [1939]) wrote: 'Economic forces are in fact political forces. Economics can be treated neither as a minor accessory of history, nor as an independent science in the light of which history can be interpreted. Much confusion would have been saved by a general return to the term "political economy", which was given the new science by Adam Smith himself and not abandoned in favour of the abstract "economics", even in Great Britain itself, till the closing years of the nineteenth century. The science of economics presupposes a given political order, and cannot be profitably studied in isolation from politics' (*ibid.*, p. 108).
21. Traditional IPE is also interested in studying how ideologies are formed, a field called 'constructivism' (cf. Cohen, 2008, pp. 131–2). In this regard, IPE Marxist scholar Robert Cox particularly emphasised the contribution of Gramsci (see Cohen, 2008, p. 90).
22. For instance, assuming that with unemployed labour wages are close to subsistence level determined on the basis of historically determined social norms.

23. Ricardo himself limited the validity of his theory of comparative advantages to the case of no capital mobility. In this light, the anti-Ricardianism of colleagues of 'the other canon', mentioned at the beginning of this chapter, appears particularly misplaced.
24. Indeed, subsequent Marxist literature took the imperialist clash between the main economic powers into great consideration; I did not consider this development here (e.g. Brewer, 1980). Rosa Luxemburg's 'external markets' as a necessity for core-capitalist countries to realise social surplus, a view recovered by Kalecki (1967) that may vindicate some mercantilist insights, was born in precisely this context.
25. Meek (1976) points out that both Turgot and Smith regarded the protection of property rights as a result of development rather than a cause of it. In a similar vein, in a recent review of Acemoglu and Robinson (2012), Diamond (2012) is also very critical of these two neoclassical institutionalists, pointing out a causal chain that goes from the emergence of a food surplus to lawful complex societies: 'The chain of causation leading slowly from productive agriculture to government, state formation, complex institutions, and wealth involved agriculturally driven population explosions and accumulations of food surpluses, leading in turn to the need for centralized decision-making in societies much too populous for decision-making by face-to-face discussions involving all citizens, and the possibility of using the food surpluses to support kings and their bureaucrats.'

References

Aspromourgos, A. (1996) *On the Origins of Classical Economics: Distribution and Value from William Petty to Adam Smith* (London and New York: Routledge).

Acemoglu, D. and Robinson, J. A. (2012) *Why Nations Fail: The Origins of Power, Prosperity, and Poverty* (New York: Crown).

Brancaccio, E. and Fontana, G. (eds) (2011) *The Global Economic Crisis. New Perspectives on the Critique of Economic Theory and Policy* (London: Routledge).

Brewer, A. (1980) *Marxist Theories of Imperialism – A Critical Survey* (London: Routledge & Kegan Paul).

Brewer, A. (1985) 'Trade with fixed real wages and mobile capital', *Journal of International Economics*, XVIII: 177–86.

Carr, E. H. (1981 [1939]) *The Twenty Years' Crisis – 1919–1939* (London: Macmillan).

Cesaratto, S. (1996) 'Alternative Perspectives on Division of Labour, Structural Change and Economic Growth', in C. De Bresson (ed.), *Economic Interdependence and Innovative Activity: An Input-output Analysis* (Edward Elgar) (rev. version: *Studi Economici*, 67, 1999).

Cesaratto, S. (2011) 'Europe, German Mercantilism and the Current Crisis', in E. Brancaccio and G. Fontana (eds), *The Global Economic Crisis. New Perspectives on the Critique of Economic Theory and Policy* (London: Routledge).

Cesaratto, S. (2012a) 'Neo-Kaleckian and Sraffian controversies on accumulation theory', University of Siena Working Paper, 650, mimeo.

Cesaratto, S. (2012b) 'Controversial and Novel Features of the Eurozone Crisis as a Balance of Payment Crisis', in O. Dejuán, J. Uxó and E. Febrero

(2013), *Post Keynesian Views of the Economic Crisis and its Remedies* (London: Routledge).
Cesaratto, S. (2013) 'The Endless Eurozone Crisis, Where Do We Stand? A Classical-Kaleckian Overview', Dipartimento di Economia Politica e Statistica, Università di Siena, Working paper 671, http://www.econ-pol.unisi.it/dipartimento/it/node/1813 (forthcoming in *Studi Economici*).
Cesaratto, S. and Stirati, A. (2011) 'Is Germany a mercantilist country? Germany in the European and global crises', *International Journal of Political Economy*, XXXIX: 56–87.
Coats, A. W. (1958) 'In Defence of Heckscher and the Idea of Mercantilism', *Scandinavian Economic History Review*, V, 173–87 (repr. in ID, *On the History of Economic Thought*, vol. I, London: Routledge, 1992).
Cohen, B. J. (2008) *International Political Economy: An Intellectual History* (Princeton: Princeton University Press).
Coleman, D. C. (1957) 'Eli Heckscher and the idea of mercantilism', *Scandinavian Economic History Review*, V: 1–25 (repr. in ID [ed.], *Revisions in Mercantilism*, London: Methuen, 1969).
Cronin, J. E. (1996) *The World the Cold War Made: Order, Chaos and the Return of History* (London: Routledge).
Crouch, C. (2008) 'What will follow the demise of privatised Keynesianism', *Political Quarterly*, LXXIX: 476-87.
Dalziel, P. C. and Harcourt, G. C. (1997) 'A note on "Mr Meade's Relation" and international capital movements', *Cambridge Journal of Economics*, XXI: 621–31.
Desai, R. (2012) 'Marx, List, and the materiality of nations', *Rethinking Marxism*, XXIV: 47–67.
Diamond, J. (2005 [1997]) *Guns, Germs and Steel: A Short History of Everybody for the Last 13,000 Years* (London: Vintage).
Diamond, J. (2012) 'What makes countries rich or poor?', *New York Review of Books*, 7 June .
Donnelly, J. (2000) *Realism and International Relations* (Cambridge: Cambridge University Press).
Ehrlich, H. B. (1955) 'British mercantilist theories of profit', *American Journal of Economics and Sociology*, XIV: 377–86
Farrell, H. and Finnemore, M. (2009) 'Ontology, methodology, and causation in the American school of international political economy', *Review of International Political Economy*, XVI: 58–71.
Franzese, R. J. and Hall, P. A. (2000) 'Institutional Dimensions of Coordinating Wage-Bargaining and Monetary Policy', in T. Iversen, J. Pontusson and D. Soskice (eds), *Unions, Employers, and Central Banks: Macroeconomic Coordination and Institutional Change in Social Market Economies* (Cambridge: Cambridge University Press): pp. 173–204.
Furniss, E. (1920) *The Position of the Laborer in a System of Nationalism: A Study in the Labor Theories of the Later English Mercantilists* (Boston: Houghton Mifflin Company).
Garegnani, P. (1983) 'Notes on Consumption, Investment and Effective Demand', in J. Eatwell and M. Milgate (eds), (1983), *Keynes's Economics and the Theory of Value and Distribution* (London: Duckworth).
Garegnani, P. (1984) 'Value and distribution in the classical economists and Marx', *Oxford Economic Papers*, XXXVI: 291–325.

Gerschenkron, A. (1962) *Economic Backwardness in Historical Perspective: A Book of Essays* (Cambridge, MA: Belknap Press of Harvard University Press).
Gilpin, R. G. (1984) 'Richness of the tradition of political realism', *International Organization*, XXXVIII: 287–304.
Gilpin, R. G. (1987) *The Political Economy of International Relations* (Princeton: Princeton University Press).
Heckscher, E. (1955) *Mercantilism*, 2nd edn (1st edn in Swedish, 1931) (London: George Allen & Unwin).
Hirschman, A. (1977) *The Passions and the Interests: Political Arguments For Capitalism Before Its Triumph* (Princeton: Princeton University Press).
Johnson, E. A. J. (1937) *Predecessors of Adam Smith* (London: P.S. King & Son).
Jonung, L. and Drea, E. (2009) 'The euro: It can't happen, It's a bad idea, It won't last. US economists on the EMU', www.ec.europa.eu/economy_finance/publications
Kalecki, M. (1967) 'The problem of effective demand with Tugan-Baranovski and Rosa Luxemburg', in ID, *Selected Essays on the Dynamics of the Capitalist Economy* (1971) (Cambridge: Cambridge University Press).
Katzenstein, P. J. (1977) 'Domestic and international forces and strategies of foreign economic policy', *International Organization*, XXXI (issue no. 4: 'Between Power and Plenty: Foreign Economic Policies of Advanced Industrial States'): 587–606.
Katzenstein, P. J. (2009) 'Mid-Atlantic: Sitting on the knife's sharp edge', *Review of International Political Economy*, XVI: 122–35.
Keohane, R. O. (2009) 'The old IPE and the new', *Review of International Political Economy*, XVI: 34–46.
Kindleberger, C. P. (1981) 'Dominance and leadership in the international economy', *International Studies Quarterly*, XXV: 242–54.
Kriesler, P. and Halevi, J. (1996) 'Asia, Japan and the internalization of effective demand', *Economies and Sociétés – Monnaie et production*, X: 301–20.
Lake, D. A. (2009) 'TRIPs across the Atlantic: theory and epistemology in IPE', *Review of International Political Economy*, XVI: 47–57.
List, F. (1841) *The National System of Political Economy* (1991 edn) (New Jersey: Augustus M. Kelly).
Marx, K. (1975 [1845]) Draft of an Article on Friedrich List's book: *Das Nationale System der Politischen Oekonomie*, *Marx-Engels Collected Works*, Vol. 4: Marx and Engels: 1844–45 (New York: International Publishers).
Marx, K. (1973 [1857]) *Grundrisse – Foundations of the Critique of Political Economy (Rough Draft)* (Harmondsworth: Penguin Books).
Marx, K. (1977 [1859] *A Contribution to the Critique of Political Economy* (Moscow: Progress Publishers).
Marx, K. (1969 [1861–3]) *Theories of Surplus Value*, Part 1 (London: Lawrence & Wishart).
Meek, R. L. (1971) 'Smith, Turgot and the Four Stages Theory', *History of Political Economy*, 3: 9–27.
Meek, R. L. (1976) *Social Science and the Ignoble Savage* (Cambridge: Cambridge University Press).
Mundell, R. A. (1957) 'International Trade and Factor Mobility', *American Economic Review*, XLVII, 321–35.
Parrinello, S. (2009) 'The Notion of National Competitiveness in a Global Economy', in D. Vint, S. Metcalfe, H. Kurz, P. Samuelson and N. Salvadori (eds), *Essays in Honour of Ian Steedman* (London: Routledge).

Riha, T. (1985) *German Political Economy: The History of an Alternative Economics* (Bradford: MCB University Press).

Robinson, J. (1966) *The New Mercantilism* (Cambridge: Cambridge University Press).

Schmoller, G. (1897) *The Mercantile System and Its Historical Significance* (New York: Macmillan).

Serrano, F. (2008) 'Kalecki y la economia americana en los anos 2000', *Circus*, I: 7–24.

Sraffa, P. (1951) 'Introduction to Ricardo's Principles', in P. Sraffa and M. H. Dobb (eds) *The Works and Correspondence of David Ricardo*, Vol. 1 (Cambridge: Cambridge University Press).

Steedman, I. (1979a) *Trade amongst Growing Economies* (Cambridge: Cambridge University Press).

Steedman, I. (ed.) (1979b) *Fundamental Issues in Trade Theory* (London: Macmillan).

Strange, S. (1970) 'International economics and international relations: a case of mutual neglect', *International Affairs*, XLVI: 304–15.

Smith, A. (1979 [1776]) *An Inquiry into the Nature and Causes of the Wealth of Nations* (Oxford: Clarendon Press).

Suviranta, B. (1923) *The Theory of Balance of Trade in England – A Study in Mercantilism* (Helsinki: Annales Academiae Scientiarum Fennicae, B-XVII).

Szporluk, R. (1988) *Communism and Nationalism: Karl Marx Versus Friedrich List* (Oxford: Oxford University Press).

Viner, J. (1948) 'Power and plenty as objectives of foreign policy in the seventeenth and eighteenth centuries', *World Politics*, I: 1–29.

11
Lessons from the Crisis: the Macroeconomics of Leverage

Paolo Leon

11.1 Introduction

Let me advance a possible explanation of the current stagnation and of the preceding crisis and crash. The 2007 crash, generally thought to be of financial origin, is imputed to speculation without regulation and/or to the inevitable eventual failure of otherwise self-fulfilling expectations: surely, there must be more to it, and an ample literature has accordingly developed. In this chapter I shall merely offer hypotheses, unaccompanied by statistical proof. Logic mistakes are likely, but my purpose is to stimulate research by others. In the following argument I will discuss events in the USA and the UK, and to a lesser extent continental Europe (perhaps Japan) – the argument is restricted to something that looks like a closed economy; and, in effect, if a China was not there, I don't know what part of the following could stand.

As anticipated, much of the argument has already been addressed by a growing literature, part of which has tried to put together inequality, leverage and crises. One such work (Kumhof and Rancière, 2010) uses a set of non-standard (i.e. Neo-Keynesian) neoclassical assumptions to obtain a basically microeconomic relationship between income distribution and the crisis, leading to macroeconomic results. Similarly, a previous attempt at reconciling household debt and inequality (Iacoviello, 2008), attributes the growth in debt to inequality, or to a decline in the bargaining power of wage earners (called 'impatient agents'). Quite apart from the method used, I consider this causation the weak link of these works, given that in the periods examined by the authors, full employment was prevalent. Many other economists have worked on the relationship between excessive leverage and the subsequent crash, but most of this literature is descriptive, and basically proposes that anything that rises too much, sooner or later must fall. To my knowledge, the only non-neoclassical work that has put forward a reasonable explanation of the crash, by linking household debt and income distribution in a macroeconomic setting, is that of Barba and Pivetti (2009), and my approach is indeed similar to theirs. However, these authors also consider that household debt results from weak bargaining power of the wage earners. I shall try to argue that it is rather precisely *the increase in household debt that determines a worsening distribution between wages and all other incomes*.

11.2 Wealth and income

One aspect of the crisis, related to the long period preceding it, is the well established fact that households (as well as enterprises) were able *to transform the increase in their wealth*, derived from ever-rising stock and residential market values, *into new revenue*: either by speculating on their assets or by increasing their debt or both, in bull times. Wealth is a stock while income is a flow: but when wealth market values increase regularly, *wealth becomes a flow*, either because it is perceived as income, or because the increase in value can be made liquid by borrowing.[1] We may call this the macroeconomics of leverage and, as we shall see below, it is not a wealth effect *à la* Pigou. When households see this flow rising, they believe it is income and increase their consumption: they do not really need to increase their savings, as their wealth continues to rise. Also, when enterprises see their wealth flow rising, they consider it as new revenue and will push both their investment and their current expenditures, because whatever liquidity constraint prevailed beforehand, it is now removed by the increase in their wealth. Banks and financial intermediaries do not need to worry about the confusion between income and wealth, because as effective demand rises, employment and GDP grow, the profitability of enterprises increases, and creditworthiness at large is not endangered. When employment grows, so will household incomes (over and above the transformed wealth) as well as their expenditures – again, their savings will not increase, because its propensity is stifled by the regular increase in the market value of their wealth. Furthermore, as soon as employment booms, home buying increases and, with it, home values, thus making everybody richer – including firms, whose propensity to invest is enhanced by the rise in the value of stocks and buildings. One might think that these circumstances are occasional, that the crisis is a one-time event, that therefore it does not entail a disruption in the established economic theory, based on automatic market correction. As we shall see, this is not so. As far as I know, the macroeconomics of leverage has not been dealt with sufficiently in economic theory (with the possible historic exceptions of Hilferding, Galbraith and, more recently, Minsky), mostly because in methodological individualism, including models with heterogeneity of agents, debt is never a macroeconomic problem, as assets and liabilities, for the economy as a whole, are always equal and wealth is always zero (in a closed economy or in the world at large). This is a truism, however, because what happens in bull times is that both assets and liabilities increase, but it is the increase in liabilities that determines the increase in assets (the market values for houses and securities), just as lending by banks determines deposits. This is, indeed, the macroeconomics of leverage.

11.3 Income distribution

In this gargantuan (i.e., milk and honey) picture, one fact stands out: the share of labour income in GDP worsens (inequality prevails). This outcome, which is considered as exogenous by the literature, is in fact unexpected, given the increase in employment and, possibly, a tighter labour market, at least following the crisis of the early 1990s. In fact, all incomes increase, but the upper deciles increase more then the lower ones: a typical Pareto improvement with worsening distribution. Some authors divide impatient agents from patient ones, with the first that need to spend more rapidly than the second, thus increasing their debt, which is financed by the second – this is, however, the ancient abstinence argument, which implies that credit and debt cancel out in the economy as a whole, leaving the growth of leverage unexplained. There are many real reasons for this long-term trend, and many derive from increased competition in the labour market: growing female participation, immigration, liberalisation in public services, flexibility in labour laws. In all these cases, as GDP and labour demand grow, more labour supply is made available, and lesser unit wages are paid (or slower growth in wages is experienced).

All the elements that produce a weakening of the bargaining power of workers were in fact present before the crisis, but nobody has explained why it was possible for such elements to continue reducing the bargaining power of workers, when full employment conditions are prevailing: *we miss a link between the bull market and labour weakness*. Galbraith proposed a simple way to explain this phenomenon before the Great Depression: speculation benefits those that own capital (the patient ones, some would say today), and during the boom distribution works against those that only earn a wage (the impatient ones). This is more a description than an explanation, because it does not account for the decline in bargaining power as the economy booms. In addition, if everyone owns some amount of capital, Galbraith's description does not square with the 2007 crash.

I propose, instead, that income distribution is linked to the transformation of wealth into income: for workers, the increase in *wage income is less needed, when the increase of wealth income takes place, and trade unions appear less powerful*. It is often forgotten that when bargaining for wages and work conditions, workers incur the risk of being fired. Individual bargaining is always risky for workers with a family, even in times of boom, because during a spell of unemployment they receive a fraction of their original income; in addition, new jobs are available, but they are not necessarily in the same location in which the family was living, with added costs. More importantly, when each worker bargains by himself, the power imbalance between him and his possible employer is strong, even in full employment, because workers are in competition with each other. Such situations should normally increase affiliation to the unions, but these organisations tend to be egalitarian in bargaining, in order to satisfy the greatest possible number of workers; as a result, the demand for affiliation becomes weaker precisely in conditions of full employment, when each worker values himself more than he values what the unions offer. It appears, therefore, that an increase in wealth that can be transformed into income is much more reassuring for the wage earner, and does not imply any confrontation with the employer. Unknowingly, enterprises *will face a lower level of conflict with their workers* for the distribution of productivity gains, which therefore do not accrue to workers in the amount needed for dynamic equilibrium. This seems a reasonable explanation of the increase in inequality, in full employment conditions. Perhaps, this phenomenon can even induce a cultural change, if workers perceive themselves as 'human capitalists' or, simply, 'rich'.

11.4 The crash

Our story can now resume. The less wages rise *vis-à-vis* the value of the workers' wealth (their home, in general), the more they must increase their debt so as to maintain their relative standard of living – but it is the maintenance of the relative distribution of income that ensures both social cohesion and sufficient effective demand. Conventional wisdom has it that whatever the generosity of the financial sector, there is always some limiting risk beyond which each household (and each firm) cannot hope to continue borrowing: an argument that can never identify precisely where the borrowing barrier stands. Rather, each historic boom shows a financial industry bent on increasing liquidity by diluting the risk of bad debts through the concoction of ever new securities reflecting some average risk, rather than the risk of any specific asset. With capital markets booming, this is a rational behaviour, because no debt is bad if the value of the corresponding asset continues to increase.

The existence of a boom barrier is possible, but it looks like no rational expectation can discover. The crash, then, must be attributed to more substantial causes. A preferable explanation has to do with the economics of leverage in a context of worsening income distribution. In a bull market, *wealth begets wealth*, until wealth market values rise: this, again, is what can be defined as leverage. Households, however, with wages that increase less than productivity and less than GDP, must maintain their relative consumption level (as well as their social status) via new debt, based on the increased value of their homes. If so, since they use part of their wealth for consumption, they can only use part of it as leverage to acquire new property: for them *wealth does not fully beget wealth*. Thus, when households can no longer buy in the housing (or in the capital) markets, housing prices and stock indices decline. As soon as the markets start dwindling, firms will also see their wealth decrease and this implies a reduction in leverage and in perceived revenues. As revenues of all kinds slow down or simply cease to increase, effective demand decreases or ceases to increase and the multiplier stops working or works negatively: an economic crisis, and not simply a financial one, sets in.

11.5 Economic policies and the crash

We have to figure out what causes the transformation of wealth into income. I think that what happened was the result of a new series of policies that, since the early 1980s, had disregarded the need to keep effective demand nearer potential output (a position similar to the pre-1929 policies), either because a self-equilibrating 'natural rate of growth' theory was used for policy purposes or, more simply, because it was believed that supply creates its own demand (Roncaglia, 2010). Simplifying, the following appear to be the tenets of the new policies:

- if a depression sets in, then labour becomes cheaper, firms' expected real incomes increase, and their demand for labour grows, restoring equilibrium. If this does not work, and will not work because the downturn in wages reduces the volume of sales, then

- liberalisation will: the transaction costs for the economy as a whole will decline, real incomes increase, and equilibrium is restored. If liberalisation does not work, and will not work because the decline in transaction costs is tantamount to a decline in employment, wages and sales, then
- labour market flexibility will: removing wage earners' rents, due to labour rigidities, will raise firms' expected incomes and profits, restoring equilibrium: a non-event, because realised profits will not increase since demand out of wages declines (or fails to rise).

All these instances are based on a well-known fallacy: if costs of production are reduced, then demand is assumed to pick up because prices should decrease and real incomes increase (this is Pigou's real balance effect); but since the cost reduction translates into lower expenditure on inputs, then whoever supplies the inputs will face lower sales and revenues. The microeconomic foundation of macroeconomics fails, and no fixed-price assumption is needed.

It is this failure to recognise the role of effective demand, and therefore the benefits of government intervention, that has prompted the alternative, through deregulation (liberalisation) policies in the financial markets:[2] if regulation on limiting the risks in the trade of securities is relaxed, banks and financial agents will provide risky paper, but will need to show that such paper can be used to leverage other paper (as shown above), producing real effects. *Speculation is the substitute for effective demand*. Moreover, deregulation is necessary, bur also weak regulation, asymmetric powers in regulation and conflicts of interest, gullible investors, simply laissez faire, will produce the toxic securities encountered in the recent crash – possibly as in 1929.

In fact, the most important deregulation that has taken place over the last thirty years is the independence of the central bank (again, a return to the pre-1929 period): a new policy occasioned by the monetarist revival of the late 1970s and early 1980s. Such independence reduced the sovereignty of governments in issuing fiduciary money: central banks were not obliged to buy government paper to cover public deficits and governments had to sell their bonds on the market. As the bond supply increased, crowding the market, interest rates rose, public deficits and debts grew and this made it more difficult for governments to sustain effective demand. At this time, the early 1980s, a serious crisis ensued, with a large increase in unemployment. Contrary to their basic tenets, governments found that their policies were wrong and that economic activity was to be sustained, but since their economic policies did not include acting on demand (or, if they did, they had to take account of their enlarged public debt), new deregulation was enacted that further reduced the role of the central bank and of the banking system in regulating money supply. Since then, nobody has really been in control of bank money, and banks lend on the basis of their own capital. Each bank is therefore pushed to create its own wealth with the help of securitisation: buying securities which can reduce the risk of underlying assets and issuing (forging, sometimes) securities bearing uncertainty and incalculable risks, banks increased their own leveraging power. The deregulation in banking spearheaded the deregulation of the financial markets. As we have seen, speculation does have real effects: the greater financial agility, by easing the issuing of purely 'fiduciary' securities (a form of endogenous money), produces wealth, leverage and effective demand – *at least until the worsening distribution of income causes a market crash*.

This story has a moral: reducing the role of the government, through deregulation, means that no economic policy aimed at preserving full employment is really feasible, with the exception of letting speculative markets grow; and the more speculative they are, the more they affect demand. This effect can only be maintained, however, if the distribution of income does not worsen: any gain in the capital and housing markets could be considered as an increase in overall productivity and, thus, could be distributed to all the workers in the economy, not just to those that operate in the financial or housing sectors. But since workers have in the meantime lost their bargaining power and the awareness of their real status (and no Wagner Act (USA, 1935) has been passed), such redistributive policy will not take place. If, as a reaction to the crash, new regulation will really discipline capital markets, then there is no possibility for speculation to raise effective demand, and for growth to resume. The economies are at or near a dead end: if they don't regulate, a crash is likely, if they do regulate, growth slows down.

11.6 Leverage and the consumption function

I hope that all this reasoning stands up. The inspiration is certainly Keynesian (not Neo-Keynesian), as witnessed by the emphasis on effective demand, on the central bank as the last resort for financing government deficits and on income distribution. The difference concerns speculation and wealth, because while Keynes thought that buying and selling existing securities has no real effects and wealth and income are separate, in my view speculation has real effects.

This is no small change to the Keynesian model: following a suggestion of Professor De Cecco, it implies that the consumption function is based not only on income, but also on leverage: and each time a bull market appears, leverage increases, consumption is influenced, and the bargaining power of wage earners declines. If these events could be applicable to the dynamic of the capitalist system, and not just to the recent boom and crash, then the rate of interest has no place in the consumption function, and therefore neither Modigliani's life cycle nor Friedman's permanent income hypotheses are relevant. Leverage is not dependent on the rate of interest, but on the deregulation of capital markets. To make this statement clearer, the financial market may show a pseudo-Pigou effect, if speculation is enhanced by lower interest rates which cause an increase in wealth, but one would need an ever-decreasing interest rate to produce ever-increasing wealth. Instead, given the rate of interest, it is the rise in stock market indices which generates first new wealth and new incomes and finally their demise: not just once, but whenever the government is kept from targeting effective demand.

Notes

1. I owe this clarification to an unknown referee.
2. It was President Clinton, a Democrat, who relaxed whatever rules were still extant in the financial sector and ran a budget surplus. As was once suggested by Gordon Brown, such policy can be defined as a mixture of Friedman and Keynes, and is one tenet of the 'third way' of the British Labour party; he forgot that it is like mixing oil and water.

References

Barba, A. and Pivetti, M. (2009) 'Rising household debt: its causes and macroeconomic implications – a long period analysis', *Cambridge Journal of Economics*, 33: 1.

Dew-Becker, I. and Gordon, R. J. (2005) 'Where did the Productivity Growth Go? Inflation Dynamic and the Distribution of Income', *Brookings Paper in Economic Activity* 2.

Iacoviello, M. (2008) 'Household debt and income inequality 1963–2003', *Journal of Money, Credit and Banking*, 40: 5.

Keys, B. J., Mukherjee, T., Seru, A. and Vig, V. (2010) 'Did securitization lead to lax screening? Evidence from subprime loans', *Quarterly Journal of Economics*, 125: 1.

Kumhof, M. and Rancière, R. (2010) 'Inequality, Leverage and Crises', IMF Working Paper, November.

Roncaglia, A. (2010) 'Le origini culturali della crisi', *Moneta e Credito*, 63(250): 107–18.