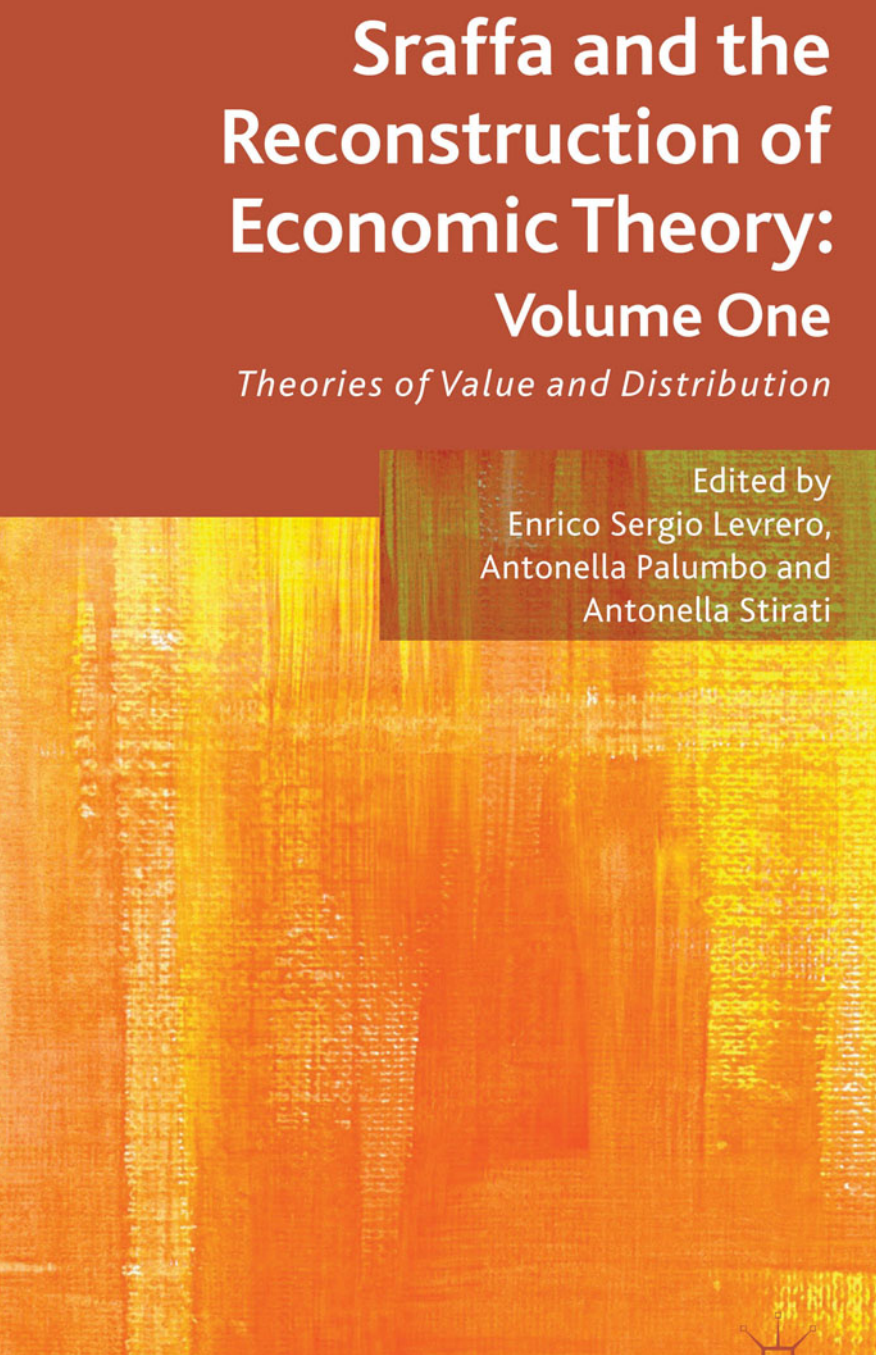
Sraffa and the Reconstruction of Economic Theory: Volume One

Theories of Value and Distribution

Edited by
Enrico Sergio Levrero,
Antonella Palumbo and
Antonella Stirati



Sraffa and the Reconstruction of Economic Theory: Volume One

Theories of Value and Distribution

Edited by
Enrico Sergio Levrero
Antonella Palumbo
and
Antonella Stirati

2013

palgrave
macmillan

Contents

List of Tables and Figures

Tables

Figures

List of Contributors

Fabián Amico, Luján University, Buenos Aires

Aldo Barba, 'Federico II' University of Naples

Enrico Bellino, Università Cattolica, Milan

Guglielmo Chiodi, Sapienza University of Rome

Leonardo Ditta, University of Perugia

Alejandro Fiorito, Luján University, Buenos Aires

Saverio M. Fratini, Roma Tre University

Pierangelo Garegnani, Roma Tre University

Harvey Gram, Queens College, City University of New York

Enrico Sergio Levrero, Roma Tre University

Sergio Parrinello, Sapienza University of Rome

Massimo Pivetti, Sapienza University of Rome

Bertram Schefold, FB, Wirtschaftswissenschaften,
J.W. Goethe-Universität, Frankfurt am Main

Antonella Stirati, Roma Tre University

Paolo Trabucchi, Roma Tre University

Introduction

Enrico Sergio Levrero

It is widely recognised that Sraffa's *Production of Commodities by Means of Commodities* (*PCMC*) is not simply a variant of a class of linear models, nor an incomplete general equilibrium model with stationary prices. Together with Sraffa's 1951 'Introduction' to Ricardo's *Works*, Sraffa (1960) revived the old classical economists' approach to value and distribution, proving that a correct price theory can spring from within that approach, and illustrating properties of the price system and of the choice of techniques which have resulted in a sharp contrast with the foundations of neoclassical (or marginalist) theory.

These two aspects of Sraffa's work have been taken up and further developed in the 50 years following the publication of *PCMC*, thus tracing a path for the development of economic theory along lines different from those that have prevailed since the end of the nineteenth century with the rise to dominance of the neoclassical approach. On the critical side, logical difficulties inherent in the lack of a notion of 'capital' generally consistent with the working of this approach, and the paradoxical results originating from that lack, undermined the determination of the wage and profit rates as the result of the relative scarcity of the factors of production (see for example Garegnani, 1960 and 1990; Pasinetti, 1966; Harcourt, 1972; Petri, 2004). On the constructive side, important theoretical results have been achieved in clarifying the analytical structure and method of analysis of the classical or surplus theory, as well as the properties of the price system – proving, for instance, the existence (on the basis of sufficiently general hypotheses) of an inverse relation between the rate of profits and the wage rate, and marking advances in the fields of rent theory, fixed capital and joint production (for an appraisal of these results see, e.g., Garegnani, 1984; Kurz

and Salvadori, 1995; Pasinetti, 1980; Schefold, 1989; Steedman, 1979). Furthermore, progress has been made in determining the wage rate, the social product and the methods of production by developing elements already discernible in the classical economists and Marx, but also in Kalecki and Keynes (see, for instance, the contributions in Bharadwaj and Schefold, 1990, and in Ciccone *et al.*, 2011). This amounts to the determination of what Garegnani (2007) called the 'intermediate data' of the surplus approach, namely magnitudes taken as given for the purpose of deductively determining the rate of profits and its changes if some of those magnitudes were to change, but determined by historically specific, more inductive considerations in the overall corpus of analysis.

Despite these advances on both the critical and reconstruction sides opened up by Sraffa, the present theoretical situation continues to be characterised by the persistent dominance of the neoclassical approach and the wide use of the latter in the field of economic policies. To a great extent, the reasons are perhaps to be sought on ideological (rather than only analytical) grounds, of the kind which, after Ricardo's death, played a role in gradually leading his conflictual and non-harmonic view of distribution to be 'submerged and forgotten' within economic theory. Yet to some extent, the situation also reflects the work which has still to be done to construct economic theory and policy analysis along classical lines. Moreover, it reflects several misunderstandings which arose during the capital controversy of the 1960s and 1970s over the implications for neoclassical theory of the phenomena of reswitching and reverse capital deepening, as well as the shift to the notions of temporary and intertemporal equilibria that occurred after Hicks (1939). As Garegnani claims (see p. 16), this shift has actually given credence to the fact that the apparatus of demand and supply may still be valid as regards consistently determining prices and outputs in a capitalist economy – and hence credence also to the opinion that even 'the admittedly imperfect previous concepts' of the neoclassical theory might be used 'as workable approximations in applied work'.

The contributions in this volume endeavour to address the current situation regarding the theory of value and distribution on both the critical and constructive sides as opened up by Sraffa (1960). Part I (The Capital Controversy and General Equilibrium Analysis) deals with aspects of general equilibrium and capital theory; Part II (The Revival and Development of the Classical Theory of Distribution) advances the analysis of the determinants of income distribution along classical lines,

together with the related issue of the contrasting methods of analysis that characterise classical and marginalist theories.

In particular, with respect to Part I, the contributions of Garegnani, Gram, Fratini and Trabucchi all analyse from different angles the problems which the treatment of capital encounters in neoclassical theory, which ultimately stem from the fact that a measurement of capital independent of distribution cannot be effected by means of a single magnitude, but rather entails using a set of magnitudes, be they the physical quantities of the several capital goods or Wicksell's dated quantities of labour and land. As regards the substitution mechanisms between goods and factors of production, relations set up between product per worker, capital per worker and rate of profits may thus happen to be different from those required by neoclassical theory. As to the supply side of the theory, taking as exogenously given the available amounts of resources gives rise to the difficulty that, on the one hand, 'capital' cannot be taken as given in *value* terms when determining prices and the rate of profits, and, on the other, fluidity as to its physical composition is required to allow the latter to satisfy the conditions of equilibrium, including uniformity of effective rates of return on the capital goods supply prices.

As recalled by Gram in his contribution, Garegnani focuses precisely on the role of a conception of capital as a single magnitude in providing a foundation for the neoclassical notion of a general substitutability among factors and for a concept of equilibrium that makes a correspondence between theory and observation possible. In contrast to those who deny such a role, Garegnani first of all reconstructs how that conception has been key to the move from the classical to the neoclassical approach. The extension to all factors of production of the Malthusian theory of rent by means of 'a *variability* of the proportion of "capital" .... to labour .... in social production .... analogous to the classical variability of the proportion of labour (plus capital) to land in agriculture' (see p. 18) necessitated in fact a notion of capital as a single quantity, because the alternative methods of production differ 'more by *the kinds of capital goods* used than by the *proportion to labour in which each of them was employed'*. Moreover, Garegnani emphasises that such a notion allowed the expression of the capital endowment in a way consistent with its homogeneity for savers – that is, with the fact that for them capital goods are perfect substitutes in proportion to their values. This is crucial since, under free competition, the physical composition of capital will tend to adjust to ensure a uniform rate of return on capital goods' supply prices, and such an endogenous

determination of the composition of capital is compatible with the neoclassical premise of a given availability of capital only if the latter is understood to be given in terms of a single magnitude, and not in terms of amounts of different capital goods, as for instance in Walras. Otherwise, as Garegnani (1960) showed precisely with reference to Walras, the rates of returns on capital goods will generally diverge, which means that the equilibrium would lack the persistence necessary to ensure some kind of correspondence between theoretical and observable variables as implied in the normal positions of the economy as conceived of by Ricardo, Wicksell or even Walras.

It is actually the need to consider the rapid changes in the prices of productive services and commodities stemming from that impermanence of the equilibrium which since Hicks (1939) has impinged on the theory taking the expectations concerning future prices or present markets for future commodities into account in the determination of general equilibrium. But in this way the neoclassical approach comes to a standstill between the Scylla of the intertemporal equilibria, where the unrealistic assumption of complete futures markets is made, and the Charybdis of arbitrary (and possibly conflicting) subjective expectations which characterise temporary equilibria. Furthermore, Garegnani stresses that these equilibria cannot be centres of gravity of the actual magnitudes. In these equilibria, in fact, the endowments of the different kinds of capital goods are taken to be arbitrary (at least at the beginning of the time horizon considered), and would thus change quickly, depriving the equilibria of their role of attractor, unlike the case of a normal position of the economy, in which the repetition of transactions on the basis of virtually unchanged data can be assumed to allow for compensation of the accidental deviations of actual prices from normal prices, the latter thus emerging as some sort of an average of the former.

The strong link between the problems in the treatment of capital in the early versions of the neoclassical theory and the shift to its neo-Walrasian versions is not clearly acknowledged in the general equilibrium literature, and this has caused several misunderstandings in the controversy on capital theory. As Garegnani underlines, the problem of capital within the neoclassical theory does not pertain only to the construct of 'aggregate production function' or to a problem of aggregation versus general equilibrium, but rather involves two kinds of general equilibria: the traditional versus the Hicksian equilibrium. And as Garegnani again points out, the eclipse of the notion of normal equilibrium present in the versions of theory by Wicksell or even Walras has helped to hinder the real undermining aspect of the applicability of the new notions of equilibria, namely that

Hicks must assume that the economy is always in equilibrium due to 'the impermanence of the new equilibria and the resulting dynamics' (p. 29).

This point is also considered by Gram, who connects this impermanence with the phenomena of indeterminacy and 'inherent' instability of intertemporal equilibrium paths which arise when the distinction between 'demand' price and 'supply' price of reproducible inputs is taken into account. Moreover, the contributions of both Garegnani and Trabucchi underline the role played by Hicks (1939) in obscuring Walras's real inconsistencies and attributing the reasons of the switch to the new notions of equilibrium to an alleged need for a 'dynamic' theory and due consideration of expectations. Hence they criticise Hicks's identification of the normal positions of Walras or Wicksell with stationary states in which the incentive to net savings has disappeared, and observe that the dependence of current actions on future prices was not overlooked by the 'economists of the past', especially when dealing with short-run phenomena. As claimed by Garegnani (p. 27), what these economists actually did was only to ignore it in 'the relevant general context' of their analyses when a discrepancy between current and expected prices was believed to be due to accidental factors, since it 'would be averaged out into the *normal price* through the repetition of transactions'; but they took it into account by comparing 'the corresponding two normal positions' of the economy when the discrepancy was the result of a change in the data of an initial normal position. All this is thoroughly analysed by Trabucchi, who in this respect also discusses Lindahl's thought and criticises the (unproven) statements of Hicks about an alleged indeterminateness of the 'traditional' (or normal) equilibrium.

Some of the misunderstandings that arise in the capital controversy and regarding the characteristics of a normal position of the economy as viewed before Hicks (1939) are central also in the work of Fratini, which concentrates on Malinvaud's interpretation of Wicksell, perhaps one of the most advanced 'traditional' versions of the neoclassical theory. In particular, Fratini criticises Malinvaud's idea that, since Wicksell's equilibrium considers a situation in which distribution variables and relative commodity prices do not vary period by period, it incorrectly omits to include a condition of zero net savings. Moreover, Fratini objects to Malinvaud's further claim that, in order to compensate for the consequent missing equation, Wicksell should have taken as exogenously given the rate of interest rather than the amount of capital as a single magnitude. Not only in fact did Wicksell actually not intend to relegate his analysis to the unrealistic case of a stationary economy (as defined by Malinvaud following in Hicks's footsteps), but one of

his principal goals was indeed to single out the determinants of the (real) rate of interest. According to Fratini (p. 109), what is thus missing in Wicksell is not the condition of zero net saving, but 'a notion of capital capable of making it work consistently'. And in this respect, Fratini's chapter also critically assesses the specific notions of 'average period of production' and 'marginal product of capital' put forward by Malinvaud, again following some of Hicks's ideas.

The other two papers in Part I are those of Bellino and Schefold. Bellino challenges the idea that in the Cass-Koopmans-Ramsey model myopic optimisation is ultimately responsible for instability and that the assumption of perfect foresight is the only way to avoid it. In the case of an optimal accumulation model with just one good, he advances a (sub-optimal, compared with the saddle path) myopic adjustment process compatible with consumer (bounded) rational behaviour which is capable of ensuring that the ensuing accumulation path converges (step by step) toward the same steady state position as the Cass-Koopmans-Ramsey path. Bellino thus criticises the idea of structural instability of the long-term accumulation path that is remediable only by assuming a transversality or perfect foresight condition – an idea which, Bellino emphasises, has erroneously permitted the main cause of crises, bubbles and other disequilibria to be attributed to consumers' inability to foresee and optimise regarding the future.

Schefold's contribution aims to provide a theoretical explanation (based on random matrices or price systems which for whatever reason have small non-dominant eigenvalues) of what is found in some empirical investigations, namely that wage curves obtained from input–output tables are approximately linear, and that, once these tables available for different countries or periods are taken as representative of all the possible sequences of technical systems, the number of wage curves appearing on the envelopes of the alternative wage curves is very small. He considers his contribution as a first step to the opening up of a field of research in which the existence of the production function and related questions can be discussed by means other than mere a priori reasoning – although on this a priori '*terrain à la Menger*' the logical possibility of reswitching and reverse capital deepening cannot be denied, as Schefold himself points out.

Schefold's contribution will certainly stimulate a debate, for instance on the true possibility of using the input–output data of different sectors and countries to compute the spectrum of the possible methods of production of the different industries. The debate will probably concentrate also upon the extent to which his analysis actually restores a 'production function'

and can possibly support neoclassical ideas, or instead simply suggests some plausibility of a Ricardian incentive to the mechanisation of production when the wage rate rises, without implying any *definite* inverse relation between the wage rate and level of employment – an inverse relation which does not seem to be confirmed by the empirical literature.

But in the 50 years since the publication of *PCMC*, the criticism of the neoclassical theory as outlined in the above contributions has been accompanied by efforts to develop the theory of distribution along classical lines, and it is this which forms the subject of Part II. In particular, two aspects of this reconstruction work are considered in this volume, and they are both related to the idea that income distribution is determined by a set of economic and social-historical circumstances affecting the strength of the 'competing parties' in wage bargaining, rather than by mechanical or semi-natural market forces as in the marginalist theory. The first aspect concerns the notion of subsistence wage which we find in the classical economists and Marx. The second is the possibility of a determination of the surplus wage as a residuum along the lines of Sraffa's suggestion of a rate of profits 'susceptible of being determined ... by the level of the money rates of interest' (Sraffa, 1960, §44).

As stressed by Parrinello – and also by Ginzburg in the third volume of this book – the analysis of these issues has to be addressed outside what Garegnani labelled the 'core' of the surplus theory (as exemplified for instance by Sraffa's price system, with its general and necessary relations between the relative prices and the distributive variables, as well as between the latter and changes in the methods of production). This is due to the fact that, unlike the general equilibrium approach, classical theory recognises 'the existence of a non uniform structure of the social system'. For this purpose, according to Parrinello (see p. 256), we need 'the formulation of causal models and structural equations and piecemeal macroeconomic modelling through a multi-stage analysis of separate subsystems', together with '[t]heoretical and applied work focused on the interfaces between the separate theoretical subsystems.' This is 'a difficult task' because an economist may be well enough equipped to deal with Sraffa's price equations but not to analyse phenomena involving institutional, political and historical aspects as well.

The notion of subsistence wage is explored in the contributions both of Chiodi and Ditta and of Levrero. Chiodi and Ditta emphasise the crucial role of this notion in the viability of an economic system, and discuss the idea (traceable back to Van Parijs) of a *Universal Basic Income*, that is, a monetary income paid by a government at a standard level and at regular intervals to each adult member of society, independently of the economic

and social condition of the people involved. Levrero concentrates upon Marx's analysis of the peculiarity of labour as a commodity, showing that the subsistence wage stems from socially determined conditions of reproduction of an *efficient* labouring class – including the amount of commodities to be given for replacing the 'wear and tear' of the labourer.

The second main constructive issue dealt with in this volume is the analysis of the determinants of the surplus wage and of Sraffa's suggestion of a monetary determination of distribution. Noted also by Chiodi and Ditta, the latter is linked to Keynes's idea (cf. Keynes, 1936, pp. 203–4) of the conventional character of the rate of interest which is ultimately subject to the policy of the monetary authorities – an idea which may represent (cf., for instance, Garegnani, 1978–9; Panico, 1988; Pivetti, 1991) the basis for a theory of distribution alternative to that in terms of the forces of supply and demand, as it is reasonable to suppose that, as a result of competition in product markets, the average rate of profit and the average rate of interest on long-term loans will tend to move in step with each other.

Of course, such a 'closure' of Sraffa's price system, which is alternative to that advanced by Kaldor and Joan Robinson (for a criticism of which see the second volume of this book), does not mean that wage bargaining would have no influence on distribution – for instance, through an effect on the decisions of the monetary authorities about the level of the money rates of interest. In his contribution Levrero argues, however, that this role of wage bargaining might indeed be central in determining income distribution. By shaping the trend of money wages, the economic and social factors which, according to the classical economists and Marx, affect the strength of workers in wage bargaining, might in fact be able to set the *real* rate of interest corresponding to any nominal rate fixed by the monetary authorities. In this scenario, income distribution would be the final result of the interaction between the policy of the monetary authorities shaping the *nominal* rate of interest, and the action of workers and their organisations in wage bargaining shaping the trend of money wages.

The determinants of income distribution are also the subject of the chapters by Pivetti and Stirati, who provide a classical interpretation of the marked shift in distribution between wages and profits that has occurred in the last 30 years. Both contributors point out the difficulties encountered in the attempt to interpret such a shift in distribution along neoclassical lines – for instance as merely due to the effects of immigration and/or of technical change – focusing rather on the serious reduction in wage earners' bargaining power that has occurred over the

last 30 years due to such factors as higher unemployment, privatisation processes, the increasing 'flexibility' of the labour market, and rising capital mobility and trade from low-wage developing countries. Stirati also critically reviews the possible effect of the financialisation of the economy advanced by several Post-Keynesian economists.

A significant aspect of Pivetti's contribution is his interpretation of the increase in top incomes, not in terms of any unprecedented wave of 'skill-biased' technological change, but of changing attitudes to inequality and the accepted levels of top-management compensation. Pivetti (see p. 179) considers this aspect part of 'the more general question of the actual channels through which the parties' relative strengths act and changes in distribution are brought about'. He states that, if the profits of enterprises in an expanding sector of the economy are raised by the presence of monopoly elements, and a shift in social norms has determined higher remunerations of top management which contribute to increasing the price level/money wage ratio in the economy, then 'a compensatory effect on this ratio will eventually have to be sought through lower interest rates, so as to ensure the majority of the working population standards of living considered indispensable for social stability'. Thus, if gross profit includes also absolute top-management compensation (see in this regard also Marx, 1867–94, III, pp. 377–9), then the reduction of interest due to lower interest rates will check the rise in profit margins and the consequent fall in real wages.

As Stirati points out, this may help to explain why in the United States for about a decade interest rates and the rates of profit on capital have not been moving in step. But she suggests (p. 213) that this phenomenon might also be the result of both real wages and interest rate which vary independently for some length of time, thus 'leaving business profits to be determined residually' – something which seems to recall Marx's ideas about the division of the rate of profits between interests and profits of enterprise.

The marked shift of distribution in favour of profits described by Pivetti and Stirati has taken place in all the main capitalist countries during the last 30 years. The chapters by Barba and by Amico and Fiorito which close this volume concentrate upon this shift in Italy and in two important Latin American economies, Argentina and Brazil.

Barba's paper explains why, in Italy, the reduction in working-class income share due both to other incomes and to labour income at the top of the pay scale has not been reflected in a worsening in the Gini measure of personal income inequality – which, however, in Italy is higher than in other rich countries. By breaking the Gini index down

into a factor-share effect, an inequality effect and the interaction between these two effects, Barba shows that delinking the trend in the wage share from the Gini index does not depend on an increase in wage earners' ownership of capital blurring the distinction between labour and property income. Moreover, he draws attention to the influences exerted on the Gini index by the growth in imputed property income as distinct from actual property income, and by permutations among income receivers along the income ladder – in particular by the widespread incidence of atypical labour contracts and the consequent spreading of labour income at the bottom of the income scale, which has partially mitigated the effect of the falling wage share and rising earning dispersion on personal income inequality.

Amico and Fiorito discuss the relationship between monetary and exchange rate policies, and the effects of the existence in several developing countries of a less dynamic primary sector that is export-oriented, and a protected and more dynamic industrial sector, whose prices are higher than the international ones and which produces for domestic consumption only, unless a special 'industrial exchange rate' is applied to exports. They show that this structural heterogeneity can curb the development process and that it shapes the distributive conflict, explaining the results in terms of inflation, distribution and growth which we have observed in the Argentine and Brazilian economies over the last few decades.

References

Bharadwaj, K. and Schefold, B. (1990) (eds) *Essays on Sraffa* (London: Unwin & Allen).

Ciccone, R. Gehrke, C. and Mongiovi, G. (2011) (eds) *Sraffa and Modern Economics* (London: Routledge).

Garegnani, P. (1960) *Il Capitale nelle Teorie della Distribuzione* (Milano: Giuffrè).

Garegnani, P. (1978–1979) 'Notes on consumption, investment and effective demand: I', *Cambridge Journal of Economics*, 2(4): 335–53; 'Notes on consumption, investment and effective demand: II', *Cambridge Journal of Economics*, 3(1): 63–82; *A Reply to Joan Robinson*, vol. 3, June.

Garegnani, P. (1984) 'Value and distribution in the classical economists and Marx', *Oxford Economic Papers*, 36(2): 291–325.

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

Garegnani, P. (2007) 'Professor Samuelson on Sraffa and the classical economists', *The European Journal of the History of Economic Thought*, 14(2): 181–242.

Harcourt, G.C. (1972) *Some Cambridge Controversies in the Theory of Capital* (Cambridge: Cambridge University Press).

Hicks, J.R. (1939) *Value and Capital* (Oxford: Oxford University Press).

Keynes, J.R. (1939) *The General Theory of Employment, Interest and Money* (London: Macmillan).
Kurz, H.D. and Salvadori, N. (1995) *Theory of Production. A Long-Period Analysis* (Cambridge: Cambridge University Press).
Marx, K. (1961–63 [1867–94]) *Capital. A Critique of Political Economy*, Vols I–III (Moscow: Foreign Languages Publishing House).
Panico, C. (1988) *Interest and Profit in the Theories of Value and Distribution* (London: Macmillan).
Pasinetti, L.L. *et al*. (1966) 'Paradoxes in capital theory: a symposium', *Quarterly Journal of Economics*, 80: 503–83.
Pasinetti, L.L. (ed.) (1980) *Essays on the Theory of Joint Production* (London: Macmillan).
Petri, F. (2004) *General Equilibrium, Capital and Macroeconomics* (Aldershot: Edward Elgar).
Pivetti, M. (1991) *An Essay on Money and Distribution* (London: Macmillan).
Schefold, B. (1989) *Mr Sraffa on Joint Production and other Essays* (London: Unwin Hyman).
Sraffa, P. (1951) 'Introduction', in *The Works and Correspondence of David Ricardo*, Vol. I, P. Sraffa (ed.) (Cambridge: Cambridge University Press, 1951–73).
Sraffa, P. (1960) *Production of Commodities by Means of Commodities* (Cambridge: Cambridge University Press).
Steedman, I. (1979) *Fundamental Issues in Trade Theory* (London: Macmillan).

Part I
The Capital Controversy and General Equilibrium Analysis

1
On the Present State of the Capital Controversy

Pierangelo Garegnani[*†]

1.1 Introduction

The post-war capital controversy seems to have had two distinguishable stages. Thanks to the unambiguous phenomena of reswitching and reverse capital deepening, the first stage was conclusive in discarding from pure theory the traditional versions of neoclassical theory that relied on the notion of capital as a single quantity. Subsequently, however, when the implications of those phenomena took centre stage in the controversy, together with the reformulations of the theory which intended to do away with the 'quantity of capital', several misunderstandings prevented, I shall contend, decisive progress in the analysis and we entered an inconclusive phase of the discussion.

These unclarified misunderstandings, I shall also contend, have then left space for the credence that, whatever their methodological deficiencies, the reformulations of neoclassical theory that have been introduced in the theoretical mainstream – essentially by Hicks's *Value and Capital*

[*]This article was originally published in the *Cambridge Journal of Economics* (2012, 36, pp. 1417–32). We thank the *Cambridge Political Economy Society* and the *Oxford University Press* for permission to reprint it (with changes and corrections).

[†]Pierangelo Garegnani died on 15 October 2011, while he was still revising the present paper. This had been accepted in principle by the *Cambridge Journal of Economics* and returned to the author with a few requests for clarification by the referees. Using the drafts he left, two of his closest colleagues (R. Ciccone and S. Levrero) have been able to transfer some of the amendments and changes he was making to the version as published here. Therefore, the form in which the paper appears, though quite definitive, may not be taken as the author's own final version.

(1939) – and which have become dominant after the first stage of the capital controversy, are immune of the inconsistencies affecting previous theory on the conception of capital. This has in turn left space for a second, no less unwarranted, consequence: a feeling that since those reformulations, and in particular general intertemporal equilibrium, would confirm at the level of pure theory the essential validity of the neoclassical demand-and-supply apparatus, they would also provide some validation for the admittedly imperfect previous concepts – foremost that of a 'quantity of capital' – as workable approximations in applied work.

To gain an understanding of the situation just outlined, it may be necessary to take a broad view of the problem, starting from the essential role that the notion of capital as a single magnitude played in originating neoclassical theory by extending the Malthusian theory of rent to cover also the division of the product between wages and profits, which classical economists had explained by a surplus principle. It is, in fact, the essential double role of providing (i) a foundation for the central neoclassical conception of general substitutability among 'factors of production' and (ii) a notion of equilibrium that makes a correspondence between theory and observation possible.

This double role will allow us to confirm, in the face of some conjectures recently advanced, the essential nature of Sraffa's critique of the neoclassical notion of capital and, more generally, of his contribution in *Production of Commodities* as comprising a rejection of the neoclassical explanation of a market economy and an opening to the alternative surplus explanation provided by the English classical economists.

Returning, then, to our main line of argument, we shall recall in Section 4 the essential terms of the difficulty of capital in neoclassical theory – i.e. the impossibility of conceiving that quantity independently of the distribution and prices it is brought in to determine – and we shall examine the way out of the problem influentially proposed in Hicks's *Value and Capital* (1939), based on Walras's old conception of capital as a physical vector of capital goods. That conception, which had gained little following in mainstream theory during the six decades since it was first advanced, will then be considered with the radical changes it renders necessary in the notion of equilibrium. They are the changes characterising the neoclassical reformulations that, we noted earlier, came to the centre of the post-war capital controversy in its later phase.

We shall then be able to proceed to the misunderstandings that, we shall contend, have marred that second phase of the controversy and which, we shall claim, characterise its present state. We shall there refer to the argument developed elsewhere (Garegnani, 2000; 2003), according

to which those reformulations of neoclassical theory also ultimately depend on the notion of capital as a single quantity, the same, as we just said, found indefensible at the level of pure theory in the early stage of the controversy.

1.2 The quantity of capital and its neoclassical role

A preliminary observation may be useful in order to get a grasp of the role of the conception of capital as a single quantity at the origin of neoclassical theory. The observation is that from the point of view of their owners, capital goods, however heterogeneous, are in fact perfect substitutes in proportion to their values.[1] As Walras had lucidly pointed out nearly 150 years ago, capital goods are demanded by savers as elements of the single commodity that he called 'perpetual net income'.[2] It is indeed the single commodity whose existence we imply when we assume competitive arbitrage that tends to realise a uniform 'effective' rate of riskless return on the price of such goods.[3] The reciprocal of that rate is in fact nothing but the price of that Walrasian commodity: if the interest rate is 10 per cent, the value of a 'perpetual net income' of one pound is 10 pounds. This definition assumes constancy of prices and the rate of interest over an infinite future, but little changes in that conception if – more in keeping with contemporary intertemporal equilibrium, its finite horizon and its changing prices – we refer at any time t to the 'income for next year' and to the price of a (gross) unit of it, i.e. $1/(1 + r_t)$.[4]

Now, it is the notion of the quantity of this single commodity, thus rooted in the experience of the wealth owners and of the firms in which they invest, that has evidently been a key to the shift from the first to the second of the two broad approaches. These approaches, at the cost of severe simplifications, can be said to have successively dominated economic theorising since its systematic inception: the classical and then the marginalist or neoclassical approach to the theory of distribution and relative prices.

We do not need a detailed distinction between these two broad approaches.[5] The first approach in order of time is the classical one and centres on the conception of a social surplus that the community can dispose of without infringing on the possibility of reproducing its outputs on a constant scale. It is the idea that – with wages linked to the subsistence of workers and therefore conceived as no less a necessity for social reproduction than are the means of production – underlies the theory of distribution and relative prices, running from the physiocrats

down to Marx via Adam Smith and Ricardo. The subsequent approach is the dominant neoclassical one, which, after a half-century of transition from classical theory, did crystallise in the 'marginalist revolution' of the last quarter of the nineteenth century. It is founded ultimately on the conception of substitutability between 'factors of production' and on the demand-and-supply functions for factors and commodities, which are taken to result from that substitutability.

What requires deeper discussion is the sense in which I consider that the notion of capital as the single Walrasian commodity 'future income' has been a key to the passage from the classical to the neoclassical approaches.

The marginalist or neoclassical approach arose essentially, we noted, out of the Malthusian theory of rent, which, when reformulated in terms of homogeneous land, could be extended to cover the distribution of the product among any number of 'factors of production' and, hence, in principle, also to the determination of the division between wages and profits – thereby replacing the notion of surplus product by which that division had earlier been explained. Now, that extension of rent theory had to be founded essentially on arguing a *variability* of the proportion of 'capital', the single Walrasian commodity, to labour (and the other non-produced factors) in social production in general, analogous to the classical variability of the proportion of labour (plus capital) to land in agriculture. It was a variability that had to descend from the alternative methods available for producing (directly or indirectly) the *same* consumption good, as well as from the methods available for producing *alternative* consumption goods.

The problem, however, was that the alternative production processes thus involved differed, generally, more by the *kinds of capital goods* used than by the *proportion to labour in which each kind of them was employed*. Changes in the proportion to labour of the same physical capital goods may in fact be possible when producing the same commodity but, as intuition suggests, these changes will generate methods of production that will be generally dominated by other methods employing different *kinds* of capital goods: a single proportion between physically specified 'factors' being generally the one that can dominate the other known methods of production of the commodity in question at some level of the distributive variables. And the same variability of the *kind* rather than the *proportion* of the capital goods will be even truer between production processes for alternative consumption goods.

The variability of the 'proportion of capital to labour' in social production on which the neoclassical theory of the division between wages and

profits is founded could therefore hardly have been conceived, had not the different kinds of capital goods required by the alternative methods of production or by the alternative consumption outputs been viewed as embodiments of quantities of the homogeneous Walrasian value commodity, which, like a fluid, underlies the demand for capital goods by savers.

The fluidity of capital necessary in order to generalise classical rent to the division between wages and profits is, however, far from being the only or, perhaps, even the main point about the role of the single savers' commodity in originating the neoclassical theory of distribution. Even more important, from a strictly analytical point of view, was the fact that such a notion allowed the expression of capital endowment – a *datum*, basically, in neoclassical theory, just as the population and the available land were for classical rent – in a way consistent with its homogeneity for savers.

I am referring here to the already mentioned tendency, under free competition, towards a uniform effective rate of riskless return on the capital goods' prices. This tendency can operate over any period of time, no matter how short, simply by means of the competitive arbitrage that will adjust the returns on relatively abundant capital goods to those on scarcer ones, by lowering the *(demand) price* of the former below their *supply prices*, thus raising their rates of return to the level of the scarcer capitals. But clearly these will be only *temporary* adjustments that will soon be followed by further adjustments requiring time. Those capital goods whose price (or demand price) when new had to fall below the respective supply price will not be produced; a tendency will be in operation at every point in time to raise that price and therefore raise the rate of return on costs for those capital goods towards the common level. The strength of this tendency will perhaps appear more evident when we realise that it is one and the same thing, with the tendency to equality between the demand prices and the supply prices (costs of production) of the (non-obsolete) capital goods, not unlike the analogous competitive tendency for the prices of any other commodity.

This means that, for the purpose of the uniformity of returns – the traditional one of the competitive uniform 'rate of profits' – the neoclassical intended generalisation of classical rent has to assume the physical composition of the capital endowment to be fully adjusted to the techniques adopted and outputs produced.[6] It has to assume, therefore, an *endogenous determination* of the physical composition of the capital endowment of the economy. But this is compatible with the basic neoclassical treatment of capital endowment as a *datum* only if the latter is conceived to be given in terms of the fluid Walrasian commodity capable of taking any physical form.[7]

Without that uniformity of effective returns, the position of the economy as determined by the theory would have been no more *persistent*,[8] under free competition, than any position of the economy with, for example, different wages for labour of the same quality or with prices of products differing from their expenses of production – effects strictly analogous to those we have just seen for capital goods when the uniformity of returns does not hold under competition.

But, then, *why* this 'persistency', leading neoclassical theory, as we have just seen, towards the troublesome notion of the given capital endowment as a single magnitude? The fact is that to such a persistency has long been attributed nothing less than the possibility of ensuring *correspondence between theory and observation* in economics.

It was the role played, even across the deep divide between classical and neoclassical theories, by what we may indicate here as the 'normal price' or, more generally, the 'normal position' of the economy – the basis of economic theorising since, at least, Adam Smith's notion of a 'natural price' as 'the central price, to which the prices of all commodities are continually gravitating' (Smith, 1950 [1776], book I, ch. VI, p. 51). The persistency of that normal price (or normal position of the economy) was in fact thought to allow for a *repetition of transactions* that, by occurring on the basis of *nearly unchanged data*, would generally suffice to compensate the temporary or 'accidental' deviations of the *actual* price from the *normal* price and thus allow the latter to generally correspond to an average of the actual prices prevailing over a sufficient interval of time.

Thus, in conclusion, it was certainly not for a matter of mere convenience or of mere simplification that capital endowment was taken as a given in terms of a *single magnitude* – a fact that characterised, with varying degrees of explicitness and with the single partial exception of Walras,[9] all mainstream expressions of neoclassical theory up to a few decades ago; up, that is, to the events we are going to discuss below. On that single magnitude there rested two key points of neoclassical theory: the plausibility of the notion of factor substitution lying at its very heart and, with the possibility of determining a 'normal position', that of a correspondence between theoretical variables and observable magnitudes – two points that, it would seem, were of primary importance to preventing the risk of the theory slipping into an intellectual game.[10,11]

1.3 On the nature of the contribution of Sraffa's *Production of Commodities*

The argument just conducted about the key role of capital as a single 'quantity' at the basis of the logical structure of neoclassical theory

raises the problem, to which we shall proceed in Section 1.4 below, that such a quantity does not exist in the terms independent of distribution and relative prices required by the theory, and that it accordingly has had to be replaced, at the level of contemporary pure neoclassical theory, by the very conception of capital as a set of distinct productive factors, whose incompatibility with a generalisation of classical rent we have just contended.

Before proceeding to that, it may however be noticed that the argument as so far developed may help to assess an interpretation of Sraffa's *Production of Commodities* (1960) advanced by Professor Sen in a comparatively recent paper (2003), which contains some stimulating suggestions as to the Gramscian nature of Sraffa's influence on Wittgenstein's philosophy.

Professor Sen's interpretation of Sraffa (1960) accepts the prevailing view – contrary to the one advanced here – that neoclassical theory need not be expounded in what he calls an 'aggregative form', i.e., with capital as a single magnitude, and that production may instead be analysed by referring only to the physical quantities of the several capital goods (Sen, 2003, p. 1246). Sraffa's critique of the neoclassical concept of capital as a single quantity would therefore be pointless, according to Sen, if meant to invalidate neoclassical theory as a 'predictive' theory, where that concept is inessential. Similarly, continues Sen, the critique cannot be intended to replace that theory with an alternative, equally causal, 'predictive' theory. Sraffa's critique must rather be viewed, Sen contends, as relating to 'interpretational' uses of the theory, which he describes as above all concerned with 'descriptive accounts of the capitalist system having a normative relevance' (*ibid.*, p. 1247).

Clearly, if we have been correct above, the reference to capital as a single magnitude, far from being a particular analytically inessential 'description', or interpretation, of neoclassical theory, lies at the very conceptual roots of it. It seems clear, then, that the critique of the concept of capital, the single magnitude, can hardly be intended to leave standing the demand-and-supply apparatus of the theory and be addressed to only some interpretations of it, detachable, so to speak, from the basic causal, predictive nucleus of the theory. In other words, the critique raises the question of the validity of the theory in its predictive purpose and, if correct, cannot but pose the question of its replacement by a better theory in the same 'causal' or 'predictive' role.

This is suggested by Sraffa himself (1960, p. v) when he refers to the classical economists as having a standpoint 'submerged and forgotten' after the advent of the 'marginal method' and, therefore, a standpoint *alternative* to that which characterises that method. It is clearly the

alternative standpoint based on the notion of surplus that Sraffa himself had outlined in his 'Introduction' to Ricardo's *Principles*, when interpreting Ricardo's early principle of the determining role of agricultural profits (Sraffa, 1951, pp. xxxi–xxxii) – the standpoint that was to be developed in *Production of Commodities*.

Sen's (2003) contention about the 'interpretational' nature of the contribution of *Production of Commodities* appears, however, to be based on a second and more basic kind of argument, besides the denial of the 'causal' relevance of Sraffa's capital critique of neoclassical theory. This second argument is essentially that which Sen had advanced already in a 1978 article on Marx's problem of the so-called 'transformation' of 'values' into 'prices of production'.

Professor Sen had referred there to Sraffa taking as data for his price equations both the real wage (or the rate of profits) and the output levels, for what Sen describes as a determination of prices which 'does not specify anything about causation' (Sen, 1978, p. 180). This is so, he argued, because by using those data 'the exercise begins at too late a stage of price determination ... to be of great use in making actual predictions about the future' (Sen, 1978, p. 181).

Previously (Garegnani, 1991), I had already objected that what Sen saw as a different kind of 'determination' was only the result of an analytical structure of classical theory radically different from neoclassical demand and supply. As implied in the surplus scheme rediscovered and developed by Sraffa (1951; 1960), the classical authors had determined the division of the product between wages and profits by referring, essentially, to a wage governed by broad social and historical forces, such as those controlling the notion of subsistence in any given society, at its stage of development, or those summarised by Adam Smith's 'progressive' or 'declining' state of that society. This is what we can readily find on reading Adam Smith and other classical authors on capital accumulation and distribution.[12] A consequence was that this left the wage and also the output levels free, so to speak, from any pre-defined functional relations with other parts of the system, relations such as those taken to constitute factor substitution in neoclassical theory. Thus, for example, output levels were left free of their neoclassical role as intermediaries of the kind of factor demand (and substitution) operating through consumption choices. Similarly, the wage was left free of any pre-defined functional dependence on alternative techniques and alternative outputs. This meant that it was natural for classical economists to determine both the real wage and the output levels *separately* from prices and other connected parts of the system – though not necessarily *independently* of them.[13]

It was that *separation* between the determination of the wage or the outputs from the rest of the system – made possible by the absence of the notion of factor substitution – that entailed, and explained, the treatment of the real wages and the output levels as givens when determining prices. With that separation, however, the system in no way lost its explanatory or, in Sen's terms, 'causal' and 'predictive' meaning. To realise that it is sufficient to think – as I recalled at the time – of Sraffa's hint at a determination of the return on capital via the rate of interest fixed by monetary authorities, or for another example at a somewhat less general level, of Sraffa's view that a change in the technical conditions of production of a basic product would change all prices and the residual distributive variable, unlike a similar change for a non-basic product (Garegnani, 1991, p. 112).

1.4 The neoclassical problem of capital

We may now return to neoclassical theory as such. In the generalisation of classical rent to cover distribution between wages and profits, the 'quantity' of that special 'factor of production', capital – required, as we noted, to allow for substitutability between factors and for the possibility of a correspondence between theory and observation – had to be measured independently of the distribution of the product between factors and independently of the relative prices which it was brought in to determine, just as the classical quantities of labour and land had to be similarly measured in determining rent. However, the commodity demanded by savers is clearly not *directly* measurable in any such independent terms, since its primary expression for savers lay, as we have noted, in the *value* of the capital goods in terms of some *numéraire*. A basic problem of the new theory was, therefore, how to measure capital, the postulated single productive factor, in terms that would be both *independent of distribution* – as the value of the capital goods is not – and at the same time appropriately related to the value quantity on which savers do make their decisions.

The 'average period of production' over which labour and, more generally, non-produced factors of production have to remain invested in order to produce the commodity according to any given technique (i.e., a set of methods of production, one for the commodity and one for each of its direct and indirect means of production) was the route along which a conciliation of the two requirements was long attempted. It was, however, an impossible task because of the necessary presence of fixed capital, of the multiplicity of non-produced factors of production

and, above all, of the necessity of the compound rate of profits in passing from the commodity to its indirect means of production.[14]

The impossibility of consistently defining a concept as basic for the intended generalisation of classical rent, as we have just argued the 'quantity of capital' was, might conceivably have led to the abandonment of that intent in favour of some return to classical analysis, as had happened, for example, after the demise of the 'wages fund' theories. However, the principle of factor substitution and the ensuing demand-and-supply explanation of distribution had apparently been rooted too deeply in mainstream economic theory for them to be extirpated. Thus, the reaction was instead to apply the principle of factor substitution to each kind of capital good taken as a distinct 'factor', with little explicit consideration of the drastic difficulties that would arise for the theory, i.e., the difficulties that are the mirror image of what we noted in Section 1.2 had made the theory, at its birth, rest on capital as a single magnitude.

As we mentioned Hicks's *Value and Capital* (1939) appears to have been the main influence bringing into the mainstream that tentative way out of the problem.[15] It was, in effect, a question of returning to the conception that Walras had advanced as early as 1877,[16] having initially failed to realise the inconsistency between it and the uniformity of returns on the supply prices of capital goods of the normal position (cf. Section 1.2 above), which he, like all his predecessors, contemporaries and successors until comparatively recent decades, had in fact originally intended to determine.[17]

The recognition by Hicks of Walras's inconsistency – a recognition that, however, remains altogether implicit in *Value and Capital* – meant that he had to accompany the adoption of that conception of capital with the abandonment of the normal position and its uniform rate of profits. But then, under competition, the tendency to such a uniform rate could but be supposed powerful and quick in bringing about appreciable changes in the prices of productive services and commodities. The *persistency* that justified the determination of the equilibrium while abstracting from changes in future prices could no longer be assumed, and the attempt had to be made to remedy that by considering the effect of future conditions on the markets for current commodities and productive services, whether through expectations concerning future prices or present markets for future commodities.

These were presumably the difficulties that, variously perceived and expressed, explained the remarkable fact that despite the fame of its author, and the well-known difficulties of the alternative conception of capital as a single magnitude, Walras's conception had failed to enter

the theoretical mainstream in the six decades that had elapsed since 1877 when it was first advanced. Indeed, even Hicks's own adoption of it, with the associated 'dynamics' of *Value and Capital*, originally had a limited impact on the mainstream: it failed, for example, to attract attention in what was then the main centre of economic theorising, the Cambridge of Marshall, Pigou, Keynes or Robertson.[18]

It was, I believe, the emergence two decades later of the striking phenomena of reswitching of techniques and reverse capital deepening, advanced in 'preview' by Joan Robinson, that rendered finally untenable the notion of capital as a single factor at the level of pure theory and opened the way to dominance for a treatment of capital on Walrasian lines, with the associated necessary reformulations of the concepts of equilibrium – marking what I have contended is a deep 'Hicksian divide' in the evolution of neoclassical theory.[19]

With this we have in fact reached the heart of the post-war capital controversy and have joined it at what we indicated above as its later stage – when the defence of neoclassical theory was conducted in terms of the reformulations of the theory to which *Value and Capital* had opened the way. We might therefore have expected that at such a stage, i.e., after the admitted failure of the notion of capital as a single magnitude, the difficulties of those reformulations would, if not take centre stage, at least emerge with sufficient clarity to be debated. However, the terms in which the reformulations in question had been introduced in *Value and Capital*, some 20 years before, made it difficult for the controversy to focus on such central questions. For those terms, and difficulties, we must therefore turn back to *Value and Capital* and to those aspects of Hicks's argument, which, I submit, have been decisive for the inconclusiveness of the later stages of the controversy.

1.5 Hicks's *Value and Capital*

Despite its title, what we find in the foreground of the book is not in any direct way the problem of capital, but, rather, the claimed need for a 'dynamic theory', accompanied by a critique of what is there called the 'static theory' of 'the economists of the past' (Hicks, 1939, p. 115).[20]

However, it is striking that when we come to a description of what such a 'static theory' consisted of, we do not find the 'normal position' that was the mainstay of those 'economists of the past'. What we find for that 'static theory' are instead two kinds of equilibria, both quite different from the normal position though having something in common with it. The first, Hicks says, is what that static theory should have

been if stated in a 'strict' way (1939, p. 115). It is represented by the equilibria analysed in Parts I and II of *Value and Capital* – those with which Hicks, using a Walrasian vector of physical capital, effectively replaces the previous notion of the capital endowment as a given single magnitude. Hicks has, however, to admit that those equilibria cannot be taken to represent the thought of the 'economists of the past' as that actually was.

For what Hicks attributes to 'those economists' in *Value and Capital* we must therefore turn to the second of the 'static equilibria' he mentions there. And there we find the *stationary position*: i.e., the position of the economy, where the *incentive* to net savings has disappeared (Hicks, 1939, p. 116). However, this is, again, a notion quite different from the neoclassical normal position, though it shares with the latter the assumption of prices constant over time. In the neoclassical normal position, the constancy of the capital endowment that, when taken together with that of the other data, results in a constancy of the equilibrium prices is merely an *abstraction* from the changes that the capital endowment and the other data are admitted to effectively undergo in the economy. It is an abstraction founded only on the *persistency* of the position and the consequent slowness of the change in its data, when compared with the time required to correct and compensate accidental deviations from the normal position, in particular the slowness of the only such change that can be construed to be endogenous to the marginalist system: that in the capital endowment.[21]

In the stationary position instead, as we noted earlier, the same constancy of the capital endowment is the endogenous result of an equilibrium condition of zero net savings, and the capital endowment is therefore an *unknown* of the equations and not the *datum* that it is in the normal position.[22] And the same is true for the proportion of capital to labour and the resulting constancy of prices of the 'steady state', which since the post-war period has become the form of stationary state most commonly contemplated in the analysis (see Section 1.6 below), where net savings are exactly what is required to equip with that same proportion as the labour already employed, the increment of labour.

The paradox of *Value and Capital* is thus that in its account of the 'usual course of economists in the past', we do not find the hallmark of that 'usual course', down to Hicks's own *Theory of Wages* (1932a), namely the normal position. That disappearance of the normal position entailed then, in *Value and Capital*, a second and even more striking paradox: it is that we do not find there any specific criticism of the normal position of those economists – the very position that Hicks in

fact proposes to reject and replace by his dynamic theory. The only criticism of that position remains the generic one of the lack of realism of assuming the constancy of prices in the definition of an equilibrium. It is a criticism that would have been more convincing if 'previous theory' had rested, as it clearly did not, on either Hicks's 'stationary states' or on the fleeting equilibria of Parts I and II of *Value and Capital*.

In fact, the dependence of current prices on future prices was all but overlooked by those 'economists of the past' – starting from Adam Smith and his dichotomy between 'market' and 'natural' prices, down to all later theorists until recent decades. To the extent to which the expected prices reflected merely accidental circumstances, or the undoing of those circumstances, their effects could be ignored in the relevant general context because they would be averaged out into the *normal price* through the repetition of transactions allowed for by the persistency of the normal position. And to the extent to which the expected prices expressed, instead, changes in the *data* of the normal position, they would be dealt with by the comparison between the corresponding two normal positions: the one before and the one after the change in question.

The real point behind the alleged past oversight of price changes – a point that remains, however, altogether implicit in *Value and Capital*[23] – was that the *persistency* allowing for the abstraction from those price changes had been made possible in the equilibria of past neoclassical theory by the treatment of the given capital endowment as a single magnitude, capable of adjusting its physical form. And this is just what the Hicks of 1939, as distinct from the Hicks of 1932, knew could not be done. The normal positions had therefore to be replaced by the 'static equilibria' of *Value and Capital*, whose fleeting character made the attempted remedy of dated prices and quantities all but inevitable. It thus appears that the 'dynamics' proposed in *Value and Capital* was the *effect* of an enforced change in the conception of capital from the single quantity to the physical vector, rather than the *cause* of that change, contrary to what Hicks seems to imply in his foreground argument in his 1939 book.

This explains, I believe, why Hicks and the neoclassical mainstream after him had to contradict and leave aside Marshall's penetrating dictum – which Hicks certainly knew very well, but failed to criticise directly – according to which 'dynamical solutions in the physical sense of economic problems are unobtainable [and] statical solutions afford starting points for such rude and imperfect approaches to dynamical solutions as we may be able to attain' (Marshall, 1898, p. 39).

1.6 The capital controversy and the present situation

The disappearance of the normal position from Hicks's (1939) argument was, however, to weigh heavily on the controversy of 30 years later. The eclipse of that notion had a series of effects on the controversy, which, I submit, converged in obscuring beyond recognition the basic terms of the question of capital in neoclassical theory.

That eclipse meant, first of all, the disappearance of the most transparent form of dependence of neoclassical theory on capital as a single magnitude, namely its ultimate use as a datum for determining the normal position. That made the previous neoclassical use of that notion a confused bone of contention rather than the simple historical fact that it was.[24] It also made it much more difficult to discern the role that the conception of the 'quantity of capital' plays at the very logical roots of the theory. As a result, and most importantly now, it made it difficult to grasp the *continuing* dependence on that conception of the reformulations of neoclassical theory that were being advanced in the later phase of the controversy, relying as they necessarily had to on a sufficient degree of factor substitution – a continuing dependence of which more will be discussed presently.

Moreover, the disappearance of the normal position was made more complete by the associated disappearance of what used to be the key long-period condition under free competition: the traditional uniform rate of profits – i.e., the uniformity of the effective returns on the supply prices or costs of capital goods (in other words, the equality between demand-and-supply prices for the non-obsolete capital goods) – ensuring the persistence of the position and the possibility of its correspondence with observation.[25] Thus, when that condition was referred to from the critical side in order to explain the rationale of the normal position and its neoclassical dependence on the capital endowment as a single fluid fund, that rationale was generally not understood and the condition was even confused at times with the altogether different condition of a uniformity in the commodity's *own rates of interest*, a condition that, however, is a mere synonym of assuming a constancy of prices in defining the equilibrium.[26] Further, these unclarified misunderstandings caused confusion at the outset of the discussion, which was made even worse by a tendency to see the neoclassical dependence on the notion of a 'quantity of capital' as pertaining to the empirical construct of an 'aggregate production function' purporting to represent the output of the whole economy as a single homogeneous aggregate, produced with a 'capital' homogeneous to it. Used for Solow's 1956 simplified

neoclassical answer to the long-term problems raised by Keynes, that notion was an initial target of criticism from some critics. Taken in isolation, however, that target was misleading, as it risked turning an inconsistency at the foundations of the neoclassical idea of a generalised 'factor substitutability' into difficulties pertaining only to an admittedly unrigorous approximation; therefore, it is presumably absent when the several productive sectors are distinguished in a general equilibrium system. It was thereby overlooked that the inconsistency was there, *whatever the number of sectors we might wish to distinguish in the economy*. In fact, the essence of the neoclassical problem of capital is not at all aggregation versus general equilibrium, but, if anything, one about *two kinds* of general equilibria: the traditional one based on normal positions, exemplified by, say, Wicksell (1962 [1906]) or even by Walras (as far as his original intentions went), versus the Hicksian one that renounces such positions in the attempt to avoid capital as the single magnitude.

If the *disappearance* of the normal position as such in *Value and Capital* had those effects of obscuring in the later stages of the controversy the essential terms of the neoclassical problem of capital, the *misinterpretation* of the normal position as a stationary state, which has been the cause of its effective disappearance, has had important *direct effects* on subsequent pure theory even beyond its indirect effects of obscuring the capital controversy. It did that by most authors taking for granted the Hicksian charge that the static method of 'previous theory' was inapplicable to the 'real world' (Hicks, 1939, p. 315) and that the kind of dynamics Hicks was propounding was, in one form or another, the only alternative to it, at the level of pure theory. In that way, Marshall's conclusions noted above were neatly overturned in practice, with no critique of them being in effect advanced (see Section 1.5 above). Hence the paradox of the rejection as unrealistic of a past analysis, which was in effect based on the very tool, the 'normal positions', that economic theorising had developed since its very beginning in order to allow for the possibility of a correspondence between theory and observation in an enormously complex field of reality. This rejection had a second paradoxical consequence: that of in effect obscuring the *true undermining* of that possibility of correspondence – the one resulting from Hicks's *own* proposal, i.e., from the impermanence of the new equilibria and the resulting dynamics, seen as the only alternative to an analysis founded on mere stationarity.

That *real* undermining of the applicability of the theory was indeed the one that Hicks himself had implicitly admitted when, in a rarely quoted passage of *Value and Capital*, he wrote that he assumed 'the economy to

be always in equilibrium' (1939, p. 131), an assumption that should have shocked the readers of *Value and Capital*: no economist had previously supposed the economy to ever actually *be* in an equilibrium position, or more generally in a position of rest, except by fluke:[27] *gravitation around such positions and not achievement of them* being what was always thought relevant for the positions of the economy in the focus of the analysis.[28] Resorting to the above argument of the economy being always in equilibrium on the part of Hicks meant, in fact, admitting that those 'equilibria' were too fleeting to be conceived as centres of attraction of the variables in question. But it certainly was no remedy for that impermanence to imagine possessing instead a theory that could *determine* a one-for-one copy of the real economy, in each of Hicks's 'weeks'.[29]

Indeed, that much of a possibility for correspondence with observation, which has been claimed by mainstream pure theory after the eclipse of the normal position, has hardly been by means of Hicks's 'temporary equilibria' or by the 'intertemporal general equilibria' with today's markets for future commodities. It has rather been by means of the 'steady states', the adaptation of the stationary state attempting to overcome the most obvious, though hardly the most important, of the contradictions between that state and reality, i.e., the fact that economies do change in size over time. On some 'dynamical' neoclassical basis, a long-term tendency has been argued or, more exactly, postulated to some such 'steady state', redefined so as to somehow include technical progress and the other phenomena that cannot but occur over indefinitely long periods of time. The results of such 'steady-state analyses' have then been compared with the observable rates of change of aggregates such as social product, capital or distributive shares, supposing that, with appropriate manipulations, these could be taken to reflect some approximately achieved steady state of that kind of analysis – and comparisons[30] between the two can allow the validity, or lack of it, of the steady-state analysis in question to be assessed.

Clearly, even at the purely methodological level it seems difficult to envisage in such 'steady states' an analysis whose results are at all capable of having a correspondence in observable phenomena, unlike what was possible at a methodological level for an analysis in terms of normal positions. It is sufficient to note that the tendency to a normal position only assumes the simple competitive tendency to uniform prices for homogeneous commodities and productive services – whereas the tendency to a steady state clearly has no such simple and clear foundation, as it already assumes to a large extent the validity of the theory whose results are being tested.

And when we leave aside the purely methodological level and move on to contents, it becomes clear that those analyses have been largely erected on no better basis than that mentioned at the beginning, i.e., on the credence that the temporary equilibrium or intertemporal general equilibrium versions of the neoclassical theory can validate the demand-and-supply apparatus as such – enough to justify the old concepts as acceptable approximations for applied work. It may therefore be contended that as far as contents are concerned, those analyses of reality owe their apparent plausibility to the multiple misunderstandings we noted above (Sections 1.5 and 1.6), which caused what, from a purely analytical point of view, was the inconclusiveness of the later phases of the controversy.

I think that neither that credence, nor, therefore, that justification for the use of essentially the old concepts, are well-founded. I have argued elsewhere (Garegnani, 2003) that intertemporal equilibrium *does not* avoid the dependence on the notion of capital as a single magnitude. Though it no longer occupies its highly visible position as a *fund* among the factor endowments, the *homogeneous commodity 'future income'* demanded by savers cannot be made to disappear from the system any more than any other commodity demanded by individuals in the economy. It can accordingly be shown to emerge as a flow, with the respective demand-and-supply functions and the corresponding markets. They are what, after Keynes, we call, respectively, (gross) savings *supply*, (gross) investment *demand* and yearly saving-investment *market*. The implications of the inconsistency of that notion of capital as a single magnitude – the same implications that enforced the abandonment of the traditional analysis in pure theory – are accordingly still there to be faced.[31]

The discussion on the matter is proceeding. However, we may already ask the following question: should we not begin to recognise that those difficulties are but the expression of a theory originally inspired by the concept of capital as an independently measurable single productive factor, which we now *all* agree does not exist?

Notes

1. Thus, for example, Bliss (1975, p. 8) rightly notes that capital 'cries out to be aggregated'. He does not however seem to notice the very simple reason for that, i.e., the *homogeneity* of capital goods for a decisive category of agents: the savers.
2. Walras (1954, paragraph 242, pp. 275–6). The internal logical conflict that the notion of capital poses for neoclassical theory is indeed well exemplified in Walras who, on the one hand, so clearly saw capital as the single value

commodity demanded by savers and, on the other hand, realised the need for a physical measurement of capital in production.

3. The *adjective* 'effective' is used here in order to remind the reader that this kind of uniformity of returns on capital is quite compatible with, and indeed *requires*, a 'nominal' difference between the commodity's own rates of interest, once changes in relative prices over the period of the loan are considered in the equilibrium (see footnote 26 below). In that case it is in fact only the *numerical expression* of that uniform *effective* rate that will differ depending on the *numéraire* adopted, not unlike what happens generally with a *numéraire* price.
4. It is the price of a unit of *gross* income, because an amount $1/(1+ r_{t+1})$ out of that unit will have to be set aside at the beginning of period $(t + 1)$ if a similar unit of gross income is to be had in $(t + 2)$.
5. Cf., e.g., Garegnani (1960).
6. The uniformity of the rates of return on capitals' supply prices of course excludes, as is generally done at the level of abstraction of the notion of normal price, the presence in the capital endowment of 'obsolete' capital goods, i.e., pertaining to methods of production presently dominated by other methods at *all* possible levels of the distributive variables. More embarrassing for a theory in which the capital endowment is a datum is the fact that the same uniformity of returns also excludes the presence in the endowment of kinds of capital goods that are not 'obsolete' in the sense above, but do pertain to methods of production other than those dominant at the prices of the equilibrium considered. (The question does not arise in Walras, who makes the special assumption that *all* methods require the *same kinds* of capital goods though in different proportions: but it reflects the general case and it reinforces the neoclassical need for capital as the single magnitude, which can take the form of any concrete capital good.)
7. Cf., e.g., Hicks (1932a, pp. 20–1).
8. 'It is to the *persistence* of the influences considered, and the time allowed for them to work out their effects that we refer when contrasting Market and Normal price' (Marshall, 1920, V, III, 6, p. 289; emphasis added). The question is discussed in Garegnani (2002).
9. When discussing capital accumulation, however, Walras, too, referred to capital as a single quantity (1954, paragraph 242) inconsistently with his specification of capital endowments.
10. It is the very risk to which Malinvaud appears to refer when he writes that 'the risk seriously exists that economics loses touch with real problems and develops on its own into a scholastic' (1991, p. 66).
11. As we may expect from the homogeneity of capital goods for savers that will tend to be reflected in *any* theory of the market economy, classical authors also often tended to treat capital as a single magnitude. However, the absence of a theory of distribution founded on factor substitution with the resulting demand-and-supply apparatus exempted them from the above two needs for thus treating capital: Sraffa (1960) exemplifies this classical immunity to the problem.
12. It would be very misleading to point to Ricardo for a different, less historically and more analytically founded view on the matter. Ricardo essentially defers to Smith's *Wealth of Nations* for the wider sociological context of his

analysis, while himself concentrating on the strict analytical points in which he differs from Smith, essentially those concerning the determination of the rate of profits. That in no way entails that his view of the scope and method of economics essentially differs from those envisaged by Adam Smith, Marx and the other classical writers.

13. In other words, the interdependencies of the outputs or the wage with the rest of the system were not denied, but rather implied to be too variable according to circumstances to allow for any useful generalised formal treatment of them involving a simultaneous determination of the system (cf., e.g., Garegnani, 1984).
14. Cf., e.g., Garegnani (1960, part I, ch. III; part II, ch. IV) for, respectively, the notion of the average period of production and its shortcomings from the viewpoint of neoclassical theory. See also Garegnani (1990, pp. 23–31).
15. It is interesting to note that the above difficulties of the Walrasian conception there adopted are not mentioned in *Value and Capital*. This is so, despite the fact that at least the question regarding factor substitutability had been prominent in a 1932 debate between Hicks and Robertson (cf. Hicks, 1932b; Robertson, 1931), when both authors stressed the necessity that the 'capital' endowment be allowed to change form in order to give rise to marginal products and, more generally, to sufficient substitutability between factors. The point returned with force in *The Theory of Wages* (1932a, pp. 20–1) where, for example, Hicks contrasts the 'full equilibrium' marginal product of labour with the 'short period' one where the 'form', as well as the 'quantity', of the capital is said *not* to change; the latter marginal product is then dismissed as something that is very doubtful, if '[it] can be given any precise meaning which is capable of useful application'. This passage regards primarily the difficulty of factor substitution, but the contrast drawn here between the 'short period' marginal product of labour and the 'full equilibrium' one appears to also imply awareness of the second deficiency of the vectorial conception of capital, i.e., the non-uniformity it entails in the effective rate of return on the capital goods' supply prices.
16. See Walras (1877, pp. 568–9), reproducing the paper Walras delivered in July 1876 at the Société Vaudoise des Sciences Naturelles. The year 1877 is also the one in which Walras published the second instalment of the first edition of the *Elements* (1874 and 1877) containing his 'theory of capital formation'.
17. Cf. Garegnani (1960, Part II, chs 2 and 3); see also Garegnani (1990, paragraphs 3–18).
18. Of course the Walrasian conception of capital had been used by mathematical economists long before Hicks, with little notice being taken of it in the mainstream literature at the time, however. Wald (1936) is a good example of that.
19. Cf. Garegnani (2002).
20. Cf. Garegnani (1976, pp. 31–6) for the traces of that deeper line concerning capital in Hicks's (1939) criticism of previous theory (cf. also footnote 23 below).
21. As, for example, Marshall wrote: 'if we are considering … the whole of a large country as one market for capital, we cannot regard the aggregate supply of it as altered quickly and to a considerable extent' (Marshall, 1920, VI, II p. 443).

22. Hicks's identification of the normal position with a stationary one was made easier by the frequent use, at the time, of the term 'stationary' to also indicate the normal position, because of its abstraction from changes in relative prices. However, Lionel Robbins (1930) had already lucidly clarified that ambiguity by his distinction between the 'static' and 'stationary' position of the economy. Hicks's attribution to 'the economists of the past' of a proper stationary state, and not of a normal position wrongly named stationary, is on the other hand made entirely clear when he writes that, in the stationary position of those economists, the 'quantity of intermediate products—the quantity of capital—will be determined through the rate of interest … fixed at a level which offers no incentive for net saving or dissaving' (Hicks, 1939, p. 118).
23. Except perhaps for what may be read into the following passage: 'Of course people used to be able to content themselves with the static apparatus because they were imperfectly aware of its limitations. Thus, they would often introduce in their static theory a "factor of production capital" and its price interest supposing that capital could be treated like the static factors … That some error was involved in their procedure would not have been denied' (Hicks, 1939, p. 116 n). We are not told, however, what that 'error' was exactly. (Cf. also for a highly misleading account of the past difficulties in the measuring of capital, the quotation in the preceding footnote).
24. Cf. e.g., 'It seems to me impossible (as a matter of intellectual history) to maintain that the possibility of perfect capital … aggregation is a neo-classical doctrine' (Hahn, 1982, p. 354). It is, however, at least equally difficult to envisage an intellectual history in which, say, Böhm Bawerk, J.B. Clark, Pigou, etc. could use an 'aggregation' of capital whose possibility they did not admit.
25. It is significant and again somewhat paradoxical that Hicks's revival of Walras's theory of capital in Parts I and II of *Value and Capital* went together with the total disappearance there of Walras's own equations of 'capital formation' (Walras, 1954, Lesson 23), which contained the condition of uniformity of returns as well as the relation equalising the demand and supply of 'net perpetual income' (see Section 1.2 above), i.e., in today's terms, savings and investment. That disappearance left a serious gap in the static theory of Parts I and II of Hicks (1939) into which we cannot, however, enter here.
26. For a telling example of this confusion, see the discussion in Garegnani (2003, pp. 153–4) of a passage in Hahn (1975), in which he used the above uniform rate of return referred to by critics in order to characterise the 'special neoclassical case' to which, in his opinion, Sraffa referred. Now, Hahn saw that case as one in which 'the equilibrium price of a good for future delivery in terms of the same good for current delivery will be the same for all goods' (Hahn, 1975, p. 360). Clearly the latter is the case of uniform commodity-own rates of interest, i.e. constant prices, which is quite compatible logically with any divergence between rates of return on capital supply prices, with which it has in effect nothing to do. Similarly, the effective uniformity of the latter rates contradicts the uniformity of the *own* interest rates whenever price changes over time are considered in the equilibrium (cf. footnote 3 above).
27. This assumption, to which Hicks is in effect led by the abandonment of the normal position, is similar to that we find in Bliss when he wrote 'it may seem more sensible to simply assume that equilibrium will prevail and to thus confine our investigations to the equilibrium state. We could regard the

object of our investigations not as "the economy" but as "economic equilibrium" … This approach may seem more attractive, if only because more tractable than the Herculian programme of constructing a complete theory of the behaviour of the economy out of equilibrium' (Bliss, 1975, p. 28). Bliss is here, so to speak, touching with his own hand the implications of that *abandonment of the normal position*, where the 'Herculian task' was instead largely left to itself by the simple Smithian device of the 'centre of gravitation', i.e. by the concentration of the analysis on persistent forces. Those implications appear to have in fact led to an impasse, such that the way out suggested by Hicks (1939) and Bliss (1975) comes close to assuming away reality. (Cf. again the passage by Malinvaud quoted in footnote 10 above.)

28. As Denis Robertson wrote with admirable simplicity and lucidity, '[i]t seems to me that anybody who rejects these two ideas, that a system can move towards equilibrium and that it may never get into it—has made it extremely difficult for himself to interpret the course of events in the real world' (Robertson, 1957, pp. 144–5).
29. We are here in conflict with the view expressed in Harcourt (1981) and often advanced by Joan Robinson (see Garegnani, 1976). Samuelson appears, on the other hand, to seriously underestimate the difficulty of determining, one to one, the actual path of the economy (what Hicks's passage reported in the text appears to imply) when in his *Foundations of Economic Analysis* (1947) he draws the analogy of a 'cannonball [that] can be held to be in equilibrium at each point on its path'. The dominant forces acting on the cannonball at each instant of time are, however, comparatively few in number and their effects on the position of the cannonball can accordingly be calculated with a degree of approximation sufficient to establish a correspondence between the theoretical and actual position of the ball at that instant. This seems to be the only meaning attributable to the idea of a cannonball being in equilibrium at each point on its path. Given instead the numberless forces of similar strength that affect the economy at each instant of time, the instantaneous position of an economy cannot even in principle be determined with any approximation: only averages of observable positions reflecting the effects of the few most *persistent* among those forces can be determined. And the accumulation of the errors would seem to make the path of the economy even less calculable than its instantaneous position is by itself. This, it seems, is what prompted Marshall to write the passage quoted above (Section 1.5, p. 27).
30. Cf., e.g., Lucas (1988).
31. Cf. e.g., the discussion of the idea, apparently subscribed to by Professor Arrow, that the adjustment between savings and investment in an intertemporal equilibrium is a perfectly consistent story that does not look any different from the story about choosing commodities today (Garegnani, 2000, p. 435 n, quoting Arrow, 1989, p. 155).

References

Arrow, K. J. (1989) 'Joan Robinson and Modern Economic Theory: An Interview', in J.R. Feiwel (ed.), *Joan Robinson and Modern Economic Theory* (London: Macmillan).

Arrow, K. J. and Debreu, G. (1954) 'Existence of an equilibrium for a competitive economy', *Econometrica*, 22(3): 265–90.
Bliss, C. J. (1975) *Capital Theory and the Distribution of Income* (Amsterdam/New York: North-Holland/American Elsevier).
Bliss, C. J. (2005) 'Introduction', in C.J. Bliss, A.J. Cohen and G.C. Harcourt (eds), *Capital Theory* (Cheltenham: Edward Elgar).
Debreu, G. (1959) *Theory of Value* (New Haven: Yale University Press).
Garegnani, P. (1960) *Il Capitale nelle Teorie della Distribuzione* (Milan: Giuffré).
Garegnani, P. (1976) 'On a Change in the Notion of Equilibrium in Recent Work on Value: A Comment on Samuelson', in M. Brown, K. Sato and P. Zarembka (eds), *Essays in Modern Capital Theory* (Amsterdam: North-Holland).
Garegnani, P. (1984) 'Value and distribution in the classical economists and Marx', *Oxford Economic Papers*, 36(2): 291–325.
Garegnani, P. (1990) 'Quantity of Capital', in *Capital Theory, The New Palgrave Series* (London: Macmillan).
Garegnani, P. (1991) 'The Labour Theory of Value: "Detour" or Technical Advance?' in G. Caravale (ed.), *Marx and Modern Economics*, Vol. I (Brookfield, VT: Elgar).
Garegnani, P. (2000) 'Savings, Investment and the Quantity of Capital in General Intertemporal Equilibrium', in H.D. Kurz (ed.), *Critical Essays on Piero Sraffa's Legacy in Economics* (Cambridge, UK: Cambridge University Press).
Garegnani, P. (2002) 'Sraffa's Price Equations: Stationary Economy or Normal Positions?' in S. Boehm, C. Gehrke, H.D. Kurz, and R. Sturm, (eds), *Is There Progress in Economics?* (London: Routledge).
Garegnani, P. (2003) 'Savings, Investment and Capital in a System of General Intertemporal Equilibrium', in F. Hahn, and F. Petri (eds), *General Equilibrium: Problems and Prospects* (London: Routledge).
Hahn, F.H. (1975) 'Revival of political economy: the wrong issues and the wrong argument', *Economic Record*, 51(3): 360–4.
Hahn, F.H. (1982) 'The neo-Ricardians', *Cambridge Journal of Economics*, 6(4): 353–74.
Harcourt, G.C. (1981) 'Marshall, Sraffa and Keynes: incompatible bedfellows?' *Eastern Economic Journal*, 7(1): 39–50.
Hicks, J.R. (1932a) *The Theory of Wages* (London: Macmillan).
Hicks, J.R. (1932b) 'Marginal productivity and the Lausanne school: a reply', *Economica*, XII(37): 296–300.
Hicks, J.R. (1939) *Value and Capital* (London: Oxford University Press).
Lucas, R.E. (1988) 'On the mechanics of economic development', *Journal of Monetary Economics*, 22: 3–42.
Malinvaud, E. (1991) 'The next fifty years', *Economic Journal*, 101: 64–8.
Marshall, A. (1898) 'Distribution and exchange', *Economic Journal*, VIII: 37–69.
Marshall, A. (1920) *Principles of Economics* (London: Macmillan).
Robbins, L. (1930) 'On a certain ambiguity on the notion of stationary state', *Economic Journal*, XL: 194–214.
Robertson, D.H. (1931) 'Wage grumbles', in Robertson, D. H., *Economic Fragments* (London: King & Son): pp. 42–57.
Robertson, D.H. (1957) *Lectures on Economic Principles* (London: Fontana).
Robinson, J. (1956) *The Accumulation of Capital* (London: Macmillan).

Robinson, J.V. (1970) 'Capital theory up to date', *Canadian Journal of Economics*, III(2): 309–17.
Samuelson, P.A. (1983 [1947]) *Foundations of Economic Analysis* (Cambridge, MA: Harvard University Press).
Sen, A.K. (1978) 'On the labour theory of value: some methodological issues', *Cambridge Journal of Economics*, 2: 175–90.
Sen, A.K. (2003) 'Sraffa, Wittgenstein, and Gramsci', *Journal of Economic Literature*, XLI: 1240–55.
Smith, A. (1950 [1776]) *The Wealth of Nations*, 2 vols (London: Dent & Sons).
Solow, R. (1956) 'A contribution to the theory of economic growth', *Quarterly Journal of Economics*, 70(1): 65–94.
Sraffa, P. (1951) 'Introduction to Ricardo's *Principles'*, in P. Sraffa (ed.), *Works and Correspondence of David Ricardo*, Vol. I (Cambridge, UK: Cambridge University Press).
Sraffa, P. (1960) *Production of Commodities by Means of Commodities: Prelude to a Critique of Economic Theory* (Cambridge, UK: Cambridge University Press).
Wald, A. (1936) 'Über einige Gleichungssysteme Matematischen Ökonomie', *Zeitschrift für Nationalökonomie*, Vol. 7: 637–70; 'On some systems of equations in mathematical economics', English translation, *Econometrica*, 1951, 19(4): 368–403.
Walras, L. (1877) *Théorie Mathématique de la Richesse Sociale* (Paris: Guillaumin).
Walras, L. (1954) *Elements of Pure Economics, or the Theory of Social Wealth* (London: George Allen & Unwin).
Wicksell, K. (1962 [1906]) *Lectures in Political Economy*, Vol. I (London: Routledge).

2

Two Strands of Thought in Pierangelo Garegnani's Capital Theory Critique

*Harvey Gram**

Two lines of thought may be discerned in the half-century arc of critical writing on capital theory which extends from Garegnani (1960) to Garegnani (2010). The first concerns the conception of the quantity of capital as an endowment to be put on a par with land and labour – the traditional three 'factors of production' (which is not to discredit Marshall's fourth factor, organisation, which, like capital, is produced). The second draws attention to a related methodological shift within neoclassical theory prompted by a seeming recognition, originally on the part of Hicks, but essentially found in Walras (as Garegnani has repeatedly pointed out), that the treatment of a given capital endowment as an *arbitrary* vector of capital goods – each one a separate resource or factor input – is inconsistent with 'the uniformity of returns on capitals' supply prices pertaining to the "normal position"' (Garegnani, 2010, p. 81). A central point in this second line of argument is that the 'normal position' was the bedrock upon which *both* classical and traditional neoclassical theory had been built.

In connection with the first line of criticism, Garegnani cites Hahn, who denies that 'perfect capital (or labour) aggregation' was ever part of neoclassical doctrine (Garegnani, 2010, p. 95, n. 25), countering that Böhm-Bawerk, J. B. Clark and Pigou did use an 'aggregation' of capital in their theories and that this was a *value* aggregate. Among neoclassical

*Helpful comments by a referee are gratefully acknowledged.

economists of influence and repute, perhaps the last to embrace the traditional notion of a given capital value, *flexible in form*, was George Stigler, who wrote in the later editions of his famous price-theory text:

> a factor may be held constant [if] its economic quantity (or value) can be held constant. We can hold the house-building tools at $2,000, say, but vary their form so that they are most appropriate to whatever quantity of labour we employ. With fewer men, we use fewer and more elaborate tools; with more men, we use more, but less elaborate, tools. (Stigler, 1987, p. 136)

In a similar passage in earlier editions, Stigler recalled 'Wage-Grumbles', a once widely read essay by Dennis Robertson (1931):

> the pure theory states that there are, say, twenty dollars' worth of shovels, and with 10 men there are 10 shovels, and with 11 men there are 11 smaller or less durable shovels of equal value. The reason for permitting the fixed service to undergo variations of form is that only with such adaptability of form will the maximum output from various proportions between 'fixed' and variable services be obtained. If the fixed service were not adaptable, we should have to assign the eleventh man to a different type of work, perhaps fetching beer, as Dennis Robertson has suggested. The marginal product of the variable service would fall more rapidly. (Stigler, 1952, p. 117)

There is a significant symmetry in Stigler's statement of the traditional view:

> Or conversely, if we are examining the marginal productivity of tools, we can hire fewer but abler workmen (with the same aggregate payroll) with fewer tools and more but less able workmen with many tools. (Stigler, 1987, p. 136)

If the symmetry were apt, the given 'aggregate payroll' would correspond, like the given value of tools, to a *capital* value, conceiving the 'workmen' as a stock of slaves whose types could be changed without altering that value. This is not the intended interpretation. Alternatively, the fixed capital endowment would correspond to a fixed budget for renting different types of tools appropriate to a variable number of workers. In this mixing of concepts – capital value versus expenditure flow – one finds an answer to those, like Hahn, who claim

that the aggregation problems encountered in defining the 'quantity of capital' as a given endowment are neither more nor less serious than those entailed in the idea of a given endowment or 'quantity of labour'. The conceptual difference turns on the fact that, under capitalism, the supply prices of the various components of the stock of capital goods are determined by the tendency under competition towards a uniform rate of profit on their value, an idea central to the very method of economic analysis as it emerged from the seminal work of the forerunners of Adam Smith – Petty, Cantillon and the Physiocrats. On the stock of human capital, no such *uniformity* of return is to be expected, except in a slave society where workers, like capital goods, are produced and owned by capitalists (with only natural resources or 'land' remaining as a primary non-produced factor). Of course, this is not to suggest that individuals investing in their own education and training do not consider the prospective rate of return on such investment when choosing a line of work – but there is no ever-present competitive pressure to ensure it is the same for all.

In connection with his second line of criticism, Garegnani finds in the lesser-read parts of *Value and Capital* (Hicks, 1939, 1946) the seeds of a once marginal enterprise, appreciated exclusively by mathematically trained theorists, but now firmly entrenched as the cornerstone of modern intertemporal equilibrium theory. At its core lies the fact – a paradox, in Garegnani's view – that Hicks does not mention in *Value and Capital* two difficulties with the Walrasian conception of capital (Garegnani, 2010, p. 94, n. 18), despite his earlier recognition that 'by Wicksell it was attacked' (Hicks, 1934, p. 345). First, the replacement of a 'quantity of capital' by a vector of capital goods undermined the possibility of 'substitution' among productive factors, as Robertson (and Stigler) recognised in assigning the eleventh (unequipped) man to fetch beer. Second, the lack of persistence of the theoretical solution, entailed in this short-run perspective, undermined the correspondence between theory and observation and so undermined the status of economics as a scientific endeavour. Garegnani finds it striking that the novelty claimed for the 'dynamic theory' set forth in *Value and Capital* is not accompanied by a critique of the 'normal position' which had long been the mainstay of 'economists in the past', held to account by Hicks for their failure to analyze the causes and consequences of *changing* prices. In its place, one finds criticism of a straw man: the stationary (or steady) state which had but one point in common with the normal position. This is the constancy of relative prices – *actually* constant in a stationary (or steady) state, but constant only as a theoretical abstraction in the

normal position, around which market prices *fluctuate* in accordance with short-run competitive forces.

It is concluded that the dynamic theory offered in *Value and Capital* was not so much a creative project as a defensive action necessitated by a change in the conception of capital – no longer a given quantity, mutable in form, but rather a given vector of capital goods inherited from the past, acting as a constraint on present output, and, in this now altered view, strictly on a par with all other 'resources'.

An important consequence of this change in the conception of capital has often been noticed, but is not, somehow, seen as destructive of economics as a scientific endeavour – namely the fragility and *impermanence* of the new dynamic equilibrium paths. It is this impermanence that creates an unbridgeable chasm between theory and observation. Garegnani has confronted the issue in recent work by focusing on 'some missing equations in contemporary treatments of intertemporal general equilibrium' (Garegnani, 2009). His analysis turns on the distinction between 'demand price' and 'supply price' for reproducible inputs, the former reflecting portfolio equilibrium insofar as 'demand price' captures the effects of arbitrage, in contrast to 'supply price', part and parcel of the normal position, once central to economic analysis.

With the abandonment of the normal position as a theoretical construct, portfolio equilibrium 'demand prices' now dominate modern capital theory, but mere inspection reveals at a glance the canker at the heart of the analysis. Modern-day theorists acknowledge that solution paths for asset prices are *inherently* unstable, when they exist, as argued most notably by Hahn (1966). Instability had earlier been hinted at by Dorfman, Samuelson and Solow (1958), who conclude an analysis of the relationship between competition and an equilibrium path of accumulation by invoking the implied authority of Smith. More than one reference to the 'invisible hand' metaphor indicate the authors' belief in a deep underlying connection between their analysis of dynamic general equilibrium and the long period method, so eloquently set forth in *Wealth of Nations*:[1]

> A *particular* efficient program is picked out by the invisible hand only if one arbitrary bit of information is added … Mathematically, this arbitrariness reflects the fact that the difference equations of intertemporal efficiency … were shown to be subject to three boundary conditions. Competition ensures that the equations will hold, and history provides two initial conditions. The remaining degree of freedom lets us pick out one more point through which the efficient path must pass.

> The truly remarkable thing about the intertemporal invisible hand is that while it results in efficiency over long periods of time, it requires only the most myopic vision on the part of market participants. Just current prices and [their] current rates of change need to be known, and at each moment long-run efficiency is preserved. But for society as a whole there is need for vision at a distance. (Dorfman, Samuelson and Solow, 1958, p. 321)

This last qualification, concerning the vision of the future, in fact concerns the value of capital: equilibrium prices for capital goods at the terminal date must be 'quoted in order that competition should lead a myopic market inevitably to the appropriate [terminal] point' (*ibid.*). A revealing final statement, offered almost as an aside, became the touchstone for much of Samuelson's later work in the theory of capital where Laws of Conservation of Value play a key role (e.g. Samuelson, 1972 and 1990):

> One interesting sidelight before we leave the subject of intertemporal pricing: Consider any efficient capital program and its corresponding profile of prices and own-rates [of interest]. *At every point of time the value of the capital stock at current efficiency prices, discounted back to the initial time, is a constant,* equal to the initial value. This law of conservation of discounted value of capital (or discounted Net National Product) reflects, as do the grand laws of conservation of energy in physics, the maximizing nature of the path. (Dorfman, Samuelson and Solow, 1958, pp. 321–2; italics in the original)

This clear statement concerning the role of a given 'quantity of capital', transmutable in form like Stigler's $2,000 worth of tools, was seldom, if ever, acknowledged in subsequent debates.

How, then, did the theory of intertemporal equilibrium manage to 'reconcile' a vector of capital goods, given at some initial date as an arbitrary list of produced goods inherited from the past, with an aggregate 'quantity of capital' that not only changes physical form over time, but also undergoes continuous revision in terms of the values of its individual components (remaining, in some specifications, an unchanging value over time)? The answer comes from the theory of optimal processes (Pontryagin *et al.*, 1962) which underlay a vast outpouring of work on the analysis of capital accumulation within the framework of a general equilibrium theory of supply and demand. Garegnani, attempting

to engage Christopher Bliss in debate, refers to a central feature of such analysis – a mathematical trick, if one may so describe it:

> Bliss meets the additional obstacle of the 'saddle' character attributed to [the] path [of capital accumulation under conditions of supply and demand equilibrium], which should be overcome by 'transversality conditions' further restricting, apparently, the possibility that such a Ramsey path to steady states [will] ever be one which the economy can walk. (Garegnani, 2010, p. 91)

Bliss acknowledges the instability of intertemporal equilibrium paths, granting that the analysis (Bliss, 1975) is only concerned with the properties of full stock-flow economic equilibrium and has no clear, or even intended, connection with economic reality. Such 'pseudo-dynamics' is nowadays universally taught, anesthetising students against any serious consideration of the problems that arise in an attempt to understand the complex interactions of stocks and flows. As for the 'transversality conditions' Garegnani is curious to have explained, they are first-order conditions under the Pontryagin Maximum Principle, ensuring, as if by fiat, that the price of each capital good is *initially set* in just such a way as to ensure that the short-run supply and demand equilibrium of each period, evolving over time as the vector of capital goods changes in consequence of the flow of investment, will eventually 'converge' in infinite time to long-run 'supply prices'. Absent 'transversality conditions', even perfectly achieved arbitrage, which establishes a uniform effective rate of return on 'demand prices' will send the system off on an unsustainable path.

This exotic theory, imported from physics, where it has innumerable and useful applications, has invidiously taken over and gradually smothered traditional economics. Few, perhaps, took seriously, as a model with relevance for an actual economy, the Arrow-Debreu general equilibrium theory with its complete set of competitive forward markets. Burgstaller ([1994] 2010) has done a great service by analyzing in detail the connections between Arrow-Debreu and dynamic analysis of the type one finds in Bliss (1975) and Burmeister (1980). Unstable saddle paths and, in the face of shocks, precise, discrete *jumps* in asset prices are everywhere to be found in Burgstaller's survey of various model types ranging from pure exchange to pure von Neumann growth. Equivalent to assuming that the economy is 'always in equilibrium' (Hicks, 1939, p. 131), this theory is evidently rejected by Garegnani

when he asks how an equilibrium saddle path can ever be one 'which the economy can walk'.

Garegnani points out that the critics of neoclassical theory too often failed to address directly the peculiarities of a theory of continuous supply and demand equilibrium with accumulation extending into the indefinite future. Notable exceptions were Eatwell (1982) and Milgate (1982), drawing on Garegnani (1976). In consequence, a once robust classical/neoclassical method of analysis gradually succumbed to a theory that is at once fragile on its own terms and incapable of incorporating an analysis of the forces of competition as commonly understood (Garegnani, 1990, p. 55).

Note

1. The 'indeterminacy' to which Dorfman, Samuelson and Solow refer in connection with their analysis of *difference* equations in a *finite* analysis, has its counterpart in the need to jump onto the convergent saddle-path characteristic of *infinite* horizon models thereby 'solving' in a strictly formal sense the problem of inherent instability in such models.

References

Bliss, C. (1975) *Capital Theory and the Distribution of Income* (Amsterdam: North-Holland).

Burgstaller, A. ([1994] 2010) *Property and Prices: Toward a Unified Theory of Value* (Cambridge: Cambridge University Press).

Burmeister, E. (1980) *Capital Theory and Dynamics* (Cambridge: Cambridge University Press).

Dorfman, R., Samuelson, P. and Solow, R. (1958) *Linear Programming and Economic Analysis* (New York: McGraw-Hill).

Eatwell, J. (1982) 'Competition', in I. Bradley and M. Howard (eds), *Classical and Marxian Political Economy* (New York: St. Martin's Press): pp. 202–28.

Garegnani, P. (1960) *Il Capitale nelle Teorie della Distribuzione* (Milan: Giuffré).

Garegnani, P. (1976) 'On a Change in the Notion of Equilibrium in Recent Work on Value and Distribution' in M. Brown, K. Sato, and P. Zarembka (eds), *Essays in Modern Capital Theory* (Amsterdam: North-Holland).

Garegnani, P. (1990) 'Quantity of Capital' in J. Eatwell, M. Milgate and P. Newman (eds) *Capital Theory* (London: Macmillan).

Garegnani, P. (2000) 'Savings, Investment and the Quantity of Capital', in H. D. Kurz (ed.), *Critical Essays on Piero Sraffa's Legacy in Economics* (Cambridge: Cambridge University Press).

Garegnani, P. (2003) 'Savings, Investment and Capital in a System of General Intertemporal Equilibrium', in F. Hahn and F. Petri (eds), *General Equilibrium: Problem and Prospects* (London: Routledge).

Garegnani, P. (2009) 'On Some Missing Equations in Contemporary Treatments of Intertemporal General Equilibrium', *Quademo di Ricerca del Centro di Ricerca e Documentazione Piero Sraffa*, no. 7, Università Roma Tre.

Garegnani, P. (2010) 'Capital in the Neoclassical Theory: Some Notes' with 'A Reply to Professor Bliss's Comment', in A. Birolo, D.K. Foley, H.D. Kurz, B. Schefold and I. Steedman (eds), *Production, Distribution and Trade: Alternative Perspectives, Essays in honour of Sergio Parrinello* (London: Routledge).
Hahn, F.H. (1966) 'Equilibrium dynamics with heterogeneous capital goods', *Quarterly Journal of Economics*, 80: 633–46.
Hicks, J.R. (1934) 'Léon Walras', *Econometrica*, 2: 338–48.
Hicks, J.R. (1939, 1946) *Value and Capital* (Oxford: Oxford University Press).
Milgate, M. (1982) *Capital and Employment* (London: Academic Press).
Pontryagin, L.S., Boltyanskii, V.G., Gamkrelidze, R.V. and Mishchenko, E.F. (1962) *The Mathematical Theory of Optimal Processes* (New York: Interscience).
Robertson, D. H. (1931) 'Wages-Grumbles', in *Economic Fragments* (London: P.S. King and Son); repr. in American Economic Association, *Readings in the Theory of Income Distribution* (Homewood, Illinois: Irwin, 1946, 1963).
Samuelson, P. (1972) 'The general saddlepoint property of optimal control motions', *Journal of Economic Theory*, 5: 102–20.
Samuelson, P. (1990) 'Two Conservation Laws in Theoretical Economics', in R. Sato and R. V. Ramachandran (eds), *Conservation Laws and Symmetry: Applications to Economics and Finance* (Dordrecht: Kluwer).
Stigler, G. ([1952] 1987) *The Theory of Price* (New York: Macmillan).

3

Only a Few Techniques Matter! On the Number of Curves on the Wage Frontier

*Bertram Schefold**†

3.1 Introduction

Wage curves have become the main tool for the analysis of technical choice, but what does their envelope look like? Joan Robinson used to say that one should expect one technique to dominate all others, independently of distribution. It sounded like a polemical remark to counteract all reference to neoclassical substitution, and she was more cautious in her writing,[1] but the drawing on the blackboard resembled Figure 3.1, w_1, w_2 representing two techniques. (The reader not familiar with the Sraffa analysis and the notation used here can pick it up in the first paragraph of Section 3.2.)

The surrogate production function, on the other hand, looked like that shown in Figure 3.2.[2] The maximum rates of profit $R_1, \ldots, R_4$ of the individual techniques were in inverse permutation to the wage rates paid at $r = 0$.

Sraffa spoke of a 'rapid succession of switches' of techniques along the envelope, and there was reswitching and reverse capital deepening (Figure 3.3).

Reverse capital deepening is opposed to the equilibrating mechanism postulated by neoclassical theory. Consider the Samuelson case and suppose the economy is in a state of full employment at P_0 in Figure 3.2.

*I should like to thank Christian Bidard, Zonghie Han and an anonymous referee for very helpful comments – the responsibility for the text is mine. I dedicate this chapter to the memory of Joan Robinson, Piero Garegnani and Paul Samuelson, with whom I discussed envelopes of wage curves in 1969, 1970 and 1973 respectively; but the results envisaged then were different.

†FB Wirtschaftswissenschaften, J. W. Goethe-Universität, Frankfurt am Main, Germany; email: schefold@wiwi.uni-frankfurt.de.

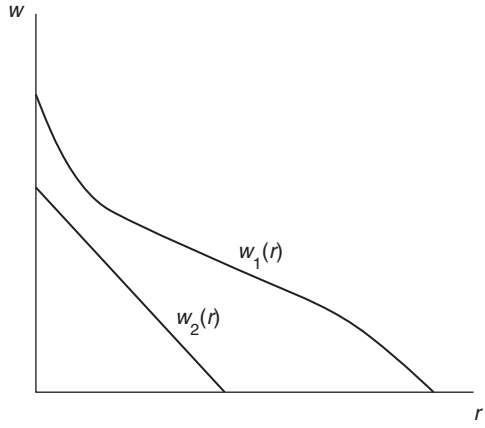


Figure 3.1 The choice of technique (Joan Robinson case)

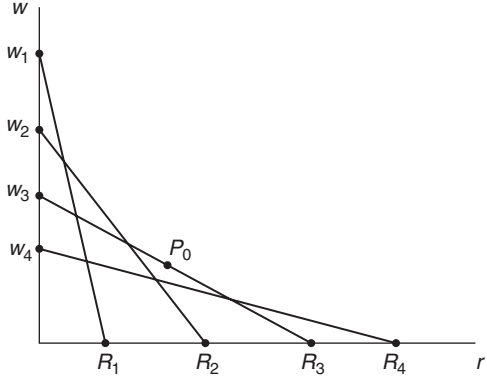


Figure 3.2 The choice of technique (Paul Samuelson case)

Suppose further that real wages are forced up by trade union action (money wages rise more than prices). It then becomes profitable to use the technique of wage curve w_2 (r) for which the intensity of capital is higher. If there is not enough accumulation of capital, unemployment results which reduces trade union power and hence wages, confirming the neoclassical view that there is one level of the real wage at which a full employment equilibrium is stable.

But if the economy is in P_0 in Figure 3.3, the same rise of the real wage would lead, if profits are maximized, to the adoption of technique $w_3(r)$, to the left of switchpoint P_3, and the intensity of capital would fall.

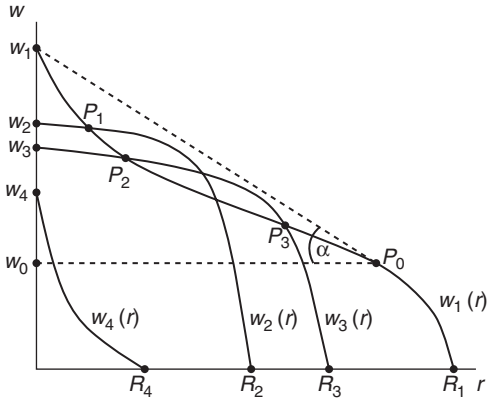


Figure 3.3 Paradox of capital
Note: There is reverse capital deepening at P_3. It would be reswitching , if $w_2(r)$ was not there so that P_2 would be on the envelope. Capital per head at P_0 equals $tg\,\alpha$.

If the amount of capital[3] employed was not reduced, labour demand would be enhanced, encouraging further rises of the wage: the change of factor prices would not stabilize the equilibrium.

The neoclassical equilibrating mechanism is already not valid in general if there is only one example of reverse capital deepening. Could it still be relevant as a rule with exceptions? Not if the cases with reverse capital deepening are frequent.

The applicability of the critique of neoclassical theory based on reverse capital deepening therefore depends on how often it occurs – if reverse capital deepening occurs very rarely, if it is only a logical possibility and not likely to be encountered in reality, the critique remains academic. How likely is it? Before asking this question, one should ask how many switches there are in the first place.

But the questions are linked. I used to believe in the 'rapid succession' (Sraffa, 1960, p. 85) of switches, as one runs down the envelope, for if only two methods are known in each of 1000 industries, there result 2^{1000} wage curves. Many of them might be inferior so that they would not appear on the envelope, like wage curve $w_4(r)$ in Figure 3.3, but many might appear several times, like $w_1(r)$ in Figure 3.3, so that 'the number of switchpoints' on the envelope could be 'at least of the same order of magnitude as the number of wage curves' (Schefold 1997 [1979], p. 279). I could believe this (not alone – I remained unopposed in many presentations of this argument), because I also believed that any two wage curves might cross several times.

There seemed to be no compromise between the idea of an envelope involving a very large number of wage curves and Robinson's postulate that one technique would be the best, independently of distribution. But we shall show that the picture may change drastically, if we admit that the wage curves are nearly straight lines as many empirical investigations by Anwar Shaikh and others on the basis of input-output analyses have shown (see Mariolis and Tsoulfidis, 2010; Han and Schefold, 2006, with the references mentioned in these papers).

The empirical investigation became possible as a result of turning to input-output analysis. In fact, to rely on input-output analysis and not on set-theoretical approaches for the representation of the spectrum of techniques was a paradigm shift, as we shall argue at the beginning of Section 3.3. Han and Schefold (2006) extended the approach to comparisons of $496 = 32 \cdot 31/2$ different envelopes of wage curves, resulting from considering pairs combined out of 32 different input-output tables; each pair represented a spectrum of techniques and gave rise to one of the 496 envelopes. We found one case of reswitching. About 3,5% of the 4389 switchpoints exhibited reverse capital deepening or another paradox of capital theory. The vast majority (more than 95%) of switchpoints were of the neoclassical type, with the intensity of capital falling as one moved across the switchpoint with a rise in the rate of profit both at the macroeconomic and the sectoral level. This investigation certainly had its difficulties; the ones I myself still regard as most important were discussed in the paper itself – the experiment should be repeated by others. But it had the advantage of giving a precise answer on the basis of numerous data to what a low probability of the most important paradox (reverse capital deepening at the aggregate level) means: less than one per cent of the switchpoints observed exhibited this phenomenon.

There was another, curious finding. It turned out, with 4389 switchpoints on 496 envelopes, that only about ten wage curves appeared on average on each of the envelopes, although each envelope was derived from two input-output tables, with 33 sectors. The book of blueprints thus consisted in each comparison of 66 methods for 33 industries so that the construction of each envelope involved $2^{33} \approx 10^{10}$ wage curves.

I have since endeavoured to explain theoretically why the wage curves must be nearly linear and why the paradoxes therefore are rare (Schefold 2008; 2013), but I had no explanation for the puzzle as to why only so few wage curves (about one in 10^9) appeared on each of the envelopes.

A model capable of explaining this finding is presented here. It seems useful to increase the number of countries compared in order to

understand its significance. But, before stating the main assumptions, it also seems useful to discuss the meaning of the comparison of the input-output tables in more detail. Joan Robinson once wrote: 'Nothing could be more idle than to get up an argument about whether reswitching is "likely" to be found in reality' (Robinson, 1979, p. 82). She insisted that two economies separated in space or one economy at two different dates had different states of technological knowledge. Hence she thought that there is 'no such phenomenon in real life as accumulation taking place in a given state of technical knowledge'. Her argument has been repeated by Neo-Keynesians as a criticism of neoclassical theories, but also in order to question the relevance of the Neo-Ricardian analysis of capital; hence the necessity to deal with it here. Comparisons of coexisting techniques are made all the time and this suggests that there is something in the idea of' 'accumulation in a given state of technical knowledge' (*ibid*.). Much public interest is focused on the question of the choice technique in the energy sector. German electricity companies are imitating Danish windmills just as once German railway engineers imitated British railway construction, and in both cases the new method replaced another, which was also known. To use wind power today means to return to an old technique, certainly with modifications, but using old knowledge, foreign knowledge and some new ideas.

Consider a larger geographical region. If there are, for example, ten countries each represented by 100 sectors, and if we assume that the entrepreneurs in any given industry and a given country have some knowledge of the methods employed by their rivals in some industry in the nine other countries, a great deal of international comparisons are made by all the entrepreneurs taken together, in a decentralized fashion, and there will be much striving according to the models set by others.

Of course, the methods cannot always easily be transferred. There are natural differences: the transport sectors of Greece (ships) and Austria (railways) are different. We abstract from this at present, because there are so many industries where the natural differences are irrelevant, both in manufacturing and in services. There are institutional differences. Trains cannot be as fast in Germany as in France, because German federalism indirectly prescribes that trains stop more often in smaller towns. The institutions can be transferred, if people are willing, but only slowly. The same is true for fixed capital. Formally, fixed capital can be reduced to integrated systems, which are akin to circulating capital systems (Schefold, 1989 [1971]); it then takes a certain number of years to build a stock of machines of balanced age composition. Or fixed capital can be treated in the manner of Leontief, and a new stock has to be built up,

if the method is copied by one country from another. Both approaches demonstrate that the transfer of methods involving fixed capital is slower than if only circulating capital is involved, but it remains feasible. Leontief's approach is more amenable to empirical analysis, since data for the stock matrices are more readily available than the data which are required if one wishes to follow von Neumann's and Sraffa's joint production approach to fixed capital (Schefold, 2012). Results by Mariolis and Tsoulfidis (2010) seem to show that the wage curves are closer to linearity, if fixed capital according to Leontief is taken into account, than if all capital is circulating. But here we will stick to circulating capital.

A further legitimate objection to the idea of transfer is that there are linkages between industries. If one country wants to follow the example of another by adopting modern windmills for electricity production, it will also need energy storage solutions, such as water reserves pumped to elevated artificial lakes for electricity generation in windless conditions. The decentralized nature of wind power generation will require a large grid, connecting areas where the wind blows strongly and regularly with centres of consumption. A country relying on nuclear energy can do with a different, smaller grid; on the other hand, it will need access to reprocessing plants. But it does not matter much for the other industrial sectors how the electricity is produced, as long as it becomes available. Hence linkages of this type do not concern the whole economy, and the most important of them will have been taken into account by aggregation. As the example of electricity generation shows, each of 100 sectors in an input-output table comprises a multitude of connected activities. I do not deny that linkages embracing the economy as a whole may exist. The most important example may be national innovation systems, which contain multiple institutions connecting public and private research. But we exclude such linkages in this chapter, except for a brief consideration in Section 3.3.

Although the products of sectors are generally not strictly homogeneous (except for electricity and a few others), their successful classification in international comparisons of input-output analysis vindicates the homogeneity postulate sufficiently for empirical and analytical purposes. How the input-output specialists do this does not concern us here, although their prior aggregation of firms to sectors may hide sizeable variations of capital–labour ratios between industries.

Finally, we have to be aware that countries, and industries within countries, are at different stages of advancement; the diffusion of known techniques is almost always associated with some improvement. Progress and the transmission of given knowledge tend to be linked, which makes

it interesting to compare the input-output tables of countries. It is true that it seems paradoxical to use the input-output tables of different countries as a book of blueprints, for if techniques are transferable, why are they different in different countries? Conversely, if techniques are different, that seems to be proof that they are not transferable so what is the point of comparing them? If we were in a stationary state, with no technical progress, with capital perfectly mobile and no natural or institutional advantages for individual countries, we should in fact expect that the most profitable techniques would have been adopted in each country so that all would use the same technique at the same rates of profits and wages in competitive conditions.

But only the rate of profit is fairly quick to adapt because of the mobility of financial capital; the methods are relatively inert and change in conjunction with progress. The envelope, which can be derived from such a book of blueprints, thus indicates the target technique towards which the entrepreneurs should look in each country; at the same time, they will try to estimate what progress could add to the productivity gain that would result from mere imitation. The knowledge both of what the coefficients of the tables indicate as known techniques and what potential improvements are likely to be feasible is dispersed. External effects in networks, communication among entrepreneurs and engineers and the flexibility of institutions will influence the outcome.

The process of imitation among developed countries thus achieves less than the envelope indicates, insofar as so many obstacles have to be overcome in copying the methods of others, but it achieves more, insofar as progress can be expected to accompany imitation. It may be a big problem for developed countries to identify best-practice techniques, which are constantly evolving. It is a lesser problem for backward countries since it does not matter so much whether they take target the really best techniques; the second or third best may still constitute a big advance relative to the position in which they are. Even the planned economy of the Soviet Union was able to move forward quickly, as long as it was in very backward state, but it got stuck when quality began to matter more. Catching up becomes the more difficult, the closer one is to the top, and a real overtaking, with a new country obtaining world leadership in technology from another, has occurred only a few times in history.

With this interpretation in mind, we return to the formal comparison of wage curves, thought to be derived from input-output tables of actual economies. I present a summary of the theoretical reasons why we may expect them to be nearly linear, in accordance with the empirical results referred to above.

3.2 Nearly linear wage curves

As usual, prices of the system for a given technique (one method in each industry) follow from

$$(1 + r)\mathbf{Ap} + w\mathbf{l} = \mathbf{p}$$

$\mathbf{A} = (a_{ij})$; $i, j = 1,...,n$; input-output coefficients, $\mathbf{l} = (l_i)$ labour vector, $\mathbf{p}$ normal prices, $\boldsymbol{w}$ wage rate, r rate of profit. Prices are normalized by means of a numéraire vector $\mathbf{d} = (d_1,... d_n)$, $\mathbf{dp} = 1$, where $\mathbf{A} \geq 0$, $\mathbf{d} > 0$, $\mathbf{A}$ indecomposable and productive. Prices in terms of the wage rate

$$\hat{\mathbf{p}} = \mathbf{p}/w = (\mathbf{I} - (1 + r)\mathbf{A})^{-1}\mathbf{l} > 0$$

rise monotonically from $\hat{\mathbf{p}}(-1) = \mathbf{l}$ via $\hat{\mathbf{p}}(0) = \mathbf{u}$ (labour values) to infinity at $r = R > 0$ (R maximum rate of profit of this system). Hence the wage rate $w(r)$ follows from $1 = \mathbf{dp} = \mathbf{d}\hat{\mathbf{p}}w$; $w = 1/\mathbf{d}\hat{\mathbf{p}}(r)$ falls monotonically, $w(r) > 0$; $0 \leq r \leq R$; $w(R) = 0$. Suppose that $\mathbf{d}$, the numéraire, is also the net product of the economy, produced at activity levels $\mathbf{q}$, $\mathbf{q}(\mathbf{l} - \mathbf{A}) = \mathbf{d}$, so that output per man employed $y = \mathbf{dp}/\mathbf{ql} = 1/\mathbf{ql} = w(0)$, $\mathbf{ql}$ employment, is constant in the stationary state. Capital per head $k = \mathbf{qAp}/\mathbf{ql}$ follows from $y = rk + w$, $k = (y–w)/r$; it varies with r along the wage curve, unless the wage curve is linear; one can read k off the wage curve; $k = tg\alpha = (w_1 - w_0)/r$, as at P_0 in Fig. 3.3. In the neoclassical case of Fig. 3.2, each wage curve with $w_1(0) > w_2(0) > w_3(0) > w_4(0)$ is associated with a unique capital intensity $k_1 > k_2 > k_3 > k_4$.

Consider the spectrum of eigenvalues of matrix $\mathbf{A}$. If we exclude imprimitive matrices, which are of interest only as special cases (see Schefold, 2008), $\mathbf{A}$ has a unique Frobenius eigenvalue μ_1, $0 < \mu_1 < 1$, such that all other eigenvalues $\mu_2,... \mu_n$ are smaller in modulus; they may be ordered $\mu_1 > |\mu_2| \geq |\mu_3| \geq \cdots \geq |\mu_n| \geq 0$. It would be possible to include eigenvalues that are semi-simple roots of the characteristic equation, using the approach of Schefold (1989 [1971]), but we exclude them in order to keep the elegance of the formulae (semi-simple roots are not generic anyway). Then we obtain the otherwise perfectly general expression, using the same approach as in Schefold (1989 [1971]), with $\mathbf{q}_i$, $\mathbf{x}_i$, being the left-hand and right-hand eigenvectors of $\mathbf{A}$; $\mathbf{q}_i\mathbf{A} = \mu_i\mathbf{q}_i$, $\mathbf{Ax}_i = \mu_i\mathbf{x}_i$; $i = 1,..., n$; and $\rho = 1 + r$, $\mathbf{q}_i(\mathbf{I} - \rho\mathbf{A}) = (1 - \rho\mu_i)\mathbf{q}_i$:

$$1/w(r) = \mathbf{d}\hat{\mathbf{p}} = (\mathbf{q}_1 + \cdots + \mathbf{q}_n)(\mathbf{I} - \rho\mathbf{A})^{-1}\mathbf{l}$$

$$= \sum_{i=1}^{n} \frac{\mathbf{q}_i\mathbf{l}}{1 - \rho\mu_i} = \sum_{i=1}^{n} \frac{\mathbf{q}_i\mathbf{x}_i}{1 - \rho\mu_i} \qquad [3.1]$$

Here we have introduced a representation of $\mathbf{d}$ and $\mathbf{l}$ as linear combinations of the $\mathbf{q}_i$, $\mathbf{x}_i$; $i = 1,...,n$; respectively, with the 'strong' normalization $\mathbf{d} = \mathbf{q}_1 + \cdots + \mathbf{q}_n$, $\mathbf{l} = \mathbf{x}_1 + \cdots + \mathbf{x}_n$ (the eigenvectors are so normalized that the coefficients in the linear combinations are all equal to one). Further, we have used that $\mathbf{q}_i\mathbf{x}_j = 0$ for $i \neq j$ since eigenvectors pertaining to different eigenvalues are orthogonal.

Formula [3.1] is general, setting aside the technical complications which might spring from the non-generic semi-simple roots. All nominators in [3.1] are positive for $1 \leq \rho < 1/\mu_1$ for those eigenvalues μ_i, $i = 2,..., n$, which are real. If μ_i is not real, there will be another corresponding conjugate complex root so that the sum will become real for ρ real. As ρ tends to $1/\mu_1$; $\mu_1 = 1/(1 + R_1)$; R_1 maximum rate of profit; w tends to zero.

The form of [3.1] confirms that wage curves can be very complicated, with n being large, but one immediately obtains the following familiar simplifications:

If $\mathbf{d} = \mathbf{q}_1$, $\mathbf{q}_2 = \cdots = \mathbf{q}_n = 0$, $\mathbf{d} = \mathbf{q}_1$ is proportional (not necessarily equal) to Sraffa's standard commodity and the wage curve becomes linear. We call $\mathbf{q}_1$ the Sraffa vector pertaining to $\mathbf{d}$, even if the $\mathbf{q}_2,...,\mathbf{q}_n$ do not all vanish.

If $\mathbf{l} = \mathbf{x}_1$, $\mathbf{x}_2 = \cdots = \mathbf{x}_n = 0$, the labour theory of value holds because $\mathbf{l}$ is the right-hand side eigenvector of $\mathbf{A}$ so that the organic compositions and the capital intensities are the same in all sectors. The wage curve is linear. We call $\mathbf{x}_1$ the Marx vector pertaining to $\mathbf{l}$, even if the $\mathbf{x}_2,...,\mathbf{x}_n$ do not all vanish.

Not yet familiar (but compare Schefold, 2008 and 2013) is the case $\mu_2 = \cdots = \mu_n = 0$. The wage curve becomes a hyperbola:

$$1 / w = \frac{\mathbf{q}_1\mathbf{x}_1}{1 - \rho\mu_1} + \mathbf{q}_2\mathbf{x}_2 + \cdots + \mathbf{q}_n\mathbf{x}_n \qquad [3.2]$$

This case looks at first as if it were only of formal relevance, but it turns out to be of great economic interest. It is discussed with more rigour and with more ramifications in Schefold (2013). Here I hope to provide a useful complement to that exposition by presenting a more intuitive and more concise argument.

If the non-dominant eigenvalues are all strictly equal to zero, $\mathbf{A}$ is a matrix of rank 1, and, being semi-positive and indecomposable, $\mathbf{A}$ must be positive and can be written as $\mathbf{A} = \mathbf{cf}$, where $\mathbf{c}$ is a positive column and $\mathbf{f}$ a positive row.[4] This may seem special, but, for $\mathbf{f} = \mathbf{e} = (1,...,1)$, $\mathbf{A}$ is the determinate limit case of random matrices, discussed in Schefold (2013). This looks even more special, but random matrices can be regarded as perturbations of $\mathbf{A} = \mathbf{ce}$ such that the individual

coefficients on any row can vary a great deal. On the other hand, the condition that $\mu_2 = ... = \mu_n = 0$ is relaxed: the $\mu_2,..., \mu_n$ are only required to be small (in modulus). The main result is as follows:

> *It can be proved that the non-dominant eigenvalues tend to disappear for large random matrices, essentially defined by the condition that the coefficients on each row are i.i.d. around a mean specific for the row.*

The coefficients on each row are thus distributed with a certain variance as is explained in more detail in Schefold (2013), with references to the relevant mathematical literature. The distribution does not exclude small or zero coefficients, but it is such that different linear combinations of many rows tend to be proportional. Any two given rows may be quite different, but, for large matrices and combinations of many rows, near-proportionality obtains.

The non-dominant eigenvalues of large random matrices only tend to disappear, they are not exactly zero, as stated. Because of the mathematical difficulties associated with the analysis of large random matrices, it is convenient to work with a deterministic analogue, i.e. non-negative matrices, for which all rows are exactly proportionate and which therefore have the property that the non-dominant eigenvalues are strictly equal to zero. Such matrices, which can be written as $\mathbf{A} = \mathbf{cf}$, are necessarily positive, if they are semi-positive. They are artificial constructs, introduced here only in order to visualize some properties of random matrices, which are more realistic.

Another, though mathematically less rigorous, way to describe random matrices is to describe them as perturbations of the elements of matrix $\mathbf{A} = \mathbf{ce}$ – a perturbation which can be large enough to introduce individual zero coefficients among the elements of **A**. Now we know that the non-dominant eigenvalues disappear also if $\mathbf{A} = \mathbf{cf}$, $\mathbf{f} > 0$, but $\mathbf{f} \neq \mathbf{e}$. It is clear, for reasons of continuity, that the elements of $\mathbf{A} = \mathbf{cf}$ can be perturbed in such a way that the moduli of $\mu_2,..., \mu_n$ will remain small, as in the case of random matrices. It is not known how far these perturbations may go – a general limit theorem, analogous to that about random matrices of the type $\mathbf{A} = \mathbf{ce}$, with perturbations obeying certain constraints, is not available, to the best of my knowledge. The mathematical theory behind such a theorem would probably be quite difficult, considering how difficult the theorems about random matrices are. But it is clear that non-dominant eigenvalues may be small, even if the distribution of the elements of the input matrix is not i.i.d. The conclusion is simply that [3.2] is approximately true

not only for random matrices, but more generally for large matrices which are random perturbations (not necessarily i.i.d.) of matrices of the form **A** = **cf**. The extent of the admissible perturbations is known as a set of sufficient conditions for **f** = **e** (random matrices), but not yet in general.

Beginning with an extreme case, we assume that $f_1 > \cdots > f_n$ and $c_1 > \cdots > c_n$. **A** = **cf** then is a technique where commodities 1,2,...,*n* are (apart from perturbations, which may be introduced) of the same declining importance in all industries (relative to the unit output of the commodity), and where the industries are hierarchically ordered relative to the unit level of activity, as in the traditional image of the industrial era where, for example, steel was the most important industry (c_1 large, enlarging all $c_1 f_i$) and steel was the most important input in other industries (f_1 large, enlarging all $c_i f_1$), and coal came second, and perhaps corn production third. The ordering is compatible with limited exceptions (because of the admissible perturbations). We call such systems hierarchic. A modern example could be an economy in which information technologies are the leading sector and play a role analogous to steel in the more traditional industrial economies. One might think that hierarchic systems were general among systems of the form **A** = **cf**, for we can always order the sectors so that $c_1 > \cdots > c_n$. But then we simultaneously define a reordering of the components of **f**, since the permutations of rows and columns must be simultaneous in single product systems, if the output matrix is to remain the diagonal (unit) matrix. The conclusion, therefore, is that for [3.2] to hold strictly or approximately, linear combinations of rows of the system must be proportional on average. Individual rows and coefficients on each row may deviate from the average to some extent, which can be defined exactly in the case of random matrices, while the general mathematical theory has yet to be worked out. In other words, the distribution of the elements on the rows may be i.i.d., or there may be another distribution; an extremely asymmetric distribution obtains if the system is hierarchic. Hierarchic systems are thus interesting as a limit case, which is intriguing, since ideas of leading sectors and followers recur in the history of economic thought.

After this long, but necessary, digression, we return to the analysis of wage curves of systems which, for whatever reason, being random or not, have small non-dominant eigenvalues so that [3.2] holds approximately. Now it turns out that much less is needed than that the *numéraire* be equal to Sraffa's standard commodity or that the labour theory of value be valid to obtain a linear wage curve, if the matrix of

the system is of rank one. Consider the vector of deviations **m** of the numéraire vector **d** from the Sraffa vector $\mathbf{q}_1$

$$\mathbf{m} = \mathbf{d} - \mathbf{q}_1 = \mathbf{q}_2 + \cdots + \mathbf{q}_n$$

and the vector of deviations **v** of the labour vector **l** from the Marx vector $\mathbf{x}_1$

$$\mathbf{v} = \mathbf{l} - \mathbf{x}_1 = \mathbf{x}_2 + \cdots + \mathbf{x}_n$$

Let $\bar{m}$ designate the mean of the components of **m** and $\bar{v}$ the mean of the components of **v.** If $\bar{m} = 0$, the deviations of the *numéraire* from the (standard) Sraffa vector are zero on average, and if $\bar{v} = 0$ the analogue holds for the labour deviations and one might say loosely, in a Marxian vein, that the labour theory of value holds on average. Now on the one hand, using the orthogonality condition:

$$\mathbf{mv} = (\mathbf{q}_2 + \cdots + \mathbf{q}_n)\ (\mathbf{x}_2 + \cdots + \mathbf{x}_n) = \mathbf{q}_2\mathbf{x}_2 + \cdots + \mathbf{q}_n\mathbf{x}_n$$

On the other hand, one has the known formula for the covariance of coefficients of the deviations (considered as random variables):

$$cov\ (\mathbf{m},\mathbf{v}) = (1/n)\mathbf{mv} - \bar{m}\bar{v}$$

There is no obvious reason for a significant correlation between **m** and **v.** The numéraire vector **d** can be chosen arbitrarily, while **l** can be assumed to be random for a quite different, independent reason: it reflects technology. Similarly, the random character of given $\mathbf{q}_1$ and $\mathbf{x}_1$ depends on the random character of the system as a technique. We are looking for the theoretical potential causes why empirical wage curves turn out to be nearly linear. The solution is first to assume that *cov* **(m, v)** = 0, so that $\mathbf{mv} = n\bar{m}\bar{v}$ and [3.2] becomes

$$1/w = \frac{\mathbf{q}_1\mathbf{x}_1}{1-\rho\mu_1} + n\bar{m}\bar{v} \qquad [3.3]$$

The wage curve of a system, which is random and/or of the form $\mathbf{A} = \mathbf{cf}$ with perturbations then is nearly linear, if the numéraire deviations are zero on average and/or if the labour theory of value holds on average:

$$w = \frac{1-\rho\mu_1}{\mathbf{q}_1\mathbf{x}_1} \qquad [3.4]$$

[3.4] is Sraffa's wage curve. We have thus made a big theoretical advance relative to a long-standing discussion. We have found that the linear wage curve results not only if one has the standard commodity or if the labour theory of value holds. *It is sufficient that either of these properties holds on average,* more formally, that $\bar{m} = 0$ and/or $\bar{v} = 0$, provided the system is random, or, more generally, that the non-dominant eigenvalues are small.

The expressions [3.2] and [3.3] are important even if linearity does not obtain, because they allow to explain the complications of the wage curves: if a wage curve is not a hyperbola, it must be due to non-dominant eigenvalues which are not zero. The work by Mariolis and Tsoulfidis on actual input-output systems has shown that most but not all eigenvalues are close to zero. If ordered according to the moduli, they seem to fall rapidly towards zero according to an exponential law. This tendency remains to be explained. Meanwhile, we can show how a wage curve with h eigenvalues (including the dominant) of significant modulus and $n - h$ eigenvalues of negligible size can be represented as the hyperbola of the form [3.2] or [3.3], with $h - 1$ terms superimposed, which cause shifts and wiggles. Extending the idea of the deviations, we define

$$\mathbf{m}_h = \mathbf{d} - (\mathbf{q}_1 + \cdots + \mathbf{q}_h) = \mathbf{q}_{h+1} + \cdots + \mathbf{q}_n$$
$$\mathbf{v}_h = \mathbf{I} - (\mathbf{x}_1 + \cdots + \mathbf{x}_h) = \mathbf{x}_{h+1} + \cdots + \mathbf{x}_n$$

Combining conjugate complex solutions, $\mathbf{m}_h$ and $\mathbf{v}_h$ are real, and, assuming zero covariance, in obvious notation $\mathbf{m}_h\mathbf{v}_h = h\bar{m}_h\bar{v}_h$. This yields

$$1/w = \frac{\mathbf{q}_1\mathbf{x}_1}{1-\rho\mu_1} = \sum_{i=2}^{h} \frac{\mathbf{q}_i\mathbf{x}_i}{1-\rho\mu_i} + h\bar{m}_h\bar{v}_h \qquad [3.5]$$

The right-hand side of [3.5] contains the terms which may cause deviations from linearity. If one compares the wage curves of many systems, using the same *numéraire* for all, it may not be assumed that $\bar{m}_h$ will be very small, since, if one assumes that $\mathbf{d} = \mathbf{q}_1$ for one of those systems, there may be a non-random drift in the transition to the other systems, but $\bar{v}_h$ could be quite small for most systems.[5] Hence we assume that the last term in [3.5] can be neglected in most cases. The influence of the first h–1 terms on the right-hand side of will be small, if the corresponding contributions $\mathbf{q}_i$ to the numéraire deviations and $\mathbf{x}_i$ to the

labour value deviations will be small. Their influence will grow as $\rho\mu_i$ approaches one,[6] but it will not become infinite:

> *Hence the possibility to explain why nearly linear wage curves will be relatively frequent, why strongly curved wage curves with considerable wiggles will be less frequent, and why the deviations from linearity are larger at higher rates of profit, as the empirical wage curves show.*

The question now is how many wage curves will make it and appear on the envelope. If that number is small, the envelope can be expected to be composed mainly of nearly linear wage curves, because *the less frequent outliers remain below.*[7] The main justification for this procedure perhaps bears repetition. We do not postulate a general new theory of nearly linear wage curves, but we propose to explain the empirical finding that wage curves are nearly linear in the relevant range, and for this explanation we do not postulate that input-output tables are generally random or that the labour theory of value holds generally on average, but that these properties hold in combination to a sufficient degree. And who could deny that there is at least some randomness in the emergence of methods of production?

We thus invoke a combination of properties to argue that the wage curves encountered on the envelope will tend to be nearly linear. By implication, the amount of capital per head 'demanded' at each rate of profit will tend to fall, as the rate of profit rises. But will the 'supply' of capital per head fall accordingly as the transition is made from one technique to another at any switchpoint? We must be brief on this point. Since technical change is piecemeal (Han and Schefold, 2006), each transition requires the replacement of one and only one method of production by another in one industry, say the first. In the usual neoclassical perspective, the amount of capital is kept constant in the transition and more labour is employed; hence the intensity of capital falls. The capital used with the technique on the left of the switchpoint can be transferred into an equal amount of capital to be used with the technique on the right of the switchpoint. The transition to a newly invented technique (to a higher wage curve) would have required acts of saving and investment, but mere substitution does not need that – the means of production are transformed, that is, the means of production of the first process used on the left of the switchpoint are sold, and with the proceeds the means of production needed on the right of the switchpoint can be bought. But how is this possible, if the means of production are

industry-specific? Think of a 'quasi-classical' example: the first industry produces cloth, the method employed on the left uses machines driven by steam engines, the method employed on the right uses hand-driven looms and employs more workers. Who wants to buy the machines, if the wage rate falls? Who will have produced the looms? Why the equality of value? Clearly, the machines would have to be written off, as if the Luddites had won, and new investment, financed by saving, would be necessary to buy the looms. Hence we would have to argue in terms of innovation instead of talking about 'substitution'. This transition would be more plausible as a movement from right to left: the intensity of capital increases with growth. The neoclassical vision of the transition as substitution cannot generally hold, as the quasi-classical example demonstrates, but the neoclassical conception of the 'supply of capital' *is* consistent under the following restrictive conditions: if the means of production are not specific for the industry, those used on the left of the switchpoint can be sold to other industries, and a somewhat different 'combination' of means of production ('factors') can be bought; activity levels will adapt. The transition can be short, if only circulating capital is involved, but it takes longer with fixed capital. Pure theory avoids the consideration of the transition and is content with the comparison of steady states. In either case, the wage curves will be nearly linear, if the conditions for the averages derived in this section hold to a sufficient degree. More precisely: industries 2,…*n* do not change and industry 1 must, apart from perturbations, be equal to a linear combination of the other industries, if we are dealing with random systems. *Hence the change of the means of production in the transition must not be larger than the admissible perturbations.* Moreover, the labour theory of value should hold on average. The conditions for the neoclassical theory to hold are essentially the same, both as regards 'demand' and 'supply' of 'capital'.

3.3 The lens of wage curves and its envelope

The question of whether reswitching is more than a fluke was first approached in terms of set theory. It was shown (Schefold, 1976a) that the set of potential methods engendering wage curves that intersect twice with the non-linear wage curve of a given system is not of measure zero in the set of potential methods engendering wage curves that interact at least once with the wage curve of the given system. But the economic content of this concept of measurement was problematic. How densely populated is the continuous space of potential methods with discrete methods that can actually be used? D'Ippolito, Petri,

Salvadori, Steedman (see Petri, 2010) and others have discussed this with interesting results but the measurement problem has remained. This is why I prefer to start here with a book of blueprints that is thought to result from the comparison of input-output tables. The empirical turn thus made possible represents a new paradigm, prepared by the many articles dedicated to the empirical analysis of wage curves. The methods are given and can be now counted. We thus bridge the gap between theory as a thought experiment and actual measurement, and we can approach the problem of the likelihood of the appearance of the paradoxes by combinatorial methods.

Let therefore k tables for k countries, each with n sectors, be given. We assume that there are no links between methods other than those within sectors. We assume away natural obstacles to the transfer of methods and suppose that, though that may take time, methods can be transferred with the associated institutional changes. We keep in mind that the concept of transition between these techniques is problematic for various reasons, especially because, whenever entrepreneurs strive to replace method α by β, which seems more profitable, they will get new ideas and end up with a method γ, of which the question is then whether it still resembles β or whether it looks like an outgrowth of α. This is true, but the wage curves and their envelopes remain important theoretical tools.

Given the book of blueprints, we obtain $s = k^n$ wage curves $w_\sigma(r)$; $\sigma = 1,...,s$. We assume them to be strictly linear, to begin with, as will be the case if all techniques are of the form $\mathbf{A}^{(\sigma)} = \mathbf{c}^{(\sigma)}\mathbf{f}$, all indecomposable, with $\mathbf{f}$ as common numéraire; $\sigma = 1,...,s$. We call this the straight lines case. Strictly speaking, two wage curves cannot have a common switch point on the envelope, if they are straight lines, except for one special case (Schefold, 2008), but it is almost obvious that we can disregard this problem here.[8] All $w_\sigma(-1) > 0$, since $\mathbf{l}^{(\sigma)} > 0$, but we do not necessarily have $w_\sigma(0) > 0$. But we can, without loss of generality, assume an ordering such that $w_1(0) > w_2(0) > \cdots > w_s(0)$. The maximum rates of profit $R_1,..., R_s$ will appear in a different ordering $R_{\sigma_1} > \cdots > R_{\sigma_s}$, where $(\sigma_1,...,\sigma_s)$ is a permutation of $(1,2,...,s)$. If technique σ is productive, we have $R_\sigma > 0$, but even if it is not, we have $1 + R_\sigma > 0$, hence $R_\sigma > -1$, since $\mathbf{A}^{(\sigma)} \geq 0$. Our s wage curves, s being a large number, will fill a concave lens with $w_{\max}(-1) \geq w_\sigma(-1) \geq w_{\min}(-1)$ and $R_{\max} \geq R_\sigma \geq R_{\min}$; $\sigma = 1,...,$ s; as in Figure 3.4. There is an upper and a lower envelope for the lens; the envelopes will look smooth (although they are composed of a finite number of straight lines) if many wage curves appear on them. *But will this be the case?*

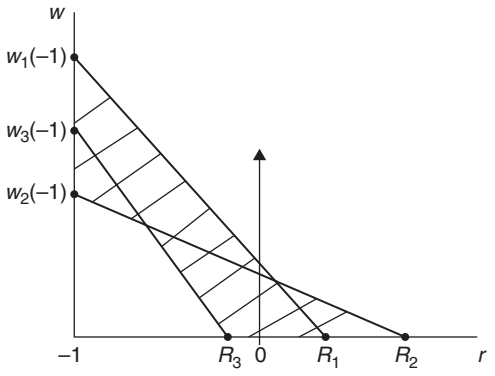


Figure 3.4 The lens composed of straight wage curves between two concave envelopes: ordering of the curves according to the level of the wage at $r = 0$.

A large number of wage curves, namely s, will be found within the lens; only three have been drawn. Observe that the lens will look more like a triangle, if $R_{\max} - R_{\min}$ is small and $w_{\max}(-1) - w_{\min}(-1)$ large. This would be the image corresponding to Kaldor's stylized facts: the capital-output ratio, represented by the inverse of the maximum rate of profit, would stay nearly constant and the capital-labour ratio would rise as one ascended the ladder of wage curves with rising productivity, in a temporal sequence (not in one given period, but in a state of rising knowledge).

What changes, if the straight lines are replaced by monotonically falling curves? The envelope would still be monotonically falling, but the lens would not necessarily be concave. The wage curves of non-productive techniques would still be not positive at positive rates of profit, like $W_3(r)$ in Figure 3.4. The deviations of the wage curves from straight lines would be most pronounced near the maximum rates of profit, according to the analysis of Section 3.2 above, hence the Kaldorian stylized facts would have to be questioned; the problems of capital theory affect Neo-Keynesian as well as neoclassical economics.

We return to the assumption of straight wage curves in order to address our main theme. We simplify by assuming that all techniques are indecomposable and productive. We number the wage curves so that $w_1(0) > w_2(0) > \cdots > w_s(0)$, and we make the decisive equal-probability assumption: in the ordering of the maximum rates of profit $R_{\sigma_1} > \cdots > R_{\sigma_s}$, all permutations $(\sigma_1,\ldots, \sigma_s)$ of $(1,\ldots, s)$ are equally likely. For why should, given such a large number of possibilities, anything else be assumed? If w_σ is relatively small so that $w_\sigma(r)$ represents a 'bad' technique for low

rates of profit, why should technique σ suddenly be 'good' at high rates? A neoclassical economist might postulate that a low w_σ (0) should be compensated by a high R_σ, but this is justified only as an observation of what perhaps happens on the upper envelope of the wage curves as a result of optimization. If w_σ (0) is low and σ has been chosen at random, R_σ will be random and thus may be high or low. Each technique σ results from the arbitrary combination of n methods, each taken arbitrarily from k tables. The quality of the technique, characterized in the linear case by w_σ (0) and R_σ, is unknown, and only optimization leads to a subset of techniques where the trade-off, high w_σ (0) implies low R_σ and *vice versa,* may be visible.

The referee suggested an argument which represents the opposite of the neoclassical supposition: since some techniques will plausibly embody more technical progress than others, it would seem more likely that a higher maximum wage will tend to be associated with higher efficiency in general and therefore with a higher maximum rate of profit, too. The argument seems to rely on linkages between all sectors, which we had excluded, assuming that linkages hold only within sectors. If, for the sake of the argument, we now admit the general linkages, related to progress, we do not necessarily arrive at the conclusion suggested by the referee. Research for innovation is often at least subjectively directed at either preferentially saving labour or certain means of production. One can also recall the different forms of technical progress with mechanization and saving of labour being fostered by objective class struggle in Marx (Schefold, 1976). They imply a tendency to raise $w(0)$ at the expense of reducing R. But the Marxian argument was strong only as long as it concerned a leading sector (textiles). A more important counterargument, in my view, is the following: Research in any sector takes its direction in a given environment, i. e. given the methods of production, and the consequent prices in other sectors, but the overall linkages are too weak to count. The spectrum of techniques consists of all $s = k^n$ methods. A successful national innovation system may render many sectors effective, but combining with methods from other countries can still increase efficiency, even in the country which leads in most, but not all, sectors. Even Mephisto will have difficulty visualizing all possible combinations, and the techniques appearing on the theoretical frontier are not likely to be adopted by any of the k countries (a point to which we shall return in the end). Most systems consist of methods of which only a small number coexist in any one country. There are k systems (the actual ones), where all methods coexist. There are $kn(k–1)$ systems, where all methods but one belong to one actual

country, and so on. What can we say about the efficiency of the many systems that combine methods from many countries? Insurmountable ignorance compels us to maintain the equal probability assumption. Hence the randomness of the process in which the actual market tends to approximate the results of an ideal market.

If we make the assumption of equal probability, together with the other, more innocent ones, which have been stated, we get at once a definite probability for the Joan Robinson case in Figure 3.1 (one wage curve constitutes the envelope). Since $w_1(0)$ is on the envelope by assumption about the ordering, $w_1(r)$ must be the envelope, and we must have $R_1 > R_\sigma$, $\sigma = 2,\ldots, s$. Since R_1 could *a priori* have been in any of the s positions with $R_{\hat{\sigma}1} > \cdots > R_{\hat{\sigma}_s}$ the probability is $1/s$. If $k = 10$ and $n = 100$, $1/s = 10^{-100}$, a very low probability. In the empirical analysis by Han and Schefold (2006), 496 envelopes of wage curves were analysed. Since $k = 2$ and $n = 33$, $s = 2^{33} \approx 10^{10}$, the occurrence of the Robinson case could not be expected. In fact, the minimum number of wage curves encountered on any envelope was 3.

Next consider the pure neoclassical case where the order of the maximum rates of profit R_σ is *exactly* inverse to that of the $w_\sigma(0)$, as in Figure 3.2. $R_1 < \cdots < R_s$ is one permutation in $s!$ possible permutations of $R_1,\ldots, R_s$, hence a probability so small as to be neglected, and the pure neoclassical ordering was, of course, not observed in the investigation by Han and Schefold.

If the ordering of $R_1,\ldots, R_s$ is not exactly inverse to that of $w_1(0),\ldots, w_s(0)$, not all wage curves will appear on the envelope, and the wage curves appearing on the envelope, if sufficiently numerous, could still constitute something like a neoclassical production function, with a certain number of inefficient techniques with wage curves totally below the envelope left out. How many wage curves do we have to expect to appear? We derive an upper bound for this expectation.

The probability that $w_1(r)$ is on the envelope equals one, since $w_1(0) > w_\sigma(0)$, $\sigma = 2,\ldots,s$. It is clear that at least one technique must be on the envelope.

The probability that $w_2(r)$ appears on the envelope equals at most 1/2, since it is necessary for the appearance that $R_2 > R_1$, and this is one of two equally probable cases: $R_1 > R_2$ and $R_2 > R_1$.

The probability that $w_\sigma(r)$ appears on the envelope equals at most $1/\sigma$, since it is necessary that $R_\sigma > R_\tau$, $\tau = 1, \ldots, \sigma - 1$ (R_σ must be in one of σ equally likely positions). It is intuitive that the probability for $w_\sigma(r)$ to appear on the envelope diminishes as $w_\sigma(0)$ diminishes.

The expected upper bound for the total number of wage curves, say $\Omega(s)$, *on the envelope is the sum of the probabilities of the cases,*[9] *hence* $1 + 1/2 + \cdots +$

$1/s$ *which tends to* $\ln s$, *as* s *increases*. About ten wage curves were found on the envelopes on average in Han and Schefold (2006), but $\ln(2^{33}) = 33 \cdot \ln 2 \approx 22.8$. The formula $\Omega(s) = \ln s$ represents in fact an upper bound, for, depending on the spacing of the $w_\sigma(0)$ and the R_σ, some wage curves and corresponding switchpoints may get *dominated,* as is illustrated in Figure 3.5, where w_2 is not on the envelope (although $R_2 > R_1$), if the third wage curve is, given $R_3 > R_2$, defined by $\tilde{w}_3$ (0), whereas w_2 is on the envelope, if w_3 (0) is sufficiently small.

Even if we neglect this domination effect,[10] the number of wage curves appearing on the envelope is surprisingly small. The share of wage curves appearing on the envelope, say Θ, is for the upper bound

$$\Theta = \frac{\ln s}{s}$$

which tends to zero for $s \to \infty$. In fact we have, if again $k = 10$ and $n = 100$, $\ln s = \ln 10^{100} = 100 \cdot \ln 10 \approx 230$ – a surprisingly low number!

And yet something remains of the idea of the surrogate production function in this example. If the greatest of the maximum rates of profit is 100%, say (so far we are only representing circulating capital), each change of the rate of profit by one percentage point induces about two changes of method, on average, and if real wages are pushed up so that the rate of profit falls by several percentage points, it becomes, in theory, profitable to make several substitutions which raise capital per head.

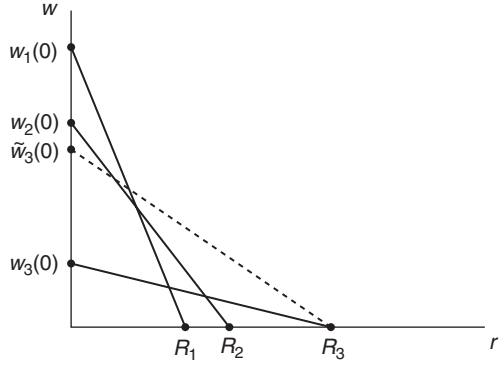


Figure 3.5 Possible domination of 'earlier' switch points by 'later' wage curves
Note: Three wage curves would be on the envelope with appropriate spacing of the $w_\sigma(0)$, R_σ, and $w_3(0)$ sufficiently small, but only two appear, if w_3 starts at $\tilde{w}_3(0)$, $\tilde{w}$ dotted wage curve.

But to get from this analysis to the elegant properties of production functions with a given and constant elasticity of substitution,[11] problematic additional assumptions would be required. In order to obtain more techniques on the envelope, given k and n, one would have to assume that the likelihood of R_σ being large increased as $w_\sigma(0)$ fell, and the $w_\sigma(0)$ and the R_σ would have to be so spaced as to obtain the curvature of the envelope which would give rise to a Cobb-Douglas or a CES production function.

We thus do not arrive at a full restoration of the production function. Another critical possibility[12] is that the elasticity of substitution of the intensity of capital to the rate of profit might be so low as to be irrelevant. If there are thirty switches, say, in the relevant range of the rate of profit, none of them associated with reverse capital deepening, but causing an overall change of capital per unit of labour of only about five percent, the increase in the demand for labour would be too low, even conceding a given 'supply of capital'. The argument again turns empirical at this point. Neoclassical authors have sought to render their argument plausible by considering the extreme cases. If wages fell really low, pre-industrial methods of production would again become profitable. Ten workers using the spade for free are cheaper than one, also working for free, but using a tractor. Only capital costs count at $w = 0$, $r = R$; labour costs nothing. The converse case is more difficult and possibly more relevant. Whether unemployment can be removed by lowering wages and mere technical substitution, without stimulating effective demand, in a closed economy, is a somewhat academic question – the main positive effects on employment of lowering wages will in practice come from increases in exports. But what about the ability of the capitalist class as a whole to reduce employment, while maintaining production, in reaction to high wage claims? Will an economy with strong trade unions end up with what some call a higher 'natural' rate of unemployment? This question, which plays a central role in the Marxian theory of accumulation (mechanization as a means of saving labour, resulting periodically in crises), would be formally the opposite of $w = 0$ and $r = R$, not if $r = 0$, but if $r = -1$! We are used to saying that wages are their maximum, if the rate of profit is zero, because we cannot really conceive of a negative rate of profit. But capital must still be paid for at $r = 0$. By contrast, at $r = R$, there is no expense on wages.[13] If the machines could be used for free and all costs were direct wage costs, as at $r = -1$, an indefinite number of machines could be used to replace labour; then it is plausible that this would happen. However, $r = -1$ is surely outside the relevant range of the rate of profit (although

we found it analytically convenient to include $r = -1$ in Figure 3.4 in order to analyse the 'lens'). Whether the elasticity of the intensity to capital is high enough to create unemployment by substituting capital for labour in the relevant range (w high, r low but positive) is again an empirical question. To this extent, the referee is surely right to ask whether the elasticity of the intensity of capital with respect to distribution is really high enough to justify the explanation of unemployment by high wages in the context of a closed economy.

At this point, the reader, reminded of economic history, will remember other problems of the choice of technique. Even if we stick to the linear wage curves, the suspicion arises that the economy will hover below the surface of the ocean among a multitude of not quite efficient techniques, in accordance with the vision discussed above (see Section 3.1) – that diffusion takes time, that the imitation of known technology is mixed up with progress and that such imitation becomes more challenging as one approaches the efficiency frontier. Hence it seems better to work with the apparatus of the wage curves and their envelope, and not with the problematic idealization, the production function.

What remains of this analysis, if the wage curves are not strictly linear? We have argued that the deviations from linearity will only exceptionally be large and that the number of wage curves on the envelope will be relatively small, hence we may expect that the wage curves deviating drastically from linearity even at small rates of profit are likely to be inside the lens of all wage curves. However, more empirical work, involving the comparison of the wage curves derived from different input-output tables, not only the wage curves derived from individual input-output tables, is needed, as well as more theory to explain the curious spectra of eigenvalues of empirical input-output tables, before we can come to safe conclusions. We hope to have opened up a field of research where the existence of the production function and related questions can be discussed by other means than mere *a priori* reasoning.

Notes

1. A reflection of her teaching is to be found in Robinson (1979, p. 267), where the better technique implies a rise of both output per head and of the maximum rate of profit. She added: 'This appears to correspond to the typical development most prevalent in modern large-scale industry', and she rejected the idea that the rise of output per head might be bought by a reduction of the maximum rate of profit (*ibid.*, p. 272).
2. The constellation shown in Figure 3.2 can also be interpreted as a succession of techniques with a rising organic composition of capital, as in Schefold

(1976), but Joan Robinson objected; see our subsequent exchange of letters in the Joan Robinson Archive and Robinson (1979, p. 272).

3. Figure 3.3 shows how the amount of capital to be employed in a steady state at a given rate of profit can be determined. This 'demand' for capital changes in the direction opposed to the hypotheses of neoclassical theory, if there is reverse capital deepening. This chapter shows that the objection is less damaging for neoclassical theory (provided it is based on empirical generalizations *à la* Schmoller and is not meant to hold *a priori à la* Menger) than I, for one, used to think. Another possible objection is that the 'supply of capital' cannot be defined meaningfully in the transition between steady states, as in this thought experiment based on Figure 3.3. Objections to the 'supply of capital' also are important; we shall come back to the supply side at the end of Section 3.2.
4. It is easy to prove that if $\mathbf{A} > 0$, $\mu_2 = \cdots = \mu_n = 0$, if and only if $\mathbf{A} = \mathbf{cf}$, $\mathbf{c} > 0$, $\mathbf{f} > 0$. Note moreover that, if $\mathbf{A}$ is given, $rk\mathbf{A} = 1$, $\mathbf{c}$ and $\mathbf{f}$ are each determined up to a scalar factor, and these factors must be inverse to each other.
5. Theoretically, $\bar{v}_h$ could be large, of course. But, if the systems are, technically speaking, random also with respect to the labour coefficients, $\bar{v}_h$ must be small, and this is our suggested explanation for why empirical wage curves in a spectrum of techniques seem to be sufficiently linear to cross only very rarely more than once.
6. Assuming $\mu_1 > \mu_i > 0$. Somewhat different results are obtained, if $-\mu_1 < \mu_i < 0$ or for conjugate complex μ_i, μ_{i+1}. The reader is invited to work out these cases for him- or herself.
7. The conclusion is obvious, as long as one has no reason to suppose that the outliers are on average technically superior.
8. To begin with the exception: let $w_1(r)$ be linear because the labour theory of value holds, and assume that another system with wage curve $w_2(r)$ becomes dominant at some rate of profit r_1 which results from the substitution of another method in one of the industries of the first system. Now linearize the second wage curve by taking the standard commodity of the second system as the common *numéraire*; w_1 will remain linear. But if a third technique appears, with a second switchpoint on the envelope, it cannot in general have equal organic composition of capital in all industries and the *numéraire* has already been fixed so that $w_3(r)$ cannot be strictly linear. The difficulty does not matter, because we are here really dealing with quasi-linear wage curves, i.e. wage curves that are nearly linear in the relevant range (usually well below the maximum rate of profit). If a system is large and its wage curve is nearly linear, and if one method of production is replaced, the resulting new wage curve will also tend to be nearly linear for reasons of continuity. In order to simplify the analysis which follows, we assume strict linearity, but it would do to assume that the curvature of wage curves is such that any two of them intersect at most once, in accordance with the empirical envelopes analysed in Han and Schefold (2006), where wage curves with a switchpoint on the envelope had at least one other intersection in common in less than two per cent of the more than 4000 cases.
9. Intuitively: imagine that you are a beggar walking in the streets of a city, and in each street σ you are given €1 with probability p_σ. Hence, if you walk in the streets 1,..., s you expect to receive € $(p_1 + \cdots + p_s)$. Now imagine that

you are walking down the envelope. The expectation of the number of wage curves is $1 + 1/2 + \cdots + 1/\boldsymbol{s} \approx \ln(\boldsymbol{s})$.

10. Is there a sufficient condition to exclude the domination effect? I guessed that equal spacing between the $w_1(0),\ldots, w_s(0)$ and between $R_{\sigma_1},\ldots, R_{\sigma_s}$ might suffice, but Christian Bidard gave a numerical example which proves that this is not the case.
11. A constant elasticity of substitution is not required for the neoclassical theory in its general form, but for most relevant applications in the modern theory of growth.
12. This was pointed out by the referee; his formulation is quoted.
13. If the subsistence wage is not counted among the means of production, as Sraffa (1960) suggests (§ 8).

References

Han, Z. and Schefold, B. (2006) 'An empirical investigation of paradoxes: reswitching and reverse capital deepening in capital theory', *Cambridge Journal of Economics*, 30: 737–65.

Mariolis, T. and Tsoulfidis, L. (2010) 'Eigenvalue distribution and the production price-profit relationship: theory and empirical evidence', *Evolutionary and Institutional Economics Review*, 8(1): 87–122.

Petri, F. (2010) '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* (London: Routledge): pp. 380–418.

Robinson, J. (1979) *Collected Economic Papers*, Vol. V (Oxford: Basil Blackwell).

Schefold, B. (1976) 'Different forms of technical progress', *Economic Journal*, 86: 806–19; repr. in Schefold (1997): pp. 257–75.

Schefold, B. (1976a) 'Relative prices as a function of the rate of profit', *Zeitschriftfür Nationalökonomie*, 36; repr. in Schefold 1997: pp. 46–75.

Schefold, B. (1989 [1971]) *Mr. Sraffa on Joint Production and Other Essays* (London: Unwin & Hyman, now Andover: Routledge).

Schefold, B. (1997 [1979]) *Normal Prices, Technical Change and Accumulation* (London: MacMillan).

Schefold, B. (2008) 'Families of strongly curved and of nearly linear wage curves: a contribution to the debate about the surrogate production function', *Bulletin of Political Economy*: 1–24.

Schefold, B. (2012) 'A comment on "The Measurement of Prices of Production" by Peter Flaschel, Reiner Franke and Roberto Veneziani', *Review of Political Economy*, 24: 469–76.

Schefold, B. (2013) 'Approximate surrogate production functions', *Cambridge Journal of Economics*, 37: 1161–84.

Sraffa, P. (1960) *Production of Commodities by Means of Commodities* (Cambridge: Cambridge University Press).

4

On the Stability of the Ramsey Accumulation Path

Enrico Bellino[*†]

4.1 Introduction

Infinite horizon neoclassical optimal accumulation theory is characterised by an analytical apparatus that now pervades several fields of theoretical analysis: capital theory, growth and value theory, macrodynamics, etc. In this context the equilibrium solutions take the form of saddle paths which are, therefore, *unstable* paths. The convergence of these solutions to the steady-state equilibrium is then ensured by imposing a transversality condition, that is, a condition which guarantees the optimality of the solution when time tends to infinity.

Sometimes, this kind of dynamics has been considered as accounting for the crises of actual economic systems. Quite recent work by Christopher Bliss supports this position:

> Question: Which Cambridge, England, paper of the 1960s or 1970s did more damage to simple neoclassical thinking than any other? I can imagine the post-Keynesians proposing one or more papers by

* The material presented in this chapter is the result of a set of discussions stimulated by Pierangelo Garegnani on the meaning of the 'transversality condition' within optimal capital accumulation models. I am grateful to him for these discussions. In addition, I wish to thank Andrea Battinelli, Carlo Beretta, Marco Bramanti, Thomas Christiaans, Roberto Ciccone, Ferdinando Colombo, Saverio Fratini, Kazuhiro Kurose, Enrico Sergio Levrero, PierCarlo Nicola, Fabio Petri, Mario Pomini, Giorgio Rodano, Neri Salvadori, Alessandro Sbuelz, Paolo Trabucchi, Gerd Weinrich and an anonymous referee for their comments and suggestions on earlier versions of this work. However, responsibility for what is written here is entirely mine. Finally, I am grateful to Micaela Tavasani for revising the English.

† Università Cattolica del Sacro Cuore, Largo Gemelli 1, 20123, Milano, Italy; e-mail: enrico.bellino@unicatt.it.

> Joan Robinson, but I would not agree. My own candidate would be Frank Hahn's 1968 paper on warranted growth paths. What Hahn did can be explained quite simply ... In general, equilibrium conditions for the various capital goods – that is, equal net returns for each good – depend upon the rate of change of prices, not just on the prices alone. Then the intertemporal dynamic equilibrium of the system is defined by a set of simultaneous differential equations. What can we say about those equations? The easiest route to the answer comes from the theory of optimal growth, as with a many-capital-good Ramsey model. From there we obtain dynamic equilibrium conditions (Euler equations), but these do not suffice by themselves. We need to add transversality conditions. The dynamic equations are of the saddle-point variety. Only transversality ensures that the system converges to the right asymptotic state. Hahn, following a suggestion from Mirrlees, realized that a non-optimized dynamic equilibrium is just like the necessary conditions for an optimized system. There will be infinitely many 'equilibrium' paths for prices, but most of these will crash into the sign-constraint barriers of the race track.
>
> In other words, a successful capitalist economy needs to have the right idea of where it should go in the long run. It would be hard to imagine a more telling criticism of the theoretical claim that an ideal decentralized capitalist system exhibits optimal properties. For if there is one thing that capitalist economies are really bad at, it is knowing where they are going in the long run and judging that destination correctly. Recent financial crises, and similar crises in the past, remind us that capitalism is like an inexperienced child that chases after the latest shiny attraction until it hurts itself and ends up in tears. (Bliss, 2010, pp. 648–9)

This quotation hints at several related issues that have become mixed up; in summary, Bliss seems to be suggesting that neoclassical optimal accumulation models aim to account for both the equilibrating *and* the destabilising forces of capitalist economies. He seems to say that crises in actual capitalist economies can be explained, at least partially, by the destabilising forces implicit in neoclassical optimal growth theory. At a deeper level, this claim would seem to be grounded on the idea that if the consumer optimises myopically, i.e., by considering only the current and the subsequent period, the ensuing dynamics is almost certain to diverge from the steady-state equilibrium. On the contrary, convergence to the steady-state equilibrium requires perfect foresight on the part of the consumer. Myopic optimisation, in other words, would be

what is actually responsible for the instabilities, while long-run foresight is the assumption that guarantees the convergence of short-run or momentary equilibria towards a steady-state. An interesting analysis of the role of consumer foresight ability in connection with the dynamic properties of the optimal accumulation path is proposed by Heller:[1]

> It is a curious fact that while most finite time horizon economic models are stable, almost all infinite time horizon models are unstable ... Hahn [*Q.J.E.*, 1966 and *R.E.S.*, 1968, e.b.] and Shell-Stiglitz [*Q.J.E.*, 1967, e.b.] cast considerable doubt for simple 'descriptive' economic models that efficient equilibrium paths are ever chosen by myopically optimizing, competitive agents when there is no end to time. In Ramsey planning models, there is a similar instability: if central planners initially choose the right path, minor external shocks could cause them to follow a non-optimal path when they operate according to myopic optimization rules ... It may be helpful to spell out this point a bit further. The infinite horizon models of Hahn and others restrict attention to momentary equilibrium paths arising from arbitragers possessing self-fulfilling expectations about price changes in the immediate future (we will call this myopic foresight). A path is in equilibrium at a moment in time if, given current prices and myopic foresight expectations about prices in the immediate future, agents make plans about the allocation of resources across consumption and various investment goods that clear markets at the current moment. Paths that are in equilibrium for non-infinitesimal lengths of time, but are not necessarily in equilibrium for all time, are called momentary equilibrium paths. Paths that are in equilibrium at every moment in time are called long-run equilibrium paths. Momentary equilibrium paths need not be long-run equilibrium paths because, for example, the stock of some capital good is run down to zero at some point, resulting in a disequilibrium at that time. (Heller, 1975, pp. 65–6)

In these pages I aim to challenge the idea that myopic optimisation is ultimately responsible for instability and that the assumption of perfect foresight is the only way to avoid it. In this work we will limit our investigation to an optimal accumulation model with just *one* good, in order to focus on the causes of instability in the simplest situation in which it emerges. We will study a *myopic* adjustment process compatible with consumer rational behaviour; the ensuing accumulation path will be proved to *converge monotonically* to the Ramsey steady-state equilibrium. The extension to the cases with more than one good, which are those considered by Hahn and referred to by Bliss, will be undertaken in a subsequent work.

4.2 Setting of the Cass-Koopmans-Ramsey model

Let us consider what the (representative) consumer has to do in a Robinson Crusoe economy in order to discover his optimal consumption path over an infinite time horizon. Consider an economic system where just one commodity is produced, consumed and employed jointly with labour as a mean of production of itself; capital depreciates at rate $\mu \in [0, 1]$. We define *period* t the half-open time interval $[t, t + 1)$ between *dates* t and $t + 1$. Consumers live forever, and they are all equal. We can thus study the behaviour of the representative consumer. To simplify, let us suppose that population remains constant. Time is considered a *discrete* variable, making it easier to analyse the situation as a sequence of events.[2] Consumer's preferences have a cardinal representation,[3] being described by an utility functional, $U = \sum_{t=0}^{+\infty} \left(\frac{1}{1+\theta}\right)^t u(c_t)$, constituted by a sum of discounted utilities achieved in each period, $u(c_t)$, where c_t is the consumption level in period t,[4] $\frac{1}{1+\theta}$ is the discount factor of future utility and $\theta > 0$ is the rate of time preference. We suppose that $u : \mathbf{R}^+ \to \mathbf{R}$ is a twice continuously differentiable, increasing and concave function; hence $u'(c_t) > 0$ and $u''(c_t) < 0$; for simplicity, let's assume also $\lim_{c_t \to 0^+} u'(c_t) = +\infty$.

Technology is described by the per-capita production function, $f(k_t)$, where $f(k_t)$ is the flow of per-worker gross output (gross of capital replacements) and k_t is the per-worker flow of services from the real capital stock. $f : \mathbf{R}^+ \to \mathbf{R}^+$ is a twice continuously differentiable, increasing and concave function; hence $f'(k) > 0$ and $f''(k) < 0$. Moreover, $f(0) = 0$, $\lim_{k \to 0^+} f'(k) = +\infty$ and $\lim_{k \to +\infty} f'(k) = 0^+$.

Now, let's analyse the 'centralised' solution, i.e., the consumption path that maximizes consumer utility over the whole period considered:

$$\max_{\{c_t, k_t\}} U = \sum_{t=0}^{+\infty} \left(\frac{1}{1+\theta}\right)^t u(c_t)$$

where

$$c_t = f(k_t) - \mu k_t - (k_{t+1} - k_t), \quad t = 0, 1, 2, \ldots \qquad (\mathrm{C}t)$$

and

$$k_0 = \bar{k}_0$$

where $\bar{k}_0$ is the (given) initial capital endowment per worker.

As regards the meaning of the physical constraint (c_t) observe that, without loss of generality, we can assume that each unit of capital provides a unitary flow of capital service within the period, so that the same symbol k_t denotes the per-worker flow of services from the real capital stock (when it appears as argument of function f) and the per-worker capital *stock* available at date t, that is, *at the beginning* of period t. On the other hand, $f(k_t) - \mu k_t$ and c_t are the *flows* of net product per worker and of consumption per worker *during* period t ('yearly' flows). Suppose that within each period a uniform flow of net output and of consumption is made available by the production process (a 'daily' flow); in this way the contemporaneity between production and consumption entailed in (Ct) does not raise problems from the logical point of view.[5] Moreover, suppose that the new capital accumulated during period t starts to generate capital services only with the beginning of the subsequent period ('time do build' assumption); in this way we can disregard the formation of new capital *during* the period.

Let's suppose that $\bar{k}_0 < \tilde{k}$, where $\tilde{k}$ is that level of capital labour ratio which makes the net product per worker equal to zero, i.e.

$$f(\tilde{k}) = \mu\tilde{k} \qquad [4.1]$$

The optimal path of consumption[6] c_t, with $t = 0, 1, 2, 3,...$ and of capital, k_t, with $t = 1, 2, 3, ...$ can be found as follows:[7] at the beginning of each period t the stock of capital, k_t, is determined by past saving decisions. After substituting the physical constraints (Ct) in the utility functional, there are just two addenda that include k_{t+1}:

$$U = \sum_{\tau=0}^{+\infty}\left(\frac{1}{1+\theta}\right)^{\tau} u[f(k_\tau) - \mu k_\tau + k_\tau - k_{\tau+1}]$$

$$= \cdots + \frac{u[f(k_t) - \mu k_t + k_t - k_{t+1}]}{(1+\theta)^t} + \frac{u[f(k_{t+1}) - \mu k_{t+1} + k_{t+1} - k_{t+2}]}{(1+\theta)^{t+1}} + \cdots$$

Derive U with respect to k_{t+1} and set the result equal to zero; after simplifying we have:

$$u[f(k_t) - \mu k_t + k_t - k_{t+1}] = u[f(k_{t+1}) - \mu k_{t+1} + k_{t+1} - k_{t+2}]\frac{f'(k_{t+1}) - \mu + 1}{\theta + 1}$$

$$t = 0, 1, 2, ... \qquad (\mathrm{U}t)$$

and re-substituting (Ct) and (C$t+1$) we obtain:[8]

$$u'(c_t) = \frac{u'(c_{t+1}) \cdot [1 - \mu + f'(k_{t+1})]}{1 + \theta}, \quad t = 0, 1, 2, \ldots \qquad \text{(UC}t\text{)}$$

To understand the economic meaning of conditions (UCt), consider the consumption allocation problem faced by the representative consumer between periods t and $t+1$. If the consumer decreases consumption in period t by $\mathrm{d}c_t$, the ensuing loss of utility is $u'(c_t)\mathrm{d}c_t$. This decrease in consumption of period t allows for more accumulation: gross output of period $t+1$ increases by $f'(k_{t+1})\mathrm{d}c_t$. Consumption of period $t+1$ may be increased by $[f'(k_{t+1}) - \mu + 1]\mathrm{d}c_t$, including what is left of the initial amount of capital invested. This will increase utility of period $t+1$ by $u'(c_{t+1})[f'(k_{t+1}) - \mu + 1]\mathrm{d}c_t$. Along the optimal path these (small) reallocations of consumption must leave welfare unchanged: the loss in utility of period t must compensate the discounted increase of utility of period $t+1$, that is:

$$u'(c_t)\mathrm{d}c_t = \frac{u'(c_{t+1})[f'(k_{t+1}) - \mu + 1]\mathrm{d}c_t}{1 + \theta}$$

After simplification we obtain equations (UCt). The left-hand and the right-hand members of (UCt) could be called 'marginal utility of consumption' and 'marginal utility of savings'. It is to be noted that the optimisation process entailed by condition (UCt) concerns just two periods at a time: t and $t+1$. The integral consumption of what remains of the additional capital transferred from period t to $t+1$ sterilises any effect of this capital accumulation for periods $t+2$, $t+3$, etc. This allows us to focus just on costs and benefits for periods t and $t+1$. No costs and benefits due to reallocations between periods t and $t+1$ arise for $t+2$, $t+3$,...; otherwise these effects should have been taken into account in identifying the optimal intertemporal allocation.

Physical constraints (Ct) and the 'equi-marginality' equalities (UCt) are necessary conditions for selecting the optimal behaviour of the representative consumer. The steady-state of the system (UCt)–(Ct) is represented by the pair (k^*, c^*), where k^* is that value of k such that

$$f'(k^*) = \theta + \mu \qquad [4.2]$$

and $c^* = f(k^*) - \mu k^*$. Equation [4.2] is called the 'modified golden rule', as it differs from the traditional 'golden rule',

$$f'(k_g) = \mu \qquad [4.3]$$

which is the condition to select the capital labour ratio, k_g, that maximises the net product per worker, $f(k) - \mu k$. As f' is decreasing we have

$$k^* < k_g < \tilde{k}$$

For further reference, it is useful to represent on two graphs the curve of net product and its derivative (see Figure 4.1).

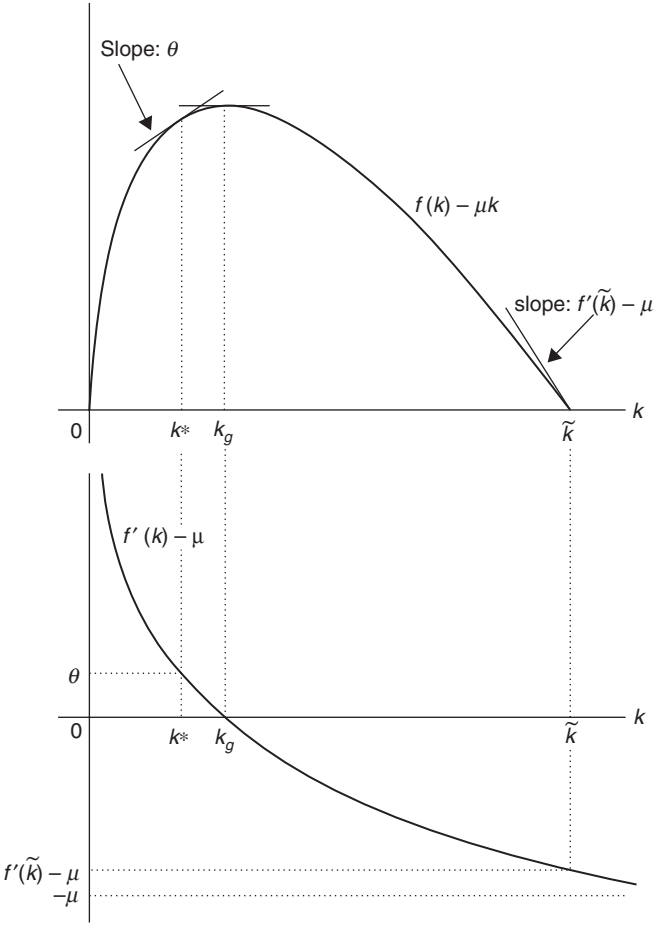


Figure 4.1 Net product, $f(k) - \mu k$, and its derivative, $f'(k) - \mu$

4.3 Excluding divergent paths: the transversality condition

The dynamics described by equations (UC*t*) and (C*t*) can be analysed by the phase portrait represented in Figure 4.2. The curves $f(k) - \mu k$ and $k = k^*$ divide the positive quadrant in four regions: 1, 2, 3 and 4. The direction of the movement in each of these regions is described by the small arrows: by (C*t*) we deduce that $k_{t+1} \gtrless k_t$ if and only if $c_t \gtrless f(k_t) - \mu k_t$; by (UC*t*) we deduce that $c_{t+1} \gtrless c_t$ if and only if $k_t \gtrless k^*$. The direction of arrows suggests that the dynamics of k_t and c_t is a saddle path, that is, an *unstable* path. The cause of this instability lies in the peculiar way in which the initial consumption level, c_0, is selected. Let us focus on this procedure, step by step. At date $t = 0$ equations (U) and (C) become

$$u'(c_0) = u'(c_1)\frac{1-\mu+f'(k_1)}{1+\theta} \qquad \text{(U0)}$$

$$c_0 = f(\bar{k}_0) - \mu\bar{k}_0 - (k_1 - \bar{k}_0) \qquad \text{(C0)}$$

$$c_1 = f(k_1) - \mu k_1 - (k_2 - k_1) \qquad \text{(C1)}$$

We have thus three equations in four unknowns: c_0, k_1, c_1 and k_2; one of them must be fixed from outside. In infinite horizon models c_0 is initially chosen arbitrarily and then one looks if this choice is compatible with utility maximisation in the long run. Suppose that $\bar{k}_0 < k^*$; hence $[1 - \mu + f'(\bar{k}_0)]/(1 + \theta) > 1$. If c_0 is initially fixed at a high level, not too far from the net product $f(\bar{k}_0) - \mu\bar{k}_0$ (like, for example, c_0' in Figure 4.2), the first member of (U0) will be quite low (since u' is decreasing); on the other hand (C0) determines k_1 not too far from $\bar{k}_0$ and thus the ratio $[1 - \mu + f'(k_1)]/(1 + \theta)$ will remain close to $[1 - \mu + f'(\bar{k}_0)]/(1 + \theta)$, and thus higher than 1. Hence, in order to align the r.h.s. of (U0) with the low level reached by $u'(c_0')$, *future* consumption must be fixed at a level higher than c_1', i.e., $c_1' > c_0'$. In other words, given a high initial consumption level, to 'justify' (rationalise) this choice future consumption must be fixed at an even higher level. Analogous adjustments, all entailing an ever increasing consumption in periods $t = 2, 3, \ldots$ will take place, up to the point where capital is totally devoured! This is what happens along the $S'\Sigma'$ path of the Cass-Koopmans diagram (see Figure 4.2): initial consumption is kept fixed at c_0' and c_t is increased (savings are decreased) in all subsequent periods; from point Z' on, besides consuming the entire net product of each period, the individual starts 'eating' away the capital stock, until it is dragged to 0.

If, on the other hand, initial consumption is fixed at a low level, like c_0'' in Figure 4.2, then $u'(c_0'')$ will be quite high and (C0) will determine k_1 at a level which is significantly higher than $\bar{k}_0$, so that the ratio $[1 - \mu + f'(k_1)]/(1 + \theta)$ will be significantly lower than $[1 - \mu + f'(\bar{k}_0)]/(1 + \theta)$. In order to keep the r.h.s. of (U0) at the high level of $u'(c_0'')$ future consumption cannot be fixed at a very high level, in spite of the large accumulation that has just taken place. This leaves a large amount of resources for accumulation, pushing the system into an over-accumulation path, like $S''\Sigma''$ in Figure 4.2, where consumption starts decreasing from point Z'' onwards where $k > k^*$, thus making the net marginal productivity of capital fall below the factor of time preference. The system is thus dragged to point $(\tilde{k}, 0)$ where the entire gross product $f(\tilde{k})$ is devoted to the maintenance of capital, $\mu\tilde{k}$: the initial error, of a too low initial consumption level, is thus corrected by low levels of future consumptions – which decrease even to zero from a certain point onwards!

It can be proved that there is only *one* level of initial consumption, c_0^*, that places the consumer on his optimal intertemporal path. All other levels of c_0 lie in an over-consumption path or in an over-accumulation path.

In both examples, initial consumption is taken as given; this is just a provisional assumption, an analytical device that shows that almost all levels of initial consumption lie on a divergent path: infinitely many divergent paths, like $S'\Sigma'$, or $S''\Sigma''$, can be obtained along which the adjustments, necessary for rationalising the 'error' of fixing an arbitrary level of c_0, are shifted to *future* changes, rather than changes in *present* consumption.

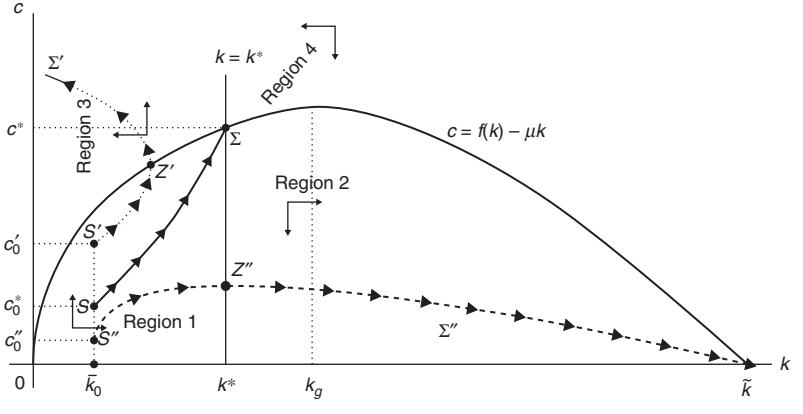


Figure 4.2 Phase portrait of the Cass-Koopmans-Ramsey model

These instability phenomena are then amended by introducing a further condition, the so-called 'transversality condition', that excludes all diverging paths like $S'\Sigma'$ or $S''\Sigma''$. In formal terms it is represented by

$$\lim_{t\to\infty} k_t u'(c_t)\left(\frac{1}{1+\theta}\right)^t = 0 \qquad \text{(T)}$$

In all divergent paths c_t would become 0, sooner or later, hence $u'(c)$ tends to infinite. This eventuality is excluded by condition (T). But, even if formally correct, this procedure seems to miss the economic substance of the problem. Why should a consumer who wants to optimise his consumption plan commit himself to keeping c_0 fixed? In his utility-maximisation problem c_0 is surely the *first* variable he will adjust. Obviously, it is not the only variable to consider; rather, he must adjust the *whole* stream of future consumptions, i.e., infinitely many consumption levels (c_0 included)! But while this problem is handy in the finite horizon case,[9] at least in principle, it seems quite difficult or even unsolvable in the infinite horizon case. The selection of the saddle path entails, from the logical point of view, the solution of infinitely many optimisation problems: for any given c_0 the whole path of pairs (k_t,c_t) satisfying conditions (UC*t*) and (C*t*) should be calculated; when we realise that it diverges from the steady-state – and this will be the case for all but one path – we have to calculate another path starting from another level of c_0. In this way we would select the unique path converging to the steady-state. But this would require almost unlimited computational power for the consumer concerning present and future consumptions and savings. In other words, it requires long-run or perfect foresight.

Moreover, this way of selecting the optimal path introduces an *instability* phenomenon which is *not* inherent to the optimisation problem we are studying. It is due both to the presumption that the consumer must face an *infinite* horizon optimisation problem and to the analytical tools available to solve such problem. The crucial difficulty consists in the fact that a rational choice of c_0 would require *at the same time* choosing the *whole future* path c_t, $t = 1, 2, 3,\ldots$. As this problem is not *directly* solvable we must resort to the *indirect* way to fix arbitrarily c_0 and check later on if the ensuing path converges or diverges. The extraordinary high computational power that is thus supposed to have the consumer makes the model extremely unrealistic. Moreover, it conveys the wrong idea of a structural instability of the long-run accumulation path, only remediable by assuming perfect foresight on the part of the consumer.

In what follows, an alternative way of selecting the optimal accumulation path is proposed, in which the set of assumed forecasting abilities on the part of the consumer is less demanding. If, realistically, we assume that in each period the consumer is able to balance the marginal effects of reallocations over a *finite* number of periods only, we obtain a path *convergent* to the steady-state that is *optimal* within the set of constraints imposed on the consumer's ability to optimise over the future. Obviously, it is necessarily sub-optimal compared with the saddle path, being the result of a set of optimisations defined over a more restrictive set of constraints but, 'better is the enemy of good', as an Italian proverb states. A set of scenarios can thus be outlined where perfect foresight is no longer necessary to exclude divergent paths: myopic optimising rules are compatible with the convergence to the steady-state.

4.4 An alternative approach: adjustments towards the optimal path

Given

$$k_0 = \bar{k}_0 < k^* \qquad [4.4]$$

one possible choice that the consumer can adopt is to consume the entire net product in each period. By (C*t*) we see that this choice is feasible and entails

$$k^\circ_{t+1} = k^\circ_t = \bar{k}_0 \text{ and } c^\circ_t = f(\bar{k}_0) - \mu\bar{k}_0 \quad \text{for any } t = 0, 1, 2,... \qquad \text{(Y0-)}$$

This is not a unique option and probably not the optimal one. It is a provisional choice[10] that can be used as a starting point to begin fixing ideas.

In this situation, consider the intertemporal re-allocation constituted by (a) and (b) below:

(a) in period $t = 0$ save and invest 1 unit of the good;
(b) the unit saved in period $t = 0$ results in

$$f'(\bar{k}_0) - \mu \qquad [4.5]$$

units of *additional net product* in *all future periods* $t = 1, 2, 3,...$.

The effects of (a) and (b) on consumer's welfare are:

(A) a loss of utility for saving 1 unit in period $t = 0$ given by[11]

$$u'[f(\bar{k}_0) - \mu\bar{k}_0]\cdot 1 \qquad [4.6]$$

(B) a utility gain ensuing from consuming the *additional net product* [4.5] in *all future periods* $t = 1, 2, 3, \ldots$. In *each period* this utility gain is

$$u'[f(\bar{k}_0) - \mu\bar{k}_0]\cdot[f'(\bar{k}_0) - \mu]. \qquad [4.7]$$

The *flow* of utility gains [4.7] arising from the additional consumption in *all future periods* $t = 1, 2, 3, \ldots$ discounted at $t = 0$ is

$$\begin{aligned}
&\frac{u'[f(\bar{k}_0)-\mu\bar{k}_0]\cdot[f'(\bar{k}_0)-\mu]}{1+\theta} + \frac{u'[f(\bar{k}_0)-\mu\bar{k}_0]\cdot[f'(\bar{k}_0)-\mu]}{(1+\theta)^2} \\
&\quad + \frac{u'[f(\bar{k}_0)-\mu\bar{k}_0]\cdot[f'(\bar{k}_0)-\mu]}{(1+\theta)^3} + \cdots \\
&\quad = u'[f(\bar{k}_0)-\mu\bar{k}_0]\cdot[f'(\bar{k}_0)-\mu]\cdot\sum_{t=1}^{+\infty}\left(\frac{1}{1+\theta}\right)^t \\
&\quad = \frac{u'[f(\bar{k}_0)-\mu\bar{k}_0]\cdot[f'(\bar{k}_0)-\mu]}{\theta} \qquad [4.8]
\end{aligned}$$

(notice that all relevant functions in [4.6], [4.7] and [4.8] are evaluated at the levels of capital services per worker and of consumption per worker planned through (Y0-) *before* re-allocation (a)–(b) takes place, that is, $\bar{k}_0$ and $f(\bar{k}_0) - \mu\bar{k}_0$).

Hence by comparing (A) with (B) we obtain

$$u'[f(\bar{k}_0) - \mu\bar{k}_0] < \frac{u'[f(\bar{k}_0) - \mu\bar{k}_0]\cdot[f'(\bar{k}_0) - \mu]}{\theta} \qquad [4.9]$$

as $[f'(\bar{k}_0) - \mu]/\theta > 1$, thanks to [4.4].

Inequality [4.9] signals that the consumer can improve his utility by saving this unit, and probably other units of the good. In order to determine how many units it is convenient to save, the consumer must solve the following problem. Let

$c_0 = f(\bar{k}_0) - \mu\bar{k}_0 - (k_1 - \bar{k}_0)$ be the reduced consumption in period $t = 0$ [4.10a]

$c_1 = f(k_1) - \mu k_1$ be the increased consumption in period $t = 1$ [4.10b]

and, therefore,

$c_t = f(k_1) - \mu k_1$, $t \geq 2$, be the increased consumption in periods $t = 2, 3, 4, \ldots$ [4.10c]

where k_1 is the solution of:

$$\begin{aligned}\max_{k_1} W_0 &= u[f(\bar{k}_0) - \mu\bar{k}_0 - (k_1 - \bar{k}_0)] + \frac{1}{1+\theta}u[f(k_1) - \mu k_1] \\ &\quad + \left(\frac{1}{1+\theta}\right)^2 u[f(k_1) - \mu k_1] + \left(\frac{1}{1+\theta}\right)^3 u[f(k_1) - \mu k_1] + \cdots \\ &= u[f(\bar{k}_0) - \mu\bar{k}_0 - (k_1 - \bar{k}_0)] + u[f(k_1) - \mu k_1] \cdot \sum_{t=1}^{+\infty}\left(\frac{1}{1+\theta}\right)^t \\ &= u[f(\bar{k}_0) - \mu\bar{k}_0 - (k_1 - \bar{k}_0)] + \frac{u[f(k_1) - \mu k_1]}{\theta} \qquad \text{(P0)}\end{aligned}$$

The first-order condition for a maximum is

$$\frac{dW_0}{dk_1} = 0: \quad \begin{aligned} &u'[f(\bar{k}_0) - \mu\bar{k}_0 - (k_1 - \bar{k}_0)] \cdot (-1) \\ &+ \frac{u'[f(k_1) - \mu k_1]}{\theta} \cdot [f'(k_1) - \mu] = 0 \end{aligned} \qquad [4.11]$$

By re-arranging [4.11] we obtain

$$u'[f(\bar{k}_0) - \mu\bar{k}_0 - (k_1 - \bar{k}_0)] = \frac{u'[f(k_1) - \mu k_1]}{\theta}[f'(k_1) - \mu] \qquad \text{(W0)}$$

(W0) is an equation in k_1. It is a particular case of equation (W*t*) (see Section 4.5 below) where parameter k_t is fixed at $k_t = \bar{k}_0$. Hence, by applying Lemma 1 below, (W0) has a unique solution, $k_1^\bullet$, such that

$$\bar{k}_0 < k_1^\bullet < k^* \qquad [4.12]$$

The second-order condition,

$$\frac{d^2W_0}{dk_1^2} < 0 \text{, i.e. } u''(\cdot) + \frac{1}{\theta}\{u''(\cdot)\cdot[f'(k_1) - \mu]^2 + u'(\cdot)\cdot f''(k_1)\} < 0$$

is satisfied at $k_1 = k_1^\bullet$ as $u' > 0$, $u'' < 0$ and $f'' < 0$.

At this point, let's suppose that the revised consumption flow

$$c_o^\bullet = f(\bar{k}_0) - \mu\bar{k}_0 - (k_1^\bullet - \bar{k}_0) < f(\bar{k}_0) - \mu\bar{k}_0$$

is actually consumed entirely during period 0.

• • •

Consider now what happens at date $t = 1$. The consumer could consume what he had planned in the previous period

$$f(k_1^\bullet) - \mu k_1^\bullet, \quad t = 1, 2, 3, \ldots \qquad \text{(Y1-)}$$

In this situation, consider the intertemporal re-allocation:

(c) in period $t = 1$ save and invest 1 unit of the good;
(d) the unit saved in period $t = 1$ results in

$$f'(k_1^\bullet) - \mu \tag{4.13}$$

units of *additional net product* in *all future periods* $t = 2, 3, 4,...$

The effects of (c) and (d) on consumer welfare are:

(C) a loss of utility for saving 1 unit in period $t = 1$ given by

$$u'[f(k_1^\bullet) - \mu k_1^\bullet]\cdot 1 \tag{4.14}$$

(D) a utility gain ensuing from consuming the *additional net product* [4.13] in *all future periods* $t = 2, 3, 4,...$ In *each period* this utility gain is

$$u'[f(k_1^\bullet) - \mu k_1^\bullet]\cdot[f'(k_1^\bullet) - \mu] \tag{4.15}$$

The *flow* of utility gains [4.15] arising from this additional consumption in *all future periods* $t = 2, 3, 4,...$ discounted at $t = 1$ is

$$\begin{aligned}
&\frac{u'[f(k_1^\bullet)-\mu k_1^\bullet]\cdot[f'(k_1^\bullet)-\mu]}{1+\theta} + \frac{u'[f(k_1^\bullet)-\mu k_1^\bullet]\cdot[f'(k_1^\bullet)-\mu]}{(1+\theta)^2} \\
&\quad + \frac{u'[f(k_1^\bullet)-\mu k_1^\bullet]\cdot[f'(k_1^\bullet)-\mu]}{(1+\theta)^3} + \cdots \\
&\qquad = u'[f(k_1^\bullet) - \mu k_1^\bullet]\cdot[f'(k_1^\bullet) - \mu]\cdot\sum_{t=1}^{+\infty}\left(\frac{1}{1+\theta}\right)^t \\
&\qquad = \frac{u'[f(k_1^\bullet) - \mu k_1^\bullet]\cdot[f'(k_1^\bullet) - \mu]}{\theta}
\end{aligned} \tag{4.16}$$

(again, all relevant functions in [4.14] , [4.15] and [4.16] are evaluated at the levels of capital services per worker and of consumption per worker planned through (Y1-) *before* re-allocation (c)–(d) takes place).

Hence by comparing (C) with (D) we obtain

$$u'[f(k_1^\bullet) - \mu k_1^\bullet] < \frac{u'[f(k_1^\bullet) - \mu k_1^\bullet]\cdot[f'(k_1^\bullet) - \mu]}{\theta} \tag{4.17}$$

as $[f'(k_1^\bullet) - \mu]/\theta > 1$, thanks to [4.12].

Inequality [4.17] signals that the consumer can improve his utility by saving this unit, and probably other units of the good. In order to

determine how many units it is convenient to save, the consumer must solve the following problem. Let

$c_1 = f(k_1^\bullet) - \mu k_1^\bullet - (k_2 - k_1^\bullet)$ be the reduced consumption in period $t = 1$ [4.18a]

$c_2 = f(k_2) - \mu k_2$ be the increased consumption in period $t = 2$ [4.18b]

and, therefore,

$c_t = f(k_2) - \mu k_2$, $t \geq 3$, be the increased consumption in periods $t = 3, 4, 5, \ldots$ [4.18c]

where k_2 is the solution of:

$$\begin{aligned}\max_{k_2} W_1 &= u[f(k_1^\bullet) - \mu k_1^\bullet - (k_2 - k_1^\bullet)] + \frac{1}{1+\theta} u[f(k_2) - \mu k_2] \\ &\quad + \left(\frac{1}{1+\theta}\right)^2 u[f(k_2) - \mu k_2] + \left(\frac{1}{1+\theta}\right)^3 u[f(k_2) - \mu k_2] + \cdots \\ &= u[f(k_1^\bullet) - \mu k_1^\bullet - (k_2 - k_1^\bullet)] + u[f(k_2) - \mu k_2] \cdot \sum_{t=1}^{+\infty} \left(\frac{1}{1+\theta}\right)^t \\ &= u[f(k_1^\bullet) - \mu k_1^\bullet - (k_2 - k_1^\bullet)] + \frac{u[f(k_2) - \mu k_2]}{\theta} \qquad \text{(P1)}\end{aligned}$$

The first-order condition for a maximum is

$$\frac{dW_1}{dk_2} = 0: \quad \begin{aligned} &u'[f(k_1^\bullet) - \mu k_1^\bullet - (k_2 - k_1^\bullet)] \cdot (-1) \\ &+ \frac{u'[f(k_2) - \mu k_2]}{\theta} \cdot [f'(k_2) - \mu] = 0 \end{aligned} \qquad [4.19]$$

Re-arranging [4.19] we obtain:

$$u'[f(k_1^\bullet) - \mu k_1^\bullet - (k_2 - k_1^\bullet)] = \frac{u'[f(k_2) - \mu k_2]}{\theta} [f'(k_2) - \mu] \qquad \text{(W1)}$$

(W1) is an equation in k_2. It is a particular case of equation (Wt) (see below, Section 4.5) where parameter k_t is fixed at $k_t = k_1^\bullet$. Hence, by applying Lemma 1 below, (W1) has a unique solution, $k_2^\bullet$, such that

$$k_1^\bullet < k_2^\bullet < k^*$$

(for the same reasons noted before, the second-order condition is satisfied). At this point, let's suppose that the revised consumption flow

$$c_1^{\bullet} = f(k_1^{\bullet}) - \mu k_1^{\bullet} - (k_2^{\bullet} - k_1^{\bullet}) < f(k_1^{\bullet}) - \mu k_1^{\bullet}$$

is actually consumed entirely during period 1.

Remark:

As soon as c_1 is revised from $f(k_1^{\bullet}) - \mu k_1^{\bullet}$ to $c_1^{\bullet}$, the optimal level of consumption planned for period $t = 0$, $c_0^{\bullet}$, is no longer optimal (in fact, the latter was determined by assuming that c_1 was settled at $f(k_1^{\bullet}) - \mu k_1^{\bullet}$, not at $c_1^{\bullet}$). However, assuming that $c_0^{\bullet}$ is entirely consumed during period $t = 0$ prevents us from any possible further re-adjustment of c_0. We will return later to this point (see Section 4.6).

• • •

Let us now consider what happens at a generic date t. The stock k_t is given; suppose

$$k_t < k^* \qquad [4.20]$$

The consumer can consume in each period

$$c_{t+\tau} = f(k_t) - \mu k_t \quad \tau = 0, 1, 2, 3,... \qquad (Yt\text{-})$$

In this situation, consider the intertemporal re-allocation:

(e) in period t save and invest 1 unit of the good;
(f) the unit saved in period t results in

$$f'(k_t) - \mu \qquad [4.21]$$

units of *additional net product* in *all future periods* $t + 1$, $t + 2$, $t + 3$... .

The effects of (e) and (f) on consumer welfare are

(E) a loss of utility for saving 1 unit in period t given by

$$u'[f(k_t) - \mu k_t] \cdot 1 \qquad [4.22]$$

(F) a utility gain ensuing from consuming the *additional net product* [4.21] in *all future periods* $t + 1$, $t + 2$, $t + 3$,... . In *each period* this utility gain is

$$u'[f(k_t) - \mu k_t] \cdot [f'(k_t) - \mu]. \qquad [4.23]$$

The *flow* of utility gains [4.23] arising from this additional consumption in *all future periods* $t + 1$, $t + 2$, $t + 3$,... discounted at t is

$$\begin{aligned} &\frac{u'[f(k_t)-\mu k_t]\cdot[f'(k_t)-\mu]}{1+\theta} + \frac{u'[f(k_t)-\mu k_t]\cdot[f'(k_t)-\mu]}{(1+\theta)^2} \\ &\quad + \frac{u'[f(k_t)-\mu k_t]\cdot[f'(k_t)-\mu]}{(1+\theta)^3} + \cdots \\ &\qquad = u'[f(k_t) - \mu k_t] \cdot [f'(k_t) - \mu] \cdot \sum_{\tau=1}^{+\infty} \left(\frac{1}{1+\theta}\right)^{\tau} \\ &\qquad = \frac{u'[f(k_t) - \mu k_t] \cdot [f'(k_t) - \mu]}{\theta} \end{aligned} \qquad [4.24]$$

(again, all relevant functions in [4.22], [4.23]and [4.24] are evaluated at the levels of capital services per worker and of consumption per worker planned through (Y*t*-) *before* re-allocation (e)–(f) takes place).

Hence by comparing (E) with (F) we obtain

$$u'[f(k_t) - \mu k_t] < \frac{u'[f(k_t) - \mu k_t] \cdot [f'(k_t) - \mu]}{\theta} \qquad [4.25]$$

as $[f'(k_t) - \mu]/\theta > 1$, thanks to [4.20].

Inequality [4.25] signals that the consumer can improve his utility by saving this unit, and probably other units of the good. In order to determine how many units it is convenient to save, the consumer must solve the following problem. Let

$c_t = f(k_t) - \mu k_t - (k_{t+1} - k_t)$ be the reduced consumption in period t [4.26a]

$c_{t+1} = f(k_{t+1}) - \mu k_{t+1}$ be the increased consumption in period $t + 1$ [4.26b]

and, therefore,

$c_{t+\tau} = f(k_{t+1}) - \mu k_{t+1}$, $\tau \geq 2$ be the increased consumption in periods $t + 2$, $t + 3$, $t + 4$,... [4.26c]

where k_{t+1} is the solution of:

$$\begin{aligned}\max_{k_{t+1}} W_t &= u[f(k_t)-\mu k_t-(k_{t+1}-k_t)]+\frac{1}{1+\theta}u[f(k_{t+1})-\mu k_{t+1}]\\ &\quad+\left(\frac{1}{1+\theta}\right)^2 u[f(k_{t+1})-\mu k_{t+1}]+\cdots\\ &= u[f(k_t)-\mu k_t-(k_{t+1}-k_t)]+u[f(k_{t+1})-\mu k_{t+1}]\cdot\sum_{\tau=1}^{+\infty}\left(\frac{1}{1+\theta}\right)^{\tau}\\ &= u[f(k_t)-\mu k_t-(k_{t+1}-k_t)]+\frac{u[f(k_{t+1})-\mu k_{t+1}]}{\theta} \qquad (\mathrm{P}t)\end{aligned}$$

The first-order condition for a maximum is

$$\frac{\mathrm{d}W_t}{\mathrm{d}k_{t+1}}=0: \quad \begin{aligned}&u'[f(k_t)-\mu k_t-(k_{t+1}-k_t)]\cdot(-1)\\ &+\frac{u'[f(k_{t+1})-\mu k_{t+1}]}{\theta}\cdot[f'(k_{t+1})-\mu]=0\end{aligned} \qquad [4.27]$$

Re-arranging [4.27] we obtain:

$$u'[f(k_t)-\mu k_t-(k_{t+1}-k_t)]=\frac{u'[f(k_{t+1})-\mu k_{t+1}]}{\theta}[f'(k_{t+1})-\mu] \qquad (\mathrm{W}t)$$

(Wt) is an equation in k_{t+1}.[12] Hence, by applying Lemma 1 below, it has a unique solution, $k^{\bullet}_{t+1}$, such that

$$k_t < k^{\bullet}_{t+1} < k^{*}$$

It is straightforward to prove that the second derivative, $\mathrm{d}^2W_t/\mathrm{d}k^2_{t+1}$, evaluated at $k^{\bullet}_{t+1}$, is negative. As before, let's suppose that the revised consumption flow

$$c^{\bullet}_t = f(k_t) - \mu k_t - (k^{\bullet}_{t+1} - k_t) < f(k_t) - \mu k_t$$

is actually consumed entirely during period t.

4.5 Convergence to the steady-state

In this section the analytical properties of equation (Wt) are studied. The results relevant from an economic point of view will be gathered in the Proposition at the end of this section. Let

$$g_{k_t}(k_{t+1}) := u'[(1-\mu)k_t + f(k_t) - k_{t+1}]$$

$$h(k_{t+1}) := \frac{u'[f(k_{t+1}) - \mu k_{t+1}]}{\theta}[f'(k_{t+1}) - \mu]$$

g is a function of k_{t+1} parameterised by k_t.

Properties of g:

Parameter k_t defines a sheaf of curves. Each of these curves is defined, continuous and strictly increasing for $k_{t+1} \in G_{k_t} = [0, (1 - \mu)k_t + f(k_t)]$ (as u' is decreasing). In the first quadrant, each of these curves has a finite and positive interception with the vertical axis, $u'[(1 - \mu)k_t + f(k_t)]$, and a vertical asymptote given by $k_{t+1} = (1 - \mu)k_t + f(k_t)$. When parameter k_t increases, the interception with the vertical axis decreases, the abscissa of the vertical asymptote increases, and curve $g_{k_t}(\cdot)$ shifts downward, that is,

$$g_{k_t}(k) > g_{k_{t+1}}(k) \quad \text{if } k_t < k_{t+1} \tag{4.28}$$

for those k where they are both defined. Hence, curves $g_{k_t}(\cdot)$ never intersect themselves; they appear as in Figure 4.3.

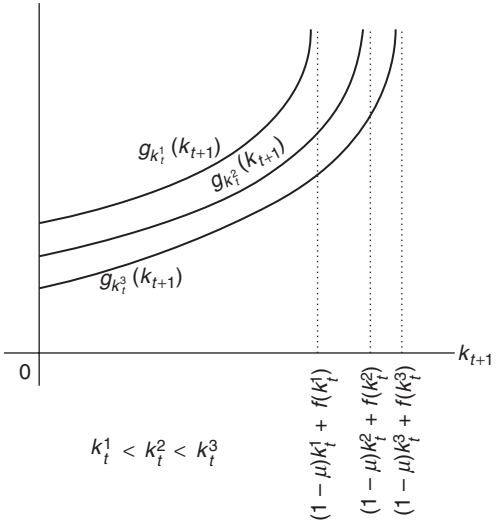


Figure 4.3 Curves $g_{k_t}(k_{t+1})$

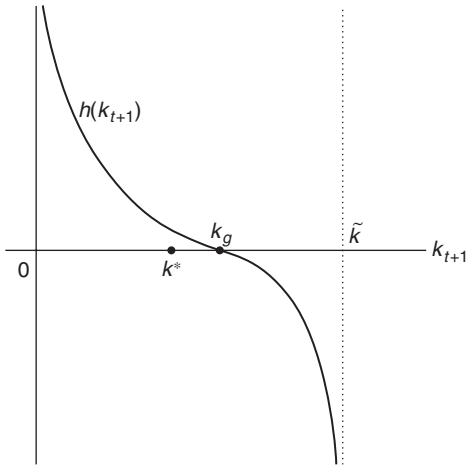


Figure 4.4 Curve $h(k_{t+1})$

Properties of *h*:

Function $h(k_{t+1})$ is defined where $f(k_{t+1}) - \mu k_{t+1} > 0$, that is, for $0 < k_{t+1} < \tilde{k}$, where $\tilde{k}$ is that level of k defined by [4.1] which makes the net product equal to zero.

Moreover,

$$\lim_{k_{t+1}\to 0^+} h(k_{t+1}) = \frac{u'[f(0^+) - \mu 0^+]}{\theta}[f'(0^+) - \mu]$$

$$= \frac{u'(0^+)}{\theta}[+\infty - \mu] = (+\infty)(+\infty) = +\infty \qquad [4.29]$$

$$\lim_{k_{t+1}\to \tilde{k}^-} h(k_{t+1}) = \frac{u'[f(\tilde{k}^-) - \mu\tilde{k}^-]}{\theta}[f'(\tilde{k}^-) - \mu]$$

$$= \frac{u'(0^+)}{\theta}[f'(\tilde{k}^-) - \mu] = (+\infty)[f'(\tilde{k}^-) - \mu] = -\infty$$

as $f'(\tilde{k}) - \mu < 0$ (see Figure 4.1). Moreover,

$$h(k_g) = \frac{u'[f(k_g) - \mu k_g]}{\theta}[f'(k_g) - \mu] = \frac{u'[f(k_g) - \mu k_g]}{\theta} \cdot 0 = 0 \qquad [4.30]$$

$$\frac{dh}{dk_{t+1}} = u''[f(k_{t+1}) - \mu k_{t+1}]\frac{[f'(k_{t+1}) - \mu]^2}{\theta} + u'[f(k_{t+1}) - \mu k_{t+1}]\frac{f''(k_{t+1})}{\theta} < 0$$

as $u'' < 0$ and $f'' < 0$ in $0 < k_{t+1} < \tilde{k}$. Curve $h(k_{t+1})$ appears as in Figure 4.4.

Lemma 1:

Given $k_t \in (0, k^*)$:

1. *there exists a unique* $k^{\bullet}_{t+1} \in (0, \hat{k})$ *which solves (Wt), where* $\hat{k} = \min\{(1 - \mu)k_t + f(k_t), k_g\}$, *that is, there exists a unique* $k^{\bullet}_{t+1}$ *which solves (Wt) on the interval where both* $g_{k_t}(k_{t+1})$ *and* $h(k_{t+1})$ *are defined and positive*;
2. $k^{\bullet}_{t+1} > k_t$;
3. $k^{\bullet}_{t+1} < k^*$.

Proof:

1. Consider equation $g_{k_t}(k_{t+1}) = h(k_{t+1})$ on the restricted domain $k_{t+1} \in [0, \hat{k}]$. For $k_{t+1} \to 0^+$ we have $g_{k_t}(0^+) = g_{k_t}(0) = u'[(1 - \mu)k_t + f(k_t)]$; hence

$$0 < g_{k_t}(0^+) < +\infty \qquad [4.31]$$

by [4.29] we have

$$h(0^+) = +\infty \qquad [4.32]$$

Hence, by [4.31] and [4.32] it follows that

$$g_{k_t}(0^+) < h(0^+) \qquad [4.33]$$

Since $k_{t+1} = (1 - \mu)k_t + f(k_t)$ is the vertical asymptote of $g_{k_t}(k_{t+1})$, we have

$$g_{k_t}\{[(1 - \mu)k_t + f(k_t)]^-\} = +\infty \qquad [4.34]$$

In order to compare g_{kt} and h at the other extreme of the domain, $\hat{k}$, three cases must be distinguished:

(a) If

$$(1 - \mu)k_t + f(k_t) < k_g \qquad [4.35]$$

then $\hat{k} = (1 - \mu)k_t + f(k_t)$ and curves $g_{k_t}(k_{t+1})$ and $h(k_{t+1})$ appear as in Figure 4.5.

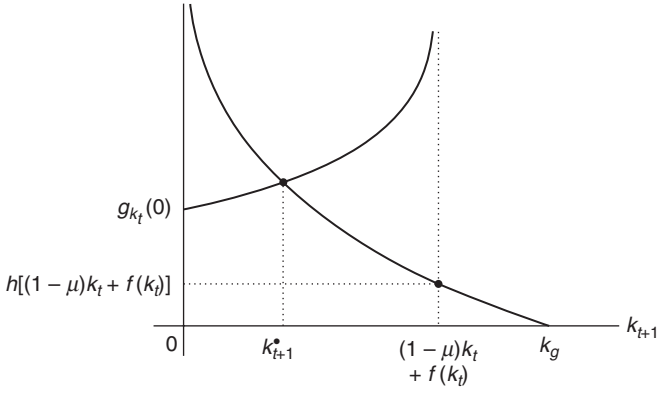


Figure 4.5 Case 1(a): $(1 - \mu)k_t + f(k_t) < k_g$

As $k_t > 0$ and from [4.35] we have that $h(k_{t+1})$ is finite and positive at $k_{t+1} = (1 - \mu)k_t + f(k_t)$, that is,

$$0 < h[(1 - \mu)k_t + f(k_t)] < \infty \qquad [4.36]$$

By [4.34] and [4.36] it follows that

$$g_{k_t}\{[(1 - \mu)k_t + f(k_t)]^{-}\} > h[(1 - \mu)k_t + f(k_t)] \qquad [4.37]$$

By continuity and thanks to [4.33] and [4.37], we conclude that there exists a unique

$$k^{\bullet}_{t+1} \in (0, (1 - \mu)k_t + f(k_t)), \quad \text{that is, } k^{\bullet}_{t+1} \in (0, \hat{k})$$

which satisfies (W*t*) (see Figure 4.5)

(b) If

$$k_g < (1 - \mu)k_t + f(k_t) \qquad [4.38]$$

then $\hat{k} = k_g$ and curves $g_{k_t}(k_{t+1})$ and $h(k_{t+1})$ appear as in Figure 4.6. By [4.38] we deduce that $g_{k_t}(k_{t+1})$ is finite and positive at $k_{t+1} = k_g$, that is

$$0 < g_{k_t}(k_g) < + \infty \qquad [4.39]$$

On the other hand, by [4.30] we know that

$$h(k_g) = 0 \qquad [4.30']$$

Hence, by [4.39] and [4.30′] it follows that

$$g_{k_t}(k_g) > h(k_g) \qquad [4.40]$$

By continuity and thanks to [4.33] and [4.40] we conclude that there exists a unique

$$k^{\bullet}_{t+1} \in (0, k_g), \quad \text{that is, } k^{\bullet}_{t+1} \in (0, \hat{k})$$

which satisfies (W*t*) (see Figure 4.6)

(c) If

$$k_g = (1 - \mu)k_t + f(k_t)$$

then $\hat{k} = k_g = (1 - \mu)k_t + f(k_t)$ and curves $g_{k_t}(k_{t+1})$ and $h(k_{t+1})$ appear as in Figure 4.7. In this case

$$g_{k_t}(\hat{k}^{-}) \equiv g_{k_t}\{[(1 - \mu)k_t + f(k_t)]^{-}\} = +\infty \qquad [4.41]$$

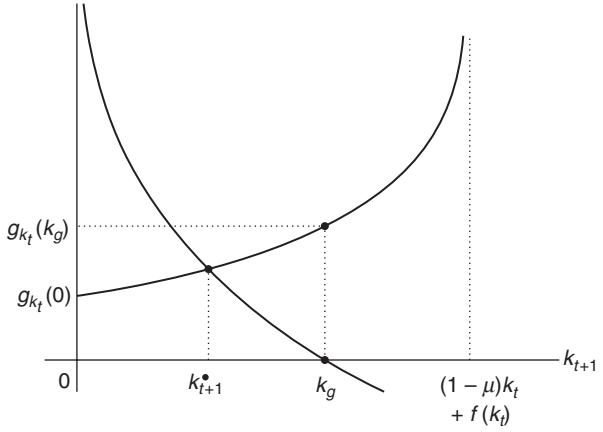


Figure 4.6 Case 1(b): $k_g < (1 - \mu)k_t + f(k_t)$

and

$$h(\hat{k}) \equiv h(k_s) = 0 \qquad [4.42]$$

Hence by [4.41] and [4.42] it follows that

$$g_{k_t}(\hat{k}) > h(\hat{k}) \qquad [4.43]$$

By continuity and thanks to [4.33] and [4.43] we conclude that there exists a unique

$$k^{\bullet}_{t+1} \in (0, \hat{k})$$

which satisfies (W*t*) (see Figure 4.7)

2. Evaluate functions $g_{k_t}(k_{t+1})$ and $h(k_{t+1})$ at $k_{t+1} = k_t$:

$$g_{k_t}(k_t) = u'[(1 - \mu)k_t + f(k_t) - k_t] = u'[f(k_t) - \mu k_t]$$

$$h(k_t) = \frac{u'[f(k_t) - \mu k_t]}{\theta}[f'(k_t) - \mu]$$

hence $g_{k_t}(k_t) < h(k_t)$ as $[f'(k_t) - \mu]/\theta > 1$ for $k_t < k^*$

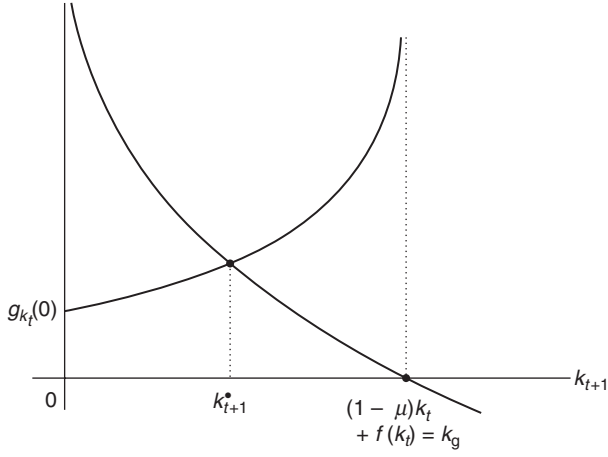


Figure 4.7 Case 1(c): $k_s = (1 - \mu)k_t + f(k_t)$

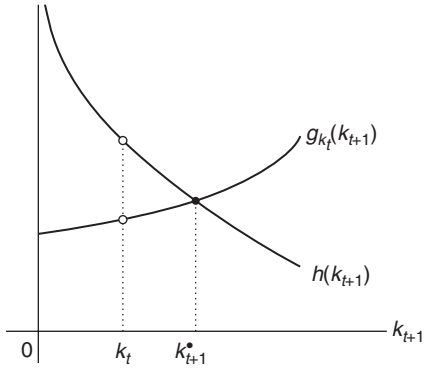


Figure 4.8 Lemma 1, item 2

Curves $g_{k_t}(k_{t+1})$ and $h(k_{t+1})$ appear as in Figure 4.8; hence the solution $k^{\bullet}_{t+1}$ of (W*t*) must thus lie *on the right* of k_t.

3. Draw curves $g_{k_t}(k_{t+1})$ and $h(k_{t+1})$ on the same graph (see Figure 4.9). Two cases must be distinguished.
   (i) If $(1 - \mu)k_t + f(k_t) \leq k^*$, by item 1 of the this Lemma 1, cases (a) or (c), we deduce that $k^{\bullet}_{t+1} < (1 - \mu)k_t + f(k_t)$; hence $k^{\bullet}_{t+1} < k^*$ (see Figure 4.9(i)).
   (ii) If $(1 - \mu)k_t + f(k_t) > k^*$, evaluate $g_{k_t}(k_{t+1})$ and $h(k_{t+1})$ at $k_{t+1} = k^*$:

$$g_{k_t}(k^*) = u'[f(k_t) - \mu k_t - (k^* - k_t)]$$

$$h(k^*) = \frac{u'[f(k^*) - \mu k^*]}{\theta}[f'(k^*) - \mu] = u'[f(k^*) - \mu k^*] \text{ due to [4.2]}$$

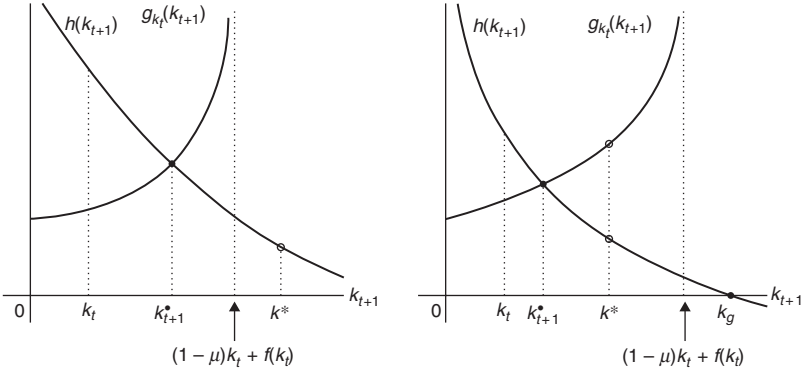
As $k_t < k^*$, then $f(k_t) - \mu k_t - (k^* - k_t) < f(k^*) - \mu k^*$; as u' is decreasing, then $g_{k_t}(k^*) > h(k^*)$. Curves $g_{k_t}(k_{t+1})$ and $h(k_{t+1})$ appear thus as in Figure 4.9 (ii); hence the solution $k^{\bullet}_{t+1}$ of (W*t*) must thus lie *on the left* of k^*.

This completes the proof. □

Now, we are going to show that if $k_0 = k^*$, equation (W*t*) defines a constant sequence: $k_t = k^*$, $t = 1, 2, 3, \ldots$

Lemma 2:

*If $k_t = k^*$, there exists a unique $k^{\bullet}_{t+1} = k^*$ which solves (W*t*).*



Case (i): $(1-\mu)k_t + f(k_t) \leq k^*$ Case (ii): $(1-\mu)k_t + f(k_t) > k^*$

Figure 4.9 Lemma 1, item 3

Proof:

Thanks to equation [4.2] it is straightforward to verify that equation $g_{k^*}(k_{t+1}) = h(k_{t+1})$ is satisfied by $k^\bullet_{t+1} = k^*$. Thus, it is enough to observe that $g_{k^*}(k_{t+1})$ is a monotonically increasing function of k_{t+1} while $h(k_{t+1})$ is a monotonically decreasing function of k_{t+1} to conclude that $k^\bullet_{t+1} = k^*$ is the unique solution of $g_{k^*}(k_{t+1}) = h(k_{t+1})$. □

Lemmas 1 and 2 entail that, given $k_0 \in (0, k^*]$, a sequence $\{k^\bullet_t\}_{t=1}^\infty$ contained in $(0, k^*]$ is univocally defined by recurrence by equation (W*t*).

Lemma 3:

$k = k^*$ is the unique steady-state of sequence $\{k^\bullet_t\}_{t=1}^\infty$.

Proof:

A *steady-state* of $\{k^\bullet_t\}_{t=1}^\infty$ is a value of k such that $k_t = k_{t+1} = k$. Substituting it into (W*t*), we obtain:

$$u'[f(k) - \mu k - (k - k)] = \frac{u'[f(k) - \mu k]}{\theta}[f'(k) - \mu]$$

which, after simplification, reduces to, $[f'(k) - \mu]/\theta = 1$, whose unique solution is $k = k^*$ (see equation [4.2]). □

Proposition:

If $k_0 = \bar{k}_0 < k^*$, the sequence $\{k^\bullet_t\}_{t=1}^\infty$ of capital/labour ratios defined by (W*t*) *converges monotonically to the steady-state k^* defined by the Ramsey modified golden rule [4.2].*

Proof:

By Lemma 1, if $\bar{k}_0 < k^*$ the sequence $\{k_t^\bullet\}_{t=1}^\infty$ is *monotonically increasing* (thanks to item 2) and *upper bounded* by k^* (thanks to item 3). Hence it must converge to some k', i.e.,

$$\lim_{t\to\infty} k_t^\bullet = k' \qquad [4.44]$$

In order to prove that $k' = k^*$ observe that, by definition, the elements $k_t^\bullet$ of the sequence satisfy equations (Wt). Consider the limit for $t \to \infty$ of (Wt):

$$\lim_{t\to\infty} u'[f(k_t^\bullet) - \mu k_t^\bullet - (k_{t+1}^\bullet - k_t^\bullet)] = \lim_{t\to\infty} \frac{u'[f(k_{t+1}^\bullet) - \mu k_{t+1}^\bullet]}{\theta}[f'(k_{t+1}^\bullet) - \mu].$$

Thanks to the continuity of functions u', f and f' we can write

$$u'\left[f\left(\lim_{t\to\infty} k_t^\bullet\right) - \mu \lim_{t\to\infty} k_t^\bullet - \left(\lim_{t\to\infty} k_{t+1}^\bullet - \lim_{t\to\infty} k_t^\bullet\right)\right]$$
$$= \frac{u'\left[f\left(\lim_{t\to\infty} k_{t+1}^\bullet\right)\right] - \mu \lim_{t\to\infty} k_{t+1}^\bullet}{\theta}\left[f'\left(\lim_{t\to\infty} k_{t+1}^\bullet\right) - \mu\right]$$

which, thanks to [4.44], can be written as

$$u'[f(k') - \mu k' - (k' - k')] = \frac{u'[f(k') - \mu k']}{\theta}[f'(k') - \mu]$$

after simplification, this equation in k' reduces to $[f'(k') - \mu]/\theta = 1$, whose unique solution is $k' = k^*$ (see equation [4.2]). This completes the proof. □

Figure 4.10 shows how $\{k_t^\bullet\}_{t=1}^\infty$ takes shape as sequence of the abscissas of the interceptions of curves $g_k(\cdot)$ with curve $h(\cdot)$. As $\bar{k}_0 < k_1^\bullet < k_2^\bullet < k_3^\bullet < \cdots < k^*$ and thanks to [4.28], curves g appear as in the diagram. Moreover, it is easy to verify that curve $g_{k^*}(k)$ crosses curve $h(k)$ at $k = k^*$: in fact, $g_{k^*}(k^*) = u'[f(k^*) - \mu k^* - (k^* - k^*)] = u'[f(k^*) - \mu k^*]$ and $h(k^*) = u'[f(k^*) - \mu k^*] \cdot [f'(k^*) - \mu]/\theta = u'[f(k^*) - \mu k^*]$.

A simplest proof of the convergence result can be given as follows if we limit to a *local* result: calculate the total differential of (Wt) with respect to k_t and k_{t+1}:

$$u''[f(k_t) - \mu k_t - k_{t+1} + k_t]\{[f'(k_t) - \mu + 1]\mathrm{d}k_t - \mathrm{d}k_{t+1}\}$$
$$= \frac{1}{\theta}\{u''[f(k_{t+1}) - \mu k_{t+1}][f'(k_{t+1}) - \mu]^2 + u'[f(k_{t+1}) - \mu k_{t+1}]f''(k_{t+1})\}\mathrm{d}k_{t+1};$$

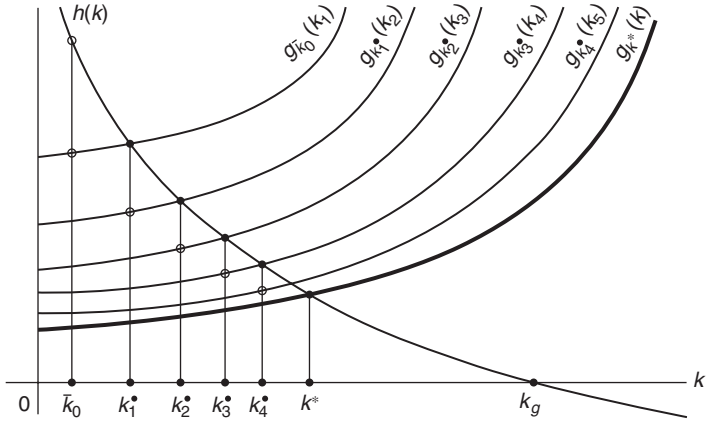


Figure 4.10 Sequence $\{k_t^\bullet\}_{t=1}^{\infty}$

evaluate at the steady-state, k^*, (where $f'(k^*) - \mu = \theta$) and re-arrange; we obtain

$$\left.\frac{dk_{t+1}}{dk_t}\right|_* = \frac{1}{1+\alpha},$$

where $\alpha = \dfrac{u'(\varphi^*)f''(k^*)}{\theta(1+\theta)u''(\varphi^*)}$ and $\varphi^* = f(k^*) - \mu k^*$. As $u' > 0$, $u'' < 0$ and $f'' < 0$ then $\alpha > 0$ and

$$0 < \left.\frac{dk_{t+1}}{dk_t}\right|_* < 1.$$

This proves the *local* stability of k^*: if $\bar{k}_0$ is taken *sufficiently close to* k^* then it converges to k^* monotonically. The Proposition on page 95 contains a global result.

A remarkable characteristic of the convergence results just seen is that they have both been obtained without assuming any transversality condition, i.e. without the need to anticipate the solution of infinitely many optimisation problems. No perfect foresight is thus needed here. For any given level of the capital/labour ratio, k_t, the consumer chooses k_{t+1} just by comparing the marginal utility of consumption in period t with the marginal utility of a *constant* flow of consumption in all future periods (from $t+1$ to infinity) discounted at t. It could be observed that the transversality condition has here been replaced with 'terminal'

conditions (Y*t*-) which, if not as unrealistic as the assumption of perfect foresight (implicit in the transversality condition), bind the consumer to a sub-optimal choice. In fact, in each period the consumer chooses his present consumption assuming zero net savings from the subsequent period onwards. This is a *myopic* behaviour: the consumer will find it convenient to revise his past decision of consuming the whole net product in each period; consequently, all conditions (Y*t*-) will be removed one after another. Hence, 'terminal' conditions (Y*t*-) will never be satisfied. Nevertheless, the transversality condition *is* satisfied by our convergent path. In fact, as k_t and c_t converge to positive and finite magnitudes, k^* and c^*, we have

$$\lim_{t\to\infty} k_t u'(c_t)\left(\frac{1}{1+\theta}\right)^t = k^* u'(c^*) \lim_{t\to\infty}\left(\frac{1}{1+\theta}\right)^t = 0.$$

But in the present setting, it is just a *consequence* of the adjustment process. It is *not a requirement* to identify the equilibrium path.[13]

4.6 'Backward' re-adjustments: towards a generalisation

As observed above in the Remark (see p. 85), as at the beginning of each period the consumer revises his present and future consumption choices taken in the previous period and provides additional savings, the consumption choices made in previous periods should be revised in order to make them coherent with the consumption-savings profile that is taking shape as time goes by. We could disregard these further re-adjustments as we assumed that by the end of each period t the consumer consumed the entire flow of consumption determined for that period, i.e., $c_t^{\bullet}$, before the beginning of period $t+1$. We could, however, imagine less simplified settings, where the consumer displays a longer foresight in anticipating his future behaviour. As already noted, in the Ramsey model the decisions to be adopted in the various periods are all connected to each other. If we want to avoid assuming that the consumer actually chooses *all* future optimal consumption levels in order to select the present one, we could imagine that he (provisionally) smoothes his consumption profile over a *finite* number of periods only. The step subsequent to that considered in the previous pages is to extend the consumer planning horizon from two periods (the present one and the entire future) to three periods, as described in the following optimisation problem:

$$\max W_0 = u(c_0) + \frac{u(c_1)}{1+\theta} + \frac{u(c_2)}{(1+\theta)^2} + \frac{u(c_3)}{(1+\theta)^3} + \cdots$$

$$\text{s.v. } c_0 = f(k_0) - \mu k_0 - (k_1 - k_0)$$

$$c_1 = f(k_1) - \mu k_1 - (k_2 - k_1)$$

$$c_\tau = f(k_2) - \mu k_2, \quad \tau \geq 2.$$

After substituting the constraints into the objective function the problem can be re-expressed in terms of variables k_1 and k_2 only:

$$\max_{k_1,k_2} W_0 = u[f(\bar{k}_0) - \mu\bar{k}_0 - (k_1 - \bar{k}_0)] + \frac{u[f(k_1) - \mu k_1 - (k_2 - k_1)]}{1+\theta}$$

$$+ \frac{u[f(k_2) - \mu k_2]}{(1+\theta)^2} + \frac{u[f(k_2) - \mu k_2]}{(1+\theta)^3} \cdots$$

$$= u[f(\bar{k}_0) - \mu\bar{k}_0 - (k_1 - \bar{k}_0)] + \frac{u[f(k_1) - \mu k_1 - (k_2 - k_1)]}{1+\theta}$$

$$+ \frac{u[f(k_2) - \mu k_2]}{\theta(1+\theta)} \qquad [4.45]$$

In the case considered in the previous sections, once given k_0, then k_1 was chosen on the assumption that the ensuing net product, $f(k_1) - \mu k_1$, was entirely consumed in each period from the subsequent period onwards.

In the present case, given k_0, variables k_1 and k_2 are chosen on the assumptions that net product $f(k_1) - \mu k_1$ is partially consumed and partially saved in period 1 and that this saving is chosen under the assumption that the entire net product $f(k_2) - \mu k_2$ will be consumed in each period from period 2 onwards.

The first-order conditions of problem [4.45], $\partial W_0/\partial k_1 = \partial W_0/\partial k_2 = 0$, entail:

$$u'[f(\bar{k}_0) - \mu\bar{k}_0 - (k_1 - \bar{k}_0)]$$

$$= u'[f(k_1) - \mu k_1 - (k_2 - k_1)] \frac{f'(k_1) - \mu + 1}{1+\theta} \qquad [4.46a]$$

$$u'[f(k_1) - \mu k_1 - (k_2 - k_1)] = u'[f(k_2) - \mu k_2] \frac{f'(k_2) - \mu}{\theta} \qquad [4.46b]$$

At $t = 2$ the consumer perceives that $c_\tau = f(k_2) - \mu k_2$, $\tau \geq 2$ is not optimal. He revises his constant consumption profile into

$$c_2 = f(k_2) - \mu k_2 - (k_3 - k_2) \qquad [4.47a]$$

$$c_3 = f(k_3) - \mu k_3 - (k_4 - k_3) \qquad [4.47b]$$

$$c_\tau = f(k_4) - \mu k_4, \quad \tau \geq 4 \qquad [4.47c]$$

where k_3 and k_4 are the solution of

$$\max W_2 = u(c_2) + \frac{u(c_3)}{1+\theta} + \frac{u(c_4)}{(1+\theta)^2} + \frac{u(c_5)}{(1+\theta)^3} + \cdots \qquad [4.48]$$

with c_2, c_3 and c_τ, $\tau \geq 4$, given by [4.47a], [4.47b] and [4.47c] respectively.

The solution of [4.48] gives rise to another pair of equations analogous to [4.46] but displaced by two periods.

In general, the dynamics of k_t is obtained by solving the following maximisation problem,

$$\max_{k_{t+1}, k_{t+2}} W_t = u[f(k_t) - \mu k_t - (k_{t+1} - k_t)] + \frac{u[f(k_{t+1}) - \mu k_{t+1} - (k_{t+2} - k_{t+1})]}{1+\theta}$$

$$+ \frac{u[f(k_{t+2}) - \mu k_{t+2}]}{(1+\theta)^2} + \frac{u[f(k_{t+2}) - \mu k_{t+2}]}{(1+\theta)^3} + \cdots$$

$$= u[f(k_t) - \mu k_t - (k_{t+1} - k_t)] + \frac{u[f(k_{t+1}) - \mu k_{t+1} - (k_{t+2} - k_{t+1})]}{1+\theta}$$

$$+ \frac{u[f(k_{t+2}) - \mu k_{t+2}]}{\theta(1+\theta)}$$

whose first-order conditions, $\partial W_t / \partial k_{t+1} = \partial W_t / \partial k_{t+2} = 0$, entail:

$$u'[f(k_t) - \mu k_t - (k_{t+1} - k_t)]$$
$$= u'[f(k_{t+1}) - \mu k_{t+1} - (k_{t+2} - k_{t+1})] \frac{f'(k_{t+1}) - \mu + 1}{1+\theta}, \qquad [4.49a]$$

$$u'[f(k_{t+1}) - \mu k_{t+1} - (k_{t+2} - k_{t+1})] = u'[f(k_{t+2}) - \mu k_{t+2}] \frac{f'(k_{t+2}) - \mu}{\theta} \qquad [4.49b]$$

Equations [4.49] implicitly describe a second-degree difference system. A detailed study of its analytical properties goes beyond the purpose of the present work.

A simple result is at hand: substituting $k_t = k_{t+1} = k_{t+2} = k$ in [4.49], both equations reduce to

$$f'(k) = \mu + \theta$$

whose unique solution is $k = k^*$. Hence k^* is the unique steady-state equilibrium of [4.49].

4.7 Concluding remarks

The present work has aimed to challenge two ideas that are normally taken for granted in neoclassical optimal accumulation theory: i) the idea that the instability phenomena of the Cass-Koopmans model (capital zeroing or capital over-accumulation) are strictly connected to a myopic foresight by the consumer in discovering his optimal accumulation path; and, in parallel, ii) that only the introduction of a transversality condition, that is, a sort of perfect foresight condition, guarantees the convergence of the system to the steady-state equilibrium.

Given the practical impossibility of solving the infinite number of optimisation problems entailed by the case of an infinite horizon problem or, similarly, placing the initial consumption exactly on the saddle path, we have proposed an adjustment process in which savings of each period are chosen optimally by assuming, provisionally, the same conventional level of consumption for all future periods. This conventional level is then revised period by period by the same bounded rational procedure. The ensuing adjustment process turns out to be monotonically convergent to the steady-state equilibrium. By accepting a 'smaller amount of rationality' for the consumer, the method of *approaching* the optimal path step by step instead of 'jumping' onto it immediately has made it possible to select a consumption level for each period without the need to immediately solve the entire set of allocation problems between any of the infinitely many pairs of future periods. The path obtained is optimal within the set of constraints assumed on the consumer's ability to optimise over the future; it is sub-optimal if viewed from a longer horizon. More refined adjustment processes, involving more than two periods in each step, should probably be less sub-optimal. Nevertheless, even the simple process presented here leads the consumer to converge to the steady-state equilibrium without the need to impose a transversality condition, i.e., without assuming perfect foresight or rational expectations.

In conclusion, diverging phenomena like those emerging in the Cass-Koopmans version of the Ramsey model appeared not to be inherent to the main problem analysed (the choice of the optimal consumption-savings path), nor directly connected to myopic optimisation. They arise from the presumption that the consumer must select all optimal future consumption levels in order to optimally select the present one.

A generalisation of the results obtained here to the case with more than one capital good still needs to be done; however, I suspect that

the roots of diverging forces operating in infinite horizon models with more than one good are of the same nature as of those found in a one-commodity model. If this is the case, they should be readily handled without the need to assume long-run perfect foresight optimising consumers. In other words, consumers' structural inability to foresee and optimise over the future should not be seen as principally responsible for crises, bubbles, etc., phenomena for which a satisfactory explanation can be sought in *other* directions, probably outside the neoclassical dynamic capital theory.

Notes

1. This work was pointed out to me by Fabio Petri.
2. For discrete time versions of the Ramsey model see, for example, Azariadis (1993, chs 7 and 13), or Stockey and Lucas (1989, ch. 2).
3. See Koopmans (1965, section I); see also Hicks (1965, ch. XXI, in particular pp. 256–7 and appendix E).
4. In this discrete time setting, 'period t' means the time interval that starts at instant t and ends at instant $t + 1$.
5. Alternatively, we could express the physical constraint in the form

$$c_{t+1} = f(k_t) - \mu k_t - (k_{t+1} - k_t), \quad t = 0, 1, 2, \ldots \qquad (C't)$$

   in this way consumption, as well as net investment, takes place *at the end* of the production period. The formulation (Ct') of the physical constraint is however uncommon in the literature; exceptions are Malinvaud (1965, pp. 311–2, in particular formulas (4) and (7)), or Mas-Colell, Whinston and Green (1995, pp. 747, item (iii)), or Nicola (2000, p. 352, formula (24.3)). Anyway, the use of constraint (Ct') does not alter results in a significant manner. We will return to this point later.
6. Throughout the work consumption is intended as consumption 'per worker'; the same can be said for net product, saving and capital. For the sake of brevity, we will omit this specification.
7. Here, I follow closely Azariadis (1993, sections 7.3 and 13.4); for the more usual continuous time case, see Blanchard and Fischer (1989, section 2.1).
8. A similar condition can be obtained for the continuous case; see, for example, Blanchard and Fischer (1989, section 2.1). By writing the physical constraint in the form (Ct'), conditions (Ut) would remain unaltered, with the only difference that now the arguments of the marginal utilities at the left-hand and at the right-hand are now c_{t+1} and c_{t+2}; consequently conditions (UCt) take the form

$$u'(c_{t+1}) = \frac{u'(c_{t+2}) \cdot [1 - \mu + f'(k_{t+1})]}{1 + \theta}, \quad t = 0, 1, 2, \ldots \qquad (UC't)$$

9. With a finite horizon, the solution can be found by fixing a terminal condition, for example, the stock of capital to be left at the end of the periods.

Thus, the relevant equations of a three-period optimisation problem, [0, 1), [1, 2) and [2, 3), are:

$$u'(c_0) = \frac{u'(c_1)[1 - \mu + f'(k_1)]}{1 + \theta}$$

$$u'(c_1) = \frac{u'(c_2)[1 - \mu + f'(k_2)]}{1 + \theta}$$

$$c_0 = f(\bar{k}_0) - \mu\bar{k}_0 - (k_1 - \bar{k}_0)$$

$$c_1 = f(k_1) - \mu k_1 - (k_2 - k_1)$$

$$c_2 = f(k_2) - \mu k_2 - (k_3 - k_2)$$

We have 5 equations in 6 unknowns: c_0, c_1, c_2, k_1, k_2, and k_3. A further equation is required to cap the degree of freedom. One possibility is to impose the total exhaustion of capital at the end of the planning period, i.e.,

$$k_3 = 0$$

The missing equation is in this case a *terminal condition*. (Alternatively, any other equation fixing the terminal capital at a given positive level could be adopted to close the system.) In the literature on the subject the transversality condition for the infinite horizon case is often introduced by passing to the limit the terminal condition of the finite horizon case; see, for example, Blanchard and Fischer (1989, p. 43).

10. In what follows, a *provisional* value assumed by a certain variable is denoted by apex ° and the *definitive* value assumed by that variable by apex •.
11. In continuity with the Cass-Koopmans-Ramsey model, the approach proposed here considers utility as a *cardinal* magnitude (see the crucial role played here by the marginal utility $u'(c)$).
12. Analogously to the case considered in note 5 above, if the physical constraint is expressed as in (C*t*′) instead of (C*t*), constraints [4.26] become

$c_{t+1} = f(k_t) - \mu k_t - (k_{t+1} - k_t)$ be the reduced consumption in period t
$c_{t+2} = f(k_{t+1}) - \mu k_{t+1}$ be the increased consumption in period $t + 1$
and
$c_{t+\tau} = f(k_{t+1}) - \mu k_{t+1}$, $\tau \geq 3$ be the increased consumption in periods $t + 3$, $t + 4$, $t + 5$, ...

but still conditions (W*t*) remain unaltered, with the only difference that now the arguments of the marginal utilities at the left-hand and at the right-hand are now c_{t+1} and c_{t+2}; consequently conditions (W*t*) expressed in terms of $c_{t+\tau}$ take the form

$$u'(c_{t+1}) = \frac{u'(c_{t+2}) \cdot [f'(k_{t+1}) - \mu]}{\theta}, \quad t = 0, 1, 2, \ldots$$

13. Curiously enough, all provisional paths resulting from any *finite* number of steps of our adjustment process – and paradoxically even the initial path from which our adjustment process starts, where c_t is provisionally set equal to for any $t = 0, 1, 2, \ldots$ – satisfy the transversality condition.

References

Azariadis, C. (1993) *Intertemporal Macroeconomics* (Cambridge, MA: Blackwell).

Blanchard, O.J. and Fischer, S. (1989) *Lectures on Macroeconomics* (Cambridge, MA: MIT Press).

Bliss, C. (2010) 'The Cambridge post-Keynesians: an outsider's insider view', *History of Political Economy*, 42(4): 631–52.

Cass, D. (1965) 'Optimum growth in an aggregative model of capital accumulation', *Review of Economic Studies*, 32(3): 233–40.

Heller, W.P. (1975) 'Tâtonnement stability of infinite horizon models with saddle-point instability', *Econometrica*, 43(1): 65–80.

Hicks, J. (1965) *Capital and Growth* (Oxford: Clarendon Press).

Koopmans, T.C. (1965) 'On the concept of optimal economic growth', *Pontificia Academia Scientiarum Scripta Varia*, 28(1): 225–87.

Malinvaud, E. (1965) 'Croissances optimales dans un modèle macroéconomique', *Pontificia Academia Scientiarum Scripta Varia*, 28(1): 301–78.

Mas-Colell, A., Whinston, M.D. and Green, J.R. (1995) *Microeconomic Theory* (Oxford: Oxford University Press).

Nicola, P. (2000) *Mainstream Mathematical Economics in the 20th Century* (Berlin: Springer).

Ramsey, F.P. (1928) 'A mathematical theory of saving', *Economic Journal*, XXXVIII (152): 543–59.

Stockey, N. and Lucas, R. (1989) *Recursive Methods in Economic Dynamics*, with the collaboration of E.C. Prescott (Cambridge, MA: Harvard University Press).

5

Malinvaud on Wicksell's Legacy to Capital Theory: Some Critical Remarks

*Saverio M. Fratini**

5.1 Introduction

Because of the neo-Walrasian shift in general equilibrium theories of value and distribution symbolically associated with the publication of Hicks's *Value and Capital*,[1] Wicksell's contribution to economic theory has been forgotten by most of the modern mainstream economists. There is, for example, no reference to him in Debreu's *Theory of Value* (1959) or Arrow and Hahn's *General Competitive Analysis* (1971).

Malinvaud's article of 2003 therefore has the unquestionable merit of attempting to focus attention on Wicksell's theory, which is one of the most important and developed versions of the marginalist explanation of income distribution and value. His analysis of Wicksell's legacy is, however, misleading on at least the three points discussed here.

The first refers to the given amount of capital in value terms that appears in Wicksell's theory. Following a certain interpretation to be found in the literature, Malinvaud presents this as connected with an alleged 'missing equation' and it is argued on this view that any variable could be regarded as given in order to fill the vacant degree of freedom. We shall instead show in Section 5.2 that Wicksell was obliged to take the value of the existing capital as a given magnitude in order to construct a marginalist theory of distribution.

*I wish to express my gratitude toward the late Pierangelo Garegnani, with whom I had the opportunity to discuss an earlier version of this chapter and from whom I learned most of what I know about capital theory. I am also grateful to E. Bellino, A. Dvoskin, P. Potestio and N. Salvadori for their comments and suggestions. The responsibility for all the remaining imperfections or errors is of course all mine.

The second point, which will be discussed in Section 5.3, concerns the 'Wicksell effect' or lack of equality between the marginal product of value capital and the rate of interest. Malinvaud instead adopts a particular notion of the marginal product of capital and demonstrates an equality between it and the rate of interest. As we shall see, however, this result has neither the same origin nor the same function as the equalities to be found when factors of production are measured in 'technical units' and not in value terms.

Finally, the third point (5.4) refers to the concept of the average period of production taken by Malinvaud from Hicks and its inverse relation to the interest rate. A very simple example will be used to show that, contrary to Malinvaud's claims, the notion of the average period of production he uses is not 'a measure of the degree of roundaboutness' of production, and its inverse relation to the rate of interest is therefore essentially irrelevant.

Even though many of the arguments used here are based on results already available in the literature, Malinvaud, surprisingly, appears to ignore them. In these cases, our original contribution will therefore be at least that of using those results in order to show the fallacy of Malinvaud's view of Wicksell's legacy. The analysis presented in Section 5.2 is largely based on Garegnani (1960 and 1990) and Kurz (2000); and in part of the Section 5.3 analysis, we generalise from arguments already developed in Pasinetti (1969) and Garegnani (1984).

5.2 The missing equation

5.2.1 Malinvaud's view of the 'missing equation' and the given quantity of capital in Wicksell's theory

Malinvaud is one of a fairly large group of economists – including Hirshleifer (1967), Sandelin (1980), Negishi (1982) and L. Samuelson (1982) – who claim that there is an equation missing in Wicksell's system of general equilibrium.[2] According to these authors, Wicksell attempted to focus attention on what Malinvaud (2003, p. 507) calls the 'production side' of capital theory – i.e., the mechanism of choice of the optimal technique and the resulting demand for labour, land and capital – and thus ended up neglecting the 'consumption side', understood as the choice between present and future consumption, which neoclassical theory sees as the basis of decisions about saving.

This neglect of the consumption side is then viewed as the cause of the under-determinacy of Wicksell's general equilibrium system. As Hirshleifer put it, Wicksell's formalisation is 'an incomplete theoretical structure' and 'to complete the Bohm-Wicksell formal system and actually determine

the rate of interest, at least one other equation is needed – specifically a time-preference relation' (Hirshleifer, 1967, pp. 191 and 197).

In other words, to be more precise, since Wicksell's equilibrium refers to a situation in which distribution variables and relative commodity prices do not vary period by period, what Hirshleifer expects to find is a condition of zero net savings (per unit of labour). Since in neoclassical theory income distribution depends (among other things) on the relative scarcity of factors of production, the distribution variables – and relative prices – can only be stationary if there is no net capital accumulation (per unit of labour). The amount of capital in use must therefore adjust so as to generate the level of the interest rate that – given the intertemporal preferences of consumers – leads to zero net savings (per unit of labour). As there is, however, no such zero net savings condition in Wicksell's analysis, an equation appears to be missing. 'The missing equation is, of course, that which follows from the behaviour of the consumption sector' (Malinvaud, 2003, p. 510).

Moreover, according to this particular interpretation of Wicksell's theory, the principal consequence of the 'missing equation' – with one degree of freedom left open – is, as Sandelin writes, 'that one central magnitude has to be determined exogenously' and therefore '[a]fter some vacillation Wicksell chooses the value of capital as an exogenous variable of his system' (Sandelin, 1980, p. 29). But since, on this view, any variable or even any proper combination of variables could be taken exogenously in order to close the system, the decision to take the value of the existing capital as a given magnitude is 'an unfortunate feature' of Wicksell's theory (Malinvaud, 2003, p. 510).

Malinvaud's suggestion, in particular, is that the degree of freedom left open by Wicksell can be filled by regarding the interest rate as exogenous (Malinvaud, 2003, p. 513). He then goes on to examine what he calls 'the production side' of capital theory – the choice of techniques and hence of the employment of capital per unit of labour – by taking the rate of interest as the independent variable. This is obviously not very satisfactory, since one of the principal goals – if not indeed the most important goal – of Wicksell's analysis, both in *Value, Capital and Rent* ([1893] 1970) and in the first book of the *Lectures* ([1901] 1967), was to explain the determinants of the (real) rate of interest.

5.2.2 The quantity of capital and the marginalist theory of value and distribution

The reconstruction of Wicksell's analysis emerging from the 'missing equation' debate is misleading, in our view, because Malinvaud and the

other authors mentioned above do not appear to attach much importance to what can instead be regarded as its fundamental characteristic, namely the fact of being a 'marginalist' theory of distribution.

In marginalist theories, as is known, the price system – which includes distribution variables, understood here as the prices of production factors – was expected to adjust to the point where excess demand is zero in every market. And the principles of decreasing marginal utility and productivity were used in order to construct excess demands for commodities and factors sensitive to changes in the price system.

Focusing attention in particular on the idea of decreasing marginal productivity, we can start by noting that, as Wicksell himself stressed (Wicksell, [1901]1967, pp. 116–17), this derived from a generalisation of the Ricardian theory of intensive rent.[3] While Ricardo made use of this principle in order to determine rents for a given wage rate, the marginalist economists tried to use it in order to determine all of the distributive variables. For such a generalisation to hold, however, capital must be regarded as capable of changing its physical form while remaining fixed in terms of quantity.

Let us consider the marginal product of labour, for example. According to a standard definition, this is the increase in output obtained from a given capital stock when an additional worker is employed. It is quite clear here that the given capital stock cannot be regarded as a vector of physical quantities of capital goods, otherwise no change would be possible in the technique used and the output obtained. Unlike labour and land, most capital goods are in fact highly specialised inputs invented and produced in order to perform a specific task in a specific way, and the switch to a more labour-intensive technique therefore entails a change in the kind of capital goods employed. If there is no change in the latter, there can be no change in the technique in use. Therefore, in order to have the marginal product of labour, capital must be conceived as a given magnitude that can take different forms.[4]

Accordingly, Jevons, Böhm-Bawerk, J.B. Clark and many other economists attempted to construct a marginalist theory of distribution by adopting a conception of capital based on the average period of production. This seemed to allow the possibility of an adjustment in the physical composition of the capital in use on the one hand and a measurement of the amount of capital independently of prices and income distribution on the other. If this idea of capital had not run up against the problems outlined below (Section 5.4), the marginalist theory would be able to determine the interest rate capable of bringing the average period of production of the technique in use into line with the amount of existing

capital, namely the wage fund available to pay the workers during the period of production.[5] And this was exactly what Wicksell, following Böhm-Bawerk, intended to do in *Value, Capital and Rent*. Even in the *Lectures*, despite his awareness of the difficulties of using the average period in general, he was interested in answering 'the question why a given amount of existing social capital gives rise to a certain rate of interest, neither higher nor lower' (Wicksell, [1901] 1967, p. 171).

The supply of capital found in Wicksell's theory, as well as many other marginalist theories, is therefore not the particular amount of capital that – given also the other data – generates the rate of interest that makes the net savings (per unit of labour) zero, as it would be if the missing equation were added. It is instead the 'existing social capital'. And since the total amount of existing capital in actual economies does not usually vary very much year by year, even without any zero net saving condition being imposed, treating it as an unchanging amount was considered a good approximation and a useful simplification.[6]

In Wicksell's theory, as a result, the zero net savings condition is not so much missing, but unwanted. Wicksell deliberately[7] omitted any such condition because he did not intend to confine his theory to the study of a hypothetical economic system in which the inducement to net capital accumulation is zero. He was instead interested in studying cases as close as possible to real life, and in real economies there is capital accumulation, even if the resulting changes in capital stock are usually very gradual.

To conclude, what is missing in Wicksell's theory, and in marginalist theories in general, is not a zero net saving condition but rather, and most importantly, a notion of capital capable of making it work consistently. In actual fact, when Wicksell became aware of the restrictions entailed by the conception of capital in terms of average period of production (cf. 5.4), he had no way to express the amount of capital other than in value terms. And this opened up two different kinds of problem.

The first, as discussed and (partially) solved by Wicksell himself, stemmed from the fact that while it is necessary, in order to have a marginal product, to measure every factor of production in its 'technical unit', value is not a technical unit of measurement of capital because, for example, '[t]he productive contribution of a piece of technical capital, such as a steam engine, is determined not by its cost but by the horsepower which develops' (Wicksell, [1901] 1967, p. 149). As is known, the solution he proposed consisted in regarding the employment of capital as the employment of saved labour and saved land (Wicksell, [1901] 1967, pp. 148 and ff.). While there is no need to pursue this point any further here, it can be observed for future reference that the dual

conception of capital – value and technical units – is the cause of the Wicksell effect considered in Section 5.3.

The second problem, which lies at the root of the search for a supposedly missing equation, concerns the fact that Wicksell included the value of the existing capital stock among the data of his theory. An amount of value thus appears in Wicksell's theory among the determinants of value and distribution and this constitutes a fatal inconsistency.[8]

It should be clear that he had no alternative, as Garegnani (1960) argued more than fifty years ago. In the first place, the principle of marginal productivity he intended to use entailed a view of capital as a single magnitude capable of taking different forms. Second, like all the marginalist economists of his time, Wicksell was interested in the determination of the rate of interest associated with the existing capital stock. Third and last, once the impossibility of using the average period of production had been established, Wicksell was left with no way other than value to express aggregate capital. It was these three points together and not any supposedly missing equation that forced Wicksell to consider capital stock in value terms among the data of his theory.

5.3 The marginal product of capital

5.3.1 Analysis of Åkerman's problem

As pointed out above, the abandonment of the average period of production left Wicksell with a dual conception of capital, aggregate capital being expressed in value terms and technical capital conceived as a series of quantities of saved labour and saved land. This duality lies at the root of what Uhr (1951) called the 'Wicksell effect', i.e., the non-equality of the rate of interest with the marginal product of (aggregate) capital.

The point is quite simple. Capital is an amount of value and its marginal product cannot therefore be something technical because value is not a 'technical unit' of measurement for capital.[9] As a result, the choice of the optimal – i.e., profit-maximising – technique does not involve the satisfaction of a condition of equality between the marginal product of value capital and the rate of interest, which is the cost of its employment. In actual fact, as Wicksell was the first to observe (Wicksell, [1901] 1967, p. 180), the two generally differ.

In this connection, Malinvaud (2003, p. 523 n.7) quotes Swan's description of the Wicksell effect as 'nothing but an inventory revaluation' (Swan, 1956, p. 355). There would thus appear to be some possibility of restoring the equality between the marginal product of capital and the interest rate by preventing such a revaluation. The idea is ascribed to

Hicks, who is credited by Malinvaud (2003, pp. 509, 517 and 519) with putting forward the right methodology for 'comparative assessments' concerning capital theory in *Value and Capital* (1946).

On the assumption of a variation in the interest rate and hence a change in the technique in use, the methodology consists in aggregating the capital employed (for a given amount of labour) with the different techniques by means of the same system of prices, namely the initial one, which is also the equilibrium one. When the variation of capital is determined in this way, the ratio of the change in the net product obtained to the change in capital employed (with fixed employment of labour) will ultimately prove equal to the rate of interest. This result has been known for a long time and its meaning has already been the object of discussion. (See in particular Swan, 1956; Bhaduri, 1966; Pasinetti, 1969; and Garegnani, 1984.)

Since Malinvaud illustrates this point with reference to the model Wicksell constructed for his analysis of Åkerman's problem,[10] we will introduce this model briefly here and return to Malinvaud's equality of the marginal product of capital and the interest rate below.

Wicksell's analysis of Åkerman's problem refers to an economy with just one consumption good produced by means of labour and capital goods (axes). In particular, with L_2 units of labour and J axes, it is possible to obtain a quantity y_2 of the consumption good, with:

$$y_2 = cL_2^{\alpha} \cdot J^{1-\alpha} \qquad [5.1]$$

Axes are produced by means of labour alone. The length of the utilisation of an axe – θ in Malinvaud's notation – depends on the amount of labour z employed in its production, with:

$$z = k\theta^{\omega} \qquad [5.2]$$

If ν and w are respectively the price of the annual services of an axe and the wage rate, profit maximisation in the consumption good sector requires:

$$w = \frac{\alpha y_2}{L_2} \qquad [5.3]$$

and

$$\nu = \frac{(1-\alpha)\cdot y_2}{J} \qquad [5.4]$$

In equilibrium, the cost of production of an axe – i.e. $w \cdot z$ – must equal the present value of the services that it will provide. Therefore:

$$w \cdot z = \nu \int_0^{\theta} e^{-\rho \cdot \tau} d\tau = \nu \frac{(1 - e^{-\rho \cdot \theta})}{\rho} \qquad [5.5]$$

Moreover, taking equation [5.5] also in consideration, we find that the lifetime θ maximising profit in the axe sector – i.e. the difference between $\nu(1 - e^{-\rho \cdot \theta})/\rho$ and $w \cdot k\theta^{\omega}$ – must satisfy the condition:

$$e^{\rho \cdot \theta} = 1 + \frac{\rho \cdot \theta}{\omega} \qquad [5.6]$$

Using the same principle adopted for the RHS of equation [5.5], we find that the price of an axe with a residual life t is $q(t) = \nu(1 - e^{-\rho \cdot t})/\rho$. Therefore, on the assumption that the axes employed in the economy are uniformly distributed for residual life (or that there are J/θ axes for every residual life t, with $0 < t \leq \theta$, which is the same thing), the value of the total employment of capital in the economy is:

$$V = \frac{J}{\theta} \int_0^{\theta} q(t) dt = \frac{J \cdot \nu}{\theta \cdot \rho} \int_0^{\theta} (1 - e^{-\rho \cdot t})\, dt = \frac{J \cdot \nu}{\theta \cdot \rho^2} (\rho \cdot \theta - 1 + e^{-\rho \cdot \theta}) \qquad [5.7]$$

Finally, since J/θ new axes must be produced in every period (or moment) in order to keep the physical capital of the economy stationary, zJ/θ units of labour must be devoted to this function. Letting L be the given supply of labour, we therefore have:

$$L = z \cdot J/\theta + L_2 \qquad [5.8]$$

We thus have eight equilibrium conditions with nine unknown variables: z, ν, w, ρ, L_2, J, y_2, V and θ. As a result, Wicksell regarded the lifetime of the axes θ as an independent variable.[11] In this way, by solving the system, he arrived at the conclusion[12] that the lifetime of the axes θ is related inversely to the rate of interest ρ, and directly to the value of the capital V and to the quantity of consumption good produced y_2. In particular, with λ_1, λ_2, λ_3 and λ_4 as four positive coefficients, the solution worked out by Wicksell ([1901] 1967, p. 289) is:

$$\rho = \lambda_1 \theta \qquad [5.9]$$

$$V = \lambda_2 \theta^{1+(1-\alpha)(1-\omega)} \qquad [5.10]$$

$$y_2 = \lambda_3 \theta^{(1-\alpha)(1-\omega)} \qquad [5.11]$$

$$\nu = \lambda_4 \theta^{-\alpha(1-\omega)} \qquad [5.12]$$

Equations [5.9]–[5.11] can easily be used to verify the Wicksell effect. In particular, we can differentiate y_2 and V with respect to θ, determine the ratio of the two results, and then – because of equation [5.9] – substitute λ_1/ρ for θ. This gives us:

$$\frac{dy_2}{dV} = \frac{\lambda_3}{\lambda_1 \lambda_2} \cdot \frac{(1-\alpha)(1-\omega)}{1+(1-\alpha)(1-\omega)} \cdot \rho \qquad [5.13]$$

Therefore, since $[\lambda_3(1-\alpha)(1-\omega)]/\{\lambda_1\lambda_2[1+(1-\alpha)(1-\omega)]\} \neq 1$ in general, the result is $dy_2/dV \neq \rho$, as Wicksell himself remarked ([1901] 1967, p. 292).

5.3.2 Malinvaud's equality between the marginal product of capital and the rate of interest

Malinvaud's approach in dealing with the model just outlined is different from Wicksell's. In the first place, he closes it by taking the interest rate ρ – rather than the lifetime θ of a new axe – as an independent variable (Malinvaud, 2003, p. 513). Second, he seeks to obtain an equality between the marginal product of capital and the rate of interest by means of Hicks's above-mentioned 'methodology'.

Malinvaud introduces his argument as follows:

> [r]easoning as Hicks ... Wicksell would have argued (i) that the volume K of capital must be defined by aggregating the numbers J/θ of the machines of various ages $a = \theta - \tau$ using proper weights, namely the respective prices $q(\theta,a)$; (ii) that for comparative assessments in the neighbourhood of a given stationary state, where by definition K = V, the physical marginal productivity of capital should not be defined with respect to infinitesimal changes in V, from one stationary state to another neighbouring stationary state, but to changes $\hat{dK}$ in K as computed with unchanged weights. [Malinvaud, 2003, p. 519]

In greater detail, Malinvaud's idea is the following. There are different kinds of capital goods – axes of different ages – in Wicksell's model. These different capital goods are aggregated into a value capital V by

means of their prices q(t) for every length of residual life t. The value capital V therefore depends: (i) on the number J/θ of axes of different ages (there is in fact the same number of axes for every age), (ii) on the lifetime θ of a new axe, and (iii) on the prices q(t). If J/θ and θ alone are allowed to change while the prices q(t) are kept at their equilibrium level, we obtain the incomplete variation $\hat{d}K$ that Malinvaud intends to use for the calculation of his marginal productivity of capital.

Let us follow Malinvaud and take equation [5.7] as our starting point. By differentiating the RHS of the equation with respect to J/θ and θ but taking ρ and ν as fixed, we obtain:

$$\hat{d}K = \frac{\nu}{\rho^2}(\rho\cdot\theta - 1 + e^{-\rho\cdot\theta})\cdot d(J/\theta) + \frac{J\cdot\nu}{\theta\cdot\rho}(1 - e^{-\rho\cdot\theta})\cdot d\theta \qquad [5.14]$$

Because of equations [5.4], [5.11] and [5.12] the variation d(J/θ) is:

$$d(J/\theta) = -\omega\frac{J}{\theta^2}d\theta \qquad [5.15]$$

Therefore, by substituting equation [5.15] into equation [5.14] and using equation [5.6], $\hat{d}K$ becomes:

$$\hat{d}K = (1-\omega)\frac{J\cdot\nu}{\rho\cdot\theta}d\theta \qquad [5.16]$$

Moreover, since from equation [5.11] we have:

$$dy_2 = (1-\omega)(1-\alpha)\frac{y_2}{\theta}d\theta \qquad [5.17]$$

we can now calculate Malinvaud's marginal productivity of capital:

$$\frac{dy_2}{\hat{d}K} = \frac{(1-\alpha)\cdot y_2}{J\cdot\nu}\rho = \rho \qquad [5.18]$$

Here we have the equality between the marginal product of capital – as 'properly defined' – and the interest rate. Malinvaud comments on this result with a certain satisfaction:

> [f]ew readers today will be surprised to find, in a discussion of capital theory under perfect competition, the assertion that the rate of interest is equal to the marginal productivity of capital when this marginal productivity is properly defined for comparative analysis.

> Some may perhaps be puzzled by the idea that the assertion follows from Wicksell's model. (Malinvaud, 2003, p. 521)

Being an already known result, as stated above, this equality is in fact not very surprising. Equally well known is the fact that it actually has very little bearing on the marginalist theory of distribution, as shown in the next section.

5.3.3 An illusory equality

Let us begin our discussion of Malinvaud's result by pointing out that the equilibrium system formed by equations [5.1]–[5.8] includes two conditions of equality between the marginal product of an input and its price, namely equations [5.3] and [5.4]. They are, as noted above, the first-order conditions of the profit-maximisation problem regarding the consumption good sector. Therefore, as usual, these conditions make it possible to determine for every level of w and ν, the corresponding demand for labour and axes in the production of the consumption good.

The equality that Malinvaud finds between his 'properly defined' marginal product of capital and the interest rate has neither the same origin nor the same function as conditions [5.3] and [5.4]. It is not in fact the first-order condition of a profit-maximisation problem and does not make it possible to determine the demand for capital in value terms associated with a certain rate of interest,[13] since this is already determined by solving the system [5.1]–[5.8] and is the amount resulting from equations [5.9] and [5.10].[14] Having established what Malinvaud's equality is not, we can now go on to see what it is.

In order to show the real meaning of Malinvaud's equality, we shall refer to an economy in which a single consumption good is produced by means of labour and n different kinds of capital goods. Assuming constant returns to scale, we denote by $y^{(i)}$ and $k^{(i)} = [k_1^{(i)}, k_2^{(i)}, \ldots, k_n^{(i)}]$ respectively the net output obtained and the vector of capital goods employed with technique (i), both per unit of labour. Given a vector $p = [p_1, p_2, \ldots, p_n]$ of prices expressed in terms of the consumption good, an interest rate ρ and a wage rate w, the (extra) profits per unit of labour entailed by the use of technique (i) are:

$$\pi^{(i)}(p, \rho, w) = y^{(i)} - \rho \sum_{\ell=1}^{n} k_\ell^{(i)} \cdot p_\ell - w \qquad [5.19]$$

Following the argument in Pasinetti (1969), if another technique (j) entails the same amount of profits per unit of labour as technique (i)

for the same system of prices and distribution variables, this means that:

$$y^{(j)} - y^{(i)} = \rho \sum_{\ell=1}^{n} (k_\ell{}^{(j)} - k_\ell{}^{(i)}) \cdot p_\ell \qquad [5.20]$$

The meaning of equation [5.20] is quite evident. If the (extra) profits are the same, and the wage rate is the same, then the difference in the net product (per unit of labour) with the two techniques must correspond to the difference in the interest paid (per unit of labour).

We can now generalise the result of equation [5.20] as follow. When the system of prices and distribution is the equilibrium one, and y^* and $k^* = [k_1^*, k_2^*, \ldots, k_n^*]$ are the net output and the vector of capital employment per unit of labour with the optimal technique, then:

$$y^* - \rho \sum_{\ell=1}^{n} k_\ell{}^* \cdot p_\ell - w = 0 \qquad [5.21]$$

and

$$(y^* + dy) - \rho \sum_{\ell=1}^{n} (k_\ell{}^* + dk_\ell) \cdot p_\ell - w = 0 \qquad [5.22]$$

for every infinitesimal variations $dy, dk_1, \ldots, dk_n$.

Therefore, as before, equations [5.21] and [5.22] imply:

$$dy = \rho \sum_{\ell=1}^{n} dk_\ell \cdot p_\ell \quad \text{or} \quad \frac{dy}{\sum_{\ell=1}^{n} dk_\ell \cdot p_\ell} = \rho \qquad [5.23]$$

and equation [5.23] is of exactly the same kind – and therefore has exactly the same meaning – as the equality that Malinvaud finds between his 'properly defined' marginal product of capital and the rate of interest for the model constructed by Wicksell in order to study Åkerman's problem.

We can conclude our discussion of the result presented by Malinvaud in his paper with a couple of remarks. First, the equality in equation [5.23] – and in equation [5.18] – derives from the fact that: (a) the prices used to aggregate the quantities of different kinds of capital goods (or axes with different residual life-periods) and the interest rate ρ are at their equilibrium levels;[15] and (b) the variations in the level of net

output and in the quantities employed of the different kinds of capital goods occur with respect to the technique that is optimal in equilibrium. Equations [5.21] and [5.22] – on which equality [5.23] is based – are in fact the consequence of conditions (a) and (b) (on this point see also Garegnani, 1984, pp. 146 and 156).

Second, far from disproving the validity of the Wicksell effect, equality [5.23] is another way to prove it. As is known, the change in value of capital (per unit of labour) in use can be broken down into a real effect and a price effect:

$$dv = \underbrace{\sum_{\ell=1}^{n} dk_\ell \cdot p_\ell}_{\text{real effect}} + \underbrace{\sum_{\ell=1}^{n} k_\ell \cdot dp_\ell}_{\text{price effect}} \qquad [5.24]$$

and since, because of equation [5.23], the real effect is $\Sigma dk_\ell \cdot p_\ell = dy/\rho$, we have:

$$dy = \frac{dy}{\rho} + \sum_{\ell=1}^{n} k_\ell . dp_\ell \qquad [5.25]$$

which implies in general $dy/dv \neq \rho$.

5.4 The average period of production

5.4.1 The average period of production: the traditional formula

We shall begin our discussion of the average period of production by elucidating the role it played – as mentioned above (Section 5.2) – in the earlier versions of marginalist theory. For this purpose, we shall consider the 'Austrian model', where the sole consumption good is obtained by the employment of labour during the T periods of time preceding the moment of output. Let u_t be the share of labour employed t periods before output is obtained, with t = 1, 2, ... T, so that $\sum_{t=1}^{T} u_t = 1$.

According to the traditional marginalist theory, the average period of production can then be defined by the following formula:

$$\bar{\theta} = \sum_{t=1}^{T} t \cdot u_t \qquad [5.26]$$

The amount of (net) output obtained per unit of labour is then assumed to be a function of this average period of production $y = f(\bar{\theta})$, with $f'(\bar{\theta}) > 0$ and $f''(\bar{\theta}) < 0$.

Moreover, if simple interest is assumed at a rate r and w is used to denote the wage rate paid at the beginning of each period in terms of the consumption good, the cost of production per unit of labour is:

$$c = \sum_{t=1}^{T} w \cdot u_t \cdot (1 + t \cdot r) = w \cdot (1 + \bar{\theta} r) \qquad [5.27]$$

For given levels of the wage rate and the interest rate, the optimal average period of production can therefore be found by solving the following first-order condition:

$$f'(\bar{\theta}) - w \cdot r = 0 \qquad [5.28]$$

At the same time, since extra-profits must vanish under the hypothesis of free competition, we have:

$$f(\bar{\theta}) - w \cdot (1 + \bar{\theta} \cdot r) = 0 \qquad [5.29]$$

Equations [5.28] and [5.29] make it possible to associate each possible interest rate r with a wage rate w and an average period of production $\bar{\theta}$. In particular, we obtain:

$$\frac{f'(\bar{\theta})}{f(\bar{\theta}) - f'(\bar{\theta}) \cdot \bar{\theta}} = r \qquad [5.30]$$

And therefore, because of $f''(\bar{\theta}) < 0$, a decrease in the interest rate involves a longer average period of production.[16] In Samuelson's words, this is 'the simple tale told by Jevons, Böhm-Bawerk, Wicksell, and other neoclassical writers', according to which, 'as the interest rate falls in consequence of abstention from present consumption in favour of future, technology must become in some sense more "roundabout", more "mechanised", and more "productive"' (Samuelson, 1966, p. 568).

As demonstrated, however, this 'simple tale' is not universally valid. To be precise, in the form presented here it is clearly based on extremely strong assumptions, such as the application of the simple interest formula and the presence of a single primary factor (labour).[17] It is

precisely because of the strong assumptions required that Wicksell, as already stated in Section 5.2, abandoned this conception of capital in his *Lectures* after initially adopting it in *Value, Capital and Rent*.

5.4.2 The inverse relationship between the Hicks-Malinvaud average period of production and the interest rate

In his discussion of Wicksell's legacy, Malinvaud takes up the idea of the average period of production introduced by Hicks in *Value and Capital* (1946) and *Capital and Time* (1973). Hicks's idea is simple: it involves using shares of cost rather than shares of labour as weights in determining the average period of production.[18]

If compound interest is assumed at a rate ρ, the cost of production per unit of labour is not as in equation [5.27] but rather:

$$c = \sum_{t=1}^{T} w \cdot u_t \cdot (1+\rho)^t \qquad [5.31]$$

where $w \cdot u_t \cdot (1 + \rho)^t$ is the part of the cost that can be ascribed to the employment of labour t periods before the output. By using the ratio of this part of the cost to the total – i.e., $u_t \cdot (1 + \rho)^t / \sum u_t \cdot (1 + \rho)^t$ – as the weight for t in the formula for the average period of production, we therefore obtain (Malinvaud, 2003, p. 516):

$$\bar{\theta} = \sum_{t=1}^{T} t \cdot \frac{u_t \cdot (1+\rho)^t}{\sum_{t=1}^{T} u_t \cdot (1+\rho)^t} = \frac{\sum_{t=1}^{T} t \cdot u_t \cdot (1+\rho)^t}{\sum_{t=1}^{T} u_t \cdot (1+\rho)^t} \qquad [5.32]$$

Now, since the weights are shares of labour in the traditional average-period formula (equation [5.26]) but shares of costs in the Hicks-Malinvaud version (equation [5.31]), while the former is completely independent of prices and distribution variables, the average period associated with a technique depends in the latter on the rate of interest. This appears to be the main concern of Hicks and Malinvaud.

In particular, there are two different effects of a change in the rate of interest on the average period of production. As the rate of interest performs two functions – a) entering into the determination of the average period associates with the techniques and b) making it possible to establish which technique is optimal and hence in use – any change in

it affects the average period of production in two ways, involving both a change of the average periods associated with the various techniques and a change of the technique in use.

Malinvaud suggests that this problem can be avoided by means of the 'methodology' used in Section 5.3 to determine the 'properly defined' marginal product of capital. In this way, he seeks to focus attention on the second effect alone, separating it from the first, and suggests that the average period associated with each technique should be kept the same in examining variations in the average period of the technique due to change in the interest rate. More precisely, with reference to equation [5.31], in Malinvaud's analysis the change in the interest rate affects the technique in use, and therefore the labour shares u_t, but is not allowed to affect the interest factors $(1+\rho)^t$, for every t = 1, 2, ..., T.[19]

Therefore, assuming a change in the interest rate, if we focus attention on the change in labour terms u_t, due to the change in the technique in use, while keeping the interest factor $(1+\rho)$ – hereafter R – constant, we obtain the change in the average period $\hat{d}\bar{\theta}$ that Malinvaud considers 'relevant for comparative analysis' (2003, p. 517):

$$\hat{d}\bar{\theta} = \frac{\sum_{t=1}^{T}(t-\bar{\theta})\cdot R^t \cdot du_t}{\sum_{t=1}^{T} u_t \cdot R^t} \qquad [5.33]$$

After a long series of mathematical operations not shown here,[20] Malinvaud arrives at the conclusion that the change in the average period $\hat{d}\bar{\theta}$ must always be opposite in sign to the change in the rate of interest. In particular, he writes (2003, p. 518):

> a decrease in the real interest rate ρ ... is associated with a lengthening of the average period of production, given what we mean by such lengthening...

and comments:

> it is interesting to know that the average period of production, *a measure of the degree of roundaboutness*, contra-varies with the interest rate. [emphasis added]

Following Hicks's 'methodology', Malinvaud thus seems to have arrived back at the simple tale of the old neoclassical writers but within

a far more general framework. His result is, however, not exactly the same as the traditional one and, as will be shown below, the Hicks-Malinvaud average period is in fact far from being 'a measure of the degree of roundaboutness' of production.

5.4.3 An example with two techniques

As shown above, the Hicks-Malinvaud average period of production associated with a given technique is generally a function of the rate of interest and can therefore change with no change in the technique in use. This fact and its possible implications are viewed by Hicks and Malinvaud as the main problem connected with the use of their idea of the average period. The problem, however, is much more serious and concerns the ranking of techniques on the basis of Hicks-Malinvaud average period, which can change with the interest rate, as we shall see.

Let us consider an example in which there are only two possible techniques, (α) and (β), and denote by $y^{(i)}$ and $u_t^{(i)}$ respectively, with $i = \alpha, \beta$, the net product per unit of labour and the share of labour required t periods before the final output is obtained with the two techniques.

For each technique, the maximum wage rate that can be paid is a function of the interest factor R:

$$w^{(i)}(R) = \frac{y^{(i)}}{\sum_{t=1}^{T} u_t^{(i)} \cdot R^t} \quad \text{with } i = \alpha, \beta \qquad [5.34]$$

By differentiating the wage rate $w^{(i)}(R)$, we obtain:

$$\frac{dw^{(i)}(R)}{dR} = -\frac{y^{(i)}}{R} \cdot \frac{\sum_{t=1}^{T} t \cdot u_t^{(i)} \cdot R^t}{\left[\sum_{t=1}^{T} u_t^{(i)} \cdot R^t\right]^2} = -\frac{w^{(i)}(R)}{R} \cdot \frac{\sum_{t=1}^{T} t \cdot u_t^{(i)} \cdot R^t}{\sum_{t=1}^{T} u_t^{(i)} \cdot R^t} \qquad [5.35]$$

and since, according to the Hicks-Malinvaud conception, the average period of production associated with technique (i), with $i = \alpha, \beta$, is:

$$\overline{\theta}^{(i)}(R) = \frac{\sum_{t=1}^{T} t \cdot u_t^{(i)} \cdot R^t}{\sum_{t=1}^{T} u_t^{(i)} \cdot R^t} \qquad [5.36]$$

equation [5.35] implies:

$$\bar{\theta}^{(i)}(R) = -\frac{dw^{(i)}}{dR} \cdot \frac{R}{w^{(i)}(R)} \qquad [5.37]$$

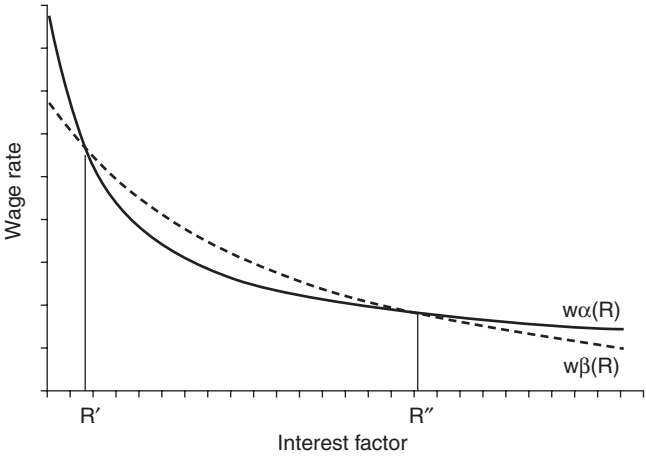
Equation [5.37] is very important in our argument. It clearly states that the average period of production associated with technique (i), with i = α, β, is equal to the elasticity of the wage rate $w^{(i)}$ with respect to the interest factor R, with the sign changed. Given a certain interest factor, the technique with the most elastic wage-interest curve is the one with the highest average period of production.

In order to show the consequences of the above result, let R′ be a switch point or, in other words, an interest factor such that $w^{(\alpha)}(R') = w^{(\beta)}(R')$. Because of equation [5.37], $\bar{\theta}^{(\alpha)}(R') > 0^{(\beta)}(R')$ if and only if $|dw^{(\alpha)}/dR| > |dw^{(\beta)}/dR|$ in R′, which means that the technique with the steepest wage-interest curve has the highest average period of production at a switchpoint.

If we assume, however, the existence of another interest factor R″, with R″ > R′, such that $w^{(\alpha)}(R'') = w^{(\beta)}(R'')$, the ranking of techniques based on the period of production calculated at R″ must be opposite to the one calculated at R′, i.e. $\bar{\theta}^{(\alpha)}(R'') < \bar{\theta}^{(\beta)}(R'')$. This result follows very simply from the observation that if the wage-interest curve $w^{(\alpha)}(R)$ is steeper than $w^{(\beta)}(R)$ at the switch point R′, then it must be less steep than $w^{(\beta)}(R)$ at the subsequent switch point, as shown in Figure 5.1. Equation [5.37] therefore implies that $\bar{\theta}^{(\alpha)} > \bar{\theta}^{(\beta)}$ at R′ and $\bar{\theta}^{(\alpha)} > \bar{\theta}^{(\beta)}$ at R″.

Moreover, when R moves in the vicinity of a switch point, the technique in use (the technique that makes it possible to pay the highest wage) for interest factors lower than the switch level is the one with the steepest wage-interest curve. The technique with the flattest wage-interest curve therefore comes into use for interest rates higher than the switch level.[21] This is Malinvaud's result, according to which an increase in the rate of interest is associated with the use of a technique with a shorter average period. And this is true at both switch points, since the technique with the highest average period at the interest factor R′ – i.e. technique α – is the one with the lowest average period at the interest factor R″. Therefore, despite the reswitching of techniques, thanks to the Hicks-Malinvaud definition, a technique with a lower average period of production is adopted at both switch points as the rate of interest increases.

This result appears, however, to have little or no significance. Contrary to what Malinvaud claims, the Hicks-Malinvaud average period does not express the 'degree of roundaboutness' or 'capital deepening' of the



Fiture 5.1 Wage-interest curves

production techniques. This is clearly proved by our simple example, where the technique with the longest average period at the first switch-point becomes the one with the shortest average period at the second.

5.5 Conclusions

To summarise, the analysis we put forward in this chapter concerns three claims that Malinvaud makes in his article of 2003: 1) 'the average period of production, a measure of the degree of roundaboutness, contra-varies with the interest rate' (p. 518); 2) 'the rate of interest is equal to the marginal productivity of capital when this marginal productivity is properly defined for comparative analysis' (p. 521); and 3) taking the value of the existing capital as a given magnitude is 'an unfortunate feature' of Wicksell's theory, since any variable could be taken exogenously in order to close the degree of freedom left open because of the 'missing equation' (Malinvaud, 2003, p. 510).

The validity of these three claims would mean complete rehabilitation of the marginalist theory of capital and hence of distribution and value. As we have shown, this is not the case.

In particular, as seen in Section 5.2, there is no missing equation in Wicksell's theory and the given amount of capital found there – like the given amounts of labour and land – is a characteristic feature of the marginalist explanation of distribution. The problem is rather the lack

of a consistent way – due to the particular structure of the theory – to express the quantity of this given capital.

With respect to Malinvaud's equality of the interest rate and the 'properly defined' marginal product of capital, as shown in Section 5.3 and contrary to the case in which the marginal product is related to factors of production measured in 'technical units', this is not a first-order condition of the (extra) profit-maximisation problem and therefore does not perform the function of determining the optimal employment of capital. It is, rather, a well-known equality found, when Wicksell price effects are ignored, for variations in the technique in use around an equilibrium position.

Finally, by means of a very simple example, it is proved in Section 5.4 that the Hicks-Malinvaud average period of production does not express the 'degree of roundaboutness' of the techniques, since the ranking of techniques in terms of it can change when the rate of interest varies and it is therefore impossible to say, independently of the rate of interest, which technique is more capital-intensive.

Notes

1. On this point, see Garegnani (1976 and 2011).
2. See Kurz (2000) for a critical survey of the views of these authors.
3. Ricardo's theory of intensive rent is grounded on the possibility of applying successive doses of labour on a fixed area of land and thus giving rise to successive but always smaller increments in the amount of produce obtained.

   Although Ricardo referred to the application of successive doses of capital on a given area of land, capital is assumed in his analysis to consist (essentially) of wages paid at the beginning of the process and the wage rate is taken as a given. Each dose of capital therefore corresponds to a dose of labour (see also Wicksell, [1901] 1967, p. 117).
4. On the conception of capital and marginal productivity, see also Trabucchi (2011).
5. See Garegnani (1960, pp. 123–34 and 147–55) for a detailed analysis of this argument.
6. As Knight wrote:

   > we assume that the fundamental conditions of economic life in the aggregate, on both supply and demand sides of the relation, remain unchanged. These fundamental conditions include (a) the total supply of productive resources ('land, labor, and capital'); (b) the 'state of the arts' or knowledge of productive methods and processes; and (c) the 'psychology', tastes and habits of the people. Significant changes in these things are generally progressive in character, in contrast to the readjustments to accidental fluctuations. (Knight, 1921, p. 311)

   Thus, as Knight stated very clearly, it is assumed that the total supply of capital, together with all the other data, remains unchanged because significant

changes in it are generally gradual or 'progressive in character', which means that they will become relevant only in the long run, while the adjustment to an equilibrium position is instead assumed to be very rapid.

It is after all common practice in every science to examine quick dynamical processes by assuming the invariance of magnitudes that vary extremely slowly with respect to the others, and the conception of equilibrium that thus emerges has been called a 'quasi-stationary state' precisely in order to stress that what is involved in these cases is an approximation.

7. Wicksell (1967, p. 171) criticised Walras for including the zero net saving condition in his equilibrium system.
8. Negishi (1982, p. 192) describes the given value of capital in Wicksell's theory as 'quite unsatisfactory' and Malinvaud (2003, p. 510) as 'an unfortunate feature'. They tend therefore to underestimate the relevance of the problem, even though it is almost clear in every paper of the 'missing equation' debate that the search for an equation to add to Wicksell's system is aimed at circumventing this difficulty. See also Potestio (1999) and Kurz and Salvadori (2001).
9. As Wicksell himself argued, the employment of capital in value terms cannot appear in a production function, since the link between it and the amount of product obtained is not technical.
10. The problem Åkerman highlighted consists essentially in verifying the validity of the fundamental relations of marginal productivity in the case with fixed or durable capital goods in which a further variable must be taken into account, namely the lifetime of capital goods employed.
11. See Garegnani (1960, p. 143 n. 51) for a critical discussion of this procedure.
12. Wicksell's aim in constructing this model was to show that an increase in the lifetime of durable capital goods does not necessarily imply – as Åkerman instead seemed to believe – a decrease in the amount of labour employed in order to keep the physical capital available unchanged. As he explicitly admitted, all the characteristics of his model were indeed selected specifically for this purpose.
13. There is a way in which the equality between the interest rate and a certain conception of the marginal product of value capital can be viewed as a first-order condition of a maximisation problem also in the case with heterogeneous capital goods. This is the case considered in Salvadori (1996).

    The procedure can be briefly summarised as follows. Given a rate of interest ρ, it is possible to determine the price vector associated with the use of a certain technique θ and then the value of the capital per worker v^θ in cases where this technique is in use. If y^θ is the net product per worker with technique θ, we have a pair (y^θ, v^θ) associated to this technique. Moreover, $w^\theta = y^\theta - \rho \cdot v^\theta$ is the wage rate that can be paid when technique θ is in use. If there are enough techniques to express y as a continuous function of v: $y = \phi_\rho(v)$, then, since the optimal technique for a given interest rate is the one making it possible to pay the highest wage rate, it is the technique that maximises the difference $\phi_\rho(v) - \rho v$. Therefore, if $\phi_\rho(v)$ is a differentiable function, $\phi_\rho'(v) = \rho$ is the first-order condition of the wage rate maximisation problem.

    It should be noted, however, that while in the present argument we have aggregated the capital goods of each technique using the prices associated with the use of that technique, in Malinvaud's case the value of capital

employed with each technique is determined by means of the same vector of prices, namely the one associated with the use of the optimal technique for the given ρ. As a result, in Malinvaud's analysis, the difference between net product per worker and interest on value capital per worker with technique θ is not, in general, the wage rate that the use of technique θ makes it possible to pay. The difference has this meaning in fact only for the optimal technique. Malinvaud's equality cannot therefore be interpreted as the first-order condition of the wage rate maximisation problem.

It can also be observed that the above equality $\phi_\rho'(v) = \rho$ does not imply a decreasing demand for capital in value terms, since the function $\phi_\rho(v)$, and then $\phi_\rho'(v)$, changes when ρ varies. (Readers are referred to Salvadori, 1996, and Kurz and Salvadori, 2010 for further details.)

14. In particular, using equations [5.9] and [5.10], we have: $V = \lambda_2 \left(\frac{\rho}{\lambda_1}\right)^{1+(1-\alpha)(1-\omega)}$.
15. Malinvaud takes ρ as an exogenous variable, even though this level of the interest rate is assumed to be the one that brings the production (demand) side of capital theory into equilibrium with its consumption (supply) side (cf. Malinvaud, 2003, pp. 510 and 513).
16. If we assume $g(\bar{\theta}) = \frac{f'(\bar{\theta})}{f(\bar{\theta}) - f'(\bar{\theta}) \cdot \bar{\theta}}$, it follows that $g'(\bar{\theta}) = \frac{f'(\bar{\theta}) \cdot f''(\bar{\theta})}{[f(\bar{\theta}) - f'(\bar{\theta}) \cdot \bar{\theta}]^2}$ and $f''(\bar{\theta}) < 0$ implies $g'(\bar{\theta}) < 0$. From equation [5.30] – i.e. $g(\bar{\theta}) = r$ – the period average $\bar{\theta}$ and the interest rate r must therefore vary in opposite directions.
17. For a discussion of the traditional average period of production, see also Garegnani (1960, pp. 123–36) and Petri (2004, pp. 99–117).
18. See Fratini (2012) for an in-depth discussion of the Hicks-Malivaud average period of production.
19. In the words of Hicks, as quoted by Malinvaud:

> if the average period changes, without the rate of interest having changed, it must indicate a change in the stream [of inputs]; but if it changes, when the rate of interest changes, this need not indicate any change in the stream at all. Consequently, even when we are considering the effect of changes in the rate of interest on the production plan, we must not allow the rate of interest which we use in the calculation of the average period to be changed. [Hicks, 1946, p. 220]

20. See Malinvaud (2003, pp. 517–18) or Fratini (2012, appendix) for the mathematical steps leading to Malinvaud's conclusion.
21. Hicks writes as follows in *Capital and Growth* (1965):

> when there is a rise in the rate of real wages (or a fall in the rate of profit) there will always be a tendency to shift to a technique with a wages curve which (in the way we have drawn our diagrams) is, at that level of wages, a curve with a slope that is less. That is to say, the new wage curve must be one on which, at that level, profits are less affected by a given rise in wages. In that sense, and in that sense only, the new technique must be

> one with a lower labour-intensity. And since the whole thing can be put the other way, it is also a technique in which wages are more affected by a given rise in profits. In that sense, and only in that sense, we can safely say that the new technique is one of greater capital intensity. [Hicks, 1965, pp. 166, 7]

The point is also considered in *Capital and Time* (1973, p. 45).

References

Arrow, K.J. and Hahn, F.H. (1971) *General Competitive Analysis* (San Francisco: Holden-Day).

Bhaduri, A. (1966) 'The concept of the marginal productivity of capital and the Wicksell effect', *Oxford Economic Papers*, XVIII: 284–8.

Burmeister, E. (1980) *Capital Theory and Dynamics* (Cambridge: Cambridge University Press).

Burmeister, E. (2006) 'A Retrospective View of *Hicks's Capital and Time: A Neo-Austrian Theory*', in H. Hagemann and R. Scazzieri (eds), *Capital, Time and Transitional Dynamics* (London: Routledge).

Debreu, G. (1959) *Theory of Value* (New York: Wiley).

Fellner, W. and Ellis, H.S. (1940) 'Hicks and the time-period controversy', *Journal of Political Economy*, XLVIII(4): 563–78.

Fratini, S.M. (2010) 'Reswitching and decreasing demand for capital', *Metroeconomica*, LXI(4): 676–82.

Fratini, S.M. (2012) 'The Hicks-Malinvaud average period of production and "marginal productivity": a critical assessment', *European Journal of the History of Economic Thought*, forthcoming.

Garegnani, P. (1960) *Il Capitale nelle Teorie della Distribuzione* (Milan: Giuffré).

Garegnani, P. (1970) 'Heterogeneous capital, the production function and the theory of distribution', *Review of Economic Studies*, XXXVII: 407–36.

Garegnani, P. (1976) 'On a Change in the Notion of Equilibrium in Recent Work on Value and Distribution', in M. Brown, K. Sato and P. Zaremka (eds), *Essays in Modern Capital Theory* (Amsterdam: North-Holland).

Garegnani, P. (1984) 'On some illusory instances of "marginal products"', *Metroeconomica*, XXXVI(2–3): 143–60.

Garegnani, P. (1990) 'Quantity of Capital', in J. Eatwell, M. Milgate and P. Newman (eds) *Capital Theory* (London: Macmillan).

Garegnani, P. (2011) 'Capital in the neoclassical theory: some notes', *Nómadas. Revista Crítica de Ciencias Sociales y Jurídicas*, MA/2011.1.

Hicks, J.R. (1946) *Value and Capital*, 2nd edn [1st ed. 1939] (Oxford: Clarendon Press).

Hicks, J.R. (1965) *Capital and Growth* (Oxford: Clarendon Press).

Hicks, J.R. (1973) *Capital and Time: A Neo-Austrian Theory* (Oxford: Clarendon Press).

Hirshleifer, J. (1967) 'A note on Böhm-Bawerk/Wicksell theory of interest', *Review of Economic Studies*, XXXIV(2): 191–9.

Knight, F. (1921) 'Cost of production and price over long and short periods', *Journal of Political Economy*, 29(4): 304–35.

Kurz, H.D. (2000) 'Wicksell and the problem of the "missing" equation', *History of Political Economy*, XXXII(4): 765–88.
Kurz, H.D. and Salvadori, N. (1995) *Theory of Production* (Cambridge: Cambridge University Press).
Kurz, H.D. and Salvadori, N. (2001) 'The aggregate neoclassical theory of distribution and the concept of a given value of capital: a reply', *Structural Change and Economic Dynamics*, XII: 479–85.
Kurz, H.D. and Salvadori, N. (2010) 'Spurious "Margins" versus the Genuine Article', in A. Birolo, D.K. Foley, H.D. Kurz, B. Schefold, and I. Steedman (eds) *Production, Distribution and Trade: Alternative Perspective. Essays in honour of Sergio Parrinello* (London: Routledge).
Malinvaud, E. (2003) 'The legacy of Knut Wicksell to capital theory', *Scandinavian Journal of Economics*, CV(4): 507–25.
Negishi, T. (1982) 'Wicksell's missing equation and Bohm-Bawerk's three causes of interest in a stationary state', *Zeitschrift für Nationalokonomie*, XLII: 161–74.
Pasinetti, L.L. (1969) 'Switches of technique and the 'rate of return' in capital theory', *Economic Journal*, LXXIX: 508–31.
Petri, F. (2004) *General Equilibrium, Capital and Macroeconomics* (Cheltenham: Edward Elgar Publishing).
Potestio, P. (1999) 'The aggregate neoclassical theory of distribution and the concept of a given value of capital: toward a more general critique', *Structural Change and Economic Dynamics*, X: 381–94.
Salvadori, N. (1996) '"Productivity Curves" in the Accumulation of Capital', in M.C. Marcuzzo, L.L. Pasinetti and A. Roncaglia (eds), *The Economics of Joan Robinson* (London and New York: Routledge). Repr. in H. Kurz and N. Salvadori (1998) *Understanding 'Classical' Economics. Studies in Long-Period Theory* (London and New York: Routledge).
Samuelson, L. (1982) 'On Wicksell's missing equation', *History of Political Economy*, XIV: 301–7.
Samuelson, P.A. (1966) 'A summing up', *Quarterly Journal of Economics*, LXXX(4): 568–83.
Sandelin, B. (1980) 'Wicksell's missing equation, the production function, and the Wicksell effect', *History of Political Economy*, XII: 29–40.
Sargan, J.D. (1955) 'The period of production', *Econometrica*, XXIII(2): 151–65.
Swan, T.W. (1956) 'Economic growth and capital accumulation', *Economic Record*, XXXII: 334–61.
Trabucchi, P. (2011) 'Capital as a single magnitude and the orthodox theory of distribution in some writings of the early 1930s', *Review of Political Economy*, XXIII(2): 169–88.
Uhr, C.G. (1951) 'Knut Wicksell – a centennial evaluation', *American Economic Review*, XLI(5): 861–78.
Wicksell, K. (1967 [1901]) *Lectures on Political Economy, Vol. I* (New York: Augustus M. Kelley Publishers).
Wicksell, K. (1970 [1893]) *Value, Capital and Rent* (New York: Augustus M. Kelley Publishers).

6

Capital and Stationary States: Considerations on the Reasons Adduced for Abandoning the Method of Normal Positions

*Paolo Trabucchi**

6.1 Introduction

As is well known, the method on which the dominant (marginalist or 'neoclassical') economic theory is based *at present* is constructed upon the notions of 'temporary' or, more frequently, 'intertemporal' equilibrium. In several respects these two notions differ widely from one another; they have, however, this in common: they both refer to positions of the economic system that do not imply a uniform rate of net return on the supply price of the capital goods existing in that system. In this crucial respect, therefore, the method employed today is, whatever its particular formulation, in stark contrast to the method used both by the classical economists and by the founders and the first systematisers of the marginalist theory who, in order to study value and distribution, always referred to positions of the economic system in which that uniformity – and more generally the uniformity of the rate of profit – was assured.

If we look at the writings in which, between the end of the 1920s and the early 1940s, Lindahl, Hicks and Hayek presented the new method for the first time, we shall see that the need to abandon the traditional method is there argued on the basis of *two* critical theses. First, the method traditionally employed by economists is interpreted as if it could be *consistently* applied only to the study of economies in a 'stationary state'. From this derives the charge of lack of realism which is levelled

*I wish to thank Antonella Palumbo and an anonymous referee for their stimulating comments on, respectively, an early draft of the first part and a more recent draft of the entire chapter.

at the traditional method. Second, it is claimed that it was only the adoption of this method that gave rise to those problems in the treatment of capital with which marginalist economists always had to some extent to struggle. The suggestion, in other words, is that these latter *theoretical* problems are of a derived, and hence *secondary*, nature as compared to the more general problems that beset the traditional *method*.

It is the purpose of the present essay to return to the two critical theses just mentioned in order to lay the ground for their systematic discussion. And though even this will cover only part of the work that needs to be done on the method now generally in use, the discussion will be of more than historical interest – as may be shown for instance by the way no less an authority than Arrow has in more recent years described the impact that, in his opinion, was made on the evolution of economic thought by the first influential exposition of the new method: 'At one stroke,' writes Arrow referring to the publication of Hicks's *Value and Capital* in 1939, 'all the *conceptual mysteries* of *capital theory* and the *confusions* about the *steady states* were wiped out' (Arrow, 1991, p. 42; italics changed from the original; for the inclusion of 'steady growth' among the cases to which the traditional method would be limited, see note 12 below).

The chapter is organised as follows. Section 6.2 will deal briefly with the nature of the traditional method. The first reason adduced for abandoning this method – its alleged limitation to the study of stationary economies – will be considered in Section 6.3; while the second reason – the idea according to which the difficulties encountered by the marginalist theory in the treatment of capital are, so to speak, a mere reflection of the adoption of the traditional method – will be the subject of Section 6.4. Section 6.5 will provide a brief conclusion.[1]

6.2 The method of normal positions

As mentioned just now, before we come to the real subject of this work, it will be convenient to take a quick look at the nature of the traditional method. This will prove useful when, in the next section, we will discuss the objections levelled against that method. We have, moreover, to clear the ground from a possible source of misunderstanding. The traditional method differs so radically from the method now generally in use that, looking at the former from the standpoint of the latter, one might almost doubt that by this method the classical and the first marginalist economists *intended* to study changing (or evolutionary) economic systems. To be sure, this charge is seldom, if ever, openly levelled at the traditional method by the founders of the new one (these authors

generally maintained that the traditional method would be logically inapplicable to non-stationary economies *despite* the intentions of the economists who actually used it). However, the charge may be thought to be implicit in some at least of the more polemical statements of these authors: as for example when, without any further explanation, Hicks wrote in 1936 of 'the static or stationary theories ... of Ricardo, Böhm-Bawerk [and] Pareto' (Hicks, 1936, p. 238).[2]

If we consider the *nature* of the prices dealt with within the traditional method (their different determination by the classical and the marginalist economists being a difference in the theory rather than in the method adopted), what we find at its basis is first of all the distinction, established for the first time in general terms by Adam Smith ([1776] 1979, pp. 72–73), between the 'market price' of a commodity and its 'natural price' (or its 'normal' price, as it was later called by Marshall, and as we shall generally do in this work). The market price, writes Smith, is 'the actual price at which [the] commodity is commonly sold'; it is therefore the price that can effectively be observed at every moment on the market, but that, as we shall see, theory is incapable of determining.[3] The normal price, on the other hand, is the price that is determined by theory as the price just sufficient to cover, at their respective normal or 'ordinary' rates, the wages and profits that must be paid to produce the commodity.

In general, that is to say at any particular moment, the market price and the normal price of a commodity will not be equal. As Smith remarks (*ibid*., p. 74), this will be the case only when 'the quantity [of the commodity] which is brought to market' exactly suffices to satisfy its 'effectual demand', i.e., only when the quantity brought to market is equal to the quantity which is demanded by 'those who are willing to pay the natural price of the commodity' (*ibid*., p. 73). When, instead, the quantity brought to market falls short of the effectual demand, the market price is driven above its normal level by competition among the purchasers; whereas when the quantity brought to market exceeds the effectual demand, it is the competition among the sellers that comes into operation, driving the market price below that level. Now, differences between the quantity brought to market and the effectual demand are continually bound to occur. In the first place, because they may be produced by a *disproportion* in the distribution of the means of production across the various sectors of the economy; and this is in turn continually bound to occur on account of the, necessarily to a large extent unforeseen, changes in the composition of demand and in the conditions of production that are normally associated with capital

accumulation. Moreover, both magnitudes are under the constant influence of *accidental forces*, such as delays, external influences on production, transitory changes in the need for a particular commodity, etc.

Together with the distinction between the market and the normal price of a commodity, we find however at the basis of the traditional method the possibility to conceive the latter, again in Smith's expression, as the 'centre' around which the market price 'gravitates' (*ibid.*, p. 75), i.e., towards which the market price is constantly tending. And this possibility is ensured precisely by the fact that, as noted before, within the traditional method the prices determined by theory refer to a position of the economic system in which a uniform rate of profit prevails. The reasons for this are well known. When the difference between the quantity brought to market and the effectual demand for a commodity – and hence between its market and its normal price – is due to a disproportion in the distribution of the means of production, this will produce systematic differences in the rates of profit realised in each sector. This will in turn entail capital transfers towards those sectors where the market price is above its normal level (for here an above-average rate of profit is realised) and away from those sectors where the opposite is true; and this will cause the quantity of the commodity brought to market in the former to increase, and in the latter to diminish, until the market price returns to its normal level. When, on the other hand, the difference is occasioned only by the influence of some accidental force, no movement of the means of production will be needed to return the market prices to their normal level (though, of course, temporary transfers of capital *may* occur if the action of these forces lasts long enough), since by their very nature these influences will tend to cease with the mere passage of time and, to a great extent, to offset one another. It is clear, however, that this second case can, in turn, be singled out only against the background of a situation where the distribution of the means of production has already adjusted to the demand structure and where, as a consequence of this, no systematic difference between the profit rates realised in different sectors is present.[4]

At this point it is important to stress that, according to Smith, the 'proportion' (*ibid.*, p. 73) between the quantity brought to market and the effectual demand of a certain commodity *cannot determine* the price at which the commodity will actually be sold. While, as we have seen, that proportion will tell us the *direction* of the divergence between the market price and the normal price, it cannot tell us in general the *extent* of the divergence. This will depend on a whole set of circumstances that may

'animate more or less the eagerness of the competition' (*ibid.*, p. 74). Now, in the first place some of these circumstances are extremely specific to each individual market. Thus, as Smith remarks, a given excess of the quantity brought to market over the effectual demand will exert different effects on the extent of the fall of the market price below its normal level according to whether it is more or less important for the sellers 'to get immediately rid of the commodity' (*ibid.*); and, other things being equal, this urgency to sell will be the stronger the more quickly perishable is the commodity in question. On the other hand, the pressure on the sellers to sell, and on the purchasers to buy, will depend in a crucial way on the *expectations* they form *during* the exchanges regarding the probable future trend in the market price. Thus even a small excess of the quantity of a commodity brought to market over its effectual demand may give rise to a considerable decrease in its market price, if both sellers and buyers become convinced that the market price is destined to decrease further in the future.

According to the traditional method economic theory cannot therefore determine the effective position occupied by the economic system at a given moment (for, in order to do so, the determination of market prices would clearly be necessary); nor, *a fortiori*, can it determine the actual path that the system will follow over a certain period of time. It is clear therefore that, according to this method, it would be impossible to formulate a *dynamic* economic *theory* in the sense that the expression is used in physics – a theory, that is to say, capable of deducing from the specific position of the system under study at a given initial point of time the position of that system at any future point. From this it would, however, be a mistake to conclude that, according to the traditional method, economic theory cannot set itself the task of studying moving economic systems; that is to say, that it cannot set itself the task of studying what we may call dynamic *problems*. Rather, we must recognise that, faced with the difficulties in grasping this movement posed by the extreme mutability and, as we saw, the elusiveness of the prices concretely observed day by day, the response of the economists was based on what, to use another expression of Marshall's ([1890] 1961, p. 366), we may call a *breaking up* of the problem. Rather than trying to replicate the constantly moving prices with theory, the method traditionally followed consisted firstly in determining a 'normal' position of the economic system – which, as we saw above, was identified with a position in which a uniform rate of profit prevails. Such a position could then be used for the *separate* solution of two quite *different* dynamic problems. On the one hand, the normal position of the system could be used to study its short-run movements

in terms of its deviations from that position and of the forces that would come into action to return the system to it, i.e., to study the *gravitation* of market prices around their normal levels. On the other hand, the *evolution* of the system could be studied in terms of the variation that its normal position would *tend* to undergo in time.

6.3 The alleged limitation of the traditional method to stationary economies

If now we turn to the first of the two theses with which the need to abandon the traditional method has been argued, it will be convenient to refer especially to some writings of Hicks ([1933] 1982; 1934; 1939). For it is Hicks who dwells in greatest detail on the reasons that would confine the applicability of the traditional method to the stationary state. As we shall see, *two* reasons can in fact be distinguished in Hicks's writings in support of this thesis.[5]

A first reason can be found in Hicks's writings from 1933 and 1934. In Hicks's opinion the limitation of the traditional method to the study of stationary economies would arise from the problems concerning what was usually called the 'determinateness' of equilibrium (for the use of the term in this context, see Kaldor, 1934, pp. 124–5). More specifically, this first reason presented itself to Hicks when, having discarded as *unrealistic* the assumption that there are no exchanges at prices that are not equilibrium prices (Hicks, [1933] 1982, p. 30; 1934, p. 342), he turned to a succession of 'market days' on which these exchanges take place and asked whether in this case too equilibrium prices are 'determinate', i.e., whether they are 'independent of the path followed to reach [them]' (Hicks, 1934, pp. 342–3), as is the case when exchanges out of equilibrium are not admitted. This, Hicks argued, would be true only in a stationary economy.

Before entering into Hicks's argument it is important to note that, with its implicit distinction between two notions of price – the prices at which exchanges begin on the first 'day', that theory is clearly unable to determine, and equilibrium prices, that are instead determined by theory and towards which the former prices should move – and with its reliance on the repetition of transactions in order that this movement should effectively occur, the problem of 'determinateness' from which Hicks started can *to a certain extent* be brought under the more general problem of the gravitation of market prices around normal prices; as, we may add, it is clearly necessary if the argument is to apply to a discussion of the traditional method.[6]

It is at any rate the very necessity to rely on the repetition of transactions on several 'market days' that, in Hicks's view, restricts the traditional method to the study of stationary states. 'This is obvious,' Hicks explains,

> as soon as one considers that the conditions of equilibrium in Period II, in an economy with net saving, cannot be the same as they were in Period I; so the prices in Period II cannot be the same. (Hicks, [1933] 1982, p. 31)

Hicks's argument appears, however, to rest on a confusion between two different, and as we shall see mutually independent, problems. The problem of the 'determinateness' of equilibrium consists in the possibility that the path followed by prices may lead to equilibrium values that are *different* from those initially determined by theory; and this may occur because equilibrium prices change *as a result* of circumstances that arise only during the process of price adjustment – indeed, it is precisely for this reason that they could not be predicted by theory. In other words, it is implied in the very terms of the problem of 'determinateness' that the influence from which equilibrium prices must be shown to be independent is that of forces of an accidental nature, as disturbances occurring during the process of price adjustment typically are. A substantially different problem is to establish whether equilibrium prices do not change significantly *during* the adjustment process under the influence of more persistent forces, so that the repetition of transactions is unable to show any tendency of market prices to move towards their normal level; and the effects of capital accumulation would seem relevant only to problems of this second order. Suppose, for example, that thanks to the absence of any accidental disturbance exchanges commenced at equilibrium prices right from the first 'day'. Clearly in this case the problems connected with the 'determinateness' of equilibrium would entirely disappear, whereas those that would arise, if any, on account of capital accumulation would remain. This distinction, which simply reflects the distinction between accidental and persistent forces that is at the basis of the traditional method, is blurred in the passage cited above by the fact that Hicks, in translating the 'days' from which he began into 'Periods' of such length that in passing from one to the other the effect of capital accumulation is appreciable, seems actually to suggest that the prices at which single transactions are held and equilibrium prices change *at the same speed*. Hicks provides no argument in support of this identification. It would represent, however, a *different* argument (as we shall see, it is

in fact at the bottom of Hicks's *second* reason for maintaining that the traditional method would be limited to the stationary state) and we shall accordingly return to it later on (see pp. 139–41).[7]

Remaining for the moment with the problem of 'determinateness', it is clear, therefore, that, *if* there is to be a connection between this problem on the one hand, and those that would arise from capital accumulation on the other, this can only be of an *indirect* kind. In particular, if, as we have just seen, the problem posed by the 'determinateness' of equilibrium necessarily stems from the influence of forces of an accidental nature that act during the process of price adjustment, the only role for the accumulation of capital in this respect would appear to be the one of conferring *cumulative force* to some accidental disturbance that, during that process, has made prices diverge from their equilibrium level. It seems however that the mere *possibility* of such cumulative effects depends not so much upon the accumulation of capital, as on the simple *existence* of capital goods (or durable goods in general), by which the results achieved in a given round of transactions (the 'market day') transmit their effects to subsequent transactions. And this emerges clearly enough from the way Hicks himself reiterated his thesis in 1934. The equilibrium prices, writes Hicks,

> will be reached, either if contracts are made provisionally or (a more important case) if people come to the market on successive 'days' with the same dispositions to trade, and there is no carry-over of stocks (or a constant carry-over) from one day to the next. (Hicks, 1934, p. 343)

As can be seen, after excluding once again the case of provisional contracts as not particularly relevant, here too, in order for equilibrium prices to be 'determinate', Hicks requires that the circumstances under which transactions are repeated remain absolutely constant from one 'day' to the next: and here too Hicks derives the limitation of the traditional method to the study of stationary economies from this condition (Hicks, 1934, p. 346). But in referring to the need that there be 'no carry-over of stocks' Hicks makes it clear that it is the very existence of capital goods that prevents the absolute constancy in the 'dispositions to trade' from being respected. For, if on a certain 'day' a capital good were to be exchanged at a price different from its equilibrium price, the alteration that would follow in the position of the individuals involved in the exchange would be transferred to subsequent 'days', thus generating the possibility of cumulative effects even in the absence of the production of new capital goods.

In conclusion, it appears that Hicks's first argument is unable to support his thesis limiting the traditional method to the study of stationary economies. Before we leave the problem of 'determinateness', we must ask however whether it is not equally bound to pose some difficulties to the traditional method. For, from what we have just seen, it may appear that the problem of 'determinateness' may supply the basis for claiming that the traditional method would be confined to the study of economies without durable goods, i.e., to the study of what are sometimes called 'atemporal' economies.[8]

In order to explore this possibility we must return to the origin of the problem discussed by Hicks in the only apparently similar problem of 'determinateness' that presented itself to Marshall in connection with his notion of 'temporary equilibrium' (on this point, see Hicks, [1933] 1982, p. 29; 1934, p. 342). For this appears to be largely responsible for the results that, at least at first sight, seem to follow from Hicks's problem.

As we saw, the problem of 'determinateness' is a single aspect of a wider problem: that of the relation between effective prices and prices determined by theory or, as we have also called it, the problem of the adjustment of the former to the latter. As Hicks noted, the point here is to ensure that these values are substantially independent of what occurs in the adjustment process itself. Now, in the context of Marshall's temporary equilibrium ([1890] 1961, pp. 331–6), the problem of price adjustment must first of all necessarily be posed in terms of the *convergence* of prices to their equilibrium values. What we have here is in fact the price at which a *given* quantity of a certain commodity would be *entirely* absorbed by the market (either by purchasers or by sellers wishing to store the commodity in order to sell it in the future); and treating the available quantity of the commodity as given so restricts the time period taken into consideration – Marshall's 'market-day' (*ibid*., p. 337) – that, when the problem of price adjustment is posed in this context, it is out of the question to refer to the gravitation of effective prices *around* the prices determined by theory. Moreover, the 'temporary' demand and supply curves whose equilibrium should give us that price (the quantity of the commodity that, at the various possible price levels, the purchasers would be willing to buy and the sellers to sell *on that same 'day'*) are potentially so unstable that the problem of showing that they are independent of the largely indeterminate prices at which exchanges begin appears not only particularly urgent, but also especially difficult to solve.

Essentially, there are two ways in which these 'initial' prices may affect the position and the slope of the 'temporary' demand and supply curves: by generating effects on the position of the individuals similar to those

we have seen in connection with the problem discussed by Hicks,[9] or by inducing changes in the expectations concerning the future trend of the price of the commodity. Owing, however, to the predominant role we must assume accidental circumstances play in determining the 'temporary' demand and supply curves – and hence the extreme sensitivity of the latter to every change, however slight, in these circumstances – the assumptions Marshall had more or less explicitly to introduce in order to prevent the onset of these two effects (the assumption of a 'constant marginal utility of money' and the implicit limitation to the case of highly perishable commodities), just as any other that may be introduced for the same purpose (as, for example, the absence of exchanges at disequilibrium prices), may not be interpreted as merely useful analytical expedients, whose removal would not substantially affect the results achieved. Nor should this come as a surprise. In the general case where the quantity of the commodity taken as given differs from the effective demand for that commodity (to be determined through the equilibrium between its *normal* demand and supply curves), Marshall's notion of temporary equilibrium assumes that the *extent* to which the market price must be above or below the normal price in order for that quantity of the commodity to be absorbed by the market is known. This notion comes, therefore, very close to an attempt at a theoretical determination of a market price, so that its lack of 'determinateness' must be seen simply as a reflection of the more general problems that beset the possibility of determining market prices (on this point, and more generally for a detailed treatment of the notion of temporary equilibrium in Marshall, see Ciccone, 1999, pp. 75–6).

Hicks showed himself well aware of all this when he suggested looking at the problem of 'determinateness' by considering a *series* of 'market days'. His error lay then in not realising that his suggestion served not to solve the same problem as Marshall's in a less arbitrary way than Marshall had done, but to demonstrate that the problem of price adjustment – and hence the more specific problem of 'determinateness' – must be posed only in respect of normal prices.

This is reflected in the hybrid nature of the problem discussed by Hicks. Although he extends his analysis to a number of 'days', Hicks appears to retain certain features of the problem that appear to be pertinent only to the case of a single 'day'. As we have seen, Hicks requires that the quantity of the various commodities that is supplied in each 'market day' remain constant. But now that, unlike what occurred in the case of Marshall's temporary equilibrium, the period of time taken into consideration is sufficiently long, not only for the quantity produced, but also for the productive capacity in each sector to change, we

must conclude that it is the owners of the means of production who *choose* not to shift their capital. Moreover, when dealing with the adjustment of prices over a fairly long period, we may ask – as in effect Hicks did – whether these prices will *reach* the prices determined by theory only if, contrary to what can be legitimately expected, it is assumed that accidental disturbances do not continually occur. It appears then that the problem discussed by Hicks actually falls under that of gravitation only when we confront the absolutely artificial case of an economy that features no disproportion in the distribution of the means of production, and in which the only reason that accounts for the difference between market prices and normal prices lies in the influence exerted by some accidental disturbance at the beginning of the period under examination.

If, instead, Hicks had pursued his suggestion to its end, he could have seen that, when the question is that of the gravitation of market prices around their normal level, one needs to look at the distinction between more and less persistent forces in order to assess the importance of those cumulative effects that appear to endanger the 'determinateness' of the prices determined by theory. For once the possibility of that distinction is admitted, there seems to be no reason why such effects cannot be dealt with in the same way as other accidental disturbances: namely as effects which, far from being absent (as Hicks appeared to believe would be necessary), continually displace the economy from its normal position and are continually corrected by the competitive tendency to return to that position. Besides, formulating the problem in terms of the gravitation of market prices, rather than their convergence, would probably have led Hicks not to over-estimate the effect that capital accumulation is likely to exert on normal prices during the adjustment process. Finally, this should have led Hicks to pose the entire question of the status of the assumptions normally used in the study of the adjustment process on a different level from that of their simple 'realism'. For all that Hicks could really have concluded from his argument is that the assumptions both of a pure flow economy and of the absence of exchanges at disequilibrium prices are merely analytical expedients for studying the adjustment process of effective prices to their theoretical values *separately* from the determination of these latter values; and hence that both assumptions are perfectly valid, and to a great extent mutually interchangeable, when the method adopted provides the basis for that separation, while both lead to inevitable contradictions when this is not the case.[10]

This brings us to the second reason that can be found in Hicks's writings in support of the alleged limitation of the traditional method to

the study of stationary economies. In *Value and Capital* Hicks claims that this limitation would derive from the failure of the traditional method to consider price expectations: 'the case of a Stationary State', Hicks argues, would in fact be the way 'to evade the issue' connected with 'the fact that supplies (and ultimately demands too) are governed by expected prices quite as much as by current prices' (Hicks, 1939, p. 117). Since, as can be easily seen by going back to the brief outline of the traditional method we have sketched in section 6.2, an important role is ascribed within that method to price expectations in the study of gravitation,[11] Hicks's statement may not be understood as the charge that the traditional method ignored the existence of price expectations altogether. His second argument must therefore rest on the possibility, mentioned above (p. 135), that, *before* the gravitation process has been able to show any tendency of market prices to move towards normal prices, the latter change *significantly* through the effect of capital accumulation. The need to include price expectations among the givens would thus ultimately stem from the impossibility of distinguishing, among the forces acting on the economic system, between forces having greater and lesser persistence; while the limitation of the traditional method would follow from the fact that the distinction appears to lose much of its importance in a stationary state, where no change in the influence exerted by some of those forces is present.[12]

It is at this point that an aspect in the writings of the founders of the new method reveals its whole importance. It is certainly notable that, although this method is variously presented as, to use Hicks's own expression (1934, p. 347), an 'extension' of the traditional method, no argument is to be found in these writings aimed at showing how 'temporary' or 'intertemporal' equilibrium prices should provide a *better* guide to the effective trend of the economic system than the one traditionally provided by normal prices. Now, this can ultimately be explained only on account of the belief that, thanks to the new method, economic theory could go beyond the distinction between two notions of price and could pass sooner or later to the determination of the actual position of the economic system. This belief is deeply rooted in that complex set of arguments we mentioned in note 1 concerning what we may call the 'dynamic' interpretation of the new method. Clearly this belief implies that the traditional method, with its distinction between more and less persistent forces, would not be *necessary* (an idea that, in turn, reinforces those doubts concerning the aims pursued with the traditional method, which we hinted at the beginning of Section 6.2); whereas what we find in the critique by Hicks we have just seen, as in

the analogous one we saw above in terms of his identification between 'days' and 'Periods', is the idea that that distinction would not in general be *possible*. But since, once again, *no argument is provided in support of this idea*, we must conclude that it is not so much an objection levelled at the traditional method, as a mere reflection of that belief concerning the 'dynamic' nature of the new method. And this emerges rather clearly in the very way in which the argument regarding the failure to consider price expectations is presented in *Value and Capital*. For if the argument appears, at least at first sight, plausible, this is because, in presenting it, Hicks actually ends by abandoning any reference, not only to the traditional method he is supposed to be discussing, but to the very notion of a theoretical price as distinguished from an actually observable price, as is witnessed by the fact that his argument concerning the relevance of price expectations is conducted in terms of the forces that would influence 'the system of prices *existing at any moment*' (Hicks, 1939, p. 116, my italics). Nor is it difficult to show how such an interpretation of the new method was bound to reveal itself as internally inconsistent. For though the new method has abandoned the determination of a position characterised by a uniform rate of profit, it is still based on the notion of equilibrium and *hence* on the distinction between two notions of price.

6.4 Marginalist theory and the treatment of capital

When it comes to the second critical thesis – the merely derived nature of the difficulties in the treatment of capital – we must distinguish between two cases. On the one hand stands the charge of lack of realism levelled at the treatment of capital in the context of the traditional method. This is, for example, what Hayek seems to refer to when he writes that 'most of the shortcomings of the theory of capital … are due to the fact that it has in effect only been studied under the assumptions of a stationary state, where most of the interesting and important capital problems are absent' (Hayek, 1941, p. 14). As the identification of the traditional method with the study of stationary states shows, this is a simple reflection of the more general charge levelled at the traditional method discussed above. But the same thesis can be argued at a deeper analytical level: according to this version, even the *logical* difficulties encountered by marginalist economists in the treatment of capital would stem from the employment of the traditional method. The relevance of this thesis to a correct understanding of the origin of the new method is immediately evident; but though the idea is repeatedly hinted at both by Hicks and by Hayek,[13] and would probably be

taken today very much for granted, it is only in Lindahl that we find an attempt to demonstrate it.

In order to understand Lindahl's argument we must first consider a particular difficulty that, in his opinion, besets the traditional method; or better, a particular consequence that he derives from the identification of the traditional method with the study of stationary states (Lindahl, [1929] 1939, pp. 310–12). Being unable to determine the path followed by the economic system in passing from one position of equilibrium to another, the traditional method, Lindahl notes, must be utilised to set up 'comparisons' between alternative positions. In order to be significant, however, comparisons must be made between positions that differ by a single circumstance, or at most by a limited number of the circumstances determining value and distribution in each. And it is this, Lindahl argues, that is only rarely possible.

Taken literally, the question raised by Lindahl resolves into enquiring up to what extent the givens chosen by theory may be considered independent of one another. Now, it does not appear that this question can ever be answered in terms of an *absolute* independence, nor that this is necessary for the validity of a theory. It is however true that the lack of independence becomes particularly marked if, as Lindahl does, the traditional method is understood as the study of stationary economies. For if we start from an economic system in which net savings are *effectively* null and allow one of the circumstances determining value and distribution to vary, we cannot, in general, expect net savings *to remain* null, so that if we seek to compare the first stationary state with a second stationary state, we will have to admit other changes in the givens as well as the one originally considered.

In Lindahl's opinion (*ibid.*, pp. 312–13) this first difficulty is 'connected' with a second one, concerned not so much with the possibility of studying the effects of the variation in one of the circumstances determining value and distribution assuming that all the others remain unvaried, as with the possibility of defining in unambiguous terms what is meant by invariance in these circumstances. And according to Lindahl the most important *example* of this second order of problems – 'which continually crop up in the treatment of static problems' – would regard the treatment of capital; or, more precisely, 'the method by which the *quantity of capital* should be measured under stationary conditions' (*ibid.*, p. 313, italics in the original).

Two methods, says Lindahl (*ibid.*, p. 313 ff.), have traditionally been used to measure this quantity. The first, adopted in particular by Böhm-Bawerk and initially by Wicksell, is the 'average period of production',

i.e., the weighted arithmetical mean of the periods of time in which the different quantities of the original factors required directly or indirectly for the production of consumers' goods remain invested, where the weights used for each period are the quantities of labour used in it. Alternatively, capital is measured as a 'sum of values'.

In Lindahl's view, the problem with these measures is that neither provides an instrument able to determine 'any fixed points for a comparison between different stationary situations' (*ibid*., p. 317); and it is in this sense that the difficulties in the measurement of capital should represent an example of the methodological difficulties in ascertaining the invariance in the givens determining value and distribution. We have seen that the problem regarding the possible lack of mutual independence among the givens from which Lindahl started out is none other than the reflection of his identification of the traditional method with the study of stationary states. But even if, to follow Lindahl's argument, we admit this identification, this does not help us to understand why the problems in measuring the 'quantity of capital' should be 'connected' with that first order of problems. All that can be gathered from the long section that Lindahl devotes to the question (*ibid*., pp. 314–17) is the *analogy* that in the case of capital can at least at first sight be drawn between problems concerning the mutual independence of the givens and those that concern their measurement: for just as the variation in one of the givens is necessarily attended by variation in the others when alternative stationary states are compared, so, Lindahl seems to suggest, when one of the givens changes, it appears that the 'quantity of capital' is changed, because its value is changed; and it may even appear that it changes in different directions according to the method used in measuring it.

This is as far as Lindahl goes – his argument is cut short by the following remark:

> The difficulties here mentioned [in the measurement of capital] are *associated* with the stationary setting of the problem. On account of its artificial and very special assumptions the static problem has little or no connection with the phenomena determining prices in the real world. Therefore the attempt must be made to build upon these foundations an improved analysis which will have more general validity. (Lindahl, [1929] 1939, p. 317, my italics)

Here the 'association' between the problems in the measurement of capital and 'the stationary formulation of the problem' – which is what

Lindahl set out to demonstrate – is simply taken for granted; the abandonment of the traditional method, identified with this 'formulation', appears then once again to be dictated exclusively by those considerations of 'realism' that we saw in the preceding section.

Lindahl appears, however, to be the first to doubt his thesis. He it was, at any rate, who, shortly before the passage cited above, had supplied elements sufficient to conclude that the *origin* of the problems in the treatment of capital should be sought somewhere else than in a simple analogy with the problems that would arise when a sufficient degree of independence among the givens cannot be assumed. 'The awkward thing', he had written in regard to the measurement of capital by the average period of production, 'is that the invested values are added together without account being taken of the interest that has accrued from the time of investment' (*ibid*., pp. 314–15); and, concerning the measurement of the quantity of capital as a sum of values, he had pointed out its 'conventionality':

> What is for instance meant … by saying that two separate communities with different populations, on different cultural levels and with different technique and *consequently quite different price relations*, have the same quantity of capital? To regard this as being the case if the circulating capital has the same value … is evidently a purely conventional idea. (*Ibid*., p. 317, my italics)

Indeed, in order for Lindahl's thesis to be valid, it would be necessary that the difficulties in measuring capital by means of the average period of production or as a sum of values should rise only within the traditional method; or, to put it differently, that *those* problems could be *solved* in the context of a different method. However, it is not clear why the perplexities aroused by the assumption (necessary in order to utilise the average period of production) that profits should be distributed according to the simple rate of interest and not, as imposed by competition, according to the compound rate, should be any less serious when we move from the traditional method to the method based on the notions of 'temporary' or of 'intertemporal' equilibrium. Nor is it clear why the circularity involved in including a sum of values among the givens determining value and distribution should not occur outside the traditional method.[14]

It is at this point that it is important to note that, when considering the methods for measuring capital, Lindahl takes into consideration *only* the methods for measuring capital as a *single magnitude*. The fact is that

when, as is distinctive of marginalist *theory*, capital is included among the givens, the traditional method *dictates* that capital be conceived as a single quantity. As is well known (see Garegnani, 1960, 1990), the reason for this lies in the fact that only in this way is it possible to determine a uniform rate of net return on the supply prices of the existing capital goods, which, as we have seen, is at the basis of the traditional method. And Lindahl showed himself quite aware of this when, a few pages before the discussion of the 'difficulties in the measurement of capital', he wrote: 'The static premiss ... implies a rather definite time-structure of … capital'. As a result of this, he added,

> only the *total amount* of the circulating capital can be regarded as *known* … The *distribution* of this capital among different types of services and among services of various ages belongs to the *unknown* factors of the problem, like the rate of interest. (Lindahl, [1929] 1939, p. 304, my italics)

6.5 Conclusions

On closer inspection, it appears then that the complex argument adduced by Lindahl, Hicks and Hayek in support of their abandonment of the traditional method requires, as originally suggested by Garegnani (1976), to be *reversed*. As we have seen, it is not possible to criticise this method for its lack of realism. To be sure, it is in this connection that the arguments aimed at criticising the traditional method appear to be more inextricably interwoven with arguments that, in effect, take the abandonment of that method very much for granted; and it is here – that is, on the 'dynamic' interpretation of the new method – that presumably most of the work of reconstruction and clarification remains to be done. Still, we have seen that it is incorrect to claim that the traditional method is confined to the study of stationary, or steady, or 'atemporal' economies. On the other hand, the logical difficulties that arise in the treatment of capital – the impossibility of conceiving capital as a single magnitude that can be measured independently of value and distribution – are 'associated' with the traditional method in a very different way from the one proposed by the three economists (it is nevertheless worth noticing that these difficulties do appear, though in a subordinate position, in their writings). For these difficulties are entirely general; that is to say, they arise whatever method is adopted. When, however, capital is included, as is distinctive of the marginalist theory, among the givens from which value and distribution have to be determined, the adoption

of the traditional method, by imposing that capital be treated as a single magnitude, prevents those difficulties from being *avoided*. Their missing solution thus makes the marginalist theory incompatible with the traditional method.

Notes

1. In the formation, and even more in the subsequent diffusion, of the new method a role by no means unimportant was played by certain considerations, suggested especially by Hicks, aimed at showing how this method ought to be conceived as the *culmination* of a number of studies that in the 1920s and 1930s had shown what is still nowadays frequently considered a *marked* interest in 'dynamic' questions. This idea is in turn connected with the need, that presented itself to the authors who were engaged in laying the foundations of the new method, of presenting it as a way of bringing the marginalist theory into *closer* contact with reality. Reasons of space forbid treatment of this aspect of the problem in the present work, notwithstanding the light it would shed even on the more strictly analytical points that arise in connection with the abandonment of the traditional method, as well as on the subsequent abandonment of the new method by the same theorists who were the first to favour its adoption. I hope to be able to return to this aspect of the question in a larger work, which, taking stock of the results achieved by several authors since Garegnani (1976) first drew attention to the methodological shift in the dominant theory (see, in particular, Milgate, 1982; Petri, 1991; Gehrke, 2003), I am preparing on the origin and the nature of the new method and the main lines of which, together with the part provided here, were presented at the Conference for the fiftieth anniversary of the publication of *Production of Commodities*.
2. A more explicit instance of this charge may be seen in Hicks (1939, p. 119), where the author writes that 'stationary-state theorists', which now include *at least* Böhm-Bawerk and Wicksell, 'naturally regarded reality as "tending" toward stationariness'.
3. More properly, what can be observed at every moment on the market will, in general, be a plurality of prices for each good. The 'market price' of which Smith speaks will not, however, lose any of its character of effective variable if we understand it as an average of those prices.
4. For a more detailed analysis of the relation between market prices and normal prices, see Vianello (1989) and Ciccone (1999).
5. Reasons of space forbid treatment of the question of how far Lindahl and Hayek shared, or in the case of Hayek even inspired, the *details* of Hicks's arguments. That they shared his main thesis appears, however, to be out of the question (see, for instance, the passages by Hayek and by Lindahl quoted respectively on pp. 141 and 143 below).
6. We shall return later (see pp. 137–9) to discuss to *what* extent this is actually true.
7. The transition from the first to the second reason (and the apparent plausibility of Hicks's statement) follow from the fact that, after suggesting looking at the question of 'determinateness' on a succession of 'days', Hicks quite

unexpectedly asks 'what are the *maintainable* prices?' (Hicks, [1933] 1982, p. 30; my italics), which, however, has clearly nothing to do with the question of 'determinateness' itself.

8. Such a limitation of the traditional method is indeed hinted at in Hayek ([1939], 1944, p. 353).
9. The only difference is that while in Marshall's case these effects may concern all kinds of goods (because by assumption all goods last till the end of the single 'day' taken into consideration), in Hicks's case the possibility concerns only durable goods, i.e., goods that last more than one 'day'.
10. It should be noted that the possibility of separating *in general* the determination of a normal position from the study of the gravitation around this position appears to imply not so much that cases in which events occurring during the gravitation process appreciably influence the normal position itself should be excluded altogether, as that they should be studied *separately* when the presence of *specific* circumstances suggests that that may occur. It remains, of course, to be seen whether the marginalist theory – in which forces of demand and supply govern both the determination of the normal position and the explanation of the gravitation around that position – is able, like the classical theory, to allow for this possibility.
11. Within the traditional method, expectations appear in the first place to have an important role in determining the *divergence* between market and normal prices. Permanent changes in the demand structure or in production conditions normally lead to an initial state of disproportion in the distribution of the means of production precisely because those changes cannot, in general, be exactly foreseen; moreover, the uncoordinated fashion of the adjustment to these changes may lead to excessive inflows or outflows of capital, while a state featuring industries either over- or underequipped with capital may also occur when temporary changes in the proportion between quantity brought to market and effective demand are taken as permanent, or, more simply, when their duration is not correctly estimated (for a more detailed discussion of these last two cases, see Vianello, 1989, pp. 93–96, and, for a partly different treatment, Ciccone, 1999, p. 73 and note 9). In the second place, expectations appear to belong to those forces, that cannot be systematically studied, that influence the *extent* of the divergence between market and normal prices, as is the case when the expectations formed by sellers and buyers during the exchanges influence 'the eagerness of the competition'. Finally, it is clear that the *correction* of wrong expectations, which is just an aspect of their endogenous formation, is an integral part of the process of adjustment that brings market prices back to their normal level.
12. Note that already in his earlier writings Hicks had mentioned the possibility of overcoming the alleged limitations of the traditional method by including price expectations among the givens of the theory (Hicks, [1933] 1982, p. 31; 1934, p. 346). Since, however, such an inclusion would be of no avail in preventing the emergence of those cumulative effects which, as we have seen, are at the origin of the problems connected with the 'determinateness' of equilibrium, this must be seen as another sign of the confusion, which is present in those writings, between two different reasons for maintaining that the traditional method would be limited to the stationary state. Note moreover that the distinction between forces of differing persistence would be of little importance

also in an economy in balanced growth. Admitting this case among those in which the traditional method would be valid (on this point, see the statement by Arrow quoted on p. 130 above) means, therefore, that the critique levelled at this method concerns, not the problem of 'determinateness', but the possibility of distinguishing among forces of differing persistence.

13. In a famous passage in *Value and Capital* Hicks, for example, writes: 'Of course, people used to be able to content themselves with the static apparatus, only because they were imperfectly aware of its limitations. Thus they would often introduce into their static theory a "factor of production" capital and its "price" interest, supposing that capital could be treated like the static factors … . That some error was involved in this procedure would not have been denied; but the absence of a general dynamic theory, in which all quantities were properly dated, made it easy to underestimate how great the error was' (Hicks, 1939, p. 116n). The origin of the difficulties in the treatment of capital ('some error involved in this procedure') lies then for Hicks in the attempt to treat capital as a single magnitude ('a "factor of production" capital'); and this in turn would be necessarily connected with the traditional method ('Thus they would often introduce…'). What we are not told is, however, why this should be so. The same can be said regarding Hayek; see, for example, the passage in Hayek (1941, pp. 4–6).
14. Lindahl's conclusion was undoubtedly favoured by the fact that he sometimes refers to the two methods for measuring the quantity of capital as, respectively, an 'un-weighted' and a 'weighted' average period of production. The weights used in the second case are not, however, the quantities of the original factors invested in each period (which we must assume are utilised also in the first, 'un-weighted', case) but 'the interest costs relevant to each period of investment, calculated at compound interest' (Lindahl, [1929] 1939, p. 314). However, this introduces a difference between the two methods that lies not merely in the weighting technique. For in the second case the 'average production period' is *defined* by Lindahl 'as the number of periods, during which the total value of the original services supplied in a given period with compound interest *at the given rate*, becomes equal to the total value of all services maturing and consumed during the same period' (*ibid*., p. 308, my italics). And only in a note added to the English translation in 1939 does Lindahl comment: '*T* [the average production period in the second version] is thus a function of *i*. Since *i* is unknown, what is given is the form of the function, not the value of *T*' (Lindahl, 1939, p. 308n.); and further on, 'the measure of capital is made dependent on ... the rate of interest–which belong[s] to the unknown factors of the problem' (*ibid*., p. 317n.). For a discussion of the concept of the average period of production in Hicks, see Fratini (2012; 2013, pp. 119–23).

References

Arrow, K.J. (1991) 'Certainty Equivalence and Inequivalence for Prices', in L.W. McKenzie and S. Zamagni (eds), *Value and Capital Fifty Years Later. Proceedings of a Conference held by the International Economic Association at Bologna* (London: Macmillan): pp. 41–63.

Ciccone, R. (1999) 'Classical and Neoclassical Short-Run Prices. A Comparative Analysis of their Intended Empirical Content', in G. Mongiovi and F. Petri

(eds), *Value, Distribution and Capital. Essays in Honour of Pierangelo Garegnani* (London: Routledge): pp. 69–92.
Fratini, S.M. (2012) 'The Hicks-Malinvaud average period of production and "marginal productivity": a critical assessment', *European Journal of the History of Economic Thought*, forthcoming.
Fratini, S.M. (2013) 'Malinvaud on Wicksell's Legacy to Capital Theory: Some Critical Remarks', in E.S. Levrero, A. Palumbo and A. Stirati (eds) *Sraffa and the Reconstruction of Economic Theory* (Basingstoke: Palgrave Macmillan): pp. 105–28.
Garegnani, P. (1960) *Il Capitale nelle Teorie della Distribuzione* (Milano: Giuffrè).
Garegnani, P. (1976) 'On a Change in the Notion of Equilibrium in Recent Work on Value and Distribution', in M. Brown, K. Sato, P. Zarembka (eds), *Essays in Modern Capital Theory* (Amsterdam: North Holland): pp. 25–45.
Garegnani, P. (1990) 'Quantity of Capital', in J.L. Eatwell, M. Milgate and P. Newman (eds), *Capital Theory* (London: Macmillan): pp. 1–78.
Gehrke, C. (2003) 'On the transition from long-period to short-period equilibria', *Review of Political Economy*, 15: 85–106.
Hayek, F.A von ([1939] 1944) 'Price Expectations, Monetary Disturbances and Maladjustments', in *Profits, Interest and Investment,* Routledge, reproduced in *Readings in Business Cycle Theory* (Philadelphia: The Blakiston Company): pp. 350–5.
Hayek, F.A. von (1941) *The Pure Theory of Capital* (London: Macmillan).
Hicks, J.R. (1933) 'Gleichgewicht und konjunktur', *Zeitschrift für Nationalökonomie*, 4 (Engl. trans. in Hicks, 1982: pp. 28–41).
Hicks, J.R. (1934) 'Léon Walras', *Econometrica*, 2: 338–48.
Hicks, J.R. (1936) 'Keynes' theory of employment', *Economic Journal*, 46: 238–53.
Hicks, J.R. (1939) *Value and Capital* (Oxford: Clarendon Press).
Hicks, J.R. (1982) *Money, Interest and Wages. Collected Essays on Economic Theory*, Vol. II (Oxford: Basil Blackwell).
Kaldor, N. (1934) 'A classificatory note on the determinateness of equilibrium', *The Review of Economic Studies*, 1: 122–36.
Lindahl, E. (1929) 'Prisbildningsproblemets uppläggning från kapital-teoretisk synpunkt', *Ekonomisk Tidskrift,* 30 (Engl. trans. in Lindahl, 1939: pp. 269–350).
Lindahl, E. (1939) *Studies in the Theory of Money and Capital* (London: George Allen & Unwin).
Marshall, A. ([1890] 1961) *The Principles of Economics*, 9th (variorum) edn, (ed.) C.W. Guillebaud (London: Macmillan).
Milgate, M. (1982) *Capital and Employment* (London: Academic Press).
Petri, F. (1991) 'Hicks's recantation of the temporary equilibrium method', *Review of Political Economy*, 3: 268–88.
Smith, A. ([1776] 1979) *An Inquiry into the Nature and Causes of the Wealth of Nations*; (general editors) R. H. Campbell and A. S. Skinner, (textual editor) W. B. Todd (Oxford: Clarendon Press – The Glasgow Edition of the Works and Correspondence of Adam Smith).
Vianello, F. (1989) 'Natural (or normal) prices. Some pointers', *Political Economy. Studies in the Surplus Approach*, 2: 89–105.

Part II
The Revival and Development of the Classical Theory of Distribution

7

Marx's Theory of Wages and the Revival of the Surplus Approach

Enrico Sergio Levrero

7.1 Introduction

This work aims to clarify some aspects of classical wage theory, looking at the form in which it was advanced by Marx. It also aims to analyse the ability of this theory to explain income distribution in the context of the present capitalist societies, especially with respect to the relationship in a fiat money economy between money wages and real wages. In particular, Sections 7.2–7.4 reconstuct the notion of the subsistence wage and the determinants of distribution according to Marx; Sections 7.5 and 7.6 consider the effects of technical changes on the secular trends in real wages; Section 7.7 focuses on the different possible forces at work in determining income distribution, and thus in 'closing' Sraffa's price system, when the wage rate happens to be above the subsistence level; lastly, Appendix A deals more exhaustively with the relationship between money wages and real wages, and the effects on price inflation and income distribution of a discrepancy between the real wage rate pursued by the workers in wage bargaining and the real wage rate targeted by the monetary authorities.[1]

7.2 The necessary price of labour

According to Marx (1867–94, I, pp. 40–1), labour-power is a commodity, but with some peculiarities. Unlike other commodities, which are first produced and then brought to market, labour-power is created only *at the moment* it is brought to market (Marx, 1867–94, I, pp. 535–6). Second, in order for labour-power to be created, the labourer – who, unlike a slave, disposes 'of his labour-power as his own commodity' – must be forced to sell it, having 'no other commodity for sale' and himself being

'short of everything necessary for the realisation of his labour-power' (Marx, 1867–94, I, p. 169). Finally, labour-power is in any case a special commodity, since there is no other commodity whose price embodies a historical and moral element and whose conditions of use (that is, the intensity and conditions of work) influence the 'exchange-value resulting from it' (Marx, 1862–63, I, p. 45) and are a terrain of conflict between sellers and buyers.

Now, if labour-power is a commodity, albeit one with special characteristics, it must exist on the market, that is, its production 'presupposes its existence', or the existence of the labourer. In fact 'there can be no labour unless the worker lives and maintains himself, i.e., receives the necessary wages' (Marx, 1862–3, II, pp. 417–18). They do not, however, include only the means of subsistence which are necessary for the mere maintenance of the labourer. Since labour-power actually 'becomes a reality only by its exercise', and thereby 'a definite quantity of muscle, nerve, brain' is wasted and must be restored, '(t)his increased expenditure demands a larger income' in order for the worker tomorrow to be able 'to repeat the same process in the same conditions as regards health and strength (Marx, 1867–94, I, p. 171; see also Marx, 1862–3, I, p. 45).

The means of subsistence to be given to the labourer must therefore at least be sufficient to maintain him *in his normal state* as a labouring individual. In this sense, subsistence is like the oil and fuel for an engine. Hence, Marx stated, in order for subsistence to be calculated, a certain length of the labourer's life, to which correspond a certain length and intensity of the working day, must be assumed. But, as the expected returns from the use of machinery are gross of its wear and tear, Marx maintains that the necessary wage must also include the amount of commodities to be given for replacing the 'wear and tear' of the labourer, that is, it must be sufficient to reproduce the labourer in accordance with the demographic and social conditions of his time. Here, however, a difference with machinery clearly emerges: in calculating the price of labour, a moral and historical element must be considered. Not only do the *natural wants* of the labourer, 'such as food, clothing, fuel, and housing, vary according to the climatic and other physical conditions of his country', but

> the number and extent of his so-called *necessary wants*, as also the mode of satisfying them, are themselves the product of historical development, and depend therefore to a great extent on the degree of civilisation of a country, more particularly on the conditions

> under which, and consequently on the habits and degree of comfort in which, the class of free labourers has been formed. (Marx, 1867–94, I, pp. 171–2, emphasis added)

From the foregoing it is clear that, as with the classical economists (see, for instance, Smith, 1776, Book V, chapter II, pp. 393, 399–401), the subsistence wage is, according to Marx, different from the minimum limit of physiological subsistence. It is also clear that Marx's definition of the natural price of labour is somewhat different from that advanced by Ricardo in Chapter V of his *Principles* ('On Wages'), although it means, according to Marx as well as to Ricardo, the *necessary* price of labour, that is the *cost of reproduction* of the labourer (Marx, 1867–94, I, p. 538). In fact, like Smith, Marx does not refer to an 'unchanged population' as Ricardo (1951–73, I, p. 93) did, to define the natural wage. Marx refers only to the conditions to be satisfied in order to reproduce the number of employed people at a given stage of accumulation, whatever the owners of the corresponding wear and tear fund will make of it. In the words of Torrens, quoted with approbation by Marx (1867–94, I, p. 172) when dealing with the definition of the natural price of labour-power, this price consists in

> such a quantity of necessaries and comforts of life, as, from the nature of the climate, and the habits of the country, are necessary to support the labourer and to enable him to reach such a family as may preserve, in the market, *an undiminished supply of labour*. (Torrens, 1815, p. 62, emphasis added)

7.3 The natural and market wage rates

When looking at Marx's definition of the subsistence wage, it is therefore apparent that, contrary to Rowthorn (1982, p. 208), there is no incompatibility for Marx between a *historically* determined subsistence wage and the need for the subsistence wage to be sufficient to enable the reproduction of the labourers.

The sense of Marx's reference to the reproduction of workers emerges clearly in the *Theories of Surplus Value*, when Marx himself tackles the definition of the natural price of labour advanced by Smith and Ricardo. Marx observes that in Smith and Ricardo the natural rate of wages 'is the value of labour-power itself, *the necessary wage*' (Marx, 1862–3, II, pp. 222 and 400). He also observes that according to Smith wages can be 'above the level of the natural rate' depending on the 'rapidity with

which capital accumulates' (Marx, 1862–3, II, p. 223), and that this might bring about a rise in *the natural or subsistence wage itself* (Marx, 1862–3, II, p. 224). Then, in considering Ricardo's mechanism of the adjustment of the market to the natural wage, Marx comes to criticise the principle of population. He notes that, according to Ricardo, in the case of all commodities, the agreement of the market and natural price depends on the facility with which the supply can be increased or diminished. Ricardo stated that in the case of gold, houses and labour, this effect cannot be speedily produced, but in principle, Marx observes, this is not a problem, as it is 'only a question of the more or less rapid or slow operation' of an economic law (Marx, 1862–3, II, p. 378). Instead the problem resides in Ricardo's emphasis on the action of such a *mechanical* element in determining the wage rate, as well as in the need of a more efficient mechanism able to keep the wage rate at the subsistence level (cf. also de Vivo, 1982).

Such a need arises, according to Marx, from the fact that a wage rate greater than the subsistence wage could not bring about an increase in population, since there is no certainty that the surplus wage gained by the labourers will be spent by them on necessary consumption instead of on enlarging their life enjoyments. As John Barton had already pointed out (Marx, 1862–3, II, pp. 581–2), there is no univocal relation between wages and population, or even an inverse relation between them. Moreover, what regulates population is more the 'facility of finding employment' than the wage level. It is thus necessary for the capitalist mode of production to create 'a peculiar law of population' (see Marx, 1867–94, I, p. 632) so that capital accumulation can be sustained, and that law is based mainly on the transformation of 'circulating' into 'fixed capital'. Through such a transformation, capitalist production provides 'for unexpected contingencies by overworking one section of the labouring population and keeping the other as a ready reserve army consisting of partially or entirely pauperised people' (Marx, 1862–3, II, pp. 477–8).

Now, this 'peculiar law of population', with the related cyclical or permanent movements in the labour reserve army, is one of the main elements at the root of Marx's distinctions between the natural and market wage rates, and thus of his analysis of the cyclical or permanent changes in the price of labour.

Like Ricardo and Smith, Marx actually distinguished temporary and permanent factors affecting wages. He thus separates 'the general law of the rise or fall in the profit rate' from the rise or fall in the rate of profit insofar as it is determined by a rise or fall in wages resulting from 'the

temporary rise or fall in the prices of necessaries' (Marx, 1862–3, III, p. 312). Moreover, when considering *cyclical* variations in labour unemployment as distinct from changes in *permanent* unemployment associated with capital accumulation, Marx (see for instance 1867–94, III, p. 437) distinguished a short-run market wage rate from a long-run market wage rate – which we may say is, like the subsistence wage, an average or 'normal' wage, to be included in the *normal* price of commodities. Thus, in dealing with changes in the average market wage, Marx (1867–94, I, pp. 640–1) tells us he will abstract from 'the great periodically recurring forms that the changing phases of the industrial cycle impress on [the surplus-population]'. In this latter cyclical context, the changes in the wage rate (relative to the average) will be due to the effect that oscillations of the demand for labour around the average or *usual* level of employment can have on the relative power of the parties involved in wage bargaining. Those effects and changes probably vanish when the actual level of employment comes back to its normal level.

In the alternative case of capital accumulation, a higher average demand for labour will be satisfied by drawing on or increasing the '*customary* supply of labour' (Marx, 1867–94, I, p. 613, our emphasis), through a reduction in permanent labour unemployment or in labour underemployment, so that the wage rate will possibly remain at least to some extent at its natural or necessary level, with no need for any change in population. However, according to Marx, the workers' improved bargaining position due to the permanent change in the reserve army of labour may lead to an increase in wages above the subsistence level, and thus also to a change in the value of labour-power, as is implicit in a historically determined notion of subsistence.

It may be noted that this rules out any basis for an interpretation of Marx's theory of wages along the lines of the so-called Canonical Classical Model (see, e.g., Samuelson, 1978).[2] The same point is also useful in clarifying that, contrary to Cottrell and Darity's (1988) argument, we do not find in Marx any suggestion that capitalist competition operates in the sphere of the labour market through changes in the *overall* supply, as is it does with other commodities. Were it to operate in the labour market as it does in other markets, it would require wages to fall as long as there is an excess labour supply. This is not consistent with Marx's emphasis on the existence of permanent labour unemployment, since it would lead to the absurd conclusion that wage rate tends to zero in the absence of a mechanism drawing the economy to full employment. Indeed, as in the classical economists, who similarly admitted permanent labour unemployment (see e.g. Ricardo,

1951–73, I, pp. 389–390; II, p. 241; IV, pp. 346, 368; Smith, 1776, Book I, Chapter VIII, p. 80), Marx viewed competition in the labour market as operating within a context of norms, laws and habits which are respected, whether consciously or not, in intentional competitive behaviour. Competition merely guarantees uniform wage rates for the same kind of labour, whose normal levels, however, are fixed *outside* the competitive process, according to the relative strength of workers and capitalists in wage bargaining (see Levrero, 2011).

7.4 Some determinants of the trend of the wage rate

But what factors, according to Marx, will in fact shape the trend of the real wage rate? In particular, what determines whether the wage rate will be at the subsistence level, or above it (thus possibly changing the subsistence or necessary level itself)?

As seen above, for Marx one element shaping the level of the average wage rate is the labour reserve army, and its changes over time due to a change in the amount of productive capacity relative to the amount of working age population, and/or to changes in labour productivity. For Marx, these changes ought to be such as to remove any obstacle to capitalist reproduction, thus confining the rise of wages 'within limits that not only leave intact the foundations of the capitalistic system, but also secure its reproduction on a progressive scale' (Marx, 1867–94, I, p. 620). In particular, due to the uncertain effects of changes in wages on either capital accumulation or the rate of population growth, Marx emphasised the tendency to introduce machinery as a reaction to a rise in wages. Over time this tendency would bring about an increase in employment that would be progressively less than the increase in total capital, thus assuring, for a given population growth rate, a stable or even increasing labour reserve army.

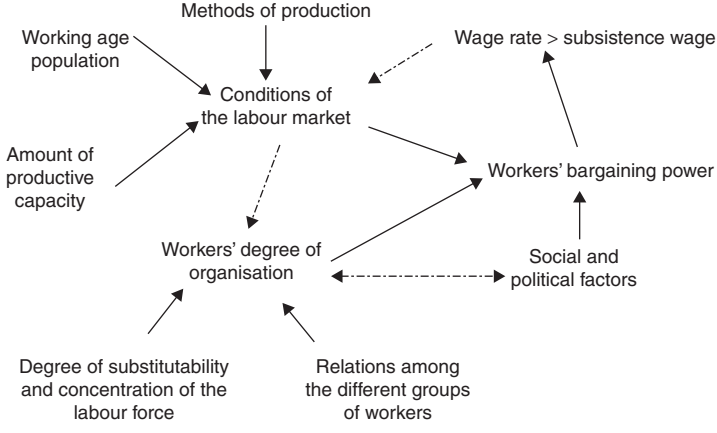
However, the introduction of machinery has further effects on the workers' bargaining position. By deskilling workers, it increases the degree of substitutability of the labour force (Marx, 1867–94, I, pp. 420–1) and throws onto the market all the members of the labourer's family, spreading 'the value of the man's labour-power over his whole family' (Marx, 1867–94, I, p. 395). Moreover, it changes the composition of the labour force and thus the weight of the unskilled and skilled workers comprising it (see for instance Marx, 1867–94, I, p. 420) and their cohesiveness in wage bargaining.

But according to Marx the workers' *degree of organisation* depends also on the 'co-operation' between employed and unemployed workers,

their degree of concentration, their class consciousness, political and cultural factors, as well as on the possibility of state interference in the labour market on behalf of one or the other of the competing parties (see, e.g., Marx, 1867–94, I, p. 640). Moreover, the workers' bargaining position is also related to the general structure of society. With respect to Ricardo's analysis of the effects of machinery, Marx noted that, in considering the use of the net product,

> [Ricardo] forgets to emphasise … the constantly growing number of the middle classes, those who stand between the workman on the one hand and the capitalist and landlord on the other. The middle classes maintain themselves to an ever increasing extent directly out of revenue, they are a burden weighing heavily on the working base and increase the social security and power of the upper ten thousands. (Marx, 1862–3, II, p. 573; see also 1862–3, II, p. 571)

Though it can acquire concreteness only with respect to a specific country and period of time, Figure 7.1 schematises some of the factors affecting wages that we have described above. As we have noted, in addition to the conditions of the labour market, Marx considers the degree of organisation of the workers, as well as social and political factors. All of them have some degree of autonomy in determining the



**Continuous arrows*: direct influence
Broken arrows: indirect or reciprocal or less definite influence

Figure 7.1 Some factors affecting the workers' bargaining position

strength of the workers in wage bargaining, and thus in determining if the wage rate will be higher than or equal to, the subsistence wage – the level of which is inherited from the past and forms a base level in wage bargaining.[3]

7.5 Relative or 'real' wages

It should now be noted that, according to Marx, it was a great scientific merit of Ricardo to have distinguished between absolute and relative wages, since the value of wages has 'to be reckoned not according to the quantity of the means of subsistence received by the worker, but according to the quantity of labour which these means of subsistence cost'. It is in fact possible that, 'reckoned in terms of use-value (quantity of commodities or money), his wages rise as productivity increases and yet the value of the wages may fall and vice versa' (Marx, 1862–63, II, p. 419; see also Marx, 1867–94, I, p. 172).[4] To consider the 'relative or real wages', that is the '*relative share* of the total product, or rather of the total value of this product, which the worker receives', is therefore the right way to evaluate the *social position* of the workers and more generally the effects of an increase in wages on the standard of living of the labourers. Indeed

> [u]p to this time, wages had always been regarded as something simple and consequently the worker was considered an animal. But here he is considered in his social relationships. The position of the classes to one another depends more on relative wages than on the absolute amount of wages. (Marx, 1862–3, II, p. 419; see also Marx, 1862–3, II, p. 404)

On the other hand, with respect to these social relationships, Marx notes that with an increase in labour productivity, 'the same number of labourers will enable the higher classes to extend, refine, and diversify the circle of their enjoyments, and thus to widen the economic, social, and political gulf separating them from their betters' (Marx, 1862–3, II, p. 572). He further observes that, according to 'the political economist', the working class should not receive any surplus wage because the only portion of the individual consumption of the labourer that is productive is that 'which is requisite for the perpetuation of the class, and which therefore must take place in order that the capitalist may have labour-power to consume' (Marx, 1867–94, I, p. 573).

Yet Marx does not overlook the possibility that the workers participate in increased productivity. It is possible, he observes, that 'owing

to an increase of productiveness, both the labourer and the capitalist may simultaneously be able to appropriate a greater quantity of these necessaries, without any change in the price of labour-power or in surplus-value' (Marx, 1867–94, I, p. 523). However, the usual situation for Marx is that the price of labour-power falls, and yet this fall will be 'accompanied by a constant growth in the mass of the labourer's means of subsistence'. Thus, Marx wrote:

> because in a given country the value of labour is falling relatively to its productivity, it must not be imagined that wages in different countries are inversely proportional to the productivity of labour. In fact exactly the opposite is the case. The more productive one country is relative to another in the world market, the higher will be its wages as compared with the other. In England, not only nominal wages but [also] real wages are higher than on the continent. The worker eats more meat; he satisfies more needs. (Marx, 1862–3, II, pp. 16–17)

But although more of their needs are satisfied in this situation, the relative social position of the workers is worsened, since the gap with the consumption possibilities of the average capitalist has widened. In fact, if you have a little house near a palace, you feel more uncomfortable and more dissatisfied, because '[o]ur wants and pleasure have their origin in society', and we 'do not measure them in relation to the objects which serve for their gratification' but 'in relation to society', so that 'they are of a relative nature' (Marx, 1884, p. 33).

7.6 The effects of technical changes

What the wage share and thus the above-mentioned relative social position of the workers will actually be will depend, according to Marx (1867–94, I, pp. 522–3), 'on the relative weight, which the pressure of capital on the one side, and the resistance of the labourer on the other, throws into the scale' on the basis of the factors outlined in Figure 7.1. Since, as we have seen, Marx believed that it is inherent in the capitalist mode of production to bring about a fall in the wage share, he believed that the balance of power would be on the side of capital.

At least in part such a prediction was linked to his belief in a tendency of the rate of profit to fall due to an ever-increasing organic composition of capital, that is, to his views regarding technical progress in the capitalist economy (cf. Meek, 1967). Marx in fact admitted that '[t]he rate of

profit could ... rise if a rise in the rate of surplus-value were accompanied by a substantial reduction in the value of the elements of constant, and particularly of fixed, capital'. But he thought that 'in reality ... the rate of profit will fall in the long run' (Marx, 1867–94, III, p. 225). That is, Marx thought the actual tendencies of the rate of surplus-value and of the organic composition of capital will lead to a fall in the rate of profit. If this happens, any increase in wages limiting the rate of surplus-value would strengthen that fall, and thus irresistible forces would be conjured up to compress wages to a minimum.

However, Marx's analysis of technical changes is not always consistent with the manifestation of that law, and thus with a necessary manifestation of a tendency to compress wages to a minimum level. Even if we were to accept that the rate of profit is determined by using the theory of labour value, the law can actually be proved only by introducing specific hypotheses regarding the prevailing technical change and the trend of the rate of surplus-value. In particular, with respect to the latter, Marx (1862–3, III, p. 310) argues that there are some limits on an increase in the rate of surplus-value (see, for instance, Marx, 1867–94, I, pp. 305, 521–2 and 530; 1867–94, III, p. 242; 1862–3, III, p. 312). Moreover, although he does not exclude capital-saving technical change and the possibility of a fall in the cost of constant capital (see Rosenberg, 1989), he assumes that capitalist development will bring about an ever-increasing physical mass of constant capital. Finally, he ultimately appeals (see Schefold, 1976) to the Ricardian increasing avarice of nature (cf. Marx, 1862–3, III, pp. 368–9; 1867–94, I, p. 506) to ensure that the *value* of the organic composition of capital will also rise, that is, that the price of capital goods will never fall to such an extent as to avoid that increase.[5]

In this way Marx can argue that at some point in time the fall in the ratio of variable to constant capital will be greater than the increase in the rate of surplus-value, thus determining a fall in the rate of profit. Yet the peculiar nature of the hypotheses underlying Marx's arguments is apparent. First, the hypotheses contrast with his belief that science will become a productive force and that economic needs shape specific forms of scientific knowledge and of technical progress. Second, it is precisely when we take into account the forms of technical change which Marx considered – that is, the division of labour, capital saving, the mechanization of production, and innovations – that it becomes evident that Marx's scenario of a fall over time in the *maximum* rate of profit can occur only in very particular cases, and that wages need not remain at a minimum in order to avoid a fall in the profit rate.

To grasp the point more fully let us write the price system as:

$$\mathbf{Ap}(1+r) + w\boldsymbol{\ell} = \mathbf{p}$$
$$\mathbf{yp} = 1$$

and derive (under the usual hypotheses) the wage-profit curve by substitution:

$$w = 1/\{\mathbf{y}[\mathbf{I} - (1+r)\mathbf{A}]^{-1}\boldsymbol{\ell}\}$$

where $\mathbf{y} = \mathbf{x(I-A)}$ is the vector of the net products, $\mathbf{x}$ that of the gross products, $\mathbf{I}$ the identity matrix, $\mathbf{A}$ the matrix of commodity inputs, r the normal rate of profit, w the wage rate, $\mathbf{p}$ the vector of prices, $\boldsymbol{\ell}$ the vector of labour inputs. Total employment L will be given by the sum of the dated quantities of labour employed in the production of the net output:

$$L = \mathbf{x}\boldsymbol{\ell} = \mathbf{y}(\mathbf{I}-\mathbf{A})^{-1}\boldsymbol{\ell}$$

If the rate of profit is equal to zero, you obtain the maximum wage rate W^* which is equal to labour productivity $1/L$, while if the wage rate is zero, the rate of profit is equal to the maximum rate of profit R^*. Drawing the wage-profit curve [W*R*] (see Figure 7.2) we see that with respect to the wage rate w_1

$$\text{tg}\,\alpha = w_1/W^* = w_1/(1/L) = w_1L/Y$$
$$\text{tg}\,ß = (W^*-w_1)/r_1 = (\Pi/L)/(\Pi/K) = K/L$$

which are respectively the wage share (with Y the net product) and the capital-labour ratio (with K the value of capital).

Now let us assume that a new method of production is discovered which allows the same (physical) net product to be obtained by less labour and a proportional increase in all the inputs, so that the capital/output ratio at zero wages increases.[6] If the wage rate in terms of the net product remains the same, the new technique (see again Figure 7.2) will bring about an increase in the rate of profit (from r_1 to r_2) and a fall in the wage share (a reduction of the tangent of angle α, now shaped by the segments OW^{**} and Ow_1). But if the wage rate increased so that the wage share and thus what Marx would have called the rate of exploitation $\Pi/W = [(1/\text{tg}\,\alpha) - 1]$ remained the same, the rate of profit would fall (to r_3).[7]

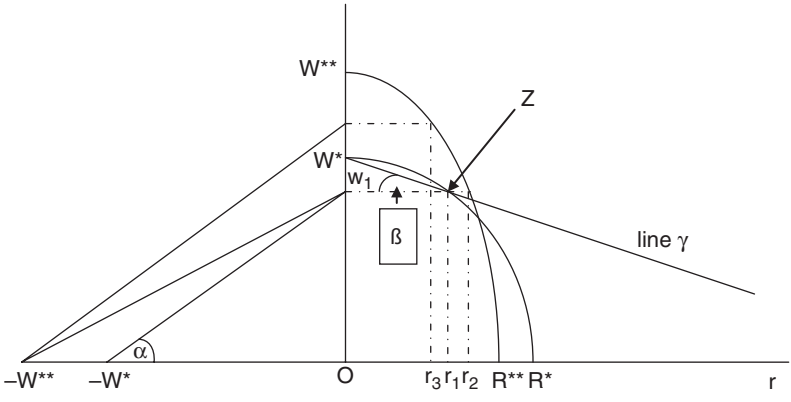


Figure 7.2 The case of mechanization of production

There is, however, no reason why technical changes should lead to a fall in the maximum rate of profit, as happens in Figure 7.2. Thus in the case where technical progress is entirely in the form of an increased division of labour you will have an increase in labour productivity for a given R, while in the case of pure capital-saving technical progress, R increases for a given labour productivity. Moreover, in many cases of innovation the whole wage-profit curve will shift to the right, and even taking the wage share and thus the rate of exploitation as given, the rate of profit increases.

Summing up, since technical change creates openings for an increase in real wages, we are no longer forced to suppose that the only scenario compatible with the reproduction of the capitalist economy over time is that of a decreasing trend in the wage share.[8] Indeed the number of possible 'regimes of accumulation' is (and always was) greater than Marx supposed. There may be, as Marx suggested, an *extensive* regime of accumulation characterised by a tendency for *absolute* surplus-value to increase, and thus by a combination of low wages, low productivity and long working hours. But the *intensive* regime of accumulation rooted in the search for *relative* surplus-value will be able to assume different configurations according to the prevailing forms of technical progress and the historically changing relative bargaining position of the workers. Of course, a classification of these different regimes requires applied analysis, such as the studies which have identified different social structures of accumulation or different regimes of regulation in the course of capitalist development (see, e.g., Boyer, 1990).

Though the variety of possible regimes of accumulation is greater than that traceable in Marx, the greater part of his analysis on the determinants of the subsistence wage and the factors affecting the workers' bargaining position remains valid, and can give us a consistent wage theory. Its relevance to an analysis of the determinants of distribution in modern capitalism rests, however, on two points which need further investigation.

The first relates to the fact that in a fiat money economy a change in money wages does not necessarily bring about a change in real wages as in the gold money economy which was considered by Marx. While in this latter case, given the methods of production, any rise in the gold money wages will actually bring about a fall in the rate of profits,[9] in the former case money wages do not determine the ratio of prices to money wages, since, given the nominal interest rate, their increase will determine a proportional increase in the price level.[10]

The second point concerns the mechanism Marx introduced in order for the process of capitalist reproduction to remove 'the very obstacles that it temporarily creates', especially with respect to the determinants of capital accumulation. The issue here is that no definite relation can indeed be advanced between the normal profit rate and investment spending, which seems to be influenced mainly by the level and changes of aggregate demand, technological innovations and political factors.

7.7 Money wages, real wages and the trend of capital accumulation

In conclusion I shall briefly seek, with reference to the latter two points, to evaluate the relevance of Marx's wage theory with respect to modern capitalist societies more satisfactorily.

As far as the second point is concerned, the idea that the rate of accumulation will fall as a consequence of a decrease in the rate of profit is the central feature of the modern Marxian profit squeeze theory (see, e.g., Duménil and Lévy, 1993; Shaikh, 1989; Goodwin, 1967). Marx, however, did not posit any necessary mechanical link between those two variables. Although he stressed that profit is the 'motive power of capitalist production' (Marx, 1867–94, III, p. 254), he also observed that 'in spite of the falling rate of profit the inducements and faculties to accumulate are augmented' (Marx, 1867–94, III, p. 260). Furthermore, Marx considered that overproduction could be solved by shrinking productive capacity (see, e.g., Marx, 1867–94, III, pp. 247–8;

and Marx, 1862–3, II, pp. 495–6), and hence that the actual trend of capital accumulation would probably be adversely affected by a fall in real wages.

Following this suggestion by Marx, as with the Keynesian premise of investment as an independent variable, the pace of accumulation would not appear to be determined by the saving rate and the rate of profit as argued in many Marxian or classical-Harrodian models on the grounds of the specific assumption of balanced growth and (consequently) of a capital–output ratio continuously equal to its normal or desired level. In actual fact, a redistribution of income to wages would increase consumption and have an uncertain effect on other components of effective demand (for instance, on exports, according to the effects on prices of the workers' claims for higher real wages). Moreover, the amount of non-residential investment seems to be influenced not so much (or not directly) by the rate of interest moving in the same direction as normal profit, but by the level and rate of growth of effective demand.[11] So, if demand increases thanks to a redistribution of income to wages, firms will be induced to expand their productive capacity, which will tend to (re-)establish *normal* profitability, irrespective of whether the latter happens to be high or low.[12] Hence, on average, not a fall, but an increase, in the *actual* rate of accumulation would occur.

Now, as far as the theory of distribution is concerned, the dependence of the rate of growth on the trends of the components of effective demand, which will shape the rate of change of income and productive capacity (see Palumbo and Trezzini, 2003), only strengthens Marx's rejection of any mechanical or natural or iron law of wages. In fact, the weight of keeping changes in wages 'within limits' capable of leaving 'intact the foundations of the capitalistic system' (Marx, *Capital*, I, p. 620) will now primarily fall, as well as on labour-saving technical progress, on the intervention of the state, directly by law in the labour market, or indirectly, by increasing the reserve army of labour through restrictive monetary and fiscal policies.[13]

But can the classical theory of wages developed by Marx be used to explain income distribution in the context of advanced capitalism, where the wage rate is probably above the subsistence level and class conflict acting on money wages would not necessarily determine a corresponding change in the real wages as in a gold money economy? In these circumstances, would not the surplus wage appear as a residuum after firms have fixed their prices on the basis of a mark-up added to their money prime costs?

As Dobb (1973) has observed, since the classical theory admitted the possibility of wages sharing in the net, or surplus, product (which is implicit in a notion of subsistence as historically determined), this possibility *by itself* does not necessarily contradict the wage-bargaining theory of Smith or Marx, as is sometimes argued on the grounds of a different theory of distribution (see Kaldor, 1956), or when analysing Sraffa's suggestion of taking the rate of profit as the independent variable in the price system (see, e.g., Goodwin, 1986).[14] Nor does there seem to be any need to disregard that wage bargaining theory when considering a fiat money economy, since a full-cost pricing rule appears to be compatible with different theories of distribution, provided that interest is included in the *normal* money costs of production, and firms equalise prices to those costs under the action of competition (see Garegnani, 1979; Panico, 1980; Pivetti, 1991). In fact, an increase in money wages could bring about a rise in the real wages, since prices initially adjust to the historical costs of capital (cf., for instance, Nordhaus, 1974) and the *real* rate of interest (that is, the opportunity cost of any capital invested in production) will happen to be lower than the initial given *nominal* rate of interest (cf. Pivetti, 1991; Stirati, 2001). If then the workers obtain *continuous* increases in their money wages, they will be assured a permanent increase in the real wage rate, provided that the monetary authorities leave the nominal interest rate on long-term riskless financial assets unchanged.

The key question, then, is what actually sets the *real* mark-up on prices. Taking the mark-up as given as in the Kaleckian tradition, on the grounds of a degree of monopoly determined by barriers to entry, the elasticities of demand and so on, leaves open the question of what happens to profits in the case of free competition, and also of how the average mark-up is arrived at when taking input–output transactions among sectors into account (cf. Pivetti, 1991; Steedman, 1992). However, given the normal profits of enterprises, it is possible that the real mark-up is directly fixed by the monetary authorities if they offset any increase in the rate of price inflation by an appropriate increase in the level of the money rate of interest.[15] But their ability in this respect would of course be greater if the trend of money wages remained unchanged over time (see Appendix A), while the pursuit of a *real* target by the monetary authorities could be in conflict with other objectives and constraints of monetary policies.[16] For instance, the reaction of trade unions to a stagnation of real wages could well lead to a lower real mark-up on prices as central banks would eventually be forced to accept it in order to stop a price spiral eroding the real value of financial assets and worsening the

balance of trade. Besides, lower interest rates might be preferred since they reduce the cost of servicing the public debt and pursing expansionary fiscal policies.

While in this framework inflation will be the result of incompatible claims on distribution which would manifest themselves through changes in money wages and in the nominal rates of interest, the actual mark-up might be seen as the final result of the whole process that determines the distribution of income, including the mechanisms and feedbacks that reconcile those claims. Income distribution will thus ultimately depend on the relative bargaining strength of the parties involved, and Marx's analysis of the elements shaping money wage trends will play a crucial role in this respect.

Appendix A

To better grasp the interactions between monetary policies and the process of wage bargaining, consider a price system of dimension n at period t = 0

$$_{n}\mathbf{p}_{1o} = {}_{n}\mathbf{A}_{n}\ {}_{n}\mathbf{p}_{1o}\,(1 + i)\,(1 + np) + {}_{n}\boldsymbol{\ell}_{1}\,w_{o}$$

where i is the rate of interest, np are the normal profits of enterprise[17] and w_o is the money wage at t = 0.

Given the money wages (and the methods of production), since $(1 + r) = (1 + i)(1 + np)$, where r is the rate of profits, if the rate of interest i increases[18] you will have

$$\mathrm{d}r = (1 + np)\mathrm{d}i$$

and the real wage rate will fall due to the increase in the prices of the n commodities. If the rise in the price–wage ratio is only slight, probably no reaction by the workers will be set in motion.[19]

Suppose now that the money wages increase by $(1 + \gamma)^{\beta}$ per cent a year, where β is the number of wage round settlements in a year and γ is the percentage increase in money wages in any round of wage bargaining. If (for the sake of simplicity) $\beta = 1$, then

$$w_1 = w_o\,(1 + \gamma)$$

If firms adjust prices to the historical cost of capital (that is, if they do not fully *anticipate* the increase in the prices of capital goods), then in period t = 1 you will have

$$\mathbf{p}_1 = \mathbf{A}\mathbf{p}_o\,(1 + i_n)(1 + np) + \boldsymbol{\ell}w_o(1 + \gamma)$$

where i_n is the nominal rate of interest, and $\mathbf{p}_o$ and $\mathbf{p}_1$ are the vectors of prices at t = 0 and t = 1. If *Qo* and *Wo* are respectively the value of gross product and the total amount of money wages at time t = 0, and $\mathbf{s} = \mathbf{1}'_n$ is the sum vector (a row vector of dimension n), we will thus have

$$\frac{\mathbf{sp}_1}{\mathbf{sp}_0} = \frac{\mathbf{sAp}_o\ (1+i)(1+np)+(1+\gamma)w_o\mathbf{s}\ell}{\mathbf{sAp}_0\ (1+i)(1+np)+w_o\mathbf{s}\ell} = (1-Wo/Qo)+(1+\gamma)(Wo/Qo) = 1+\gamma(Wo/Qo)$$

that is $\Delta p/p = \gamma(Wo/Qo)$, where *Wo/Qo* is the wage share in the gross product at t = 0. Hence the inflation rate $\Delta p/p$ will be greater, the greater (for a given γ) the initial 'wage share' (see also Stirati, 2001), or the greater (for a given ratio *Wo*/*Qo*) the rate of change in money wages γ (see Figure 7.3).

The increase in money wages will thus determine at t = 1 an increase in the real wages w_r equal to

$$dw_r/w_r = dw/w - dp/p = \gamma - \gamma(Wo/Qo) = \gamma(1-Wo/Qo) > 0$$

which is greater, the lower *Wo/Qo*, since the weight of the *unanticipated* increase in the price of capital goods will be greater. On average this change in income distribution will be 're-absorbed' if no other change in money wages occurs. However, even at t = 1 the change in income distribution will not happen if firms instantaneously adjust prices to the reproduction costs of capital, or if the monetary authorities change the nominal rate of interest satisfying the relation

$$(1 + i_n)(1 + np) = (1 + i_r)(1 + \gamma)(1 + np)$$

which implies

$$\mathbf{p}_1/w_1 = [\mathbf{Ap}_o\,(1 + i_r)(1 + np)(1 + \gamma) + \ell\ w_o(1 + \gamma)]/w_o(1 + \gamma) = \mathbf{p}_o/w_o$$

namely, that $dp/p = dw/w = \gamma$.

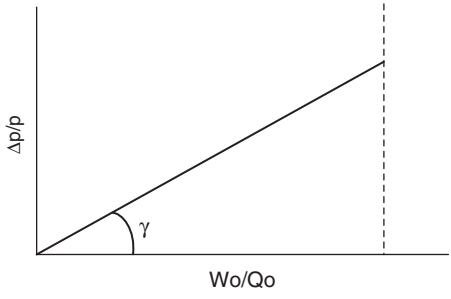


Figure 7.3 The price inflation for any given initial wage share

Consider now a repeated increase in money wages by γ per cent per year until t = j. In this case no change in distribution will happen only if

$$\mathbf{p}_j = \mathbf{p}_o\,(1+\gamma)^{\mathrm{j}} = \mathbf{A}\mathbf{p}_o(1+\gamma)^{(\mathrm{j}-1)}(1+i_r)(1+np)(1+\gamma) + \ell\,\mathrm{w}_o(1+\gamma)^{\mathrm{j}}$$

that is, again, if $(1+i_n) = (1+i_r)(1+\gamma)$. But if the real wage rate corresponding to the rate of profits $[i_r + np(1+i_r)]$ is not equal to the real wage w_r^w aimed at by the workers according to their relative strength in wage bargaining,[20] this can lead to an increase in money wages by ϕ per cent greater than the rate γ *expected* and *anticipated* by the workers, that is to

$$\mathrm{d}w/w = \Omega(w_r^w - w_r) + \gamma = \phi.$$

It will again bring about a change in distribution[21] unless the monetary authorities do not make a further upward adjustment in the nominal rate of interest in order to maintain a desired real rate of interest i^T_r, that is, unless

$$i_n = \phi + (1+\phi)i_r^T.$$

The discrepancy between the real wage rate w_r^T corresponding to the real rate of interest targeted by the monetary authorities, and the real wage rate w_r^w aimed at by the workers, may thus bring about a wage–price spiral, which may only be attenuated by a fall in the absolute value (for given differences between sectors) of the normal profits of enterprise ***np***,[22] or by an increase in productivity and/or an improvement in the terms of trade and the exchange rate (especially if the share of imported goods in the value of gross product is substantial) – since these changes will lead to an increase in w_r^T for a given target real rate of interest. However, on average, in normal conditions, an adjustment of w_r^T and w_r^w to the same value will be achieved and be driven by the fact that monetary policy is not set in a vacuum, and that the price trend is one of the elements affecting the decisions of the monetary authorities.[23] Moreover, anti-inflationary fiscal and monetary policies may affect the bargaining position of the workers by increasing the rate of unemployment, thus adding it to price inflation as a means of limiting the increase in real wages.

A further analysis of the process of adjustment to a normal position of the economy is beyond the scope of the present work, and would require specifying not only the (sometimes conflicting) objects, constraints and channels of the monetary policy, and its indirect effect on variables (like the unemployment rate) which affect the bargaining position of the workers, but more generally, the interactions between income distribution and the trend in effective demand. Here it is merely noted that the signs and weights of these interactions may change according to the circumstances, and be influenced by institutional and social factors.

Notes

1. Sections 7.2–7.6 are an abridged and revised version of the paper 'Marx on Absolute and Relative Wages and the Modern Theory of Distribution' published in the *Review of Political Economy*, January 2013, while Section 7.7 and Appendix A are new and integrate what was argued in that paper on the relation between money wages and real wages. I thank an anonymous referee for his comments and suggestions. I also thank Francis & Taylor for permission to reprint Figures 7.1 and 7.2.

2. Hollander (2008) makes precisely this error when he argues that we find in Marx a *physiological* subsistence wage compatible with zero population growth, and a different subsistence wage for any required increase in population equal to that of capital. The Canonical model, on the other hand, has been criticised even with respect to its ability to interpret the wage theories of Smith and Ricardo, on the grounds that, in them, the adjustment of population to capital growth does not imply full employment, and no univocal relation between the wage rate and capital and population growth rates can actually be traced in their works (cf., for instance, Garegnani, 1990).
3. Note that this does not preclude that circumstances may arise that can reduce subsistence itself, as when there is a large-scale immigration of workers accustomed to a lower standard of living, or the economy falls into a retrograde state. On this point and the classical wage theory, see Levrero (2012).
4. For instance, if only corn is the wage good, it may happen that the wage rate in terms of corn increases, but the quantity of labour V embodied in the amount of corn given to the labourers decreases, if labour productivity in the production of corn rises more than the corn wage rate. In this case the rate of surplus value S/V (where S is the surplus labour) will rise, and the wage share V/L = [1/ (1 + S/V)] will fall (where the living labour L is the value of the 'modern' net product in terms of embodied labour). Thus, if 100 tons of corn is now produced by 1.5 workers instead of 2, and each worker continues to receive 25 tons of corn, the wage share will pass from 50% to 37.5% (which is also equal to the wage rate measured in embodied labour). The value of wages remains unchanged only if the wage rate in terms of corn changes to 33.3 tons of corn, with an increase equal to that in labour productivity.
5. Note that Marx assumes that technical progress will lead to a fall in the maximum rate of profit *R*, that is, of the ratio of living to dead labour, or $(S + V)/C$, where *S* is the surplus-value, *V* the variable capital and *C* the constant capital. Then, since $C/V = (1 + s)/R$, where *s* is the rate of surplus value, an increase in *s* and a decrease in *R* must necessarily be associated in Marx with an increase in *C/V*.
6. This case resembles the mechanisation of production as defined by Schefold (1976), which in turn is similar to technical progress as considered by Marx when arguing a fall in the maximum rate of profit.
7. We are assuming here that, as in Marx, both $1/R$ and the organic composition of capital $\omega = K/w_1L = \text{tg}ß/r_1Z = 1/r_1T$ (where $ß$ is the angle formed by the segments w_1Z and ZW^*, while T is the point on the *r* axis determined by its intersection with the line γ) increase with the introduction of the new technique. However, unlike what was advanced by Marx, the new *dominant* technique might be characterised by both a lower labour productivity and a lower maximum rate of profit. In this case, since with the new technique both the rate of profit and the wage share increase, an unlikely fall in the wage rate (and a further increase in r) would have to occur to keep the wage share unchanged.
8. It is worth noting here that Sraffa viewed Marx's arguments on the law of the falling rate of profit as referring to cases in which no proper technical progress (i.e., inventions) is taking place (see Gehrke and Kurz, 2006).
9. Of course it holds in the case of Marx's reference to the labour theory of value according to which there will be a strict proportionality between the changes in money and real wages when taking as given the amounts of labour embodied in the money commodity and in the wage basket. But it holds also if the labour theory of value does not hold. In this case the price level may change

when the money wage changes. Nevertheless, the change in 'money' wages will always lead to a variation in the same direction in real wages.

10. Conversely, in a gold money economy, a substantial and persistent change in the price level may occur only if a change occurs in the technical conditions of production of 'gold' or in those of the other commodities. In this respect it is also worth noting that, following Steuart, Smith and Tooke, Marx criticised the quantity theory of money, maintaining that, for a given value of 'gold', it is the mass of circulating money (or alternatively, its velocity of circulation) which will rise if prices rise, and not vice versa (see for instance Marx, 1867–94, I, pp. 117–22).
11. It does not overlook the fact that a fall in the profits rate might lead to a 'strike of capital' if that rate falls below some minimum level or below that prevailing in other countries, or if that fall is accompanied by a general loss of power in society.
12. It also explains (see Garegnani, 1992; and Vianello, 1985) why the *normal* rate of profit is not determined by the rate of accumulation as suggested in the Post-Keynesian models: an increase in the *actual* pace of accumulation will be able to be 'financed' by an increase in output per unit of capital and in the amount of productive capacity, given the wage rate and the methods of production.
13. On the other hand, the dependence of the rate of growth on effective demand does not imply the rejection of Marx's crucial notion of an inverse relationship between wages and profits. The idea that room exists *in the long run* for co-operation between capital and labour thanks to changes in the average realised rate of profit brought about by changes in the average degree of capacity utilisation (see, e.g., Amadeo, 1986; Cassetti, 2005; Dutt, 1990) does not take into account that a permanent change in effective demand is not necessarily unexpected, that gross investment is guided by the rate of profit expected to be earned on the *new* installed productive capacity (and thus by the *normal* rate of profit corresponding to a normal degree of capacity utilisation), and that the desired degree of capacity utilisation is usually shaped by the experience of many cycles, thus changing slowly over time.
14. The fact that, when the wage rate increases above subsistence level, it would end up being measured by an abstract standard of value does not negate the possibility of taking as given the wage rate in the price system if the forces of distribution are seen as acting primarily on the wage rate. Since in fact changes in distribution are never drastic, we can always divide the wage rate into its subsistence and surplus components, and measure the surplus wage by the bundle of commodities corresponding to the standard consumption of the workers in the initial situation. The composition of this bundle of commodities will arise both from the hierarchies of needs and the distinction between necessary and superfluous goods (comprising Smith's luxuries of the poor) that are observable in any given initial situation, as well as from the processes of imitation of the consumption of the upper classes which are characteristic of consumption behaviour. See Levrero (2000) for further considerations on this point.
15. Given the target *real* interest rate r^T and the expected rate of inflation p^a, the money interest rate i^* should thus be such that $(1 + i^*) = (1 + r^T)(1 + p^a)$. This implies that, if the actual inflation rate p is equal to the expected one, and $i = i^*$, then the real rate of interest $r = (i^* - p)/(1 + p) = r^T$. Of course, if the workers' target real wages happen to be incompatible with the target rate r^T, then a change

in the inflation rate will be set up if no change occurs in the normal profits of enterprise or the technical conditions of production. Note that a different money interest rate i^* should be viewed in the light of the price index chosen by the monetary authorities to calculate the (actual and expected) inflation rate.

16. It is significant in this respect that the variance in the *real* rate of interest is usually greater than that in the *nominal* rate (see Moore, 1988, p. 257). It should also be noted that, if the subsistence wage is included in the methods of production, the maximum rate of profits R may be viewed as a constraint on monetary policy, since, if the interest rate leads to $r > R$, an inflation barrier will be set up.
17. For the sake of simplicity I consider *np* as uniform among industries and as a percentage of gross interest. The condition $np = u/(1 + i)$ could be put forward to ensure that it will be equal to that calculated by the percentage u on the anticipated capital alone. Note that, if different profits of enterprise among industries are considered, the price vector must be multiplied by the row vector of dimension n $(\mathbf{1} + \boldsymbol{np})$, thus taking into account the different 'risk and trouble' involved in investing in the various industries, as well as the presence of monopoly elements, if any.
18. Indicating as i^{BC} the interest rate on long-term riskless financial assets as influenced by the policy of the Central Bank, we may put $i = hi^{BC}$, where $h \geq 1$ is a parameter set by the banking sector.
19. In the case of a decrease in the rate of interest, it can be usuful to set the condition $dw/w = \max(\theta dp/p; 0)$, where θ is a parameter which reflects the strength of the workers in wage bargaining. Note that, contrary to Kalecki and Eichner, a change in the interest rate here affects the rate of profit, and not the normal profits of enterprise. Moreover, contrary to many Post-Keynesian or Kaleckian models (see Lavoie, 1995), it affects the rate of profits directly, not indirectly through the effect, if any, of the change in the rate of interest on the rate of accumulation.
20. You could consider, for instance, that $w_r^w = \varepsilon_0 + \varepsilon_1 U + \varepsilon_2 n^{TU} + \varepsilon_3 \pi$, where U indicates the conditions of the labour market (for instance, the rate of unemployment), n^{TU} the share in the total labour force of the employment in the unionised and leading sectors in wage bargaining, ε_0 a parameter summing up the complex social and political factors that can affect the bargaining position of the workers, and $\varepsilon_3\pi$ the share of productivity growth aimed at by the workers (expressed, for instance, by a fraction of labour productivity).
21. In this respect the timing of wage bargaining, wage indexation, as well as the rapidity of adjustment of prices by the firms, and of the *nominal* rate of interest by the monetary authorities, will influence the actual course of the real wage rate. In actual fact, however, the ability of the monetary authorities to change the nominal interest rates usually proves to be greater than that of the workers to change the money wages.
22. Of course it involves the relations between industrial and financial capital. On these relations (and their conflictual nature especially in time of crisis, or when the wage rate happens to be at subsistence level) see Argitis (2001), who also reconstructs the analysis of such a conflict by Marx, from whom the suggestion might be taken that if a change occurs in the aliquot part of the rate of profits constituted by the normal profits of enterprise (a share, as noted by Massie, often fixed by custom) there may be pressure to restore it,

also by impinging on those possible components of this part of the profits which do not merely reflect the 'risk and trouble' of employing capital productively – like the 'wages of management', which, according to Marx (see, e.g., 1867–94, III, pp. 377–82), have been fictitiously separated from the profits of enterprise since the rise of the joint-stock company.

23. For instance, monetary policy is influenced by international capital movements, which are led by the difference between the nominal rates of interest and expected variations in the exchange rate.

References

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

Argitis, G. (2001) 'Intra-capitalist conflicts, monetary policy and income distribution', *Review of Political Economy*, 13(4): 453–70.

Boyer, R. (1990) *The Regulation School. A Critical Appraisal* (New York: Columbia University Press).

Cassetti, M. (2006) 'A note on the long-run behaviour of Kaleckian models', *Review of Political Economy*, 18(4): 497–508.

Cottrell, A. and Darity, W.A. Jr (1988) 'Marx, Malthus and wages', *History of Political Economy*, 20(2): 17–36.

de Vivo, G. (1982) 'Notes on Marx's critique of Ricardo', *Contributions to Political Economy*, I: 87–99.

Dobb, M. (1973) *Theories of Value and Distribution since Adam Smith – Ideology and Economic Theory* (Cambridge: Cambridge University Press).

Duménil, G. and Lévy, D. (1993) *The Economics of the Profit Rate* (Aldershot: Edward Elgar).

Dutt, A.K. (1990) *Growth, Distribution and Uneven Development* (Cambridge: Cambridge University Press).

Garegnani, P. (1979) Notes on consumption, investment and effective demand: II, *Cambridge Journal of Economics*, 3(1): 63–82.

Garegnani, P. (1990) 'Sraffa: Classical versus Marginalist Analysis', in K. Bharadwaj and B. Schefold (eds), *Essays on Piero Sraffa* (London: Allen & Unwin).

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).

Gehrke, C. and Kurz, H.D. (2006) 'Sraffa on von Bortkiewicz. Reconstructing the classical theory of value and distribution', *History of Political Economy*, 38(1): 91–149.

Goodwin, R.M. (1967) 'A Growth Cycle', in R.M. Goodwin, *Essays in Economic Dynamics* (London: Macmillan).

Goodwin, R.M. (1986) 'Swinging along the turnpike with von Neumann and Sraffa', *Cambridge Journal of Economics*, 10: 203–10.

Hollander, S. (2008) *The Economics of Karl Marx: Analysis and Application* (New York: Cambridge University Press).

Kaldor, N. (1956) 'Alternative theories of distribution', *Review of Economic Studies*, 7: 1–27.

Kurz, H.D. and Salvadori, N. (1995) *Theory of Production. A Long-period Analysis* (Cambridge: Cambridge University Press).

Lavoie, M. (1995) 'Interest rates in post-Keynesian models of growth and distribution', *Metroeconomica*, 46: 146–77.
Levrero, E.S. (2000) 'Crescita e Distribuzione in Von Neumann e l'Analisi di Sraffa', in M. Pivetti (ed.) *Piero Sraffa. Contributi per una Biografia Intellettuale* (Roma: Carocci).
Levrero, E.S. (2011) 'Some Notes on Wages and Competition in the Labour Market', in R. Ciccone, C. Gehrke and G. Mongiovi (eds), *Sraffa and Modern Economics* (London: Routledge).
Levrero, E.S. (2012) *Four Lessons on Wages and the Labour Market* (Roma: Aracne).
Marx, K. (1961–3 [1867–94]) *Capital. A Critique of Political Economy*, Vols I–III (Moscow: Foreign Languages Publishing House).
Marx, K. (1975 [1865]) *Wages, Price and Profit* (Peking: Foreign Languages Press).
Marx, K. (1978a [1862–3]) *Theories of Surplus Value*, Vols I–III (Moscow: Progress Publisher).
Marx, K. (1978b [1884]) *Wage, Labour and Capital* (Peking: Foreign Language Press).
Meek, R.L. (1967) *Economics and Ideology and Other Essays* (London: Chapman & Hall).
Moore, B.J. (1988) *Horizontalists and Verticalists. The Macroeconomics of Credit Money* (Cambridge: Cambridge University Press).
Nordhaus, W.D. (1974) 'The Falling Share of Profits', *Brookings Papers on Economic Activity*, 1974: 169–208.
Palumbo, A. and Trezzini, A. (2003) 'Growth without normal capacity utilisation', *European Journal of History of Economic Thought*, 10(1): 109–35.
Panico, C. (1988) *Interest and Profit in the Theories of Value and Distribution* (London: Macmillan).
Pivetti, M. (1991) *An Essay on Money and Distribution* (London: Macmillan).
Ricardo, D. (1951–73) *The Works and Correspondence of David Ricardo* (Cambridge: Cambridge University Press).
Rosenberg, N. (1989) *Inside the Black Box* (Cambridge: Cambridge University Press).
Rowthorn, B. (1980) Marx's Theory of Wages', in B. Rowthorn, *Capitalism, Conflict and Inflation. Essays in Political Economy* (London: Lawrence & Wishart).
Samuelson, P.A. (1978) The canonical classical model of political economy, *Journal of Economic Literature*, 16(4): 1415–34.
Schefold, B. (1976) 'Different forms of technical progress', *Economic Journal*, 86: 806–19.
Shaikh, A. (1991) 'Wandering around the Warranted Path: Dynamic Nonlinear Solutions to the Harrodian Knife-Edge', in E. Nell and W. Semmler (eds), *Kaldor and Mainstream Economics: Confrontation or Convergence (Festschrift for Nicholas Kaldor)* (London: Macmillan).
Smith, A. (1976 [1776]) *An Inquiry into the Nature and Causes of the Wealth of Nations* (Chicago: University of Chicago Press).
Sraffa, P. (1960) *Production of Commodities by Means of Commodities* (Cambridge: Cambridge University Press).
Steedman, I. (1992) 'Questions for Kaleckians', *Review of Political Economy*, 2: 125–51.
Stirati, A. (2001) 'Inflation, unemployment and hysteresis: an alternative view', *Review of Political Economy*, 14(4): 427–51.
Torrens, R. (1815) *An Essay on the External Corn Trade* (London: J. Hatchard).
Vianello, F. (1985) 'The pace of accumulation', *Political Economy*, I(1): 69–87.

8

On Advanced Capitalism and the Determinants of the Change in Income Distribution: A Classical Interpretation

*Massimo Pivetti**

8.1 Changes in distribution and the question of technological change

It is useful to start our analysis with the question of technological change, since it is in these terms that the general increase in inequality experienced by advanced capitalism over the last 30 years has, by consensus, been explained.

Technological change, though paramount in the dominant theoretical approach to distribution in terms of the relative scarcity of factors, also plays a significant role in the alternative classical surplus approach. In the neoclassical approach, capital is kept relatively scarce by labour-saving technical changes, so that, in spite of its rapid accumulation over time by the 'capitalist saver', a positive profit on it continues to be obtained. In the words of Wicksell, 'the capitalist saver is thus, fundamentally, the friend of labour, though the technical inventor is not infrequently its enemy' (Wicksell, 1901, p. 164).[1] In the alternative approach, which considers the parties' relative strength, technical changes, induced primarily by rises in wages, contribute to checking changes in distribution in wage earners' favour through technological unemployment and the continuous re-creation of an 'industrial reserve army'. In Marx's words, 'on the one hand, the additional capital formed

*University of Rome La Sapienza. A preliminary version of this chapter has been presented also at the Conference 'Crisis de la teoría ecónomica y políticas alternativas ante la crisis global', held at the Universidad Nacional Autónoma de México, Facultad de Economía, 7–11 November 2011. I wish to thank A. Barba, I. Perrotini Hernandez, F. Serrano and other participants in that conference for comments. I wish also to thank an anonymous referee for helpful suggestions.

in the course of accumulation attracts fewer and fewer labourers in proportion to its magnitude. On the other hand, the old capital periodically reproduced with change of composition, repels more and more labourers formerly employed by it' (Marx, 1887, p. 589).

The fact is, however, that neither for the USA nor for Europe can an *acceleration* of technological change be clearly detected over the last 30 years, whether in terms of changes in labour productivity or any other possible measure of technical change.[2] This fact of itself makes it rather difficult to account along neoclassical lines for the marked shift in distribution experienced by advanced capitalism, unless one is prepared to exclusively stress the increase that has occurred in the rate of immigration as the ultimate (and easily measurable) source of the postulated change in the relative scarcity of factors. The relative scarcity approach to distribution has no difficulty in principle, of course, in acknowledging that immigration, by increasing the supply of labour, tends to depress wage rates (cf., for example, Samuelson, 1964, p. 552). But the unconditional and faith-based approval with which neoclassical economics in general looks at all aspects of globalisation has produced a substantial amount of literature over the last 20 years or so tending to deny that immigration actually has significant distributional effects on the destination economies.[3]

The alternative perspective, on the other hand, sees immigration, though important, as simply one among several other important determinants of a serious breakdown of wage earners' bargaining power, a breakdown that does not need any acceleration of technological change to be fully accounted for.

8.2 The increase in wage inequality and the phenomenon of the 'working rich'

Before considering these determinants, however, it is worth pointing out that the consensus explanation that the increase in inequality is due to an acceleration of technological change has been developed in the USA not so much with respect to the change in the distribution of income between wages and profits, but rather with respect to the increase in *wage* inequalities.[4]

According to a recent report by the US Congressional Budget Office (CBO), it is largely thanks to the increased concentration of labour income that the share of after-tax household income for the 1 per cent of the population with the highest income has more than doubled since the end of the 1970s, rising from nearly 8 per cent in 1979 to 17 per cent in 2007.

The share of the highest income quintile grew from 43 per cent of after-tax household income in 1979 to 53 per cent in 2007 – that is, the after-tax income of the richest fifth of the population now exceeds the income of the other four-fifths. The population in the lowest income quintile received about 5 per cent of after-tax income in 2007, down from about 7 per cent in 1979, while the middle three income quintiles all saw their shares of after-tax income decline by 2 or 3 percentage points between 1979 and 2007. The CBO report points out that the equalising effect of federal taxes was smaller in 2007 than in 1979, as 'the composition of federal revenues shifted away from progressive income taxes to less-progressive payroll taxes' (Congressional Budget Office, 2011, pp. 4–5).

Now, as a result of the surge in top wage incomes, the latter have ended up constituting a larger fraction than in the past *also* of top incomes in their entirety.[5] Attention has thus been focused in the USA on the 'working rich' – essentially top management – whose income has risen so astonishingly over the last three decades. According to the consensus explanation, an unprecedented wave of 'skill-biased' technological change has rapidly raised the earnings of individuals with more skills, as measured, for example, by education.[6]

Here, again, an explanation for the increase in wage inequalities along these lines would require some independent and unambiguous evidence that 'skill-biased' technological change has actually accelerated markedly in the last three decades relative to earlier periods. But such evidence is simply not there,[7] and it appears most likely that the idea of a marked acceleration of skill-biased technical change in the period since 1980 has been *inferred* from the very observation of the marked increase which occurred in top wage shares. On this, Piketty and Saez (2003) have pointed out that the increase in top compensation in the United States can hardly be the consequence of technical change, both because the increase is very large and concentrated among the highest income earners, and because such a huge increase in top wage shares 'has not taken place in most European countries which experienced the same technical change as the United States'. They suggest an interpretation of the phenomenon of high compensation at the top in terms of non-market mechanisms such as 'social norms regarding inequality and the acceptability of very high wages'. According to this interpretation, social, fiscal and union pressure to contain the fast growth of top compensations observed over the past 30 years would have been significantly reduced, which would have greatly enhanced the high wages earners' ability to increase their compensations above pre-war levels in relative terms (cf. Piketty and Saez, 2003, pp. 34–5).

8.3 Ultimate determinants versus channels of changes in distribution

In the classical surplus approach to distribution, the question of the level of top-management compensation is part of the more general question of the actual channels through which the parties' relative strengths act and changes in distribution are brought about. Within the framework of the 'monetary' explanation of distribution, for example, it is maintained that interest rate decisions are taken according to a wide range of policy objectives and constraints, among which the level of real wages is regarded as especially relevant (see on this Pivetti, 1991, pp. 28–30). If profits of enterprise (business income) in a rapidly expanding sector of the economy – the financial services industry – are kept abnormally high by the presence of monopoly elements,[8] and a shift in social norms regarding the acceptability of very high compensation has resulted in overly generous remunerations for top management which contribute to increasing the price level/money wage ratio in the economy, then a compensatory effect on this ratio will eventually have to be sought through lower interest rates, so as to ensure the majority of the working population standards of living considered indispensable for social stability, as well as to sustain activity levels.

Consider the following expressions for value added per unit of labour (1), for the part of value added accruing to gross profits and top-management remuneration (2), and for the gross profit margin, here defined as the ratio of value added per unit of labour to the money wage rate (3):

$$(1)\quad p \cdot a = w + k(i + r_e) + m_r$$

$$(2)\quad p \cdot a - w = k(i + r_e) + m_r$$

$$(3)\quad \frac{p \,.\, a}{w} = 1 + \frac{k(i + r_e) + m_r}{w}$$

where p is the unit price of output (a composite commodity representative of the gross output of a closed economy), a is output per unit of labour, w is the money wage rate, k is capital per unit of labour, i is gross pure profit per unit of capital (comprehensive, that is to say, of depreciation per unit of capital), r_e is profit of enterprise per unit of capital, and m_r is top-management compensation per unit of labour. In (3), the gross profit margin $(1 + \mu = \frac{p \,.\, a}{w})$ includes, besides the gross remuneration of capital and profit of enterprise, also top-management compensation – a magnitude assumed to be given in absolute terms, independently of

the amount of capital employed in production. Now, given w and a, by lowering interest rates, a policy-controlled variable, it is possible to reduce ki, thereby checking the rise in $\frac{p \cdot a}{w}$ and the consequent fall of the real wage due to increasing levels of r_e and m_r. I am convinced that significant aspects of the US pre-crisis experience can be accounted for from this perspective. The Federal Reserve post-1995 policy of progressive lowering of interest rates, which on the one hand helped to protract for a few years the macroeconomic sustainability of a massive process of substitution of loans for wages (see on this Barba and Pivetti, 2009), did on the other actually succeed in also checking the fall of real wages.[9]

An explanation of distribution based on the parties' relative strengths requires, first, an analysis of the social-institutional channels through which they act – which should in turn enable the distributive variable on which they discharge themselves in the first place, in each concrete situation, to be singled out; and second, that the main factors upon which those relative strengths depend are selected and discussed. The monetary view of distribution, according to which the course of the normal rate of profit is governed by that of the long-term rate of interest, is of course connected with the former of these two questions, chiefly with its solution. But in what follows my attention will be focused on the ultimate determinants of the parties' relative strength – on the main causes, that is to say, of the breakdown over the last 30 years of wage earners' bargaining power throughout advanced capitalism, rather than on the actual channels through which the parties' relative strengths have acted and the changes in distribution have been brought about.

8.4 The general state of employment and a complex net of interconnections

The breakdown of wage earners' bargaining power over the last 30 years can be traced back to several factors: from higher unemployment levels and a declining union density, to rising 'flexibility' of labour markets and the almost complete liberalisation of trade as well as of labour and capital movements. The following is a tentative more detailed list of these factors. Without attempting any precise quantitative assessment, let us in Table 8.1 simply indicate the direction in which each one of them is widely acknowledged to have moved over the last 30 years.

Between these factors and wage earners' bargaining power a complex net of interconnections tends to establish itself. At its centre one may reasonably put the general state of employment – that is, employment

Table 8.1 Factors that have weakened wage earners' bargaining power since the early 1980s

Factor	Direction
Unemployment rates	↗
Privatisations	↗
Overall labour-market flexibility	↗
Restrictions of rights to strike	↗
Relative number of immigrant workers	↗
International capital mobility, interest rates/growth rates differentials and primary surpluses	↗
Trade from low-wage developing countries	↗
Direct investment flows (investment in plant and equipment) to low-wage developing countries	↗
Employment shift to lower-paying service-producing industries	↗
Incidence of full-time employment	↘
Rates of unionisation	↘
Minimum wages	↘
Indirect and deferred wages	↘
Overall progressivity of tax systems	↘

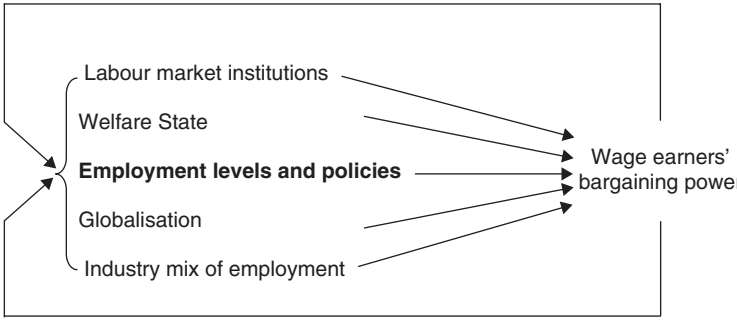


Figure 8.1 Network of interconnections formed by factors in Table 8.1 and wage earners' bargaining power

levels and policies. This net of interconnections is shown in Figure 8.1, in which feedbacks from wage earners' bargaining power to the state of employment and the other determinants of wage earners' bargaining power occur through the course of real wages, both direct and indirect wages. In Europe, as well as in the USA, household income of low- and middle-income households is almost entirely made up of wages and salaries, so that any change in their level affects correspondingly aggregate

demand and employment, thereby creating the conditions for further changes in wage earners' bargaining power and in income distribution.

As we shall see presently, there are linkages also between the factors appearing on the left-hand side of the figure – that is, between the state of employment and the welfare state, between the state of employment, globalisation and labour-market institutions, as well as between globalisation and the industry mix of employment. Many of these linkages have been acknowledged in the literature, one way or another. This is the case as regards the linkages between employment levels, unions' bargaining power and rates of unionisation, as well as those between employment levels, the incidence of collective bargaining arrangements and of full-time employment (see for example Salverda and Mayhew, 2009; Bosch *et al.*, 2010; Dew-Becker and Gordon, 2007, pp. 172–5). The linkages between globalisation and the industry mix of employment have also been widely acknowledged, together with those between globalisation and labour markets' flexibility. As to the former, attention has been drawn to the continuous shift over the last 30 years from goods-producing to lower-paying service industries, resulting from trade and deindustrialisation;[10] as to the latter, to the overall enhanced disposition on the part of workers to grant wage and other concessions to their employers due to the mere threat of direct foreign competition or of the relocation of part or all of a production facility (cf. Mishel *et al.*, 2009, pp. 186–95). Here I should like to draw the reader's attention to a few logical/historical aspects of the above interconnections which, while generally overlooked, seem to me especially relevant for a deeper understanding of the breakdown of wage earners' bargaining power throughout advanced capitalism.

8.4.1 Employment levels and policies, and labour-market institutions

A close connection can be detected, especially in the European context, between the abandonment over the last 30 years of demand management policies aimed at high employment, and the progressive relaxation of so-called labour-market rigidities: from collective bargaining arrangements and inclusive wage-setting institutions to employment protection legislation; from legal rules concerning hiring, dismissal and working hours to various forms of insurance against unemployment. The point is that such benign labour-market institutions are hardly viable and conceivable in the absence of policies persistently oriented towards full employment. Indeed, capitalism is ultimately incompatible with an advanced set of institutions designed to protect wage earners if there is no economic policy in place to neutralise the market's inability

to ensure adjustment of effective to potential output. Even in the face of rising demand, firms do not hire if they are not free to dismiss at will, nor to choose whom and how to hire, unless, in the light of actual experience, they can reasonably count upon a *stable* growth of the demand for their products and hence of their activity levels. Benign labour-market institutions and advanced employee protection legislation were possible and did spread in Europe for 30 years after World War II because employers were becoming aware that the primary objective of economic policy was the maintenance of high levels of employment, regarded by governments in their turn over those three decades as indispensable for social stability (see on this Pivetti, 2004).

8.4.2 Employment levels and policies, and the welfare state (indirect and deferred wages)

The historical record of advanced capitalism over the last 60 years also reveals the existence of significant connections between the general state of employment and that proportion of wage earners' standard of living which is determined by some relevant components of public spending (health, pensions, education, transport and residential construction) and by forms of taxation. Especially in Europe, the development of the welfare state was an outstanding aspect of 'incomes policies' – of the so-called social pact or exchange. It was a question of ensuring that, in spite of persistently low unemployment and strength relations that were consequently favourable to wage earners, rises in money wages occurred as much as possible in the medium to long run within the limits set by rises in productivity, so as not to jeopardise the profitability of private investment or hamper the competitiveness of domestic production. In practice, through the development of the welfare state, strong bargaining powers for wage earners resulted to a great extent in the expansion of collective services, rather than impinging directly upon business income and the profitability of capital. But as unemployment started rising at the end of the 1970s, and the bargaining power of the unions weakened, the preservation of a generous but fiscally onerous (for the wealthier sections of the population) welfare state was increasingly regarded as less necessary. It can therefore be affirmed, in the light of actual experience, that rises in unemployment that remain unchecked by policy, or a positively deflationary policy stance, not only tend to cause wages to increase systematically less than productivity, but are also accompanied by reductions in social spending and the overall progressivity of tax systems, as well as by increases in the incidence of levies on labour income. Naturally, the process tends to

feed upon itself. Both reduced tax progressivity and lean social spending exert a negative impact on aggregate demand and employment: reduced social spending is matched by more private spending on health, pensions, education, etc., so that household income available for other expenditure is correspondingly reduced.

8.4.3 Employment levels and policies, and globalisation

In the three decades following the Bretton Woods Agreements, great attention was devoted by the governments of the major capitalist countries, within the framework of their full-employment policies, to their economic and financial relationships with the rest of the world and to the balance of payments constraints on growth. In those years, national full-employment policies were represented to a significant extent by various forms of regulation of foreign trade, strict immigration policies, and, most importantly, capital controls. Indeed, in 1944 the International Monetary Fund (IMF) Agreement accorded to every member government the explicit right to control all capital movements, and contemplated even the possibility of *requiring* member countries using the resources of the Fund to exercise the control of the outflow of capital: 'If, after receiving such a request, a member fails to exercise appropriate controls, the Fund may declare the member ineligible to use the general resources of the Fund' (from Article VI of the Agreement, Section 1). Keynes could thus state before the House of Lords that what in the pre-war system, in the field of international capital movements, 'used to be a heresy, is now endorsed as orthodox' (Keynes, 1944, p.17).[11]

But having abandoned full employment as the primary policy objective at the end of the 1970s,[12] national economies were exposed to a strong acceleration of the internationalisation process, involving each of its main dimensions: trade, immigration and capital mobility. This acceleration has played a significant role in the change in distribution experienced by the USA and Europe over the last three decades.

Deindustrialisation brought about by the rise in trade from low-wage developing countries and increased investment in plant and equipment there has affected the bargaining power of wage earners, both in the USA and in Europe, through (i) its impact on the level and composition of output; (ii) the pressure exerted on money wages by decreasing international prices of an increasing number of consumption goods and their inputs; and (iii) the mere threat of production relocations – a threat rendered entirely credible by those that had already occurred, thanks to the new regime of full capital mobility.

Two sets of circumstances contributed to a considerable degree to the increase in the relative number of immigrant workers experienced over the last three decades by advanced capitalism. The first was the crisis and final collapse of the Soviet Bloc and of 'real socialism', which from about 1985 onwards brought an unprecedented surge in the supply of cheap *educated* labour to Western Europe. The second was the so-called Washington consensus: the systematic imposition of very heavy conditions on the governments 'benefiting' from the interventions of the international financial organisations – privatisations, monetary austerity, drastic reductions in primary spending. These conditions must have favoured over several years the expulsion of the workforce from underdeveloped and developing countries, thereby contributing to creating within advanced capitalism the conditions for a practically unlimited supply of cheap labour.

It is now worth stressing that, especially in continental Europe, the free mobility of capital has also affected distribution through a significantly reduced possibility of governing domestic interest rates. Interest rates are ultimately dictated to most countries that have given up capital controls by the need to check outflows of funds incompatible with the exchange rate policy and regime chosen by their authorities. This loss of national monetary sovereignty is likely to bring about also a loss of budgetary sovereignty – of the freedom of each nation, that is to say, to determine its public expenditure levels and priorities, as well as its forms of taxation (see Pivetti, 1999). Faced with interest rates higher than the rate of output growth, governments anxious to check the rise in the ratio of public debt to gross domestic output will necessarily aim at the formation of primary surpluses, which will have to be the higher, relative to GDP, the higher interest rates and the accumulated stock of debt happen to be. And since the free mobility of capital also compels the avoidance of measures likely to cause its flight, a 'capital-friendly' taxation system becomes a must: this entails budgetary constraints that target in particular cuts in public pensions and the other chief categories of social spending. As already observed, reductions in tax progressivity and social spending tend to depress employment, thereby contributing to the containment also of direct wages.

To convince ourselves of the relevance of the transmission channel just depicted, suffice it to recall that from Bretton Woods up to the end of the 1970s, maintaining national sovereignty in the monetary and fiscal fields through capital control was regarded as and actually constituted the very foundation of the full-employment and redistributive policies then pursued by the major capitalist countries (cf. Pivetti, 1993).

Full liberalisation of capital movements, especially in the European context, should therefore be seen as the absolute epitome of the dismissal of those policies.

8.5 Implications for long-run growth

What are the implications of the above analysis for the growth prospects of advanced capitalism and the world economy?

It will readily be acknowledged that correcting the contributory factors to the breakdown of wage earners' bargaining power would be a lengthy process, even if there were widespread acceptance that distributive conditions opposite to those of the past three decades are important for growth. This is far from the case, however.[13] Though the response to the 2007–08 crisis and the recession that followed was initially somewhat more promising in the USA than in Europe, as time went by the US response became increasingly uncertain, with President Obama's Democratic administration eventually choosing to co-operate with rather than confront the opposition, and to increasingly yield to the well-off section of the population and to Wall Street.

In spite of the crisis and the fact that in most countries households persist in being heavily burdened with debt, governments continue to have recourse to 'austerity' packages of an unprecedented rigour. Although it is occasionally acknowledged that this 'austerity' wave is likely to exert a negative impact on domestic demand and activity levels in the short run, the conviction apparently persists that it cannot fail to eventually bring about the conditions for the resumption of stable growth. Such a forecast is simply riding on faith. It could hardly be more in contrast with the real world, first, because the German solution – growing through exports, while at the same time checking the growth of domestic demand[14] – is obviously not viable for advanced capitalism (nor for the European Union) as a whole; second and most importantly, because the fall in employment caused by 'austerity' will produce a further contraction of demand through both the reduction in the number of wage earners and its negative impact on the course of wages. 'Austerity', in other words, tends to amplify the distributive imbalances that have brought about the crisis and the recession (see Barba and Pivetti, 2011).

All in all, distributive conditions within advanced capitalism seem to be converging towards those of the major developing economies, a convergence likely also to eventually hinder the latters' growth. Countries like Brazil, China, India and Russia are bound to find it increasingly hard to transform their exportable surpluses into effective net exports.

Export-led growth, both within the developing and the developed world, will probably end up being looked at as a sort of 'paradise lost' – a bygone smart solution, though not a generalizable one, to the central contradiction of capitalism, that between its tendency to impede continuing relative prosperity of the working class and its persistent need to generate adequate levels of aggregate demand. Over the past few years, the economic crisis in advanced capitalism notwithstanding, major developing countries have continued to enjoy fast growth thanks to massive domestic investments in infrastructures and/or higher consumption spending by the poorest sections of their populations. This has helped to avoid negative rates of growth in a large part of Europe and in the USA. In the longer run, however, it seems unlikely that worldwide stagnation can be avoided through Brazilian-type alms packages, such as the so-called *Bolsa Familia,*[15] or Chinese-type huge public investments in collapsing dams and crashing high-speed trains.

Notes

1. On the neoclassical relationship between technological change and income distribution, the following observation by Stigler is especially noteworthy: 'Economists are generally agreed that inventions on balance increase the marginal productivity of capital. It is difficult to determine whether this conclusion that most inventions are labour-saving is based upon an *independent view* of technological processes or whether it is an inference from the observation that capital has long accumulated at a rapid rate and that the interest rate has shown no persistent tendency to fall.' (Stigler, 1947, p. 327, italics added)
2. Actually, US data for the private non-farm business sector (excluding government enterprises) show for the period 1978–2007 a significant *slowing down* of the rise in output per hour (and, for what it's worth, also in the so-called multi-factor productivity) compared to the previous 30-year period (see US Bureau of Labor Statistics, Office of Productivity and Technology, May 2011, tables XG 2b and XG 4b).
3. For a recent overview of this 'empirical' literature, see Bodvarsson and Van den Berg (2009). G.J. Borjas, however, an acknowledged neoclassical authority in the field, appears to have gradually changed his mind on the matter, eventually admitting that 'the evidence consistently suggests that immigration has indeed harmed the employment opportunities of competing native workers' and that 'it has a sizable effect on the wage of competing workers at the national level' (Borjas, 2003, pp. 1336 and 1354).
4. 'From 1979 to 2006 – it has been recently recalled – the bottom fifth of the labor force gained [in the USA] 11 per cent, … the upper fifth minus the top 1 per cent gained 55 per cent, while the top 1 percent gained 256 per cent. Exclamation mark.' (Sharpe, 2011, p. 122)
5. Cf. Piketty and Saez (2003); Kopczuk and Saez (2004); Kennickell (2009); and Atkinson *et al.* (2001).

6. According to a representative version of the consensus explanation for the general increase in wage inequality, improvements in information and communications technologies (ICT) since the early 1980s would have increased 'the ability of the most talented workers to handle more work or to scale their ideas by working with more production inputs. … The intuition is that individuals who have less decreasing returns to scale will operate at a greater scale.' Thus the rise of ICT would have allowed the most skilled 'to manage more workers and capital, to entertain more people, or to write more papers', thereby causing the incomes of the highest paid to raise (Parker and Vissing-Jorgensen, 2010, pp. 3 and 48).
7. See on this Card and DiNardo (2002); Autor *et al.* (1998, pp. 1180, 1185, 1203); DiNardo and Pischke (1997); and Mishel *et al.* (2009), where it is pointed out that a 'technology story' explaining the growth in wage inequality is not convincing, since, '[i]n fact, during the entire period since 1980 the relative demand for college graduates grew no faster than during the prior 30 years' (*ibid.*, pp. 214–15).
8. The very nature of finance and insurance tends to make them a 'monopoly business', as has long been acknowledged (cf. Bagehot, 1873, ch. X, pp. 171–5. I am grateful to Aldo Barba for suggesting this reference). Over the last 30 years, hardly any new entry occurred in the group of the 'happy few' which make up the core of the US financial services industry, and from about 1985 onwards the number of banks actually fell significantly (on the increased concentration experienced by the US banking industry since the mid-1980s, see Tregenna, 2009, pp. 609–13). As to the increased weight of the industry and its profits, the share of US GDP accruing to finance and insurance, according to the Bureau of Economic Analysis, has risen fairly steadily from 2.3 per cent in 1947 to around 8 per cent in 2007; the rise in the industry's profits, however, has been both astonishing and concentrated in time: at the beginning of the new century they had risen to 40 per cent of total corporate profits, from around 13 per cent of total corporate profits in 1985.
9. Cf. Joint Economic Committee (2003, p. 16); see also Juhn *et al.* (2003), and Mishel *et al.* (2003).
10. Since 1980, with the employment shift to the service-producing industries, while female median earnings increased in the USA by almost 28 per cent in real terms (albeit remaining about 23 per cent lower than male earnings), median earnings of full-time, year-round male workers (15 years and older) never rose above the level reached at the end of the 1970s (cf. US Department of Commerce, 2011, Fig. 2, p. 12).
11. As a matter of fact, at Bretton Woods it had not been an easy task for Keynes, in his negotiations with the Americans, to reach an agreement that was not too distant from his most cherished convictions on 'economic entanglement between nations', as expressed in a well-known 1933 paper of his: 'Ideas, knowledge, art, hospitality, travel – these are the things which should of their nature be international. But let goods be homespun whenever it is reasonably and conveniently possible; and, above all, let finance be primarily national' (Keynes, 1933, p. 236).
12. As evidence of the awareness existing at the time of the epoch-making shift away from full employment to reducing inflation, one may refer to the inquiry promoted at the end of the 1970s by the British Treasury and Civil

Service Committee of the House of Commons into the economic and social impact of the 'change in the objectives of economic policy' of the major industrial countries (the questionnaire, drafted by the Committee, was to be answered by institutional and academic witnesses; cf. House of Commons, 1980). Full political and economic awareness of the implications of that policy shift could hardly be rendered more perspicuous than by the following statement by Sir Alan Budd, a top Treasury official who became Margaret Thatcher's chief economic adviser: 'The Thatcher government never believed for a moment that [monetarism] was the correct way to bring down inflation. They did, however, see that this would be a very good way to raise unemployment. And raising unemployment was an extremely desirable way of reducing the strength of the working classes. … What was engineered … has allowed the capitalists to make high profits ever since' (quoted in Wade, 2011, p. 34). While in England and the USA the policy shift away from full employment and benign labour-market institutions occurred openly and 'frontally' between the end of the 1970s and the first half of the 1980s, in continental Europe it developed in a more gradual and indirect way, essentially through a process of progressive emasculation of national economic sovereignties (see Pivetti, 2011).

13. While rising income inequalities within advanced capitalism have been widely acknowledged by international institutions, the question of their impact on growth has been hardly considered at all in the reports prepared by those institutions (see, for example, OECD, 2008 and 2011).
14. Let us recall that over the last decade German consumption expenditure rose totally by only 2 per cent in real terms, owing to wage stagnation and the rise in households' saving rate linked to the distributive shift in favour of the wealthier classes.
15. Owing to one of the world's highest interest rates, government spending on the service of its domestic debt was in 2009 more than thirteen times the amount allotted to this programme, Brazil's *governo popular* chief social programme (cf. Rocha, 2010, p. 21; for a somewhat more optimistic view of the Brazilian case, see Anderson, 2011).

References

Anderson, P. (2011) 'Lula's Brazil', *London Review of Books*, 31 March: 3–12.

Atkinson, A.B., Piketty, T. and Saez, E. (2011) 'Top incomes in the long run of history', *Journal of Economic Literature*, 49: 3–71.

Autor, D.H., Katz, L.F. and Krueger, A.B. (1998) 'Computing inequality: have computers changed the labor market?', *Quarterly Journal of Economics,* November: 1169–1213.

Bagehot, W. (1978 [1873]) *Lombard Street*, in N. St John-Stevas (ed.) *The Collected Works of Walter Bagehot*, Vol. 9 (London: *The Economist*): pp. 171–82.

Barba, A. and Pivetti, M. (2009) 'Rising household debt: its causes and macroeconomic implications – a long period analysis', *Cambridge Journal of Economics*, 33: 113–37.

Barba, A. and Pivetti, M. (2011) 'Changes in Income Distribution, Financial Disorder and Crisis', in E. Brancaccio and G. Fontana (eds), *The Global*

Economic Crisis. New Perspectives on the Critique of Economic Theory and Policy (Abingdon and New York: Routledge).
Bodvarsson, O.B. and Van den Berg, H. (2009) *The Economics of Immigration. Theory and Policy* (Heidelberg, London and New York: Springer).
Borjas, G.J. (2003) 'The labor demand curve is downward sloping: re-examining the impact of immigration on the labor market', *Quarterly Journal of Economics*, 118: 1335–74.
Bosch, G., Mayhew, K. and Gautié, J. (2010) 'Industrial Relations, Legal Regulations, and Wage Setting', in J. Schmitt and J. Gautié (eds), *Low-Wage Work in the United States and Europe* (New York: Russell Sage).
Card, D. and DiNardo, J.E. (2002) 'Skill-biased technological change and rising wage inequality: some problems and puzzles', *Journal of Labor Economics*, 20(4): 733–83.
Congressional Budget Office (2011) *Trends in the Distribution of Household Income between 1979 and 2007,* CBO Summary (Washington DC: US Congress, October).
Dew-Becker, I. and Gordon, R.J. (2007) 'Selected issues in the rise of income inequalities', *Brookings Papers on Economic Activity*, 2: 169–90.
DiNardo, J.E. and Pischke, J.-S. (1997) 'The returns to computer use revisited: have pencils changed the wage structure too?', *Quarterly Journal of Economics*, 112: 291–303.
House of Commons (Session 1979–80), Treasury and Civil Service Committee, *Memoranda on Monetary Policy*, 2 Vols (London: HMSO).
Joint Economic Committee (2003) *Economic Indicators*, prepared for the Joint Economic Committee by the Council of Economic Advisers (Washington DC: USGPO, March).
Juhn, C., Murphy, K.M. and Topel, R.H. (2002) 'Current unemployment, historically contemplated', B*rookings Papers on Economic Activity*, 1: 79–116.
Kennickell, A.B. (2009) 'Ponds and Streams: Wealth and Income in the United States, 1989 to 2007', Federal Reserve Board Finance and Economics Discussion Paper 2009-13.
Keynes, J.M. (1933) 'National Self-Sufficiency', in D. Moggridge (ed.), *The Collected Writings of John Maynard Keynes*, Vol. XXI (London: Macmillan, 1982): pp. 233–46.
Keynes, J.M. (1980 [1944]) Speech before the House of Lords, 23 May 1944, in D. Moggridge (ed.), *The Collected Writings of John Maynard Keynes*, Vol. XXVI (London: Macmillan): ch. 1.
Kopczuc, W. and Saez, E. (2004) 'Top wealth shares in the United States, 1916–2000: evidence from estate tax returns', *National Tax Journal*, 57: 445–87.
Marx, K. (1887) *Capital: A Critique of Political Economy*, Vol. I (London: Lawrence & Wishart, 1954).
Mishel, L., Bernstein, J. and Boushley, H. (2003) *The State of Working America, 2002–03* (Ithaca NY: Cornell University Press).
Mishel, L., Bernstein, J. and Shierholz, H. (2009) *The State of Working America, 2008–09* (Ithaca NY: Cornell University Press).
OECD (2008) *Growing Unequal? Income Distribution and Poverty in OECD Countries* (Paris: OECD Publications).
OECD (2011) *Divided We Stand: Why Inequality Keeps Rising* (Paris: OECD Publications).

Parker, J.A. and Vissing-Jorgensen, A. (2010) 'The increase in income cyclicality of high-income households and its relation to the rise in top incomes shares', *Brookings Papers on Economic Activity*, Fall: 1–55.

Piketty, T. and Saez, E. (2003) 'Income inequality in the United States, 1913–1998', *Quarterly Journal of Economics*, CXVIII(1): 1–39.

Pivetti, M. (1991) *An Essay on Money and Distribution* (London: Macmillan).

Pivetti, M. (1993) 'Bretton Woods through the lens of state-of-the-art macro-theory and the European Monetary System', *Contributions to Political Economy*, 12: 99–110.

Pivetti, M. (1999) 'High Public Debt and Inflation: On the "Disciplinary" View of European Monetary Union', in G. Mongiovi and F. Petri (eds), *Value, Distribution and Capital. Essays in honour of Pierangelo Garegnani* (London and New York: Routledge).

Pivetti, M. (2004) 'Capitalismo "Incontrastato" e Benessere Sociale: Analisi di un Improbabile Binomio', in B. Jossa (ed.), *Il Futuro del Capitalismo* (Bologna: Il Mulino).

Pivetti, M. (2011) 'Le Strategie dell'Integrazione Europea e il loro Impatto sull'Italia', in L. Paggi (ed.), *Un'altra Italia in un'altra Europa. Mercato e Interesse Nazionale* (Roma: Carocci).

Rocha, G.M. (2010) 'Le Brésil à l'Heure du Bilan Social. Bourse et Favelas Plébiscitent "Lula"', *Le Monde Diplomatique*, September, 21.

Salverda, W. and Mayhew, K. (2009) 'Capitalist economies and wage inequality', *Oxford Review of Economic Policy*, 25(1): 126–54.

Samuelson, P. (1964) *Economics: An Introductory Analysis*, 6th edn (New York: McGraw-Hill).

Sharpe, M. (2011) '"Big money", A review of J.S. Hacker and P. Pierson, *Winner-Take-All Politics: How Washington Made the Rich Richer and Turned its Back on the Middle Class* (New York: Simon & Schuster, 2010)', in *Challenge,* 54(1): 121–5.

Stigler, G.J. (1947) *The Theory of Prices* (New York: Macmillan).

Tregenna, F. (2009) 'The fat years: the structure and profitability of the US banking sector in the pre-crisis period', *Cambridge Journal of Economics*: 609–32.

US Department of Commerce, Economics and Statistics Administration, US Census Bureau (2011) *Income, Poverty, and Health Insurance Coverage in the United States: 2010* (Washington DC: USGPO, September).

Wade, R. (2011) 'The economy has not solved its problems', *Challenge*, 54(2): 5–41.

Wicksell, K. (1934 [1901]) *Lectures on Political Economy, Vol. 1, General Theory* (London: Routledge & Kegan).

9

Alternative 'Closures' to Sraffa's System: Some Reflections in the Light of the Changes in Functional Income Distribution in the United States

Antonella Stirati

9.1 Changes in income distribution and alternative approaches to economic theory

The last 30 years have witnessed a dramatic change in income distribution, with the wage share[1] falling in all major industrialised countries (Figure 9.1).

These changes do not appear to be even partly attributable to an increase in the ratio of the value of the capital stock to value added, which, apart from Japan and, to a lesser extent, France, is not evident in the data for the major economies (Figure 9.2).[2] An increase in that ratio would obviously entail, for any given profit rate, an increase in the profit share (and a fall in the wage share); however, this would *not* reflect an increase in the return on capital and hence a change in distribution as it is properly understood, but rather a structural change in the economy.[3]

Note that the wage-share data tend to underestimate the change in distribution, since the wage share comprises the *salaries* of top managers that are included in firms' payrolls,[4] which rose sharply, particularly in the Anglo-Saxon countries, and which may be regarded as capturing part of business profits rather than as labour incomes. In fact personal income distribution data (cf. Piketty and Saez, 2006, among others) and the trends in wages of non-managerial workers suggest a sharper change in income distribution in the USA than is shown by income shares data. The *level* of real hourly earnings of production workers in the private sector has actually fallen in the USA since the mid-1970s (see Figure 9.3), despite a steady increase in productivity in the same period (GDP

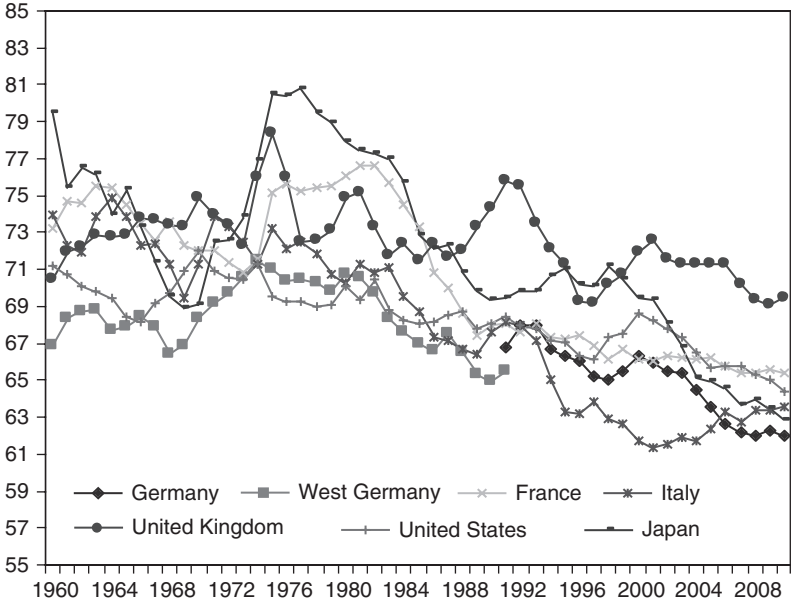


Figure 9.1 Adjusted wage share as a percentage of GDP at factor costs
Source: Ameco database (2012).

at constant prices per person employed rose by 70 per cent between 1973 and 2012; see also Pollin, 2003, p. 43).

The interpretation of such changes in income distribution represents a challenge to economic analysis. Mainstream analyses, including New Keynesian ones, which retain, at least in the long run, the notion of factor substitution leading to a 'factor intensity'[5] inversely related to its rate of return, have encountered some difficulties in the interpretation of the changes in income distribution we have described. In these models a fall in the equilibrium real wage rate would have to be associated with a rise in the labour-to-capital and labour-to-output ratios determined by optimisation in consumption and production.[6] This is particularly important in the context of the explanation of changes in income shares, since if factor substitution was supposed to take place in the manner described by marginal theory, a fall in wages (or in the proportion between wage and product per worker), owing to the increase in the above-mentioned ratios between labour and output and labour and capital, would

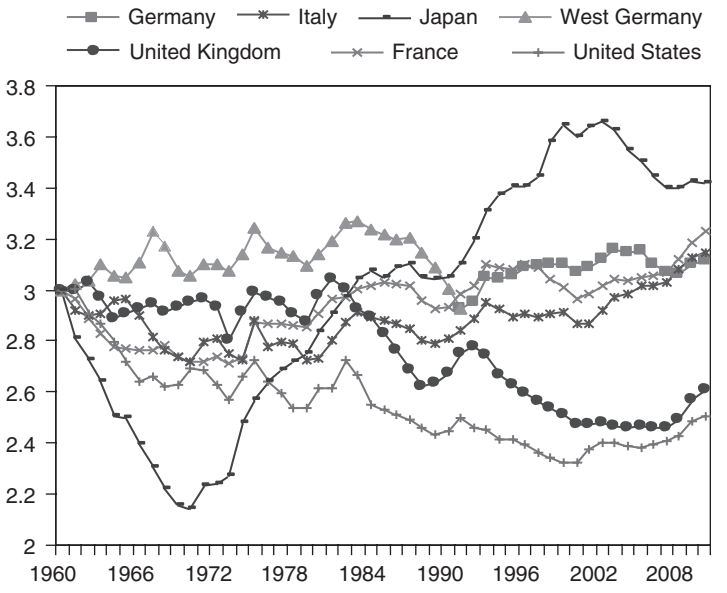


Figure 9.2 Ratio of net capital stock to GDP
Source: Ameco database (2012).

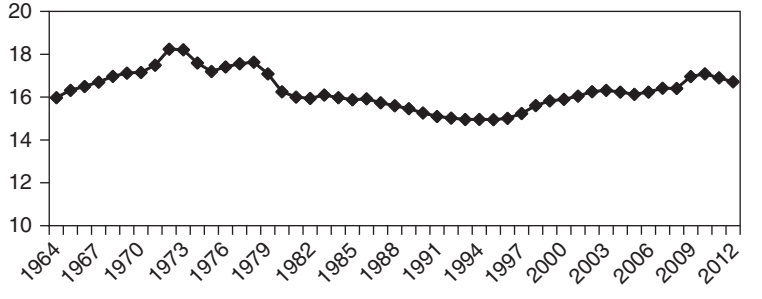


Figure 9.3 Hourly wage level of production and non-supervisory workers in the US private sector, 1964–2012, 2005 dollars
Sources: Bureau of Labor Statistics (2012) for nominal wages; Ameco database (2012) for consumer price index (harmonized CPI, all items).

probably give rise to small or nil (depending on the values of elasticity of substitution) changes in income shares (Rowthorn, 1999). This, by itself, creates an underlying difficulty in all mainstream attempts to explain the observed changes in income shares, which are usually overcome by attributing such changes to labour-saving technical innovation. These

technical changes are, however, hard to identify empirically, and the methods adopted in applied works have been criticised, among other things, as often amounting to *assuming* that changes that cannot be otherwise explained according to the theory *must* be due to technical innovations of the appropriate type (Stockhammer, 2009, pp. 19–22).

Besides the role of labour-saving technical innovations, another, more appealing, explanation of changes in income distribution lies in the role of globalisation in increasing unskilled labour supplies through various channels (immigration, off-shoring of intermediate production, imports of products from emerging economies). Again, however, when this is treated within mainstream economic models that retain 'well-behaved' factor substitution and the consequent tendency to full employment of factors, the expected results are at variance with facts in some important respects. In particular, while according to mainstream trade theory, globalisation should lead to a fall in relative unskilled labour incomes in advanced economies, it should also improve the relative wages of unskilled versus skilled workers in emerging ones, a phenomenon which is not generally observed (IMF, 2007, p. 176). In addition it must be noted, as pointed out by Krugman (2007), that what has been witnessed in industrial countries and particularly the USA, in recent decades, is not a change in the relative positions of skilled and unskilled workers, but rather the inability of *both* groups of workers to benefit from increased productivity, to the advantage of a tiny minority of the population.

The advantage of the classical-Keynesian approach to the phenomenon of distribution and employment is that it entails no *a priori* connections between the changes in distribution and the changes in the proportion between labour and output, or between the value of the capital stock and the value of output. Income distribution is analysed separately from output and employment levels which depend, even in the long run, on the principle of effective demand.[7] Therefore if, for example, a change in the bargaining strength of the parties affects income distribution, the effects on income shares may be significant.

Yet empirical observation may pose some questions also on the analyses of distribution that have been advanced within this approach. In the following sections I will be exploring the ways in which the classical-Keynesian framework of analysis might be used to interpret the changes in distribution described above. My objective, which is primarily to frame the questions rather than finding the answers, is based both on theoretical considerations and empirical evidence, and is not intended to provide a 'test' for theory, but rather to contribute to an initial reflection on the open questions.

9.2 The different 'closures' to the classical price equations

We can start considering possible interpretations of changes in income distribution in the light of the revival of the classical approach by looking at the system of relative price equations:

$$\mathbf{P} = \mathbf{PA}(1+ i) + \mathbf{PA}\rho + \mathbf{l}w$$

where **P** is the relative price vector expressed in terms of a *numéraire*, **A** is the matrix of the production coefficients, i is the interest rate, **l** is the labour inputs vector, w is the real wage and finally ρ is the diagonal matrix of the rate of profits of enterprise or business profits that remunerate risk and illiquidity associated with productive activity. Note, however, that while risk and liquidity premia can be supposed proportional to the capital stock, this is not necessarily the case for top-managerial remuneration, though the latter may be regarded as a component of business profits adding to risk and illiquidity premia.

The business profit term is there to represent the fact that the return on capital invested productively must be higher than the interest rate obtained on safe financial investments (such as long-term bonds) to compensate for risk and illiquidity associated with productive investments, which are likely to differ across industries. Thus while competition must equalise the pure remuneration of capital across industries, total profit rates, including both interest rate and business profits, do not necessarily equalise. In addition, the diagonal matrix ρ may also contain monopoly extra-profits earned in some industries because of barriers to entry.[8]

In the classical tradition, the relation between these distributive variables has been conceived in different ways. Most of the old classical economists thought, like Ricardo, that the real wage was the given variable, determined by socially established subsistence requirements and by the relative bargaining position of the workers (Stirati, 1992; 1994). Thus, profits would be determined residually. Given the normal rate of business profits, the resulting pure remuneration of capital employed in production would ultimately govern the rate of interest in credit markets (since a rate of return on capital higher than the rate of interest would stimulate a demand for loans too high with respect to savings and hence a tendency for the rate of interest to increase – and *vice versa* in the opposite case – cf. Ricardo, 1821, p. 364). It was recognised, however, that the interest rate could differ from the remuneration of capital determined in production for some length of time: according

to Ricardo, for example, '*for the last twenty years*, the Bank ... have lent money below the market rate of interest' (1821, p. 364, my italics – by market rate is meant the remuneration of capital obtained in production). Thus the causal links went from real wages to the profit rate and from the latter (with given business profits) to the interest rate. Marx's conception was different, though not always spelt out entirely clearly and consistently: he retained the notion of a given real wage rate determined by necessary consumption and class relations, and hence a rate of profit in its entirety determined residually but saw the rate of interest also as a magnitude that could be independently determined, leaving the business profits to be treated as the ultimate residual variable. This conception indicates the possibility of a tension between the interests of financial and productive capital (for a discussion of these approaches see Pivetti, 1991, pp. 61–9; Panico, 1988, chs 1–3).

In the current revival of the classical approach to distribution two main lines have been pursued, which will be briefly described below: that of determining the rate of profit as a whole by means of the so-called Cambridge equation and that of accepting the broad view of the classical economists who saw income distribution as the outcome of class conflict, but attributing the role of the 'independent variable' mainly to the interest rate.

9.3 The rate of growth as a determinant of the rate of profit

One very controversial way of determining distribution within the classical-Keynesian approach is the so-called 'Cambridge equation' (Kaldor, 1956; Pasinetti, 1962) which establishes a link between the rate of growth of the economy and the rate of interest. Under the simplifying assumption of zero savings out of wages the equation takes the form

$$G = s_p r$$

where G is the rate of capital accumulation, s_p the propensity to save out of profit income and r the profit rate.

The analytical criticism of this way of determining the rate of profits that has been voiced is that, if G is the actual rate of accumulation, r appearing on the right hand side is not the normal, but the *ex post* actual profit rate reflecting the actual degree of utilisation of capacity. The *ex post* profit rate changes with capacity utilisation, totally irrespective of changes in the real wage, given the techniques in use. This

contrasts with the nature of the rate of profit that appears in the price equations, which is the normal profit rate, associated with i) the dominant technique, described by the technical coefficients in the equation, and ii) a normal or desired degree of capacity utilisation. Only under the latter conditions the normal rate of profits has a definite, inverse relation to the real wage rate, and can therefore, if given, univocally determine it, as in Sraffa (Ciccone, 1990a; Garegnani, 1992; Garegnani and Palumbo, 1998; Aspromourgos, vol. 3 of this book). The equation is therefore irrelevant to the theory of distribution.

Of interest here is that although the Cambridge equation is still used in Post-Keynesian formal macro-models, to my knowledge no one has attempted to interpret the changes in income distribution in recent decades with reference to changes in the rates of accumulation or in the share of investment in value added, as would be in the logic of the model under discussion. Indeed, such a suggestion would be at variance with observation. In major industrial countries the decline in the wage share has in fact gone *together* with a decline in the rates of accumulation and GDP growth. It may be noted in this regard that even the *ex post* actual profit share and profit rate that are measured by the national accounting statistics do not in the long run exhibit the relation expressed in the equation. The explanation lies in the short-run nature of the connection between the actual rate of growth and the ex-post rate of profits, reflecting changes in capacity utilisation. The relation is therefore visible in the data for changes in the rates of GDP growth over the cycle (e.g., the profit share tends to fall during recessions),[9] but not necessarily so when looking at longer-term tendencies, since situations of under-utilisation or over-utilisation of capacity tend to be corrected – albeit not necessarily ever fully eliminated – by changes in capacity; though some decline in the profit share associated with lower than desired capacity utilisation might actually turn out to be rather persistent when the economy is stagnating or declining over relatively long periods of time.

9.4 Bargaining power and social relations

Most economists adopting the classical-Keynesian perspective actually tend to regard changes in income distribution as determined by changes in the bargaining position of the parties, in turn associated with changes in the economic, social and institutional set-up. Even from such a broad perspective, however, the main channels through which the changes tend to occur may be regarded as acting primarily directly on real wages or on the interest rates.

As described above, the old classical economists and Marx tended to regard the real wage rate as the given variable, determined by social and economic conditions (among which the unemployment and underemployment rates are of primary importance), with the profit rate determined residually. Some contemporary economists, however, have argued that this point of view is ill-suited to represent the process of determination of real wages, since in fiat money economies workers can only bargain over nominal wages, while firms can and will generally pass increased money costs of production on to prices (Pivetti, 1991, pp. 36–7). On this basis, and following the suggestion by Sraffa (1960, p. 33) that the rate of profit may be regarded as determined by the rate of interest, it has been argued that, in general, distribution may be seen as determined primarily via interest rate policy, acting on the minimum competitive return on capital appearing in the price equations. According to this view, which is close to the Keynesian tradition of regarding the interest rate as a 'conventional' and institutional variable, the interest rate is independently determined by the monetary authorities subject to a number of objectives and constraints. Since business profits, regarded as risk and illiquidity premia, must also be conceived as given magnitudes in a competitive economy, the residually determined variable is the real wage. This remains true even when business profits earned above the interest rate include, besides risk and illiquidity premia, an extra-profit which is due to the existence of obstacles to free competition and barriers to entry, since these extra-profits cannot be regarded as arbitrary, but as a definite addition to the competitively determined benchmark. The mechanism through which changes in the interest rates would cause a change in real wages is the price level. For a given nominal wage rate and productivity, and given the rate of business profits, an increase in interest rates causes an increase in the monetary costs of production, and hence in the price level vis-à-vis the money wage (Pivetti, 1991). As a consequence, monetary authorities are regarded as generally capable of governing income distribution. This way of looking at the role of the interest rate has gained acceptance also in wider 'Post-Keynesian' literature, which sees the interest rate as a component of the 'mark-up' charged by firms over the costs of production (see for example Hein and Schoder, 2009).

In regard to the connection between the rate of profits and the rate of interest, a somewhat more cautious and detailed argument has been made by Garegnani. He agrees that competition in product markets tends to ensure, *over sufficiently long periods of time*, that the real rate of interest and the rate of profit move in step, and that, if the rate of

interest is a 'conventional' variable, largely determined by monetary authorities, then it can determine the normal rate of profit and the corresponding real wage rate. However, according to Garegnani, 'The policy of the monetary authorities is not conducted in a vacuum, and the movement of prices and of the money wages determined in the wage bargain will be amongst the most important considerations in the formulation of that policy' (Garegnani, 1983, p. 63). This more cautious position is also supported by Ciccone (1990b) and Stirati (2001), who argue that the reaction of money wages to a price increase may affect the *real* interest rate, for a given money interest rate established by the central bank and financial markets. While a once-for-all increase in nominal wages would have – with a lag – a permanent effect on the price *level*, with no consequences for the real interest rate, a continuous increase in nominal wages would cause price inflation and affect the real interest rate (see Stirati, 2001, for a formal analysis). The causation between the interest rate and real wages therefore cannot be conceived too mechanically, and the direction of causation may actually vary according to circumstances.

It is perhaps also worth recalling at this stage that Sraffa explicitly refers to the *surplus component of wages*, since the subsistence requirements, according to Sraffa, are necessary costs and as such should most appropriately be treated in the same manner as other production costs, and included in the coefficient matrix (1960, pp. 9–10). If so, there must be limits within which wages may be affected by monetary policy, depending on the size of the surplus component of wages above the subsistence requirements. The terms 'subsistence' and 'necessaries' are also used by the old classical economists, whose standpoint Sraffa in his preface claims to be adopting. 'Necessaries' therefore has the same meaning as it did for them, including, to quote Smith 'those things which the established rules of decency have rendered necessary' (1776, V.ii.k.3). Thus, even in advanced economies, the notion of a subsistence floor continues to be relevant, and may not be so far distant from current wages. On the other hand, the tendency in contemporary economies towards a continuous increase in productivity creates scope for increases in the rate of profits that do not require a fall in real wage levels.

9.5 What is the role of globalisation and financialisation?

Recent discussions and applied analyses of the changes in income distribution have tended to emphasise the role of globalisation and financialisation in affecting income distribution over recent decades.

Some of the arguments that have been advanced, particularly in non-mainstream literature, will be summarised and commented on here with a view to establishing whether, and how, they can be integrated into the classical-Keynesian framework of analysis as factors that have, broadly speaking, affected class relations and economic institutions and hence the bargaining position of the parties.

The phenomenon of 'globalisation', involving increased competition from emerging economies in product markets, off-shoring and immigration, can be expected to directly affect the bargaining position of workers in advanced economies by increasing job losses and unemployment, or by representing a powerful threat that jobs will be lost (because of de-localisation or imports from other countries) as a consequence of higher wages or improved work contracts and conditions.[10] In addition, one might argue that the impact of globalisation works through other channels as well; of particular importance are the constraints on macroeconomic policies that are imposed by free capital mobility (for example on public budget and public debt management), which can then affect the formation of aggregate demand. All in all, it thus would appear that many of the factors that are synthesised as 'globalisation' are likely to *directly* affect the bargaining position of workers through their impact on employment growth and unemployment and the threat of de-localisation of production and job losses. Several observers, for example, have referred to job insecurity as a main factor in subdued nominal wage dynamics in the USA even in the 1990s, a period of low and falling unemployment: Greenspan (then governor of the Federal Reserve) referred to 'traumatized workers' and their job insecurity as a factor in explaining low inflation even in a period of sustained growth and low unemployment rates (Greenspan, 1997; see also Pollin, 2003, pp. 50–6; Choi, 2001) and Robert Gordon in 1997 wrote in this context 'The 1990s have been a time of labour peace, relatively weak unions, a relatively low minimum wage' (1997, p. 30) and pointed to intensified world competition in product and labour markets, and the increased inflows of immigrant labour in the USA as likely explanations.

Many discussions of changes in functional income distribution carried out by non-mainstream, Post-Keynesian economists point to the role of financialisation (see Hein, 2011, for a survey of the literature). Yet my impression is that in this regard the analysis of the causal mechanisms envisaged, apart from the effects of higher interest rates discussed above, are not yet entirely clear, or satisfactory, though some of the econometric exercises carried out confirm a negative correlation between financialisation of the economy and the wage share (Stockhammer, 2009).

In terms of the classical-Keynesian approach considered here, the following considerations arise. First, it is perhaps worth clarifying that, from a macroeconomic point of view, the increases in the wealth of top managers or investors in financial markets in the form of capital gains cannot be regarded as affecting the distribution of income. The moment any individual agent wishes to use that wealth to finance current consumption the assets must be sold, so that the aggregate effect on the private sector as a whole is nil. On the other hand, if increased wealth obtained in the form of capital gains can be used as collateral to obtain additional credit (as was largely the case in the recent past) the outcomes of this for the distribution and uses of current output differ according to whether it is assumed that the latter is (on the average over economic cycles) given and equal to potential output or is instead considered elastic, even in the long run, to changes in aggregate demand (Garegnani, 1992). Under the latter assumption, which is consistent with the approach taken here, the transformation of capital gains into credit is the same thing as any credit-financed increase in autonomous expenditure, and would not therefore encroach upon someone else's income and consumption. A similar conclusion can be drawn with regard to the effects of increased asset value on the propensity to consume.

However, there may be a sense in which the high earnings of *individual* managers and firms in the financial sector, even in the form of capital gains, might have affected income distribution – i.e., by affecting the social norms concerning income levels, and the opportunity costs and expected earnings of individuals and firms operating in other sectors of the economy. In this regard, however, the question within the present framework of analysis is whether and to what extent such change in the social norms concerning the legitimate remuneration of managers may actually be regarded as an *independent* factor in affecting distribution, that is, whether it can be regarded as a *cause* rather than as a *consequence* of changes in the bargaining position of the workers. Another connection between financialisation and a higher normal profit rate might lie in the possibility that the high returns (mostly in the form of capital gains) realised by means of financial or real-estate transactions have produced a change in the rate of return on capital regarded as the minimum competitive benchmark, from the interest rate on riskless long-term public bonds to the (higher) returns on some different portfolio of assets. A third argument connecting financialisation with higher normal profitability emphasised in Post-Keynesian literature concerns the changes that it is believed to have brought about in the governance of firms towards stock-holder value orientation, focussing on short-term

realisation of high profits. Institutional changes such as the possibility of hostile takeovers and the performance-related remuneration of managers would have changed the latter's priorities from the growth of the firm (financed by the re-investment of retained profits) to distributing high dividends to stock-holders and in general would have led to a greater power of stock-holders in defining the objective of the firm. However, this is more an argument about the utilisation of profits than about an increase in profitability; though there are some discussions in this literature about the effect of the changes in the governance of firms on employment and workers' bargaining positions (see Hein, 2011; Stockhammer, 2004 for discussion of the literature).

On the other hand, the enormous increase in financial transactions and leverage that have been made possible by the deregulation and expansion of financial markets seems likely to have increased the profits of financial corporations for any given capital stock, out of commissions and interest payments, thus capturing an increasing share of national income. If, however, the financial sector is largely a monopolistic sector, and if, as has been argued, many of the financial products and services can be regarded as luxury, non-basic goods[11] (Barba and de Vivo, 2012), then such high profits would not have affected the normal rate of profit in the other sectors of the economy, though they would show in national accounting data for aggregate income shares in proportion to the weight of the financial sector. According to national accounting data, however, in the USA between 1987 and the first half of the 2000s the proportion of gross operating surplus to value added in the financial sector increased from 40 per cent to 45 per cent, roughly in line with what happened in other sectors (see Figure 9.6, p. 207) – so that from the point of view of the trends shown in aggregate national accounts the financial sector has not provided a determining contribution to the upward trend in the 'profit' share.[12]

Summing up, the arguments about the effects of financialisation on changes in manager compensation norms, the change in the benchmark return on capital and shareholder value, would imply within the approach outlined above that for any given real long-term interest rate there will tend to be an independent increase in normal profitability. On the other hand, increased monopoly profits in the financial sector would not affect normal profitability in the economy at large, but would in principle result in a fall of the aggregate wage share (and a mirror increase in operating surplus) proportional to the weight of the financial sector in the economy. US data, however, do not indicate that this was a major cause of the changes in income shares.

9.6 Some evidence from the US economy over the last 50 years

The purpose of this preliminary examination of empirical evidence and historical reconstruction is to provide some insights into two questions. The first is whether the evidence is consistent with the idea that real interest rates and profit rates actually tend to move in step. The second is whether interest rate policy may be regarded as an independent factor in determining changes in income distribution. Recalling the analysis summarised in Section 9.4, an increase in nominal interest rate, according to the modern revival of the classical approach, can affect distribution by causing a rise in the price level *vis-à-vis* the nominal wage. On the other hand, nominal wage inflation might be able either to resist the distributive effect of higher nominal interest rates or to determine an autonomous push towards higher real wages. However, it appears rather difficult to disentangle these different directions of causation by looking directly at nominal variables. I will therefore simply try to assess a) whether real wage increases are indeed associated with periods of higher nominal wage inflation (which would suggest that bargaining over nominal wages matters for income distribution); and b) whether there are elements in the overall institutional, economic and labour market conditions that suggest the relevance and priority of conditions directly affecting wage bargaining with respect to autonomous changes in the policies of the monetary authority. The reconstruction, however, is descriptive and is not intended to represent a 'test' of any particular theory, but rather aims at stimulating questions and further reflections on these complex issues.

Naturally statistical data cannot provide information on the normal profit rate as defined above, but only on the actual rate affected by capacity utilisation and earned on the existing capital stock, which is not necessarily of the kind associated with the dominant technique and includes quasi-rents on economically or technically obsolete plants as well as rents or extra-profit connected to elements of technical monopoly. However, data on long-term trends in income shares (that is, on the proportion of real wages to output per worker) and ex-post profit rate may provide an acceptable indication of the direction and order of magnitude of the changes.

The choice of the USA reflects that country's leading role and the fact that it is accordingly the least constrained in fixing interest rates. It also reflects the observation that in the USA (unlike some European countries)[13] changes in distribution do not appear to be so closely connected

with parallel changes in unemployment, suggesting perhaps a greater role for monetary policy than in other countries.

The long-term real interest rate on US government bonds began to increase sharply in 1980 and reached a peak in 1983, as a result of the change in nominal interest policy beginning in 1979 (Figure 9.4). It subsequently fell, stabilising at around 4.5 per cent in the period 1989–1997, about two points above the values observed in the early 1960s and about 3.5 above the – historically low – average values[14] in the period 1969–1979. After 1997 real interest rates began to fall and have remained below 2 per cent since 2003, despite a moderate recovery between 2005 and 2007.

If we now look at the long-term changes in income distribution over the same period, we see that the wage share for the entire economy tended to decrease since the early 1970s, with two counter-movements: the first in the severe recession of 1980–82, which caused a fall in output per worker (see note 9 above); the second in 1997–2001, which by contrast were years of rapid GDP growth and falling unemployment associated with a rise in real wages (Figure 9.5).

Further disaggregated information shows that these changes reflect similar trends in most sectors of the economy. We can start by looking at the manufacturing sector data over the same period (Figure 9.5). Here too there was a fall in the wage share in 1975–79, and a continuously

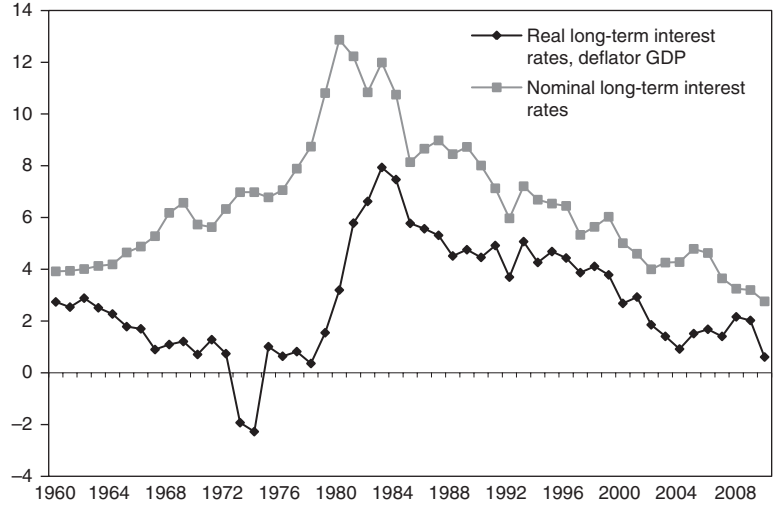


Figure 9.4 Nominal and real long-term interest rates, USA 1960–2011
Source: Ameco database (2012).

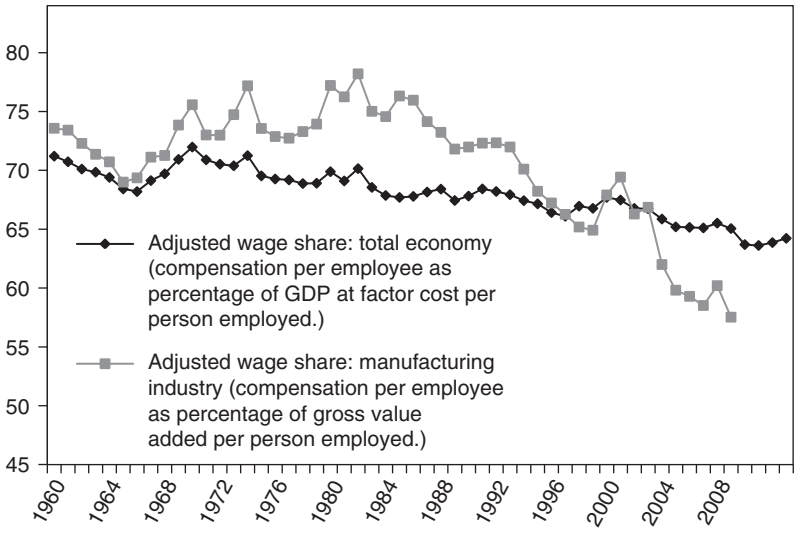


Figure 9.5 Wage share in the economy and in manufacturing, USA 1960–2011
Source: Ameco database (2012).

decreasing trend is clearly visible after 1985. The decline accelerates both in the economy as a whole and in manufacturing after 2000.

Data on gross operating surplus as a share of value added in various subsectors of the economy (available as a continuous series only since 1987) also confirm that operating surplus and mixed income[15] share of output increased – albeit to a different extent – in all sectors of the economy (Figures 9.6 and 9.7; note that although the mining industry exhibits a very sharp increase in the gross operating surplus, its weight in the economy is limited (1 per cent in 2009) so that it is not a major factor in the overall trend).

Consistently with the picture just outlined, the national accounting data on the net returns on the net capital stock (Figure 9.8) indicate, after the 1980–83 fall associated with the recession, an increasing trend which reverses the decline experienced between the mid-1960s and the end of the 1970s. The increasing trend in the net return on the net capital stock is interrupted only in 1997–2000, for the reasons described above, and then during the 2007–09 crisis.

Though the return on fixed capital increased, it could be wondered whether this has to do with an increase in the cost of intermediate inputs (circulating capital), particularly primary commodities, per unit of output.

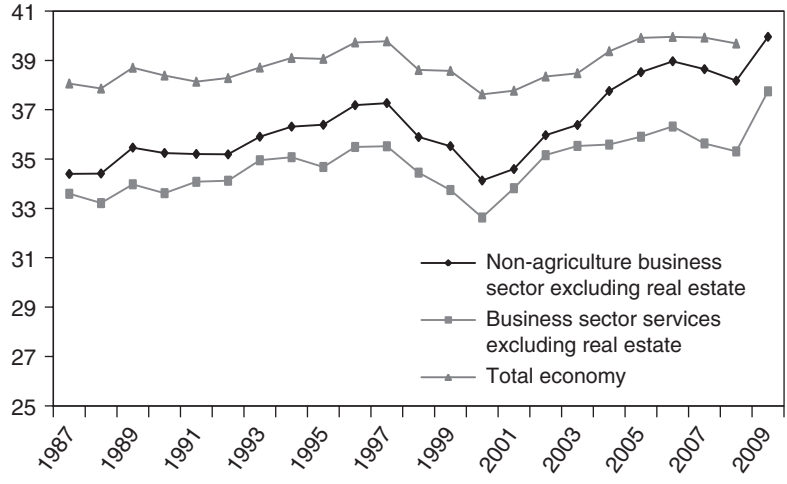


Figure 9.6 Gross operating surplus and mixed incomes as a percentage of value added at factor costs, 1987–2009
Source: OECD structural statistics.

Again, some prudence is necessary when evaluating data on intermediate inputs, since on the one hand, they depend also on the degree of vertical integration in production, on the other, and more importantly, the speed at which intermediate inputs circulate is certainly much higher than the accounting period (one year). This means that the actual proportion of circulating capital over value added on average during the year amounts to a much lower figure than that in the data (that is, if the velocity of turnover for intermediate inputs were one month, the proportion based on annual data should be divided by 12 to obtain the average proportion of advanced circulating capital on value added during the year). It should also be considered that turnover rate may vary significantly over time and across sectors owing to several circumstances, including demand conditions. Figure 9.9 shows that the proportion of intermediate inputs to value added has fluctuated over the period without reflecting on income shares. After 2005 the proportion rises somewhat above previous historical peaks up to a maximum of 2 percentage points – this does not seem to be a relevant figure in the light of the above considerations.

Thus, overall the change in net returns on net capital stock between the 1970s and the period 1995–97 appears broadly consistent with a gradual adjustment of income distribution to the rise in the average value of the real interest rate – data not shown in the figure show that

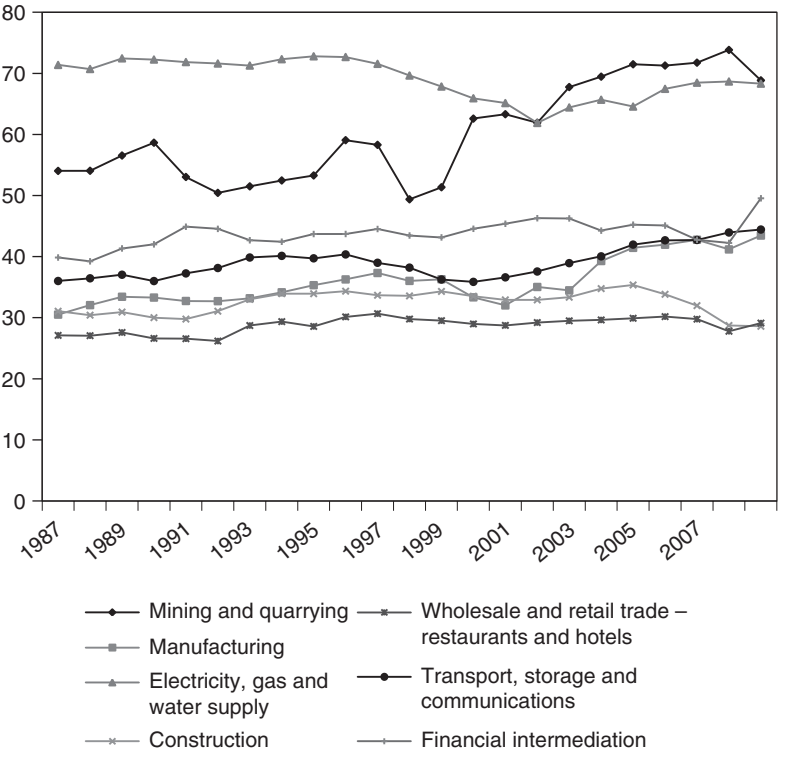


Figure 9.7 Gross operating surplus and mixed incomes as a proportion of value added, 1987–2009
Source: OECD structural statistics (2012).
Note: Value added at factor costs.

the return on capital increased from about 5.4 on average in the 1970s to an average 7.5 in 1995–97 (Ameco database, 2011), somewhat less than the increase in real interest rate. These data, however, must be considered in view of the fact that the large increases in top managers' *salaries* are included here in the aggregate of the compensation of employees, and tend therefore to underestimate the redistribution of income away from non-managerial labour (see Section 9.1).

Since 2000, however, real interest rates begin to decline significantly and after 2003 tend to stabilise at historically low values (below 2 per cent), while on the other hand the net return on net capital stock stabilises and even increases after the 2008–09 crisis.

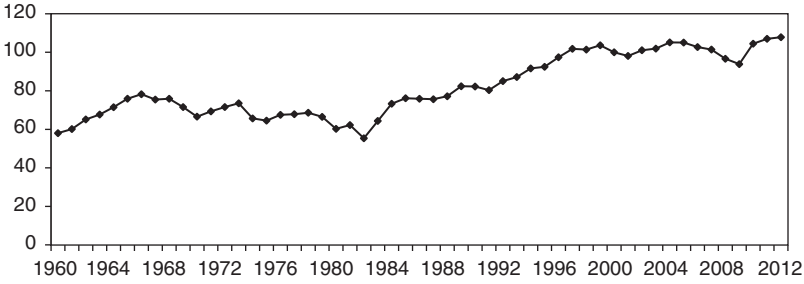


Figure 9.8 Net returns on net capital stock, US total economy, 1960–2012 (2000 = 100)
Source: Ameco database (2012).

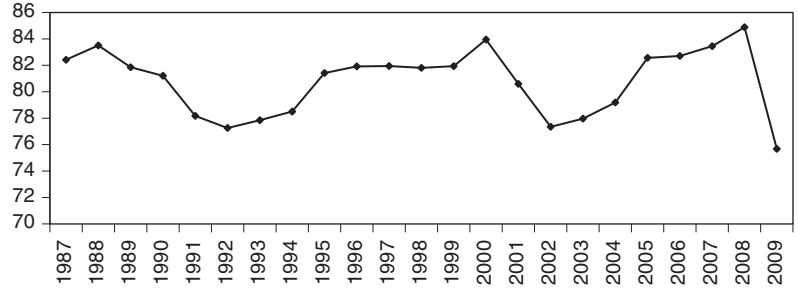


Figure 9.9 Intermediate inputs as a percentage of value added, current prices, 1987–2009
Source: OECD structural statistics (2012).

The evidence for the last decade might thus be consistent with an exogenous increase in the 'mark-up' such as could have been brought about by 'financialisation', as argued in some of the literature (see above, Section 9.5), but also, alternatively, with the view that both real wages and interest rate can vary independently for some length of time, thus leaving business profits to be determined *residually*. In the period under discussion, nominal wage dynamics remained the same as in previous years, fluctuating between rates of increase of 2 per cent and 4 per cent according to the cycle (Figure. 9.10), and was likely affected by the same economic and institutional framework: the continuing adverse effects of globalisation, an increase in unemployment from 4 per cent to 6 per cent between 2000 and 2003, and a Republican presidency may all have contributed to the insecurity of workers noted by several observers

during the 1990s. At the same time, for a number of reasons, in the last decade the Federal Reserve successfully pursued low *nominal* and *real* interest rates. An interest rate lower than the returns on other forms of investments may have contributed to the acceleration of financialisation, since it provided an incentive to borrow and reinvest in financial markets or real estate (much less so in production, since aggregate productive investments are influenced, and constrained, by aggregate demand), and provided a 'rational foundation' for the expectation of increasing stock market and real estate values, favouring the then self-reinforcing 'bubbles'. It is generally recognised that a major objective of the Federal Reserve's low interest rate policy was precisely that of sustaining financial markets. This, however, leaves open the question of why a tendency towards a lower proportion of the price level *vis-à-vis* nominal wage did not manifest itself as a consequence of competition in product markets and decreased monetary costs of production. With subdued wage dynamics and a central bank aiming to establish low real interest rates, such a tendency would be expected to manifest itself *via* price deflation, whereas in fact price dynamics remained very much the same as in the 1990s.

As far as the second question is concerned, that is, whether it is possible to say something about the respective roles of wage bargaining and monetary policy as independent causes of changes in income distribution, the first observation is that periods of higher nominal wage dynamic often are also periods in which real wages rise, as can be seen from Figure 9.10. This is true of the early phase up to 1980, except when a significant exogenous inflationary impulse came from the oil shocks of 1973 and 1979; but also between 1996 and 2001 and to a lesser extent in 2006–07 when the growth in real wages is clearly associated with a more sustained nominal wage dynamic. On the other hand, between 1984 and 1991 real wages fell at a roughly constant rate of about 1 per cent a year, independently of the fluctuations in nominal wage growth rate. It is of some interest that in the second half of the 1990s the (moderate and short-lived) reversal of the trend in income distribution took place in the last part of the long Clinton boom, characterised by sustained GDP growth and fall in the unemployment rate (Figure 9.11), as a result, it would seem, of conditions directly affecting the labour market and wage bargaining.

With regard to the phase beginning in the early 1980s, it is quite clear that after 1979 there was a determined change in monetary policy aimed at implementing high nominal and real interest rates, and the data are broadly compatible with a causation going from higher long-term interest

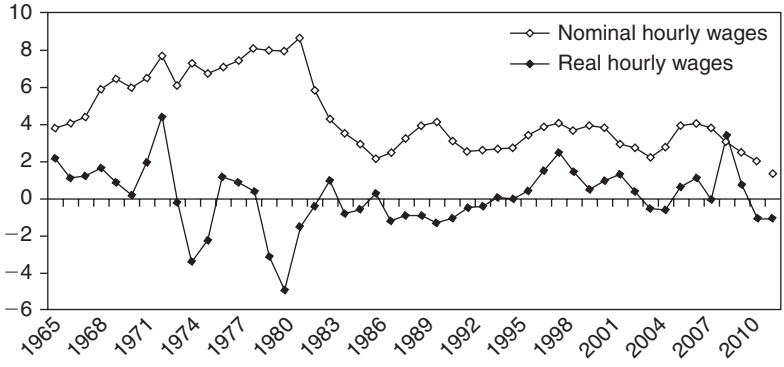


Figure 9.10 Nominal and real hourly wages of production and non-supervisory workers in the private sector, 1965–2012, percentage annual rate of change
Sources: BLS (2012) for nominal wages; Ameco database (2012) for harmonized CPI.

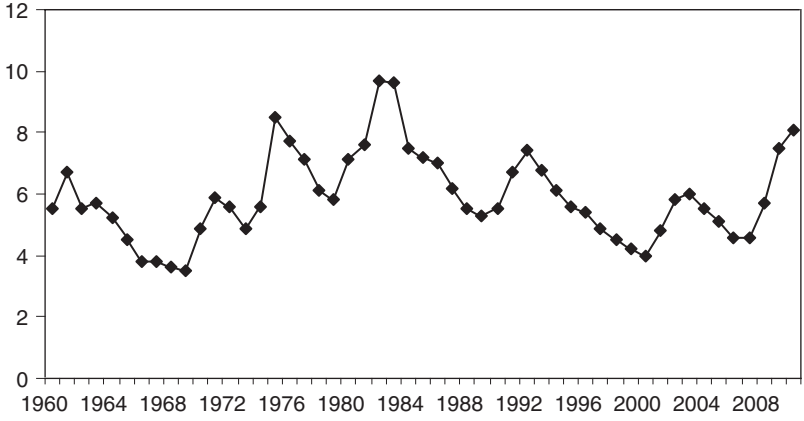


Figure 9.11 Unemployment rate, USA total, 1960–2010 (Eurostat definition)
Source: Ameco database (2012).

rates to changes in income distribution. Yet things might be more complex. The increase in interest rates was preceded by firm-level reactions to the enhanced bargaining position of workers. In the early 1970s firms were already undertaking restructuring processes featuring de-localisation of production and 'union avoidance' (Bluestone and Harrison, 1988). In fact in the USA the first signs of a change in labour relations can be dated from the reversal of the trend of production workers' real wages in the early to mid-1970s (Pollin, 2003, pp. 42 ff.; see Figure 9.3) and a fall

in wage share. In parallel with this, the unemployment rate, which had been falling steadily between 1960 and 1969, initiated a reversal of the trend. Thus, it could be suggested that Volcker's policy of high nominal and real interest rates was actually made possible by indications that a weakening of labour was *already* under way. In addition, the change in interest rates and monetary policy after 1979 was accompanied by a huge surge in the unemployment rate, which was in turn conducive to further institutional changes affecting unions and the labour market. There are therefore elements that indicate that labour market conditions and the institutional set-up directly affecting the bargaining position of workers may be important in determining not only nominal but also real wages.

9.7 Conclusions

The classical-Keynesian approach appears better equipped than mainstream theory (in all its variants) to provide a consistent explanation of the dramatic changes that have taken place in recent decades. This is due to the combination of criticism of the analytical foundations of decreasing factor demand curves, the consequent revival of the classical separation between the determination of employment and distribution and the classical approach to distribution as determined by norms, institutions and power relations. By contrast, these changes in income distribution add to the collection of empirical observations that collide with mainstream theory – unsurprisingly, in view of the flawed theoretical foundations of the approach.

Even some mainstream New Keynesian economists now appear to acknowledge that not only personal, but also functional income distribution essentially depends on norms and institutions, and that only changes in the latter can explain, first the absolute and relative improvement in labour earnings until the mid-1970s, and subsequently, the reversal of that trend (Krugman, 2007). Note that this is a novelty since, despite the emphasis on labour market institutions, in standard New Keynesian macro-models the equilibrium real wage is determined residually, given the mark-up. This in turn is determined by the elasticity of product demand curves in imperfectly competitive markets. Thus, changes in institutions in these models shift the equilibrium unemployment rate, but have a limited effect on the equilibrium wage rate (as a proportion of output per worker) only if decreasing labour marginal product is assumed.

Within the classical approach, however, different emphases can be found as to whether the changes in institutions and power relations affect the distribution of the surplus (over and above the necessary requirements

of the workers) by acting primarily on interest rate determination or wage bargaining. In this regard, this chapter's evidence from the USA seems to suggest that Garegnani's cautious position may be the most appropriate: in principle either variable may be subject to be determined 'residually' according to circumstances, and monetary policy concerning interest rates may be influenced and constrained, among other things, by power relations and institutions in the labour market, so that the latter can be very important in determining distribution, even in contemporary, fiat money economies. In the US experience, the circumstances directly affecting wage bargaining appear to have been important, and are often emphasised in narrative accounts (see, among others, Bluestone and Harrison, 1988; Krugman, 2007; Pollin, 2003), even though in the USA, after the strong recession and high unemployment of the early 1980s, wage trends are related mostly to economic and institutional changes and less to the unemployment rate than has been the case in European countries (Levrero and Stirati, 2006; Stirati, 2010, pp. 132–9).

Finally, US data suggest that for about a decade now interest rates and the rates of profit on capital have not been moving in step – as expected (in the long run) if the working of competition tends to equalise the pure remuneration of capital in different uses, with given normal rates of business profits. In principle this could be explained by exogenous changes in normal business profits, exogenous in this context meaning independent of economic and institutional factors (e.g., the weakening of unions) that directly affect wage bargaining. The origins of such changes in turn would need to be carefully identified. Alternatively, it should be acknowledged that interest rates and profitability can actually diverge, at least for some time, owing to independent forces acting at the same time and in the same direction on the interest rate and real wages, leaving business profits to be determined residually. In this case, the circumstances that may have prevented or delayed the decrease in the proportion between prices and nominal wages as a result of competition, and the consequences of the gap between profitability and interest rates, should be further investigated.

Notes

1. The data reported here refer to the wage share adjusted by imputing to the self-employed a labour income equal to that of employees; by this definition, the share is not affected by changes in the proportion between employees and self-employed workers. GDP at factor costs is net of taxes on production. The adjusted share therefore is the same as the ratio of average labour income per worker (or standard labour unit) to average product per worker (or standard

labour unit): Adjusted wage share = [(Compensation of employees/employees) x total employed] / Value added at factor cost.

2. Some caution is necessary on this point, since statistics on the capital stock are notoriously unreliable and difficult to construct. In addition, the capital-to-output ratio reflects not only structural/technical changes but also changes due to variations in the degree of utilisation of capacity.
3. In discussing the data I am taking the classical rather than neoclassical view. Therefore I am not supposing any pre-determined relation between the desired capital output ratio (which will depend on economic structure and relative prices) and the rate of profit. What I am considering here is merely the accounting fact that for a given rate of profit, an increase in the value of the capital stock would entail increased total profits. There is sometimes ambiguity in Post-Keynesian literature, since changes in income shares may, in a sense, always be treated as changes in income distribution from a macroeconomic point of view – but the underlying causes may be different in nature. For example, in the formal analysis presented in Panico *et al.* (2012) the change in income shares is entirely caused by a change in the proportion between the value of capital and output, and not by a change in the rate of profit. The analysis therefore does not address the interpretation of the changes that have actually been taking place in most countries in recent decades.
4. The share includes compensations in the form of salaries, but not payments in the form of stock options, or as compensation of professional services.
5. Sraffa (1960) and Garegnani (1970), among others, have shown that the notion of factor intensity itself is devoid of meaning, since – for any given underlying technique – factor intensity is not independent of relative prices and distribution.
6. This is the case when, in the familiar model of a price equation and a bargained wage equation, the former is decreasing in the real wage-employment space because it reflects a decreasing marginal product of labour. If labour marginal product is assumed constant, as in some textbooks, a change in the bargaining strength of the workers would shift the bargained wage equation, determining lower equilibrium unemployment, but no changes (given productivity) in the equilibrium wage. In this case therefore the model could not be used to explain changes in distribution.
7. The main analytical premise of these views is to be found in the criticism of decreasing factor demand functions (Garegnani, 1970). In contemporary macro-models the tendency of the economy to potential output is generally attributed to the so-called real balance effect and 'Keynes effect'. However the former cannot by itself be regarded as capable of ensuring that tendency (Patinkin, 1987), while the inverse relationship between aggregate investment and the interest rate (the Keynes effect) must in the end rely precisely on well-behaved factor substitution (see Petri, 2004, chap. 7).
8. Extra-profits may be earned in some firms/sectors thanks to better than dominant technique, that is to some form of more or less transitory technological monopoly allowing reduction of costs below those associated with the dominant technique. Such extra-profits would have the nature of 'rents' and are not price-determining, that is, they are not included in the term ρ in the price equations which must refer to the dominant technique. On the other hand such rents are necessarily part of the empirically observed size of the operating surplus (see note 9 below for a definition).

9. Operating surplus or 'profits' in national accounts are a residual concept – what remains of net (or gross) domestic income after subtracting the (adjusted) wage share. If the wage share is adjusted for self-employed imputed labour incomes, it measures by definition the ratio between the real wage (in terms of the GDP deflator) and the value added at constant prices per person employed (labour productivity). The latter is empirically known to move pro-cyclically (Okun's law) and therefore *for a given real wage*, the wage share would tend to rise and the 'profit share' to fall during a recession. However, there is much empirical evidence that real wages also tend to vary pro-cyclically, so that in actual fact the wage share may rise or fall in recessions according to which effect prevails. Generally however, particularly in the initial phase of a recession, one can observe a fall in the profit share. Note that these *empirical* considerations have little to do with mainstream debates over changing mark-ups over the cycle, which confront problems and use concepts and empirical methods (such as increasing marginal costs, the neoclassical aggregate production function, total factor productivity) that are entirely embedded in mainstream theory.
10. At the same time, in contrast with predictions deriving from mainstream theory, 'globalisation' does not *necessarily* improve the *relative* position of unskilled vs skilled workers in emerging economies, owing to large labour reserves and/or to general political and institutional conditions in those countries.
11. That is, they do not figure in the costs of production of wage goods or other basic goods, so that the interest and fees paid to obtain them may be conceived as expenditures in luxury consumer goods or services, such as gambling games.
12. The financial sector has increased its weight in GDP between 1987 and 2005 from 6 per cent to 8.3 per cent, and since its operating surplus as a proportion of value added is somewhat higher than in most other sectors, it has contributed to the increase in the 'profit share' through a composition effect, which is however of modest impact.
13. See Stirati (2010).
14. The average real interest rate of the period 1969–1979, calculated excluding the two years in which interest rates became negative, was 0.9 per cent.
15. Longer historical series for compensation of employees, operating surplus and capital stock by industry are available from the US Bureau of Economic Analysis. However BEA statistics are based on national definitions that differ from the internationally accepted system of national accounts and are not comparable with the data of other countries. I have therefore chosen to use only the statistics for the USA provided by international sources such as the OECD, which follow the internationally comparable definitions, at the cost of unavailability of the industry capital stock series and the availability of other series only since 1987. The operating surplus in Figures 9.5 and 9.6 is not adjusted (see note 1 above) and hence comprises the *entire* incomes of self-employed workers (mixed incomes).

References

Barba, A. and de Vivo, G. (2012) 'An "unproductive labour" view of finance', *Cambridge Journal of Economics*, 36(6): 1267–90.

Bluestone, B. and Harrison, B. (1988) *The Great U Turn – Corporate Restructuring and the Polarizing of America* (New York: Basic Books).
Bronfenbrenner, K. (2001) *The Effects of Plant Closing or the Threat of Plant Closing on the Right of Workers to Organize*, Report to the US Department of Labour.
Ciccone, R. (1990a) 'Accumulation and Capacity Utilization: Some Critical Considerations on Joan Robinson's Theory of Distribution', in K. Bharadwaj and B. Schefold (eds), *Essays on Piero Sraffa – Critical Perspectives on the Revival of the Classical Theory* (London: Unwin Hyman): pp. 417–29.
Ciccone, R. (1990b) 'A Comment to Pivetti', in K. Bharadwaj and B. Schefold (eds), *Essays on Piero Sraffa – Critical Perspectives on the Revival of the Classical Theory* (London: Unwin Hyman).
Choi, M. (2001) *Threat Effect of Foreign Direct Investment on Labor Union Wage Premium*, (Amherst MA: PERI Working paper no 27), http://www.peri.umass.edu/fileadmin/pdf/working_papers/working_papers_1-50/WP27.pdf, accessed 25 January 2013.
Garegnani, P. (1970) 'Heterogeneous capital, the production function and the theory of distribution', *Review of Economic Studies*, July, 37(3): 407–36.
Garegnani, P. (1983) 'Notes on consumption, investment and effective demand', *Cambridge Journal of Economics*, 1978, 2: 335–53, repr. in J. Eatwell and M. Milgate (eds), *Keynes's Economics and the Theory of Value and Distribution* (London: Duckworth).
Garegnani, P. (1992) 'Some Notes for an Analysis of Accumulation', in J. Halevi, D. Laibman and E. Nell (eds) *Beyond the Steady State* (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, UK and Lyme, USA: Edward Elgar).
Gordon, R.J. (1997) 'The time-varying NAIRU and its implications for economic policy', *Journal of Economic Perspectives*, 11(1): 11–32.
Greenspan, A. (1997) 'Testimony of Chairman', *The Federal Reserve's semi-annual monetary policy report, Before the Committee on Banking, Housing, and Urban Affairs, U.S. Senate,* July 22, 1997, http://www.federalreserve.gov/boarddocs/hh/1997/july/testimony.htm.
Hein, E. (2011) 'Distribution, "Financialisation" and the Financial and Economic Crisis – Implications for Post-Crisis Economic Policies', MPRA Paper No. 31180, http://mpra.ub.uni-muenchen.de/31180/
Hein, E. and Schoder, C. (2009) *Interest Rates, Distribution and Capital Accumulation – A Post Kaleckian Perspective on the US and Germany*, Institute for International Political Economy, Working Paper No. 04/2009.
IMF (2007) *World Economic Outlook. Spillovers and Cycles in the Global Economy*, April, Washington, http://www.imf.org/external/pubs/ft/weo/2007/01/pdf/text.pdf.
Kaldor, N. (1956) 'Alternative theories of distribution', *Review of Economic Studies*, 23: 83–100.
Krugman, P. (2007) *The Conscience of a Liberal* (New York-London: Norton & Co.).
Levrero, E.S. and Stirati, A. (2006) 'The Influence of Unemployment, Productivity and Institutions on Real Wage Trends: The Case of Italy 1970–2000', in E. Hein, A. Heise and A. Truger (eds) *Wages, Employment, Distribution and Growth – International Perspectives* (Basingstoke: Palgrave Macmillan).

Panico, C. (1988) *Interest and Profit in the Theory of Value and Distribution* (Basingstoke: Macmillan Press).
Panico, C., Pinto, A. and Anyul, M.P. (2012) 'Income distribution and the size of the financial sector – a Sraffian analysis', *Cambridge Journal of Economics*, 36(6): 1267–90.
Pasinetti, L.L. (1962) 'Rate of profit and income distribution in relation to the rate of economic growth', *Review of Economic Studies*, XXIX (4), October: 267–79.
Patinkin, D. (1987) 'Real Balances', in J. Eatwell, M. Milgate and P. Newman (eds) *New Palgrave: A Dictionary of Economics* (London: Macmillan).
Petri, F. (2004) *General Equilibrium, Capital and Macroeconomics* (Cheltenham: Edward Elgar).
Piketty, T. and Saez, E. (2006) 'The evolution of top incomes: a historical and international perspective', *American Economic Review*, 96(2): 200–5.
Pivetti, M. (1991) *An Essay on Money and Distribution* (London: Macmillan).
Pollin, R. (2003) *Contours of Descent – U.S. Economic Fractures and the Landscape of Global Austerity* (London & New York: Verso).
Ricardo, D. (1951 [1821]) *The Principles of Political Economy*, in P. Sraffa (ed.) with the collaboration of M. Dobb, *Works and Correspondence*, Vol I. (Cambridge: Cambridge University Press for the Royal Economic Society).
Rowthorn, R. (1999) *Unemployment, Capital-Labour Substitution and Economic Growth*, IMF Working Papers, no 99-43.
Smith, A. (1976 [1776]) *An Inquiry into the Nature and Causes of the Wealth of Nations*, R.N. Campbell, A.S. Skinner and W.B. Todd (eds), *The Glasgow Edition of the Works and Correspondence of Adam Smith* (Oxford: Oxford University Press).
Sraffa, P. (1960) *Production of Commodities by Means of Commodities. Prelude to a Critique of Economic Theory* (Cambridge: Cambridge University Press).
Stirati, A. (1992) 'Unemployment, institutions and the living standard in the classical theory of wages', *Contributions to Political Economy*, 11: 41–66.
Stirati, A. (1994) *The Theory of Wages in Classical Economics. A Study of Adam Smith, David Ricardo, and their Contemporaries* (Aldershot: Edward Elgar).
Stirati, A. (1999) 'Ricardo and the Wages Fund', in G. Mongiovi and F. Petri (eds), *Value, Distribution and Capital: Essays in Honour of Pierangelo Garegnani* (London: Routledge).
Stirati, A. (2001) 'Inflation, unemployment and hysteresis: an alternative view', *Review of Political Economy*, 13: 427–51.
Stirati, A. (2010) 'Changes in Functional Income Distribution in Italy and Europe – Service Sector Prices, Labour Market Conditions and Institutional Change', in E. Brancaccio and G. Fontana (eds), *The Global Economic Crisis: New Perspectives on The Critique of Economic Theory and Policy* (London: Routledge).
Stockhammer, E. (2004) 'Financialisation and the slowdown of accumulation', *Cambridge Journal of Economics*, 28(5): 719–41.
Stockhammer, E. (2009) *Determinants of Functional Income Distribution in OECD Countries*, IMK studies, September.

10

Sraffa and Keynes: Two Ways of Making a 'Revolution' in Economic Theory*

Guglielmo Chiodi and Leonardo Ditta

10.1 Introduction

This chapter aims to show how a true 'revolution' in economic theory during the last century can be attributed jointly to Sraffa and Keynes. Despite following different research paths, both went deeply into the foundations of established economic theory, to the point of subverting, within their own framework, traditional ways of thinking about and representing the functioning of an economic system. We will also reflect on the persistence, despite Keynes's and Sraffa's corrosive critiques, of the very uncomfortable state of contemporary economic theory.

Having set up a consistent logical framework firmly rooted in the classical tradition, Sraffa laid great emphasis on *non-market institutions* and *non-market values,* and it is with this aspect of Sraffian theory that the first part of the chapter will be mainly concerned, specifically with the crucial role played by the notion of subsistence in the viability of a system, an analysis of its Smithian roots and the importance of that notion in present times.

Keynes's focus on the determination of aggregate output also represents a 'revolution', contrasting sharply with the then prevailing (Robbins's) view of the supposed irrelevance of the social product, a central concept in the analysis of the old classical economists. Keynes's approach is also 'revolutionary' for the centrality of effective demand he assumed in determining the level of aggregate output – the relevant

*We are grateful to an anonymous referee for valuable comments on an earlier version.

magnitudes being fixed *outside the market mechanism*. The decisions to be taken in a system characterised by the pervasive influence and power of monetary institutions can never be regarded as conducive to an 'equilibrium' of the type contemplated by traditional theory. The second part of the chapter will be addressed to this peculiarity of Keynes's contribution and to all its possible links to the classical and to Sraffa's approach.

Both Keynes and Sraffa developed their approach starting from a powerful critical refutation of the marginalist orthodox theory; they chose, however, to attack two different, although both highly crucial, aspects of this theory, namely the determination of the level of the social product (national income or total output) and its distribution among the factors of production.

Keynes rejected the idea that a market capitalistic economic system could of itself achieve the full employment of productive resources, labour and capital stock. There is no natural tendency toward full employment; on the contrary, in a modern economy, that is in a *monetary production economy*, where processes are time-consuming and decisions must be taken in conditions of absolute uncertainty, unemployment is the natural outcome of *laissez-faire*.

Sraffa devoted his efforts to criticising the marginalist theory of value and distribution, thus taking as given the quantities of goods comprising the social output. He criticised and rejected the notion of marginal productivity of factors, showing that there is no way of determining the share of social income accruing to factors of production independently of prices; and, demonstrating the absence of a monotonic inverse relationship between profit rate and capital intensity, he undermined the basis of the marginalist theory.

Moreover, both authors stressed the relevance of the monetary rate of interest. Sraffa suggested that the latter could be taken as the 'given' for the closure of the system. Keynes contested the marginalist view of the rate of interest as a 'real' variable, capable of balancing the demand for investment with the supply of saving, stressing the monetary and institutional nature of that rate; he was concerned with the necessity of bringing the rate of interest down so that unemployment could be avoided in the long term. He also maintained that no 'objective' motive justifies a high rate of interest and the related return on capital, given that a strong effective demand rather than the rate of interest provides the stimulus for accumulation of capital.

We maintain overall that in the orthodox theory the free operation of markets is at the core of both the determination of output level and

the distribution of income, whereas neither in Keynes nor in Sraffa does the market play such a determinant role.

10.2 Political, social and ethical aspects in Sraffa's *PC*

As is well known, the subtitle of Sraffa's *Production of Commodities by Means of Commodities* (*PC*) is *Prelude to a Critique of Economic Theory*.

A great deal of work has been done along the lines suggested by Sraffa to address all the *logical* faults of the traditional analytical framework and to reconstruct an alternative economic theory with the old classical economists' approach as its frame of reference. The 'first phase' – to use Walsh's (2000, p.5) terminology – of the revival of the old classical political economy has already been completed, and its ultimate aim has surely been achieved, *viz*. that of showing how feeble were the logical premises on which traditional theory based its propositions and prescriptions.[1]

The neoclassical paradigm, however, is not only still alive but also as dominant as ever, in spite of the contrast between world economic facts and its theoretical properties.[2] The inevitable suspicion arises that the *ideological* content embedded in and conveyed by the neoclassical paradigm is far stronger than the *analytical framework* which supports that paradigm itself. We want to argue, in particular, that the ideological content of neoclassical theory might be concisely expressed by strong confidence in the market mechanism and the belief that the market is the most suitable institution to regulate all the relevant relations among people.

The first part of this chapter aims to explore Sraffa's prelude to a critique of economic theory from a different perspective from that taken during the first phase of the revival of classical theory, aiming to better accomplish that revival by examining Sraffa's neglect of *the market* as the pivotal institution for the working of an economic system. We shall explore aspects of Sraffa's approach underlying its analytical 'core' that are mainly concerned with political, social and ethical aspects of the life of a community – issues generally supposed to be *outside* the realm of 'economic' consideration. We firmly believe that these aspects are indissolubly intertwined with the 'economic' ones, as did the old classical economists (Smith, 1970; Ricardo, 1970; and also Marx, 1973 and 1974).

The present contribution should thus be considered as naturally belonging to what Walsh (2000) once again termed the 'second phase' of the classical theory revival.[3]

10.3 Two fundamental features of *PC*: subsistence and self-reproduction

Sraffa's *PC* opens with two fundamental features in the form of tables of numbers representing quantities of commodities used and produced in the production processes:

(i) the inclusion of the sustenance for the workers among the commodities used for production;
(ii) the self-reproduction of the system as a whole, i.e., as a complex social as well as economic organism, as the ultimate objective to be attained.

In fact, throughout the book, workers' subsistence is completely *unrelated* to any sort of result of the production processes. This provides the strongest link to the classical *theory* of value and distribution, and is very helpful in seeing the Sraffian framework as an *alternative* paradigm to the neoclassical one. Sustenance for the workers is *expressed*, commodity by commodity, in physical terms. Its quantity and quality, however, are *determined* neither simply nor univocally – whether by reference to strictly physiological conditions or, more generally, to social conditions. On reflection, in fact, the commodity set which constitutes that sustenance for the workers can only be expressed through *value judgements*. Thus, for example, there is no unequivocal *a priori* solution to the problem of determining the most 'appropriate' diet for a human being, of a given community at a given moment in time. Is the diet appropriate for a 'long' or for a 'short' life? What are the 'strictly necessary' consumption goods? What is the most 'appropriate' ratio of 'biologically indispensable' goods to 'spiritually indispensable' goods?[4]

10.4 The representation of the economy as involving value judgements

Sraffa's representation of the economy does not seem value neutral, but on the contrary, the outcome of a process involving profound value judgements. Hence it is impossible to regard the quantities of the commodities he presents merely as 'observable' data. In other words, to borrow Putnam's (2003) terminology 'facts' are not separated from 'values', since both are unavoidably 'entangled'.[5]

Within the neoclassical paradigm, one of the effects produced by the belief that 'facts' could be so easily separated from 'values' has been

that of facilitating an analogous process of a neat separation of the 'economy' *as such* from the more general system. An immediate consequence of this process has been the further belief that the 'economy', as an *autonomous entity*, could be governed by laws of its own, uncontaminated by any value judgement.[6]

A second, no less important effect, has been the belief that 'economics' has the same sort of status as the so-called 'hard' sciences, in which a separation between 'facts' and 'values' was *generally* thought to be possible – which greatly increases the disparity between Science and Ethics, as any value judgement automatically lay *outside* the bounds of theory because it was thought to belong to the *non*-scientific sphere. Since the early 1930s, when it was tenaciously pursued by Robbins, turning economics into a value-free discipline has pervasively permeated the entire neoclassical paradigm.[7]

10.5 The crucial notion of 'viability'

We can now begin to see that Sraffa's representation of the economy is far from being as 'simple' as it first appears, not least because it is *not* value-free, with points (i) and (ii) above (see p. 221) serving to illustrate its inner complexity.

As regards the inclusion of workers' subsistence, it should be noted that the *whole* set of commodities used and produced that represent the economy cannot be separated conceptually into distinct and independent parts. The *quantitative relations* among those commodities *as a whole* establish whether the system is 'viable' or not, that is, whether it is *capable* of self-replication.

The notion of 'viability' in Sraffa is of crucial importance, as will shall see, for it allows, in the first instance, the unambiguous selection of those systems worthy of analytical consideration.[8] The key factors in a system's 'viability' are what Sraffa concisely calls 'methods of production and productive consumption', comprising workers' subsistence, the means of production and the quantities of commodities produced. Consequently, the 'viability' of a system, which depends on the same set of factors as workers' subsistence and the means of production, is but a reflection of value judgements.

Among 'viable' systems, Sraffa makes a distinction between those producing for subsistence and those producing with a 'surplus'. A 'surplus', which a given community could have at its disposal at any time, cannot be viewed as a 'gift' of nature, nor, even worse, as the result of some 'magic'. More seriously, a 'surplus', like 'viability', should be viewed

as a *historically* determined set of commodities, in the sense already explained above.[9]

10.6 Two features of the 'core'

The reproducibility of a system as the economy's ultimate objective is the other crucial aspect of Sraffa's *PC*. Viability essentially means the *possibility* of human beings *surviving* and reproducing themselves *as a community*. Thus, viability means *social reproduction*. Remarkably, Sraffa is *explicit* at the very beginning of his work that this is the *natural* objective of any community. Note that neither von Neumann (1937) nor Leontief (1941) seems to emphasise this important aspect.

For the system to be viable, the commodities produced need to be exchanged in certain proportions. The primary role of the exchange ratios within a specific economy is to make viability *effective*. The classical economists and Marx as well as Sraffa called them 'necessary' or 'natural' because of their close links to viability, which is the *natural* end of any community of human beings, and here two fundamental features of the 'core' of Sraffa's system should be noted.

One is that the economist's task is not only performed *within* the boundaries of the 'core'; it must also be performed *outside* them. From a different perspective, Sraffa's 'core' cannot be considered simply as a set of *logically* consistent relations between variables, which could obviously be formally analysed by any mathematician. The economist, as a *social* analyst, also has to perform a *semantic* task of *interpreting* and *shaping* the very structure of the 'core'. Take initially any *viable* system characterised by given methods of production. For the system to be designed so as to produce with *no* 'surplus', the workers' subsistence needs to be *determinant* of the structure of the system *jointly and simultaneously* with the methods of production. That same system, however, could alternatively be designed to produce a 'surplus'. In such a case, however, workers' subsistence must be downgraded to the point of being both quantitatively and qualitatively worse than in the former case, and in such a way as *to be accommodated* to the methods of production. In this way workers' subsistence is viewed as a *determined* variable instead of being a determinant of viability – a view which cannot be reconciled with the classical theory of value and distribution and *a fortiori* with Sraffa's. These alternative perspectives, therefore, can only reflect *alternative* value judgements.[10]

The second feature of the 'core' is that it does not contemplate *the market* in any form or function whatsoever. Within that framework it

appears an *unnecessary institution* – in the precise sense in which Sraffa intends the context in which production prices appear 'which contains no reference to market prices', (Sraffa, 1960, p. 9).[11] This should be strongly contrasted with the neoclassical paradigm for several reasons.

Sraffa was in fact conceiving and writing *PC* at a time when Logical Positivism was smoothly flowing from philosophy to economics, until it assumed its most popular form in Robbins' 1932 essay. Over that same period, a refined mathematical work on the Walrasian General Equilibrium Theory was in progress, culminating in several important works,[12] with the 1939 elaboration by Hicks in *Value and Capital* making the most robust form of the neoclassical paradigm accessible to a far wider range of scholars. When *PC* was finally published, the basic structure of the Walrasian General Equilibrium Theory had already assumed its definitive form through the works of Arrow and Debreu (1954) and Debreu (1959).

Thus, Sraffa's pervasive and influential alternative paradigm appeared at the high point of the neoclassical paradigm, and around the time that a harsh and successful attack on Keynes's *General Theory* had already been launched.

The idea that a *market economy* must be considered the most 'efficient' system seemed absolutely triumphant until the 1960s. Moreover the notion of 'equity' could also be inferred from the notion of 'efficiency', in so far as the allocation of the given resources was supposed to follow 'objective' criteria and the distribution of income was supposed to take place according to the *contribution* of the respective 'factors' of production.[13]

10.7 The 'core' as only a part of a wider social system

Sraffa's alternative paradigm, with its *new conception* of the economy as a *self-reproducing system* thus presented a very great challenge to the then orthodoxy. Self-reproduction of the system means the reproduction of human beings as a society, with no room for the market, so it is by definition *holistic* or *macro*. The seeds of Sraffa's 'revolution' lie in this very *new conception* of the economy and are reflected in the make-up of the 'core' of his theory.

The analytical framework of *PC*, based (certainly not by chance) on the old classical economists' approach, precisely reflects Sraffa's new 'vision', so different from the neoclassical one. The neoclassical view of the economy as an aggregation of 'monads' with boundless wants, each pursuing some individual 'maximising' objective, requires an institution

capable of making each individual plan compatible with any other, so that exchange relations exist solely for the achievement of each individual's objective, and the market is supposed to determine the exchange ratios by following the fundamental criterion of the relative *scarcity* of each commodity. The neoclassical 'economy' is completely separated from any other part of the system, having its own laws and its own self-regulating mechanisms, with no reason to refer to any value judgements, and without even an indirect connection with *the life* of a society of human beings.

In contrast, Sraffa's 'core' depicts an economic system which is only *a part* of a wider social system. There is nothing within the 'core' that can be determined *unless* 'something else' of importance has been already determined from *outside* the system (and, in some circumstances, even *before* the production processes get started). The most representative example, and one of the most important prerequisites for production and thus for the viability of the system as a whole, is workers' subsistence, which is thus completely *unrelated* to any result of the production processes.

Even in those systems producing with a 'surplus', the 'surplus' wage, which the workers might possibly get in addition to their subsistence, can only be conceived of in value terms, and it must ultimately be the end result of some sort of 'conflict' between the classes sharing with the 'surplus'.

It is important to note the 'openness' of Sraffa's analytical framework, which enables alternative value judgements to be incorporated in the analysis. For example, a claim could be made on the 'morality' of a certain state of income distribution or on certain ethical aspects of it. This links directly to Adam Smith, who in the final chapters of *The Wealth of Nations* refers explicitly to ethical and moral elements, connected to the *freedom* and *dignity* of any one labourer as a *person*:

> By necessaries I understand not only the commodities which are indispensably necessary for the support of life, but whatever the custom of the country renders it *indecent for creditable people, even of the lowest order, to be without*. A linen shirt, for example, is, strictly speaking, not a necessary of life… But in the present time, through the greater part of Europe, a creditable day-labourer would be ashamed to appear in public without a linen shirt. (Smith, 1952, pp. 351–2)[14]

Also, a new list of 'priorities' for commodities entering workers' subsistence can be defined – in contrast with the artificial postulate of

commodity 'substitution' typical of the neoclassical theory. In this way, 'conflicts' within the community can be viewed as a *natural* means of making claims and explicit value judgements on the rules and the procedures for achieving definite goals.

10.8 Some generalizations

The two features of Sraffa's *PC* mentioned at the beginning of this chapter, namely (i) the inclusion of sustenance for the workers among the commodities used for production and (ii) the self-reproduction of the system as the ultimate objective of the economy, can both be generalised within Sraffa's analytical framework.

The starting point for such a generalisation might be that of examining more deeply the notion of self-reproduction of the system. On further reflection, self-reproduction implies *continuation*; and for the system to *continue over time* it is absolutely necessary, in the first instance, to provide for the replacement of the labour force. This obviously requires children to be brought up *today* so they are fit to enter the labour force *tomorrow*. Ricardo defines the wage as '[t]he natural price of labour … that price which is necessary to enable the labourers … *to subsist and to perpetuate their race* … [it] depends on the price of the food, necessaries, and conveniences required for the support of the labourer and *his family*' (Ricardo, 1970, p. 93, italics added).

Moreover, since all may grow old or become disabled, there are strong ethical grounds for 'feeding' these people also – regardless of the fact that they respectively do not, or cannot, belong to the labour force. Therefore, in addition to the workers directly employed in the production processes, workers' families have *a right* to sustenance. This right, however, though expressible *individually,* refers of its own nature to the *social* requirement of the self-reproduction of the system. Were it disappointed, the viability of the system would be jeopardised.

A further generalisation might even be to introduce into Sraffa's framework the possibility of granting a Universal Basic Income (UBI).

The UBI is a monetary income paid by a government at a uniform level and at regular intervals to each adult member of society (Van Parijs, 2001, p. 5).[15] It is given *independently* of the economic and social condition of the people involved, as well as *independently* of their willingness to work. It is thus a *universal* and *unconditional* income, with no link whatsoever to work and production.

Its roots dated back to Thomas More's *Utopia* of 1516, in which the idea of any person being provided with her/his own sustenance is explicitly

stated. The crucial event, however, is the well-known Speenhamland Law, introduced in England in 1795 during the Industrial Revolution. This law laid down the principle of assuring a minimum income to the poor *irrespective of their earnings*, in an environment in which a dramatic transformation was taking place for labourers: their labour force was becoming a *commodity* in the market. The immediate and obvious implication of this was the neat separation between the wage, on the one hand, and the necessaries of life, on the other – a separation which had no reason to exist before the Industrial Revolution.

Sraffa's analytical framework and his whole conception of the economic system, as we have seen, can accommodate and give strong support to such generalisations – which, by contrast, can find no place whatsoever within the neoclassical paradigm, which generally ignores the notions of 'subsistence' and the 'right to live'.[16] And, worse than this, the crucial step which distances the neoclassical theory still further from the classical and Sraffian approach is the view of the labour force not only as a commodity priced on the market on the basis of its relative *scarcity*, but also as a commodity having its own exchange value strictly linked to its corresponding *contribution* to production, which – as Sraffa has definitely shown – is in any case logically and conceptually *indefinable*.

10.9 'Classical' features of Keynes's *GT*

> The composition of this book has been for the author a long struggle of escape, and so must the reading of it be for most readers … a struggle of escape from habitual modes of thought and expression … The difficulty lies, not in the new ideas, but in escaping from the old ones, which ramify, for those brought up as most of us have been, into every corner of our minds. (Keynes, 1936, p. viii)

It is quite paradoxical that today, with the world economy deeply immersed in the worst economic and social crisis since the 1930s, the remedies are sought in the 'old ideas' Keynes struggled to escape from when writing *The general theory of employment, interest and money* (hereafter *GT*) against the orthodox economic theory of his time.

We do not dispute that there may be many flaws in *GT*, or that new and different problems have arisen since its publication. Nonetheless the 'Keynesian revolution', although it can be 'unaccomplished' (borrowing Pasinetti's terminology),[17] can still provide a guide to the economic problems of the real world. It can provide a better guide indeed if compared to the still prevalent orthodoxy based on axioms and

unrealistic assumptions in which the market is idolised as the indisputable institution governing economic systems.

The present section of this chapter is connected with the previous section in two ways: i) one of the main features of Keynes's *GT* is that all the relevant variables are determined *outside the market and independently of any market mechanism*; ii) ethical, philosophical and social elements of economic life make it impossible to avoid value judgement, thus the discipline of economics cannot share the same (value-free) neutrality as is attributed to the 'hard sciences'.

To begin with, we attempt, paraphrasing Keynes's own words reported above, to escape not only from the modes of thought that preceded the *General Theory*, but also from the immense literature that has followed its publication. Our intention is to focus on the peculiarities of *GT* as they would have appeared to our eyes if we had read it in 1936. We are aware that this sort of experiment is unlikely to succeed, because try as we will, in reading the *GT* today we cannot erase all our previous readings and experiences from our mind. We aim, however, to focus solely on *GT*, disregarding criticisms and interpretations put forward by generations of economists over a period of 75 years. The point of view taken by Keynes at the start of his investigation is very 'revolutionary', if we recall that in the 1930s Keynes turned his own ideas regarding economic theory on their heads. At around this time, in his milestone *Essay on the Nature and Significance of Economic Science,* Lord Robbins confined the subject matter of economics to the narrow boundaries of individual choices under given constraints in a world of scarcity, also limiting the scope of the classical concept of Social Product, and of any social aggregate. Indeed, given the nature of the price system stemming from a set of ordered preferences of single agents, Lord Robbins wrote:

> the addition of prices or individual incomes to form social aggregates is an operation with a very limited meaning. (Robbins, 1932, p. 57)

It can be safely stated that Lord Robbins' book paved the way to a methodological turn by which the 'scientific approach' of physics was to be transferred to economics. The method, which gained increasing momentum, consisted essentially of deducing implications from certain postulates such as utility maximisation or the equalisation of supply and demand through the operation of market forces, so stressing the fundamental role of the market. The laws of physics, however, have no counterparts in economics, because of the very nature and uncertainty of economic interactions involving human beings.[18] Despite Keynes's

declared intention to contrast 'classical theory' (a note on p. 3 of *GT* informs us that 'classical' includes 'the *followers* of Ricardo, J.S. Mill, Marshall, Edgeworth and Prof. Pigou'), his first step is instead to adopt Social Aggregates and particularly the Social Product as the main subject of his investigation, like the old classical economists. Indeed we can consider Keynes's macroeconomic approach as the economics of the Social Product, as opposed to the economics of individual choices. Keynes's abandonment of the traditional way of thinking of economics as centred upon exchange in the market in favour of a vision in which production and employment were the central elements of an economic and *social* system, as in classical political economy, implies a change of paradigm. As Pasinetti puts it:

> The hard decision … was taken to break with orthodoxy. The change consisted … in shifting the whole body of economic elaborations away from the traditional 'exchange paradigm' foundations on which they had been laid to the alternative foundations of a 'production paradigm'.[19]

A further consideration, in our view, is that Keynes's emphasis in *GT* on the importance of unemployment can be viewed in relation to the reproducibility of the system as a social organism. The following quotation makes this point clear:

> The outstanding faults of economic society in which we live are its failure to provide for full employment and its *arbitrary and inequitable distribution of wealth and income*. (*GT*, p. 372, italics added)

Increasing unemployment and inequality in income distribution and wealth are both elements that can lead to social unrest, as well as being responsible for destroying the social contract. Thus full employment must be centre stage and the main objective of economic policies.

10.10 Keynes's effective demand and the absence of any market mechanism

The second step contrasts the belief that economic phenomena can be explored through models imported from the hard sciences, like physics, with the pervasive uncertainty, in the sense stressed by Keynes himself, which prevents these phenomena from being reduced to any kind of probabilistic model.

The aim of our brief excursus through the *GT* is to emphasise that the 'revolutionary' aspect of Keynes's theory consists, in our view, in his strong belief that the determination of the level of aggregate output depended crucially on effective demand, with the relevant magnitudes believed to be fixed *outside the market and independently of any market mechanism*. This view was in contrast to the one brought about by the growing importance, at that time, of market equilibrium analysis, where all the variables are simultaneously determined. The decisions to be taken in a system characterised by the pervasive influence and power of monetary institutions can never be thought of as conducive to an 'equilibrium' of the type contemplated by traditional theory.

We can find relevant and meaningful passages of the *General Theory* to support these statements. In chapter 3 of book I, devoted to the celebrated Principle of Effective Demand, Keynes lays down the basis of his theory:

> The amount of labour N which the entrepreneurs decide to employ depends on the sum (D) of *two* quantities, namely D_1, the amount which the community is expected to spend on consumption, and D_2, the amount which it is expected to devote to new investment. D is what we have called above the *effective demand*. (*GT*, p. 29, emphasis added)

Having stated that increases in employment will raise consumption expenditure D_1, but not to the level required to match the 'aggregate supply price of the corresponding output (Z)', (giving a falling propensity to consume) he concludes by saying that 'the greater the volume of employment the greater will be the gap between aggregate supply and consumption. So the amount of investments D_2 must increase to fulfil the increasing gap Z-D_1' (*GT*, p. 30). Thus

> – except on the special assumptions of the classical theory according to which there is some force in operation which, when employment increases, always causes D_2 to increase sufficiently to fill the widening gap between Z and D_1 – the economic system may find itself in stable equilibrium with N at level below full employment … The propensity to consume and the rate of new investment determine between them the level of employment … If the propensity to consume and the rate of new investment result in a deficient effective demand, the actual level of employment will fall short of the supply of labour potentially

> available at the existing real wage … The insufficiency of effective demand will inhibit the process of production in spite of the fact that the marginal product of labour still exceeds in value the marginal disutility of employment. (*GT*, pp. 30–1)

10.11 The rate of interest as a purely monetary phenomenon

These concise sentences clearly lay down the gist of the Keynesian new approach. In the subsequent books III and IV (chapters 8–13 of *GT*) the propensity to consume and the inducement to invest are further explored.

The well-known result of this discussion is that the aggregate consumption expenditure C can be considered a quite stable function of aggregate income Y, with $0 < \Delta C/\Delta Y < 1$. On the other hand, the level of investment fluctuates and is difficult to predict, as it depends upon expectation that cannot be easily introduced into the scheme. In *GT* (chap. 11) the scale of investment is deemed to depend on the relation between the rate of interest and the schedule of marginal efficiency of investment, which in turn depends on the relation between the supply price of a capital asset and its expected return. Two chapters (12 and 13) are devoted respectively to long-term expectations and to the theory of interest. In chapter 13 the liquidity preference appears as a key element of the general theory of interest.

The interest rate is defined as 'the reward for parting with liquidity', and

> the rate of interests is not the 'price' which brings into equilibrium the demand for resources to invest with the readiness to abstain from present consumption. It is the 'price' which equilibrates the desire to hold wealth in the form of cash with the available quantity of cash. (*GT*, p. 167)

Thus, Keynes says, the rate of interest, in given circumstances, is determined by liquidity preference and the quantity of money. In this way money is brought in as a main factor influencing real variables. As is well known, Keynes rejected both the characteristics which orthodox theory attributed to money, namely neutrality and dichotomy.

Keynes's next step is to explain why, given that the rate of interest is positive, anyone should find it convenient to hold his balance in cash instead of in a form that yields interest. Here Keynes introduces another

'revolutionary' concept, stating the *necessary condition* for the existence of liquidity preference:

> There is, however, a necessary condition failing which the existence of a liquidity-preference for money as a means of holding wealth could not exist.
>
> This necessary condition is the existence of *uncertainty* as to the future of the rate of interest, i.e. as to the complex of rates of interest for varying maturities which will rule at future dates. (*GT*, p. 168, italics in the original)

Later on, in chapter 15, Keynes explores the incentives to liquidity preference or demand for money, both from the psychological and the business perspectives. In his analysis *there is no market mechanism* able to generate an equilibrium rate of interest (as the pre-Keynesian theory would postulate) to make full employment correspond to the level of effective demand. As Shackle puts it:

> Liquidity preference destroys the hydraulic theories of money and interest, and shows how and why *valuations* can change without the occurrence of *transactions*. (Shackle, 1973, p. 517, italics in the original)

So there are three independent parameters behind the determination of effective demand: i) the propensity to consume; ii) the marginal efficiency of investment; and iii) the liquidity preference and bank policy, jointly influencing the expected rate of interest. These three parameters relate to psychological or behavioural aspects of human life.

10.12 Income and employment as determined by elements lying outside the market

Without entering into the countless controversies and discussions over all these very important issues, we want to stress the common feature underling all of them: that is, that none of them can be juxtaposed to the operation of the markets.

Thus, the conclusion is that income and employment determination do not depend on any form of *market mechanism*.

It is also important to note that in Keynes's construction, all the relevant variables are characterised by instability, although each to a different degree.

This becomes clearer in chapter 18, 'The General Theory restated'. Here Keynes summarises his theory as presented in the previous chapters: i) the physical conditions of the capital goods industries; ii) the state of confidence regarding the expected returns; and iii) the propensity determining the liquidity preference and the money supply, all determine the rate of new investment (*GT*, p. 248).

A change in the rate of investment, through a change in the level of income, will produce a change in the same direction in aggregate consumption of income but not of the same amount. The increase (decrease) as a proportion of aggregate income is given by the operation of the investment multiplier. But this will also reflect on the schedule of liquidity preference, because 'an increment (or decrement) of employment is liable, however, to raise (or lower) the schedule of liquidity preference' (*ibid.*).

Thus the demand for money tends to change, even at stable wages unit and prices:

> but in addition the wage unit itself will tend to rise as employment improves, and the increase in output will be accompanied by a rise of prices (in terms of the wage unit) owing to increasing cost in the short period.
>
> Thus the position of equilibrium will be influenced by these repercussions; and there are other repercussions also. Moreover, there is not one of the above factors which is not liable to change without much warning, and sometimes substantially. Hence the extreme complexity of the actual course of events. (*GT*, p. 249)

This shows how difficult is to precisely describe and grasp in a theoretical model the variables involved in economic life. However Keynes is not suggesting giving up any attempt to understand the economic phenomena. In fact he adds:

> Nevertheless, these seem to be the factors which it is useful and convenient to isolate…and our practical intuition (which can take account of a more detailed complex of facts than can be treated in general principles) will be offered a less intractable material upon which to work. (*Ibid.*)

The remainder of the chapter is devoted to a discussion of the hypothetical propensities that could lead to a stable system. It finishes with

a strong statement of the need to attempt to control such tendencies as they are not 'a necessary principle which cannot be changed' (*GT*, p. 254).

10.13 Heterodox aspects of Keynes's *GT*

So far we have a theory that explains the level of output and employment as determined by exogenous parameters; full employment equilibrium cannot thus be restored by simply removing rigidities of any kind. A further point concerns the ethical dimension of economic life.

Keynes's approach to this point was already very explicit in his early writings on economics. He was concerned with the evolution of the economy over time and with the ethical implications of this evolution; to him economic growth was not an end in itself, but just a means. His approach thus differed from the utilitarianism-based one of neoclassical economics.

In 'Am I a Liberal?' (in *Essays in Persuasion*, 1931, originally published in 1925) Keynes wrote:

> The transition from economic anarchy to a régime which deliberately aims at controlling and directing economic forces in the interests of social justice and social stability, will present enormous difficulties both technical and political. I suggest, nevertheless, that the true destiny of New Liberalism is to seek their solution. (Keynes, 1972, p. 305)

In 'Liberalism and Labour' in the same volume he added:

> The political problem of mankind is to combine three things: Economic Efficiency, Social Justice, and Individual Liberty. The first needs criticism, precaution, and technical knowledge; the second, an unselfish and enthusiastic spirit which loves the ordinary man; the third, tolerance, breadth, appreciation of the excellencies of variety and independence, which prefers, above everything, to give unhindered opportunity to the exceptional and to the aspiring. (*Ibid.*, p. 311)

The last chapter of *GT* (chapter 24), which to our knowledge has been generally ignored by our profession, is entitled 'Concluding Notes on the Social Philosophy towards which the General Theory Might Lead'. It opens with the strong statement of ethical nature quoted on page 227 regarding unemployment and distribution.

He notes the relevance of the theory presented in the previous chapters to the first fault mentioned, unemployment. Two aspects of the

theory, however, are also relevant to the second one. The first relates to the orthodox belief according to which the accumulation of capital, and thus the growth of output, is favoured by a high proportion of saving out of income, i.e. on 'the savings of the rich out of their superfluity'. But we saw that, on the contrary, the arguments presented in *GT* are conducive to the conclusion that, unless we are in full-employment equilibrium, investment is not encouraged by a low propensity to consume but is impeded by it. A strong dynamic of demand is probably, all other things being equal, to strengthen the inducement to invest. So, as Keynes adds:

> [M]easures for the redistribution of incomes in a way likely to raise the propensity to consume may prove positively favourable to the growth of capital … Thus our argument leads towards the conclusion that in the contemporary world the growth of wealth, so far from being dependent on the abstinence of the rich, as is commonly supposed, is more likely to be impeded by it. One of the chief social justifications of great inequality of wealth is, therefore, removed. (*GT*, p. 373)

The second aspect, much more fundamental as far as the distribution of income is concerned, derives from the theory of interest presented in *GT*. In the orthodox theory a relative high rate of interest is justified by the need to provide a sufficient stimulus to save. But, as we have just seen, savings are bound to be determined by investments and not the other way around. As the inducement to invest is also influenced by a low rate of interest, as long as we are not in a full-employment position, increases in saving (lower consumption expenditure) are not necessary to achieve a higher level of investment. So the very conclusion of the old theory is completely reversed and there is no 'objective' incentive for a high rate of interest and the related return on capital.

A further strong attack on the ethical basis of neoclassical theory of distribution follows with the following propositions:

> [T]he return from them [capital instruments] would have to cover little more than their exhaustion by wastage and obsolescence together with some margin to cover risk and the exercise of skill and judgement. In short the aggregate return from durable goods in the course of their life … just cover their labour-cost of production *plus* an allowance for risk and the costs of skills and supervision. (*GT*, p. 375)

This can be viewed as a 'revolution', in that it strongly questions the essence of the marginal theory of income distribution, according to which the rewards of the 'factors of production' are objectively determined, in their own markets, by the law of marginal productivity.

As we share the view that one structural cause of the present crisis is an increasing disparity in income and wealth distribution, the above quotations seem all the more important if we look at how the overall propensity to consume is affected by income distribution. An increase in inequality lowers consumption propensity. The reduction of consumption demand that follows rising income inequality causes a downward pressure on aggregate demand, thus on output and employment – unless some exogenous sources of demand (such as government spending or export demand) compensate for it.

We would like to conclude with two quotations; one from *GT* and the other from Shackle's (1973) article. The first stresses the ethical imprinting of Keynes's book:

> [T]his state of affairs would be quite compatible with some measure of individualism, yet it would mean the euthanasia of the rentier, and, consequently, the euthanasia of the cumulative oppressive power of the capitalist to exploit the scarcity-value of capital. (*GT*, pp. 375–6)

Finally, the following proposition by Shackle reflects, in our opinion, the very essence of *GT*:

> [I]n human affairs the future is different from the past, different in nature and essence (for one is figment and the other is, in some sense, 'fact'), and different in form, for what we imagine as happening may be quite betrayed by history, when the present has moved from the beginning to the end of the interval. (Shackle, 1973, p. 517)

Notes

1. Among the many contributions, cf. Garegnani (1970; 1984) and Pasinetti (1977; 2000).
2. On this aspect cf. Mainwaring (2010) and Chiodi and Ditta (2010).
3. Sen (1987) is one of the most outstanding authors in this respect.
4. It is worth noting that recourse to value judgements is still inevitable for the determination of the means of production. This might seem less evident than in the case of subsistence, because the means of production appear only *indirectly*

related to the livelihood of the people (or, in any case, they might seem related to the livelihood of the people in a different fashion from subsistence). Yet, on further reflection, the means of production are also the end result of the same complex historical process which determines subsistence, and therefore they share with subsistence the same set of determinant factors.

5. Without entering into endless philosophical debates over the presumed dichotomy between 'facts' and 'values' (that dichotomy seems to have been definitely thrown out, having been found logically faulty), suffice it to point out that the distinction between 'facts' and 'values' paradoxically still exists and persists in most of the economic literature that makes use of the neoclassical paradigm, together with a strong resistance to the inclusion of 'values' within the boundaries of the discipline – notwithstanding Sen's reiterated efforts to show the logical and factual faultiness of this view in economics, a discipline which should essentially be devoted to tackling problems of people's well-being. See Sen (1967 and 1984).
6. On the economy as an autonomous entity cf. Polanyi (1957).
7. In this connection it is well worth reading the criticism raised by Hicks (1959; 1983a).
8. On the notion of viability from a different perspective, see Chiodi (1992 and 1998).
9. Very interesting considerations on the notion of 'surplus', in this respect, can be found in Pearson (1957).
10. Cf. Chiodi (2010). Another example showing the impossibility of separating the 'core' from the system as a whole can be given in relation to the choice of the *numéraire*. By taking as a reference a system with a 'surplus' shared by the workers, we are left with a system of equations with one degree of freedom. In such a case, *any* unknown variable can *formally* be fixed from outside the system of production to solve the system of equations, but it would be meaningless from the *economic* viewpoint. For the solution to be sensible, a *reasoned* choice of the variable is also required, based on considerations lying outside the 'core'.
11. In this connection, it seems an over-statement that the uniform rate of profits is the end result of a process implying movements in *market* prices.
12. See the works of Neisser (1932), Zeuthen (1932), von Stackelberg (1933), Wald (1933–4 and 1934–5).
13. Graaff (1957) presents a very lucid and sharp critique as regards the possibility of the market providing the required information for making a rational choice.
14. On Adam Smith's notion of subsistence the considerations of Stabile (2008), pp. 16–18 and 61–6, are of particular interest.
15. See also Van Parijs (1992 and 1995).
16. An exception, however, is Pigou (1920, p. vii, and chapters I and II).
17. Pasinetti (2007).
18. In fact we don't believe it is possible to fit human behaviour into any sort of model; where human life is concerned it is impossible to postulate simple relationships or to attribute to only one cause the outcomes of human and social interactions. On this see Moog (1964).
19. Pasinetti (2007, p. 23).

References

Arrow, K.J. and Debreu, G. (1954) 'Existence of an equilibrium for a competitive economy', *Econometrica*, July: 265–90.

Birolo, A., Foley, D., Kurz, H.D., Schefold, B. and Steedman, I. (eds) (2010) *Production, Distribution and Trade: Alternative Perspectives* (London: Routledge).

Chiodi, G. (1992) 'On Sraffa's notion of viability', *Studi Economici*, 46: 5–23.

Chiodi, G. (1998) 'On non-self-replacing states', *Metroeconomica*, February: 97–107.

Chiodi, G. (2010) 'The Means of Subsistence and the Notion of "Viability" in Sraffa's Surplus Approach', in S. Zambelli (ed.), *Computable, Constructive and Behavioural Economic Dynamics. Essays in Honour of Kumaraswamy (Vela) Velupillai* (London: Routledge): pp. 318–30.

Chiodi, G. and Ditta, L. (2004) 'Economia e processi di sviluppo: la sorprendente lezione di Vianna Moog', in *Sociologia e Ricerca Sociale*, 74: 12–36.

Chiodi, G. and Ditta, L. (2010) 'Can Sraffa Point Us to a Better Future? A Comment on Mainwaring', in A. Birolo, D. Foley, H.D. Kurz, B. Schefold, I. Steedman (eds), *Production, Distribution and Trade: Alternative Perspectives* (London: Routledge): pp. 136–40.

Cohen, J. and Rogers, J. (eds) (2001) *What's Wrong with the Free Lunch*? (Boston: Beacon Press).

Debreu, G. (1959) *Theory of Value. An Axiomatic Analysis of Economic Equilibrium* (New Haven and London: Yale University Press).

Garegnani, P. (1970) 'Heterogeneous capital, the production function and the theory of distribution', *Review of Economic Studies*, XXXVII: 407–36.

Garegnani, P. (1984) 'Value and distribution in the classical economists and Marx', *Oxford Economic Papers*, June: 291–325.

Graaff, J. de V. (1957) *Theoretical Welfare Economics* (repr. 1967) (London: Cambridge University Press).

Hicks, J.R. (1939) *Value and Capital. An Inquiry into some Fundamental Principles of Economic Theory* (London: Oxford University Press, 2nd edn (1946)).

Hicks, J.R. (1959) 'A Manifesto', in J.R. Hicks (1981) *Wealth and Welfare. Collected Essays on Economic Theory*, Vol. I (Oxford: Basil Blackwell): pp. 135–41.

Hicks, J.R. (1981) *Wealth and Welfare. Collected Essays on Economic Theory*, Vol. I (Oxford: Basil Blackwell).

Hicks, J.R. (1983a) 'A Discipline not a Science', in J.R. Hicks (1983b) *Classics and Moderns*, Vol. III (Oxford: Basil Blackwell): pp. 365–75.

Hicks, J.R. (1983b) *Classics and Moderns*, Vol. III (Oxford: Basil Blackwell).

Keynes, J.M. (1972 [1931]) *Essays in Persuasion* (London and Basingstoke: Macmillan).

Keynes, J.M. (1936) *The General Theory of Employment, Interest and Money* (London: Macmillan & Co. Ltd).

Leontief, W.W. (1941) *The Structure of American Economy, 1919–1939* (Cambridge, MA: Harvard University Press).

Mainwaring, L. (2010) 'Can Sraffa Point Us to a Better Future?', in A. Birolo, D. Foley, H.D. Kurz, B. Schefold and I. Steedman (eds), *Production, Distribution and Trade: Alternative Perspectives* (London: Routledge): pp. 119–35.

Marx, K. (1973 [1932]) *Economic and Philosophic Manuscripts of 1844*, edited with an introduction by D.J. Struik (London: Lawrence & Wishart).

Marx, K. (1974 [1867]) *Capital*, Vol. I, Frederick Engels (ed.) (London: Lawrence & Wishart).
Moog, V. (1964 [1954]) *Bandeirantes and Pioneers*, trans. from the Portuguese by L.L. Barrett, with an Introduction by A.A. Berle (New York: George Braziller).
Neisser, H. (1932) 'Lohnhöhe und beschäftigungsgrad im marktgleichgewicht', *Weltwirtsch. Arch.*, 36: 413–55.
Pasinetti, L.L. (1977) *Lectures on the Theory of Production* (London and Basingstoke: Macmillan Press).
Pasinetti, L.L. (2000) 'Critique of the neoclassical theory of growth and distribution', *Banca Nazionale del Lavoro Quarterly Review*, December: 383–431.
Pasinetti, L.L. (2007) *Keynes and the Cambridge Keynesians* (Cambridge: Cambridge University Press).
Pearson, H.W. (1971 [1957]) 'The Economy Has No Surplus: A Critique of A Theory of Development', in K. Polanyi, C.M. Arensberg and H.W. Pearson (eds), *Trade and Market in the Early Empire. Economies in History and Theory* (Chicago: Gateway Edition, Henry Regnery Company): pp. 320–41.
Pigou, A. C. (1920) *The Economics of Welfare* (London: Macmillan).
Polanyi, K. (1971 [1957]) 'Aristotle Discovers the Economy', in K. Polanyi, C.M. Arensberg and H.W. Pearson (eds), *Trade and Market in the Early Empire. Economies in History and Theory* (Chicago: Gateway Edition, Henry Regnery Company): pp. 64–94.
Polanyi, K., Arensberg, C.M. and Pearson, H.W. (eds) (1971 [1957]) *Trade and Market in the Early Empire. Economies in History and Theory* (Chicago: Gateway Edition, Henry Regnery Company).
Putnam, H. (2003) 'For ethics and economics without the dichotomy', *Review of Political Economy*, July: 395–412.
Ricardo, D. (1970 [1817]) *On the Principles of Political Economy and Taxation*, P. Sraffa (ed.) (London: Cambridge University Press).
Robbins, L. (1932) *An Essay on the Nature and Significance of Economic Science* (London: Macmillan & Co.), 2nd edn (1935).
Sen, A. (1967) 'The nature and classes of prescriptive judgements', *Philosophical Quarterly*, 17(66), January: 46–62.
Sen, A. (1984) 'The living standard', *Oxford Economic Papers*, November: 74–90.
Sen, A. (1987) *On Ethics and Economics* (Oxford: Blackwell).
Shackle, G.L.S. (1973) 'Keynes and today's establishment in economic theory: a view', *Journal of Economic Literature*, June: 516–19.
Smith, A. (1970 [1776]) *The Wealth of Nations* (London: J. M. Dent & Sons Ltd).
Sraffa, P. (1960) *Production of Commodities by Means of Commodities. Prelude to a Critique of Economic Theory* (Cambridge: Cambridge University Press).
Stabile, D.R. (2008) *The Living Wage* (Cheltenham and Northampton, MA: Edward Elgar).
Van Parijs, P. (ed.) (1992) *Arguing for Basic Income. Ethical Foundations for a Radical Reform* (London & New York: Verso).
Van Parijs, P. (1995) *Real Freedom for All* (Oxford: Oxford University Press).
Van Parijs, P. (2001) 'A Basic Income for All', in J. Cohen, J. Rogers (eds) *What's Wrong with the Free Lunch*? (Boston: Beacon Press): pp. 3–26.
von Neumann, J. (1937) 'Über ein Ökonomisches Gleichungssystem und eine Verallgemeinerung des Brouwerschen Fixpunktsatzes', in K. Menger (ed.), *Ergebnisse eines Mathematischen Kolloquiums (1935–1936)*, Heft 8 (Leipzig und

Wien: Franz Deuticke), pp. 73–83; English trans. in *Review of Economic Studies*, vol. 13(1945): 1–9.
von Stackelberg, H. (1933) 'Zwei kritische bemerkurgen zur preistheorie Gustav Cassel', *Zeitschrift für Nationalökonomie*, 4: 456–72.
Wald, A. (1933–34) 'Über die Eindeutige Positive Lösbarkeit der Neuen Produktionsgleichungen (I)', in K. Menger (ed.) *Ergebnisse eines Mathematischen Kolloquiums*, Heft 6 (Leipzig und Wien: Franz Deuticke) (1935): pp. 12–18.
Wald, A. (1934–35) 'Über die Produktionsgleichungen der Ökonomischen Wertlehre (II)', in K. Menger (ed.), *Ergebnisse eines Mathematischen Kolloquiums*, Heft 7, Leipzig und Wien: Franz Deuticke) (1936): pp. 1–6.
Walsh, V. (2000) 'Smith after Sen', *Review of Political Economy*, 12: 5–25.
Zambelli, S. (ed.) (2010) *Computable, Constructive and Behavioural Economic Dynamics. Essays in Honour of Kumaraswamy (Vela) Velupillai*, (London: Routledge).
Zeuthen, F.L.B. (1933) 'Das prinzip der knappheit technische kombination und ökonomische qualität', *Zeitschrift für Nationalökonomie*, 4: 1–24.

11

Causality and Structure in Piecemeal Macroeconomic Modelling

Sergio Parrinello

11.1 Introduction

This work defends a number of assertions concerning some basic features of post-Keynesian macroeconomic models.

The assumption of profit maximisation under free competition should be preserved as one of the building blocks of a theory of distribution, regardless of whether the neoclassical theory of distribution is adopted or rejected. This makes it possible to assess two different types of Keynesian macroeconomic model. While it is widely acknowledged that post-Keynesian models share a distinctive feature with Keynes's General Theory in that they admit a *short-period* equilibrium of the economy with involuntary unemployment (the equilibrating role of the real balance effect being dismissed as empirically unimportant), they are divided into at least two groups that address the *long period* in different ways. One group, represented by Kaldor, advocates macroeconomic growth models that reject the neoclassical theory of distribution but are characterised by full-employment equilibria. Another, including most of the economists who have developed Sraffa's critique of marginalism, extends Keynes's principle of effective demand to the long period while preserving the notion of unemployment equilibrium.

The received causal interpretation of post-Keynesian models should be reconsidered in the light of the recent resumption of the structural equation approach and accounts of causality in econometrics and analytical philosophy. According to a common interpretation of post-Keynesian as opposed to neoclassical and new classical models, the arrow of causality goes from investments to savings and (although this characterisation is not accepted by all post-Keynesians) from the price level to an endogenous money supply instead of the other way round.

Piecemeal macroeconomic modelling is better able than a general equilibrium approach to take into account the existence of a non-uniform structure of an economic system. This assertion concerns another division among post-Keynesian economists: one group (the majority) maintains, or at least does not question, the *method* of general equilibrium theory expressed by a comprehensive system of algebraic equations; another seeks to break that system down into subsystems endowed with different structural characteristics.

The overviews of theories of distribution put forward by Sen (1963) and Foley-Michl (1999) are taken as points of reference in the following discussion, the former for the short period and the latter for the long period and growth.

11.2 A taxonomy of macroeconomic models

Let us start from the overdetermination approach adopted by Sen (1963) to construct a taxonomy of theories of distribution. The equations of the overdetermined macroeconomic model are as follows:

(a) the production function: $X = F(N,K)$
(b) labour marginal productivity = wage rate: $PF_N = W$
(c) output value = total income: $PX = rK + WN$
(d) investment = total savings: $PI = S$
(e) the saving function: $S = s_k rK + s_w WN$
(f) given investment: $I = \underline{I}$
(g) full employment: $N = \underline{N}$
(h) full capital utilisation: $K = \underline{K}$

where

X: output	P: nominal output price	s_k, s_w: saving ratios out of profits and wages respectively
N: employment	W: nominal wage rate	I: total real investment
K: real capital	r: nominal rental on capital	S: total nominal savings

The model is overdetermined because eight independent equations constrain seven unknowns: *X, N, K, I, S, W/P, r/P*. Let us call (a)–(h) 'model O'. Sen's line of reasoning develops in three steps: 1) model

Table 11.1 Over-determination differently resolved

Neoclassical theory	**Keynesian theory**	**Neo-Keynesian theory (Kaldor)**	**Johansen's theory (Rattsø)**
Drop (e): autonomous I	Drop (f): full employment	Drop (b): wage = marginal product	Add an endogenous government surplus F and replace (d) with $I - S = F$

O is taken as the starting point; 2) consistent alternative theories are derived by dropping appropriate equations; 3) a causal interpretation of each theory is formulated on the basis of the distinction between exogenous and endogenous variables. Sen derives three main alternative theories of distribution – neoclassical, strictly Keynesian and neo-Keynesian – to which Rattsø (1982) adds a fourth attributed to Leif Johansen (Table 11.1).

In our view, a more useful alternative to the overdetermination approach[1] would start from an *underdetermined* or *open/incomplete* model that can be taken as a common subsystem subject to alternative integrations. The natural candidate for this role is a production system in the ideal case of competitive profit maximisation. If we ignore the criticism of the aggregate production function, the equations that form such subsystem in Sen's classification approach are:

(a) the production function: $X = F(N, K)$
(b) labour marginal productivity = wages: $PF_N = W$
(c) output value = total income: $PX = rK + WN$

Let us call this underdetermined model with the five variables X, N, K, W/P, r/P and the three equations (a), (b), (c) 'model U'. This can be used as a benchmark to classify theories of distribution. We shall start by isolating the theories that encompass model U from the others and then go on to consider different additional relations that can be combined with model U and create a determinate model.

11.3 Profit maximisation

The neo-Keynesian theory described by Sen does not satisfy model U, because it violates equation (b), a condition of profit maximisation. As regards the short period, the neoclassical and Keynesian models can

instead be consistently obtained by adding the following groups of equations to the common sub-model U.

Neoclassical theory	**Keynesian theory**
(d) $PI = S$	(d) $PI = S$
(e) $S = s_k rK + s_w WN$	(e) $S = s_k rK + s_w WN$
(g) $N = \underline{N}$	(f) $I = \underline{I}$
(h) $K = \underline{K}$	(h) $K = \underline{K}$

It is interesting to compare two passages taken from Kaldor and Marglin with regard to the formulation of the neo-Keynesian model for the long period. The former makes this somewhat cryptic remark:

> We have seen how the various "models "of distribution, the Ricardo-Marxian, the Keynesian and the Kaleckian are related to each other. I am not sure where "marginal productivity" comes in in all this – except that in so far as it has any importance it does through an extreme sensitivity of v [the capital output ratio] to changes in P/Y [the share of profits]. (Kaldor, 1955–56, p. 96)

Marglin makes the following critical observation:

> Kaldor resolved the overdetermination by dismissing the marginal productivity relationship altogether. But Kaldor's peremptory dismissal derives from the illegitimacy of a production function based on aggregate capital (l956), not from dissatisfaction with the assumption of competitive profit maximization. If the production function and its derivatives are assumed to represent physical relationships, one cannot so easily ignore the implications of profit maximization. Rather, the issue should be faced head on. (Marglin, 1987, p. 226)

I agree with Marglin. The assumption of profit maximisation should be preserved in both the short- and the long-period versions of the macroeconomic model regardless of whether the aggregate production function is dismissed. A basic Keynesian growth model for the long period should be consistent both with a profit-maximising choice of techniques and unemployment equilibria. A Kaldorian model, in which the demand for labour is at least explicitly disconnected from profit maximisation, instead appears to accommodate a full-employment equilibrium by

relying on some alternative employment function, perhaps of an empirical nature.

11.4 A preliminary look at causality

Difficult analytical-philosophic questions have to be answered or circumvented by resorting to some primitive concept if a causal interpretation is to be established of a macroeconomic model in general and of its Keynesian features, the saving–investment and price–money relations, in particular. These questions will now be formulated with respect to the savings–investment relation, even though the arguments presented below are general in character and encompass the price–money relation. Let us assume that at least one of the variables *S/P, I* is an *ex ante* magnitude, which means that *S/P* and *I* are not identical by definition.

What is meant by saying that *I* causes *S/P*? It obviously means something more than a correlation between *I* and *S/P*. In terms of the present-day philosophical debate, there appear to be two primary candidates for the role of this 'something more'.

The first, suggested by Hume (1777) alongside his regularity concept of causation, and developed by Lewis (1972), hinges on the idea of *counterfactual dependence*. In 'Causation' (1972), Lewis puts forward the idea that a particular (token) event *B* is counterfactually dependent on another particular event *A* if and only if both *A* and *B* occurred and the counterfactual 'Had A not occurred, B would not have occurred' is true. Event *C* thus causes event *E* if and only if there is a chain $C, D_1, \ldots, D_n, E$ such that each link (except *C*) is counterfactually dependent on the one before it. In our case, the causal *relata* are variables or *type* events instead of *token* events. A basic problem is the selection of the relevant causes out of the host of causes that satisfy Lewis's definition. One way out of this difficulty (regarded by J.S. Mill as insurmountable) would be by setting certain background conditions as distinct from causes.[2] This causation in terms of counterfactual dependence applies to causal models formalised by causal functions,[3] where by convention the variable on the left side of each causal equation is an effect of the variables on the right side and the respective inverse function, if it exists, does not preserve a causal meaning. In our case, *I* causes *S/P* if there is a chain of variables $I, D_1, \ldots, D_n, S/P$ such that each link in the chain (except *I*) is counterfactually dependent on the one before it. In this case, the background conditions are implicit in the structure of the model. In short, a *difference* in *I* makes a *difference* in *S/P*, whereas a difference in *S/P* may not make a difference in *I*.

On the alternative view, shared to differing extents by a heterogeneous group of philosophers and scientists,[4] causation is *production* associated with *processes, mechanisms and causal powers*. '*I* causes *S/P*' thus means that *I produces, brings about or drives S/P*. This seems to correspond to the common meaning more than the account in terms of counterfactual dependence. It is also in line with the idea of a production process of commodities that is so familiar in economics, in particular a simple labour process or a process defined as a vector of dated inputs (causes) and outputs (effects). Production becomes the fundamental concept, whereas counterfactual dependence *may* be a derived property. Despite such appealing features, causality in the sense of production will remain a primitive concept until the black box of the underlying causal mechanism is opened up and explained without falling into a vicious circle by reintroducing the idea of cause.

The two views outlined above are alternatives. A counterfactual account describes a property of an explicit or implicit causal model and the truth of the counterfactual statement is relative to the model chosen. Production is instead a property of the real world. Even if the model is a good description of the world, the accounts of causality based on the two views are distinct and may diverge. Each of them has to face some limited applicability and may clash with common sense. For example, let I_P, I_G denote private and public investments. It is possible to say that I_P produces, brings about or drives *S/P*, but at the same time *S/P* does not depend counterfactually on I_P if a difference ΔI_P triggers a difference ΔI_G of opposite sign, through another causal link, and this maintains the same level *S/P*. By contrast, it is possible to say that *S/P* depends on I_P, but I_P does not produce *S/P* if initially $I_P = 0$ and a difference ΔI_P brings about a difference ΔI_G that makes a difference $\Delta(S/P)$.[5]

11.5 Causality and interventions

A choice must be made here among the theories of causality available in the philosophical literature and mentioned in the previous section.[6] Subject to some philosophical provisos and to recognition of the more fundamental theory based on the concept of production, we shall adopt a *manipulationist* theory that can be interpreted as a subspecies of the counterfactual account of causality. An excellent overview of this theory, which is accepted by many scientists (but few philosophers), can be found in Woodward (2008). Roughly speaking, causality from a variable *X* to another variable *Y* is seen as the existence of a *possible manipulation* or *control* of *X* that makes a difference in *Y*. Woodward and Pearl (2000) develop

this through similar notions of *intervention* applied to the variables of causal models. Woodward defines an intervention on variable X with respect to Y under appropriate *ceteris paribus* assumptions. Pearl (2000) applies a similar but simpler concept to recursive structural models, envisaging a sort of *surgical* intervention that sets the value of X and removes the structural equation where X depends on its parent variables in order to ascertain a counterfactual dependence of Y on X. In formal terms, Pearl introduces the $\mathbf{do}(\boldsymbol{X}=\boldsymbol{x})$ operator to represent an intervention that assigns the value x to X and generalises the previous deterministic definition of causality by adopting a probabilistic theory of causality and Bayesian-network methods of causal inference. His basic idea is that X is a cause of Y if, for some values x, x', $x \neq x'$, the conditional probability $P(Y = y \mid do(X = x'))$ is different from the conditional probability $P(Y = y \mid X = x)$, where x is the observed value of X.[7] In short, both Woodward and Pearl combine the notion of intervention with that of counterfactual dependence.

According to the manipulationist approach, '*I* causes *S/P*' therefore means that there exists some manipulation of *I* through a possible intervention that would change or make a difference in *S/P*.

The following objections raised by some philosophers to this manipulationist theory of causation should be noted. 1) The theory does not escape a certain degree of *anthropomorphism*, as an intervention appears to be related to an agency theory of causation. 2) The theory is vitiated by *non-reducibility* or *circularity* because the very notion of intervention is causal. 3) The application of the theory to structural models requires the assumption of *modularity*.[8] These criticisms will not deter us from adopting the theory for our purposes. First, the contamination of the agency theory of causality and the associated anthropomorphism do not appear so restrictive in the case of macroeconomic models, where human agents act behind the aggregate variables. Second, Woodward (2008) has convincingly defended his approach against the charges of circularity and agency. Third, while the assumption of modularity is certainly a restriction, it can be accepted as a useful simplification in the same way as a basic economic theory was developed by assuming the non-existence of externalities. Further problems for a causal interpretation of post-Keynesian macroeconomic models do remain, however.

11.6 Additional problems

A preliminary problem is encountered on attempting to convert a general equilibrium model, taken as a starting point, into a model with an unemployment equilibrium. In general terms, if the price of good j is set

in terms of another good, nothing can be found *within the model* that allows us to drop the equilibrium equation on the j market instead of the market of good k≠j in order to eliminate the overdeterminacy. It is instead necessary to resume the assumption of some causal mechanism, pertaining to disequilibrium analysis and external to the equilibrium model, if a one-to-one correspondence is to be established between a fixed price of a commodity and the equation of its market equilibrium to be removed. This argument applies in particular to the labour market.

More importantly, causality in the macroeconomic models discussed so far cannot be revealed by the distinction between exogenous and endogenous variables and the corresponding closures. The status of exogeneity, without a qualification of the principle underlying the choice between endogenous and exogenous variables (see the distinction drawn by Koopmans in Section 11.7), is not necessary and may not be sufficient for the attribution of a causal role to a certain variable, in particular to the investment I in the Keynesian models. This argument raises two related questions. The first is also the title of an article by Hausman (1982) – 'Are there causal relations among dependent variables?' – where the author argues in favour of an affirmative answer. Causal relations can exist among the dependent variables of a complete model if we adopt an account of causality in terms of interventions applied to a causal model. The second is *whether an exogenous variable can be non-causal*. It should be noted that in a sense, an intervention converts an endogenous variable into an exogenous (*exogenised*) variable and deletes the corresponding causal function while leaving the rest of the model unchanged. The same question can therefore be addressed by asking whether the variable is exogenous to the original model or made exogenous by an intervention. It is obvious that a non-causal relation can exist between an exogenous variable X and an endogenous variable Y because of the structure of the model. A different situation arises if the structure of the model admits the possibility of Y being manipulated through a change in X but *intervention on X is not possible*. The impossibility of intervention is a relative concept. It may be impossible if it violates some law of nature or is impossible for human beings. In particular, intervention on X may be impossible if X is defined as a potential instrumental variable in a policy-oriented model.

Let us clarify the problems at issue in the Keynesian model with endogenous investments and a uniform saving ratio s out of total income ($s_k = s_w = s$). Equations (e) and (f) are rewritten as:

$$(e') \; S = sPX$$
$$(f') \; I = f\,(r/P)$$

Let us assume that (e′) and (f′) are reinterpreted as *causal functions* describing two distinct causal mechanisms. Let us further assume an intervention such that $I = \underline{I}$ and equation (f′) is eliminated without affecting the other equations. 'Investments *I* cause savings *S/P*' means that manipulation of *I* through an intervention makes a difference in *S/P* through changes in income (the theory of the multiplier), and this holds despite the fact that *I, S, P* are endogenous variables with respect to the complete model. But why not set the real savings and drop equation (e′) – instead of (f′) – as the result of an alternative intervention on *S/P* and say that savings 'cause' investment through a change in the rate of interest? On the basis of the model, it may be just as legitimate to say that consumption causes both savings and investments, as the model does not preclude an intervention on consumption, which is implicitly defined $C = PX - S$ in Sen's classification. As a matter of fact, the form of the equations and the distinction between endogenous and exogenous variables of the model do not reveal the existence of a causal relation running from investments to savings instead of the other way round even if $S(.)$, $I(.)$ are interpreted as causal functions. Stressing the causal role of investments within the received post-Keynesian models can, however, be justified by some external assumption about the existence of *possible* and impossible policy interventions associated with instrumental variables in Tinbergen's sense. This idea may be in the back of the minds of many post-Keynesians who are in favour of a causal interpretation of their equilibrium models. They may be thinking of an implicit experimental *policy* design where investments can be manipulated through public investments but savings cannot be controlled in such a direct way. This is a plausible interpretation, but the form of the model reveals nothing in this sense. Some older distinctions between exogenous and endogenous variables can offer further support for our argument.

11.7 Useful earlier views on exogeneity

A seminal discussion of the choice between exogenous and endogenous variables in economic models can be found in the early econometric literature during the years just before and after World War II in the circles of the Cowles Commission and until the mid-1960s. After decades of comparative acquiescence, the question has recently been taken up by some econometricians (Hoover, 2001) engaged in a discussion with analytical philosophers on causality and causal models. Pure economists too often appear to be satisfied with a purely logical distinction

between the variables of a model (endogenous, i.e., determined by the model itself, and exogenous, i.e., determined outside it) and with what Koopmans called *the departmental principle* (see below). I instead believe that economists can benefit from both earlier and more recent contributions, especially for the assessment of the choice of alternative *closures* and corresponding choices of *exogenous* and endogenous variables.

Koopmans (1950) observed that three different principles have been implicitly or explicitly applied in determining which variables are exogenous in the economic literature, namely the *departmental principle*, the *causal principle* and the *purpose of exposition* (which is mentioned only in passing).

> The *departmental principle* treats as exogenous those variables which are wholly or partly outside the scope of economics, like weather and climate, earthquakes, population, technological change, political events. The *causal principle,* which does not always lead to the same result, regards as exogenous those variables which influence the remaining (endogenous) variables but are not influenced thereby … Third principle or consideration: the *purpose of exposition*. At a certain stage of the analysis, variables are often treated as exogenous to facilitate understanding of the model studied, reserving for later elaboration their inclusion among the endogenous variables (Koopmans, 1950, pp. 393–5)

Koopmans argues that only the causal principle can be relevant for the choice of the exogenous variables from a statistical-econometric point of view, in particular with reference to the maximum-likelihood method of estimation. It is in fact clear that the departmental principle is of little use to an econometrician who has no reason to defer to a conventional definition of economics. The 'purpose of exposition' instead requires further discussion, as it appears to be related to or overlap with a fourth principle, namely the principle of open or incomplete models. This point is discussed at greater length below.

Simon and Rescher (1966) add a further interpretation of the exogenous variables of static equilibrium models from a dynamic perspective. Assume that a dynamic system is described in canonical form and the variables observed over a period are divided into three classes:

1. Variables that have changed so slowly that they can be replaced by constants for the period under observation, deleting the corresponding mechanisms from the system.

2. Variables that have adjusted so promptly that they are always close to (partial) equilibrium, hence their first derivatives are always close to zero…
3. All other variables. (Simon and Rescher, 1966, pp. 334–5)

This suggests that a static equilibrium model approximates a dynamic system over a period by assuming that variables of type 1 are exogenous. We shall leave aside some new concepts of exogeneity adopted in econometrics (e.g., super-exogeneity) and focus instead on some important contributions of early econometricians related to the problem considered here.

11.8 Causal relations and general laws

Haavelmo (1944) stresses the need to specify some hypothetical-potential *experimental design i*n order to attribute empirical content to general models. Only then can the relations among the variables of the model be interpreted as causal relations. In particular, some general economic 'laws' are not causal *per se* in the same way as most general laws in physics. Let us illustrate this analogy by reference to two physical laws and two economic laws:

1. Ohm's Law states that $I = \frac{V}{R}$, where I is the current intensity (measured e.g. in ampères), V the potential difference (e.g. in volts), and R a parameter called the resistance, (e.g. in ohms).
2. The ideal gas law states that $PV = NkT$, where P is the absolute pressure of the gas (measured e.g. in atmospheres), V the volume of the gas (e.g. in litres), N the number of molecules in the gas (e.g. in moles), k a constant (called Boltzmann's constant), and T the absolute temperature.
3. Accepted by economists for a long time as a general law of economics, the equation of the quantity theory of money states that $MV = PQ$, where M is the total amount of money, V its velocity of circulation, P the general price level, and Q the total amount of commodities subject to transactions.
4. A less simple economic example is the price equation $(1+\mathrm{r})\mathbf{A}\mathrm{p} + \mathrm{w}\mathbf{l} = \mathbf{B}\mathrm{p}$, where $\mathbf{p}$ is a price vector, r the uniform rate of profits, w the wage rate, and the triple: $\mathbf{A}$ (commodity input matrix), $\mathbf{l}$ (labour input vector), $\mathbf{B}$ (commodity output matrix) the technology in use.

Let us compare the four equations:

$$I = \frac{V}{R} \tag{1}$$

$$PV = NkT \tag{2}$$

$$MV = PQ, \tag{3}$$

$$(1+\mathrm{r})\mathbf{Ap} + \mathrm{w}\mathbf{l} = \mathbf{Bp} \tag{4}$$

None of the equations, considered as such and regardless of its form in terms of dependent and independent variables, expresses a causal relation. Each equation is an open model, and dependent and independent variables can be freely chosen in order to analyse the properties of the relations. A causality nexus emerges only if a hypothetical experiment is specified and the model is closed by setting some variable equal to a parameter that can be controlled from outside the model. This entails a causal interpretation of the choice between exogenous and endogenous variables.

In the case of law (1), it is possible to imagine an electric circuit made up of a two-cell battery, a conductor (e.g., a wire of a certain metal, diameter and length), and an ammeter. The battery can be replaced with a one-cell battery, *ceteris paribus*, and the ammeter will measure a 50% decrease in intensity. It is then possible to say that a change in V (exogenous variable) caused a change in I. Alternatively, it is possible to change the metal, diameter or length of the wire. In this case, I would again change and this change would be *caused* by a change in R (exogenous variable), under *ceteris paribus* assumptions including the battery. Similar considerations apply in the cases of laws (2) and (3). It is only on the basis of a hypothetical experiment that a causal relation can be specified among P, V, T in equation (2) or M, V, P, Q in equation (3) subject to a *ceteris paribus* clause. Without a potential experiment, the equations (1), (2) and (3) might at most *define* a variable in terms of the others or enter accounting identities, but without a causal interpretation. For example, the ratio of the intensity I over the potential V defines the resistance R according to equation (1). The ratio $V \equiv \frac{PQ}{M}$ defines the velocity of money circulation in equation (3). On these interpretations, it is possible to extract a pair of variables from each equation that become *causally identical even though they are conceptually distinct*, as pointed out in Hoover (2011). Equation (4) will be discussed in the next section.

11.9 The price equation (1+r)Ap + wl = Bp

In addition to the common features of the four laws outlined above, it is important to stress some specific characteristics of the price equation.

1. $(1+r)\mathbf{Ap} + w\mathbf{l} = \mathbf{Bp}$ represents a general law as regards its *form*, whereas the values of its parameters can change across time and space. Even the form written above can be interpreted as a special case of a more general law that describes a system of production with land, exhaustible natural resources, taxes and foreign trade.
2. $(1+r)\mathbf{Ap} + w\mathbf{l} = \mathbf{Bp}$ must hold also as accounting relations among *ex-post* magnitudes of a corresponding economic subsystem observed in a state of long-period competitive equilibrium.
3. In the case of production with only circulating capital, a mathematical solution to $(1+r)\mathbf{Ap}+w\mathbf{l} = \mathbf{Bp}$, combined with the choice of an appropriate standard of value (the standard commodity), can be written either $w = \frac{R-r}{R}$ or $r = (1-w)\,R$, where R is the maximum rate of profit that depends on the cost-minimising techniques A^*, l^*, B^*. If the convention is adopted that the caused variable is written on the left side of each equation, the two expressions, even though they are not structural relations, describe opposite causal relations between *w* and *r*.
4. The equation $(1+r)\mathbf{Ap}+w\mathbf{l}=\mathbf{Bp}$ should be interpreted as the result of a cost-minimising choice of techniques with respect to a set of linear available techniques A_j, $\mathbf{l}_j$, $\mathbf{B}_j$, j = 1, 2, …, n. The corresponding subsystem of price/cost inequalities can be a substitute for the price equation in terms of marginal productivities derived from an aggregate production function in Sen's classification of theories of distribution.
5. The model represented by $(1 + r)\ \mathbf{Ap} + w\mathbf{l} = \mathbf{Bp}$ is open and not *causal*. The absence of causal relations among its variables does not preclude an underlying causal microeconomic model. The equations are consistent with a deterministic model where agents act to maximise their profits and parametric prices 'cause' their choices of techniques. Alternatively, they may derive from a process of natural selection or a probabilistic mechanism where a multitude of microstates conform to the laws of thermodynamics and some regularity emerges only at the level of macro states (at the industry level in our case).
6. $(1 + r)\ \mathbf{Ap} + w\mathbf{l} = \mathbf{Bp}$ can be closed by setting either $w = \underline{w}$ or $r = \underline{r}$, where $\underline{w}$ (in terms of some standard of value) and $\underline{r}$ are respectively parameters potentially under control. It may not, however, be said that w, measured in terms of a given standard of value, causally

> depends on *r* only because the latter variable is chosen as exogenous. Nor can it be said that *r* causally depends on *w* only because their causal role is reversed. In actual fact, the choice of the exogenous variable may correspond to what Koopmans called the purpose of exposition. In this case, if r is chosen as exogenous rather than w, the purpose might simply be the quest for analytical simplicity obtained by dealing with a system of linear rather than non-linear equations.

The price equation $(1+r)\mathbf{Ap} + w\mathbf{l} = \mathbf{Bp}$ will be reconsidered in a separate section because of the key role it can play in modelling an economic system through distinct stages of analysis. A piecemeal theoretical construction, where 'piecemeal' is not used in a pejorative sense, is defended by bringing various ideas both old and new together by means of a common thread, namely the recognition of non-uniform structures.

11.10 Non-uniform social structures

A social system is a *complex* system that is not characterised by a uniform structure in a number of overlapping senses. The system may lack a uniform structure because of a) differences in the *persistency and autonomy* of economic relations in the sense of Frisch and Haavelmo; b) differences among relations in terms of *invariance* and *modularity* in the sense of Woodward and Hausman; and c) the existence of *hierarchies* in the sense of Simon. The notions of *causal order*, *decomposability* and *near decomposability*, as introduced by Simon, are properties of systems of equations that represent non-uniform structures within causal models. Let us expand these notions by quoting significant passages from the works of the above authors.

Frisch and Haavelmo strongly emphasise the existence of *persistent* and *autonomous structural* relations in the face of the difficulty of exploiting controlled experiments of economic processes and the passive observation to which economists are necessarily confined. Their ideas and general principles are especially useful with a view to characterising the empirical content of different macroeconomic models and their potential for policy interventions. While *persistency* is self-explanatory, *autonomy* requires a more technical definition.[9] A system of autonomous equations is characterized by the property

> that it is possible that the parameters in any one of the equations could in fact change ... without any change taking place in any of the parameters of the other equations. (Frisch, 1948, preface)

Modularity has already been defined in Section 11.3 (note 5). As regards invariance:

> The general idea of invariance is this: a generalization describing a relationship between two or more variables is invariant if it would continue to hold – would remain stable or unchanged – as various other conditions change. The set or range of changes over which a relationship or generalization is invariant is its domain of invariance. (Woodward, 2000, p. 205)

The concepts of causal order and decomposability are well known and their definition is omitted here. *Near decomposability* and *hierarchies* describe non-uniform structures from the viewpoint of the evolutionary dynamics of *complex* systems. The major innovative idea can be found in Ando and Simon (1961) and its subsequent developments in Iwasaki and Simon (1994) and Simon (1996). Simon is interested in a general characterisation of social, biological and physical systems,[10] and his primary results are summarised in the following extract:

> we may move to a theory of *nearly decomposable systems*, in which the interactions among the subsystems are weak but not negligible … At least some kinds of hierarchic systems can be approximated successfully as nearly decomposable systems. The main theoretical findings from the approach can be summed up in two propositions: 1) in a nearly decomposable system the short-run behavior of each of the component subsystems is approximately independent of the short-run behavior of the other components; 2) in the long run the behavior of any one of the components depends in only an aggregate way on the behavior of the other components. (Simon, 1996, p. 197)

11.11 The piecemeal construction of causal macroeconomic theories

A theory of an economic system should acknowledge the existence of non-uniform social structures in the various senses illustrated above, whereas the traditional general equilibrium approach in economics seems to presuppose a uniform structure. If the aim is to describe a capitalist economy ruled by free competition, the production system, represented by a subsystem of equations (the price equations), should stand as a subsystem with a high degree of autonomy and persistence with respect to the other subsystems that form a hierarchic social system.

More importantly, the modelling of each subsystem may require a different analytical approach and expertise in different social sciences in addition to economics and including historical disciplines. This echoes the classical approach and the idea of the 'core' of the surplus theory of value and distribution adopted in Garegnani (1984):

> The surplus theories have, so to speak, a core which is isolated from the rest of the analysis because the wage, the social product and the technical conditions of production appear there as already determined. It is in this 'core' that we find the determination of the shares other than wages as a residual: a determination which, as we shall see in the next section, will also entail the determination of the relative prices of commodities. Further, as a natural extension of this, we shall find in the 'core' an analysis of the relations between, on the one hand, the real wage, the social product and the technical conditions of production (the independent variables) and, on the other hand, the shares other than wages constituting the surplus, and the relative prices (the dependent variables). (Garegnani, 1984, p. 296)

11.12 Conclusions

This paper suggests the need for an extension of the classification of growth models presented in Sen (1963) and Foley-Michl (1999) and a critical appraisal of the said models.

First of all, a comprehensive classification should include a classical growth model with long-period unemployment equilibria and profit-maximising choice of techniques. Furthermore, our cursory inspection of causal models reveals a common limitation of the macroeconomic models examined in Sen (1963) and Foley-Michl (1999). Their form, in terms of algebraic functions and equilibrium equations, is not such as to reveal a causal nexus running from investments to savings and the non-uniform causal structure of a complex economic system. The choice of exogenous variables and the corresponding closures is not sufficient to describe a causal nexus among the variables. Moreover, the simultaneous equations of the general equilibrium approach cannot cope with the existence of a non-uniform structure of the social system. For this purpose, we need the formulation of causal models and structural equations and piecemeal macroeconomic modelling through a multi-stage analysis of separate subsystems. This appears to be a sound methodological principle. Theoretical and applied work focused on the interfaces between the separate theoretical subsystems remains necessary. This is a difficult task

because, for example, an economist well-equipped to deal with Sraffa's price equations and their extensions may not be so capable of dealing with phenomena outside that analytical context and involving specific institutional, political, historical and philosophic analyses as well. On this issue, which concerns the prospects for progress in political economy, attention can be drawn to an analogy in the use of words in two different areas of enquiry. Some interpretations have used the term *technical progress* for the statistical residual measured by the shift of the aggregate production function in the Solow-Denison approach to growth performances. It has been pointed out, however, that the term in this context is just another name for our ignorance. In a sense, the words 'institutional factors' in the received post-Keynesian theory of prices have a similar connotation. In point of fact, economists tell plausible stories to justify, for example, the choice of an exogenous variable chosen from among the rates of profit, accumulation and real wages. Sraffa has suggested that institutional factors affect the rate of profit, for example, and in particular that the central bank can act on this rate through control over the rate of interest. Similarly, Joan Robinson has described the 'animal spirits' of entrepreneurs as the prime mover of accumulation, and institutional and historical factors have been invoked to justify the choice of an exogenous real wage. While these suggestive terms indicate plausible alternatives to the neoclassical theory of distribution, they are at the same time indicative of an incomplete theory (ignorance would perhaps be too strong an expression) calling for the analysis of other subsystems, especially as regards consumers, the behaviour of firms, and social norms.

Notes

1. It should be noted that the disaggregated approach of Pasinetti (1965) exhibits a certain similarity with the one adopted in Sen's classification, in that both rest on overdetermined models. While Sen's purpose is to classify the major theories of distribution, Pasinetti appears to be intent on showing that a disaggregated formulation of a growth model of the Harrod type displays greater potential flexibility of the economy – by comparison with its aggregate ancestor – to keep within a path of full employment. This flexibility, which might be embedded in a variety of different descriptive or normative models left to future research, should obviate what Harrod calls a discrepancy between the *natural* and the *warranted* rates of growth.
2. See Mackie (1974).
3. The meaning of the word 'counterfactual' is similar but not identical in Lewis and in the literature of causal models. In the latter context, it is not really appropriate to call variables counterfactual because all *relata* can be purely hypothetical instead of either factual or counterfactual.

4. As exemplified by Russell (1918), Salmon (1984) and Cartwright (2007).
5. The contrast and ambivalence of the two rival accounts of causality are illustrated by the typical counterexamples of *pre-emption*, *late-emption* and *overdetermination* in the philosophical literature.
6. A discussion of causes as difference-makers related to theories of explanation and understanding can be found in Parrinello (1999).
7. It should be noted that $P(Y = y \mid X = x)$ may differ from $P(Y = y \mid do(X = x'))$ even if $x = x'$ because an intervention $do(X = x')$ converts an endogenous to an exogenous determination of same value of X.
8. 'Modularity involves a stronger invariance condition that also applies between equations. It says that each structural equation in a system of structural equations that correctly captures the causal relations among a set of variables is invariant under interventions that disrupt other equations in the system by setting the values of their dependent variables within some limited range' (Hausman and Woodward, 1999, pp. 542–3).
9. This definition is close to the definition of modularity attributed to Pearl and Woodward.
10. Simon defines 'effective' hierarchies among subsystems in social systems: 'Almost all societies have elementary units called families, which may be grouped into villages or tribes, and these into larger groupings, and so on. If we make a chart of social interactions, of who talks with whom, the clusters of dense interaction in the chart will identify a rather well-defined hierarchic structure. The grouping may be defined operationally by some measure of frequency of interaction in this socio-metric matrix' (Simon, 1996, p. 185).

References

Ando, A. and Simon, H. A. (1961) 'Aggregation of variables in dynamic systems', *Econometrica*, 29 (2): 111–38.

Cartwright, N. (2007) 'Causal Powers: What Are They? Why Do We Need Them? What Can Be Done with Them and What Cannot?', *Centre for Philosophy of Natural and Social Science Contingency and Dissent in Science. Technical Report 04/07*, London School of Economics.

Foley, D. and Michl, T. (1999) *Growth and Distribution* (Cambridge, MA: Harvard University Press).

Frisch, R. (ed.) (1948) *Autonomy of Economic Relations* (Oslo: Universitets Økonomiske Institutt).

Garegnani, P. (1984) 'Value and distribution in the classical economists and Marx', *Oxford Economic Papers*, 36(2): 291–325.

Haavelmo, T. (1944) 'The probability approach in econometrics', *Econometrica*, 12 (suppl), July: iii–115.

Hausman, D. (1982) 'Are there causal relations among dependent variables?', *Philosophy of Science*, 50: 58–81.

Hausman, D. and Woodward, J. (1999) 'Independence, invariance and the causal Markov condition', *British Journal of Philosophy of Science*, 50: 52l–83.

Hoover, K.D. (2001) *Causality in Macroeconomics* (Cambridge: Cambridge University Press).

Hoover, K.D. (2011) 'Causal structure and hierarchies of models', Electronic copy available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1735588, accessed 25 January 2013.
Hume, D. (1975 [1777]) *A Treatise of Human Nature*, L.A. Selby-Bigge (ed.), 2nd edn revised by P.H. Nidditch (Oxford: Clarendon Press).
Iwasaki, Y. and Simon, H. A. (1994) 'Causality and model abstraction', *Artificial Intelligence*, 67(1): 143–94.
Kaldor, N. (1955–56) 'Alternative theories of distribution', *Review of Economic Studies*, 22: 83–100.
Koopmans, T. (1950) 'When Is an Equation System Complete for Statistical Purposes?', in T. Koopmans (ed.), *Statistical Inference in Dynamic Economic Models*, Cowles Commission Research (New York/London: John Wiley & Sons).
Lewis, D. (1972) 'Causation', *Journal of Philosophy*, 70: 556–67.
Mackie, J.L. (1974) *The Cement of the Universe. A Study of Causation* (Oxford: Oxford University Press).
Marglin, S. (1987) *Growth, Distribution and Prices* (Harvard Economic Studies).
Parrinello, S. (1999) 'Explaining and understanding economic events by contrasting alternatives', *Metroeconomica*, 50(3): 325–50.
Pasinetti, L. (1965) 'A new theoretical approach to the problems of economic growth', *Pontificia Academiae Scientiarum Scripta Varia*, 28, 571–696 (repr. in *The Econometric Approach to Development Planning* (Amsterdam: North Holland, 1965).
Pearl, J. (2000) *Causality: Models, Reasoning, and Inference* (Cambridge: Cambridge University Press).
Rattsø, J. (1982) 'Different macroclosures of the original Johansen model and their impact on policy evaluation', *Journal of Policy Modeling*, 4(1): 85–97.
Russell, B. (1918) 'On the Notion of Cause', in *Mysticism and Logic* (London: George Allen & Unwin).
Salmon, W. (1984) *Scientific Explanation and the Causal Structure of the World* (Princeton: Princeton University Press).
Sen, A. (1963) 'Neoclassical and Neo-Keynesian theories of distribution', *Economic Record*, 39: 53–64.
Simon, H. (1953) 'Causal Order and Identifiability', in W.C. Hood and T.C. Koopmans (eds), *Studies in Econometric Method*, Cowles Commission Monograph 14 (New York: Wiley, 1953): pp. 49–74. Page numbers refer to the reprint in Simon, *Models of Man* (New York: Wiley, 1957): pp. 10–36.
Simon, H.A. (1996) *The Sciences of the Artificial* (Cambridge, MA: MIT Press, 3rd edn).
Simon, H.A. and Rescher, N. (1966) 'Cause and counterfactual', *Philosophy of Science*, 33(4): 323–40.
Woodward, J. (2000) 'Explanation and invariance in the special sciences', *British Journal of Philosophy of Science*, 51(2): 197–254.
Woodward, J. (2003) *Making Things Happen: A Theory of Causal Explanation* (Oxford: Oxford University Press).
Woodward, J. (2008) 'Causation and Manipulability', *Stanford Encyclopedia of Philosophy*, http://plato.stanford.edu/entries/causation-mani, accessed 25 January 2013.

12
On the Link between Functional and Personal Distribution in Italy

*Aldo Barba**

12.1 Introduction

Italian wage and salary earners have experienced a critical change in their position in recent decades. This change has two dimensions. On the one hand, since the mid-1970s, Italian employees have obtained a lower share of value added.[1] On the other hand, since the late 1980s, wider disparities in workers' pay structure have meant that a small segment of the salaried workforce has not seen its share decline.[2] The proportion of income allotted to the bulk of the working class has thus been reduced not only by the expansion of other incomes, but also by the expansion of the income of employees at the top of the pay scale.

Surprisingly, the alterations in both wage share and wage dispersion have not been mirrored by a *long-lasting* deteriorating trend in personal income inequality as measured by the Gini index: only during the 1991–93 recession is a substantial widening of the latter recorded. In an attempt to explain this, the gap between perceived distributive tendencies and real distributive tendencies has been investigated, thus confining the phenomenon to the field of errors of perception and disappointed expectations.[3] And when real distributive tendencies have been under consideration, it has been suggested that, although employees suffered a worsening of their position, self-employed people's status improved, thus generating *counterbalancing* tendencies.[4] In the light of

*I wish to thank an anonymous referee, M. Pivetti, G. de Vivo, A. Trezzini and other participants in the 2010 Sraffa conference for comments and suggestions. Address for correspondence: Dipartimento di Economia, Management e Istituzioni, via Cinthia-Monte S. Angelo 80126 Napoli (Italy); email: albarba@unina.it.

these results, the Italian experience of the last two decades would seem to confirm that

> [O]ne of the most widespread fallacies in this area is the belief that one can go readily from the functional to the personal distribution. Wages and salaries, it is believed, are the income of the 'poor'; interest, dividends, rents and earnings of individual business, are the income of the 'rich'; hence anything that raises wage rates relative to other factor returns will tend to render income less diverse, and conversely. Fortunately, or unfortunately, this conclusion is false. (Friedman, 1962, p. 253)

The main purpose of this chapter is to discuss the link between movements in functional income distribution and the Gini index in Italy over the past twenty years, investigating why personal income distribution appears rather stable and independent of the underlying distribution of national income by factor shares. Section 12.2 presents data on this from the Bank of Italy's Survey on Household Income and Wealth (SHIW). Section 12.3 discusses the relationship between functional distribution and the SHIW/Gini index identifying some reasons why there could be a de-link between these two dimensions. In Section 12.4 we elaborate evidence on this by decomposing the Gini index by factor components. Section 12.5 discusses how the distributive outlook in Italy has been influenced by the increase in *imputed* property income as distinct from *actual* property income. Section 12.6 assesses the role that permutations among income receivers along the income ladder might have exerted in stabilising the Gini index. Section 12.7 concludes.

12.2 Confronting movements in functional and personal distribution

According to the SHIW, the share of employees' compensation net of social security contributions and direct taxes on labour income fell about 10 percentage points between 1987 and 1998[5] when measured against gross household disposable income.[6] Since then, it has tended to be stationary , and in 2008 the wage share had changed little from its value in 1998. If we compare this fall in the wage share with the behaviour of the Gini index over the same time span, however, this measure of inequality does not show a corresponding prolonged worsening.[7]

It was only from 1991 to 1993, in fact, that the total Gini increased. As a matter of fact, from 1987 to 1991, the Gini index *fell,* in spite of

a fall in the wage share of about 4 percentage points. And from 1993 to 1998, notwithstanding an additional fall in the wage share of about 5 percentage points, the Gini index remained at its 1993 level. In the period from 1998 onwards, changes in the Gini index and in the wage share moved rather consistently, though showing no clear trend. With respect to the stationarity in the wage share over the last decade, however, it should be noted that this movement is somewhat unusual. According to data from both National Accounts and SHIW, in 2008 average real hourly earnings were at their 1991 level. Since the end of the 1990s, the fall in the wage share has been halted by stagnating or falling *absolute* levels of labour productivity, a very peculiar phenomenon that is causing some concern about the reliability of statistical records (more on this in Section 12.5).

Let us have a closer look at the movements of disposable income behind the trend in the Gini index shown in Figure 12.1.

After the 1991–93 recession, the bottom 20% of the distribution recorded a 2-point fall; given that the share of total disposable income

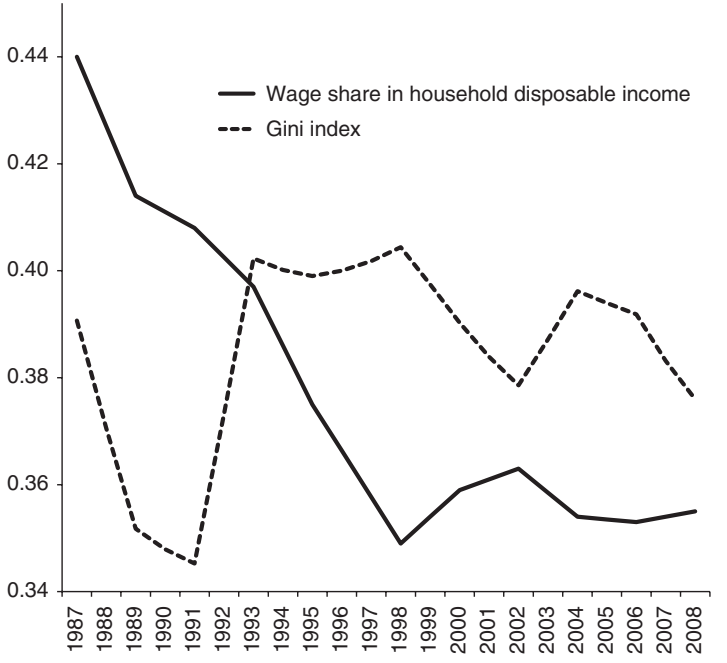


Figure 12.1 Functional income distribution versus personal income distribution

Table 12.1 Shares in household disposable income

	Percentages share in household income				
	1989	**1993**	**1998**	**2004**	**2008**
Quintiles of the distribution					
1st	6.6	4.8	5.0	5.4	5.8
2nd	12.3	10.7	11.0	11.3	11.8
3rd	16.7	16.3	16.1	16.0	16.3
4th	22.1	22.8	22.3	21.9	22.3
5th	42.3	45.3	45.7	45.5	43.9
80–90	15.0	15.7	15.3	15.1	15.3
90–100	27.3	29.5	30.4	30.4	28.6

Source: SHIW.

of this quintile was about 6 per cent, the loss amounted to one third (Table 12.1). The second quintile recorded a fall of 1.5 percentage points, a retrenchment of more than a tenth of its share (12 per cent) of total income. The 3rd quintile did not benefit from this; a contraction was also recorded. For the poorest 60 per cent of the population taken as a whole, 4 percentage points of disposable income were lost: 1 percentage point accrued to the 4th quintile; the remaining 3 points to the 5th (and in particular to the richest decile: 2 points). In 1998, shares of household disposable income were overall almost unchanged, apart from a further rise in the share of the richest decile at the expense of the first four quintiles. Since then there has been a slight recovery at the bottom of the distribution. In 2008 the 1st and the 2nd quintile regained 1 point each, mostly from the richest decile.

The poor correlation between the wage share and the Gini index recorded over the last two decades becomes more striking if one notes also that the share of dependent labour income accruing to the poorest 20 per cent of dependent workers has shrunk since the end of the 1980s (see Table 12.2).

The year 1989 is the turning point of a downward trend in the dispersion of income from wages.[8] Since then, the distribution of annual earnings (net of taxes) has widened, and in 2008 the level of wage inequality was back to that recorded in 1980. Between 1980 and 1989, the share of the wage bill accruing to the poorest 20 per cent of employees rose by about 3.6 percentage points, at the expense of the remaining quintiles. Between 1989 and 2008, on the contrary, the share of the lowest quintile fell about 2.3 percentage points, with the totality of this loss

Table 12.2 Shares in wage bill

	Percentages share in total wages		
	1980	**1989**	**2008**
Quintiles of the distribution			
1st	7.2	10.8	8.5
2nd	16.3	16.1	15.4
3rd	20.1	19.0	18.9
4th	23.5	22.0	22.4
5th	32.9	32.1	34.8
80–90	13.7	13.1	13.4
90–100	19.2	19.0	21.4

Source: SHIW.

appropriated by the richest quintile, in particular by those above the 90th centile (+2.2 points).[9]

The change in the slice of the wage bill accruing to the bottom 20 per cent of the working population has been of such magnitude that it has generated a fall in the *absolute level* of the wage bill accruing to that segment of workers. In fact, the overall growth (at 2000 constant prices) of the total wage bill from 1989 to 2008 has been about 14.5 per cent, well below the 20 per cent change in the share of the poorest quintile of the working population. Given that the feeble growth rate of the wage bill of the whole employed population was unable to compensate for the enlarged wage dispersion, the poorest fifth of the employees in 2008 received as wages about 10 per cent less than their wages at the end of the 1980s. This flies in the face of the belief that wider income disparities are not such an evil in a society where the most disadvantaged members of the population enjoy rising absolute incomes, even though their relative position worsens. Indeed, although this chapter is mainly concerned with relative changes in distribution, it is important not to lose sight of the fact that in Italy such changes have been recorded in a situation of stagnating *absolute* levels of wages, thus making the international comparison with even larger increases in wage dispersion – such as the one recorded in the USA[10] – rather meaningless.

12.3 From functional to personal distribution

Before analysing the disconnection between the increase in wage inequality in the two dimensions under scrutiny and changes in personal

income distribution among the whole population, let us recall a peculiarity of the Italian experience over recent decades. It is a well-established fact that absolute levels of income inequality in Italy are the highest amongst rich countries.[11] Moreover, when the whole population is taken into account, Italy exhibits a more unequal distribution of wages and salaries than the USA (this reflects both Italy's unequal distribution of annual earnings among the employees and her traditionally low percentage of labour force in employment). Capital income is distributed as unequally as in Germany and the USA. Finally, as we have already noticed, some widening in the Gini measure of income inequality is recorded, but it is only between 1991 and 1993 that it rose sharply. This 'fortunate' occurrence, to use Friedman's expression, seems to support the widespread idea that functional distribution should not be listed among the central themes in the current revival of the distributive debate. But what sort of transformation would have made changes in the Gini index only indirectly connected with changes in distribution between capital and labour?

Apart from statistical disputes, there are basically three reasons why it is not possible to go readily from changes in shares of income (after transfers and taxes), to changes in the Gini index. First, wages and incomes from other sources could be evenly spread among the population. This implies that, whatever change is recorded in income shares, its effect on personal distribution is limited. According to this view – the theoretical counterpart of which is the representative agent framework – the wage earner is a property income earner as well, and thus a fall in the wage share is necessarily counterbalanced by a corresponding increase in a similarly distributed income source. Sources of income are, so to speak, 'neutral' in terms of personal distribution. Second, a fall in the wage share could leave personal income inequality unchanged even if wages are more equally distributed than the other factor sources, as long as this effect is compensated for by a decrease in wage dispersion. Third, an equalising effect on personal distribution could come about even if wage dispersion among wage and salary earners is exacerbated, provided there is an increase in the number of employees at the bottom of the *total* income ladder.

12.4 The decomposition of the Gini index by factor sources

It may prove useful in a discussion of the different sources of the link between functional distribution and the SHIW/Gini measure of personal

distribution to refer to the following decomposition of the Gini index for total income inequality (G):[12]

$$G = \sum_{k=1}^{K} S_k C_k \qquad [12.1]$$

where total personal income is divided into K income sources (wage income, entrepreneurial income, rents, and so on), S_k is the share of source k in total disposable personal income, and C_k is the pseudo-Gini index (or the concentration index). The pseudo-Gini of a given income source differs from its pure Gini because in the latter the income earners are arranged in ascending order according to that single source of income, while in the former they are arranged according to their *total* income.[13] Based on decomposition [12.1], the total Gini coefficient can be interpreted as a weighted average of the pseudo-Gini coefficients of each source, the weights being the factor income shares. If changes in functional distribution involved shifts among income sources with similar pseudo-Gini indexes this change would exert no influence on total income inequality. Yet, if the change affects an income source with a lower pseudo-Gini index, a narrowing in this share of total income will not be neutral, leading to an increase in total income inequality, unless the retrenchment is counterbalanced by a decrease in its pseudo-Gini index.

12.4.1 Shares of factor sources

Using the SHIW dataset, Italian personal disposable income is divided into six different factor components: compensation of employees, income from self-employment and entrepreneurial income, pensions and other transfers, interest income, other property income (accrued rents and income from financial assets), and imputed rents. Table 12.3 shows the share of each factor income source from 1987 to 2008, for the whole Italian population aged 15 or over.

The chief source of personal income is employee compensation. However, as we saw in Section 12.2, the wage share dropped from 44 per cent in 1987 to about 35 per cent in 1998, since when it has remained the same. Self-employment income share shows a decrease from 1987 to the mid-1990s. From 1995 to 2006, however, it gained ground. The share of pensions and other transfers grew by 6 points over the period 1987–95 and has since remained the same. Imputed rent share grew about 7.5 points over the whole period, but net interest share fell from 1993 onwards. This was due to both a reduction in interest rates and

Table 12.3 Factor shares

	Years					
	1987	**1991**	**1993**	**1995**	**1998**	**2008**
Wages and salaries	44	40.8	39.7	37.5	34.9	35.5
Income from self employment and individual businesses	22.5	20.1	17.1	16	18.2	17.6
Pensions and other transfers	18.6	20.1	21.6	23.5	23.4	24.6
Actual rents and income from financial assets	1.5	1.7	2.8	2.4	4.1	2.2
Interest income	1.7	3.5	3.6	3.2	2.8	1.0
Imputed rents	11.7	13.8	15.3	17.4	16.6	19.1

Source: SHIW.

an increase in household mortgage debt. The share of accrued property income grew between 1987 and 1998 from 1.5% to 4%. Since then, the trend in property income proper has reversed, losing 2 percentage points over the period 1998–2008.

As a whole, factor shares display a major shift from wage income towards imputed rent and pensions and transfers. While over the period 1987–95 the income share of both employees and self-employed fell, from 1995 onwards the former continued to fall or to stagnate, while the latter started to regain its initial level. Adding together wage income and pension and other transfers, the rough measure of the direct and deferred wage share thus obtained is 2.5 percentage points lower in 2008 than in 1987. Property income proper, however, did not gain from this fall. In fact, the share of actual rents, income from financial assets and interest income grew until 1993, but subsequently decreased. From 1993 to 2000, both the wage share and the non-imputed property income share fell.

12.4.2 Pseudo-Gini indexes

The *pure* Gini coefficients for selected years are shown in Table 12.4. Data include zero-income subjects. The six different income sources can be grouped into two main clusters, according to their inner concentration. The well-distributed cluster includes wages, pensions and transfers, and imputed rents; the unequally distributed cluster comprises interest income, gross entrepreneurial incomes and accrued rents and financial incomes.

Wage income is the most equally distributed factor source – i.e., with the lowest Gini. Pensions and transfers also display a relatively low Gini

coefficient. The third source of the well-distributed cluster – imputed rents – is also equally distributed (its Gini coefficient is just above the Gini coefficient of pensions and transfers and well below the Gini coefficients of non-imputed property income). This reflects the fact that in Italy the percentage of home owners is very high. At the opposite end of the distributive spectrum, actual capital income is the most unequally distributed income source. Self-employed income also displays a relatively high Gini coefficient. Within the highly concentrated cluster, interest income emerges as the better-distributed property income source.[14]

When considering the movements of the inner distribution through time, the increasing trend in the pure Gini coefficient of wage income occurs only until the mid-1990s. In other words, taking into account the unemployed who have become employees, wage differentiation increases until the 1995 survey, decreasing from 69.7 in 1995 to 67.2 in 2008. For pensions and transfers there is no substantial change in the pure Gini coefficient. Entrepreneurial and self-employed income, as well as income from financial assets, shows no clear trend for pure Gini coefficients. The Gini index for imputed rent and interest income decreases.

As we have seen, the inner distribution of each factor income source does not affect the total Gini directly. The total Gini, in fact, can be interpreted as a weighted average – with weights given by the factor income shares – of a Gini measure of each source obtained by ranking individuals not over each factor source they receive, but over their total

Table 12.4 Gini indexes of each source and total Gini (zero-income subject included)

	Years					
	1987	**1991**	**1993**	**1995**	**1998**	**2008**
Wages and salaries	63.5	64.5	69.0	69.7	69.5	67.2
Income from self employment and individual businesses	91.6	87.8	90.0	90.0	90.2	90.9
Pensions and other transfers	72.2	72.3	72.0	71.4	71.2	72
Actual rents and income from financial assets	97.9	97.7	97.3	97.1	96.2	96.7
Interest income	95.5	86.3	87.8	88.0	84.7	87.9
Imputed rents	74.4	73.9	74.4	73.9	72.7	68.9
Total Gini	39.1	34.5	40.2	39.9	40.4	37.6

Source: SHIW.

Table 12.5 Pseudo-Gini indexes of each source (zero-income subject included)

	Years					
	1987	**1991**	**1993**	**1995**	**1998**	**2008**
Wages and salaries	36.3	30.9	40.3	38.1	33.9	26.7
Income from self employment and individual businesses	76.0	56.3	56.9	54.5	61.4	64.8
Pensions and other transfers	–12.5	–0.3	0.6	12.6	13.1	18.9
Actual rents and income from financial assets	81.1	84.1	85.5	84.1	87.1	82.1
Interest income	71.7	62.3	65.0	65.3	61.7	60.3
Imputed rents	50.5	55.3	55.1	56.3	54.7	50.5

Source: SHIW.

disposable income. Table 12.5 shows the effect of this reordering[15] for each income source in selected years.

Apart from the 1991–93 episode, the pseudo-Gini index of wage income decreases. Pensions and transfers display the opposite tendency. Despite the fact that this source shows an unchanged inner distribution, the pseudo-Gini index rises. In other words, transfer receivers are becoming more and more concentrated in the medium and upper sectors of the income ladder. Income from self-employment and individual businesses shows a reduction in the pseudo-Gini until 1995, followed by an increase over the last decade. The overall picture that emerges from Tables 12.4 and 12.5 rather qualifies the results from Table 12.2. While Table 12.2 definitively contradicts the idea of a more equally distributed wage share, Table 12.5 indicates that this income source – even though its distribution is less egalitarian *among workers* – is increasingly accruing to the poorest individuals.

In Figures 12.2 and 12.3 we compare the pseudo-Gini curve for wage income and pension and transfer income in the years 1989 and 2008.[16]

As can be seen in Figure 12.2, if the population is ranked in ascending order according to total disposable income, the share of wages and salaries attributed to the first quintile grew from 2.2 per cent in 1989 to 4.4 per cent in 2008, while the share of the wage bill of the poorest 40 per cent of the population rose from 15 per cent to 19 per cent. Pensions and other transfers display the opposite tendency (see Figure 12.3). While in 1989 more than 50 per cent of pensions and transfers accrued to the poorest 40 per cent of the population, in 2008 that share fell to 28.7 per cent.

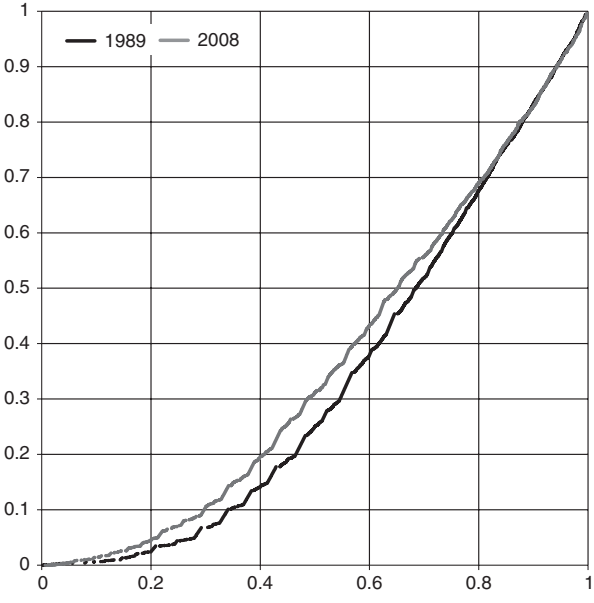


Figure 12.2 Pseudo-Lorenz curve of wage income

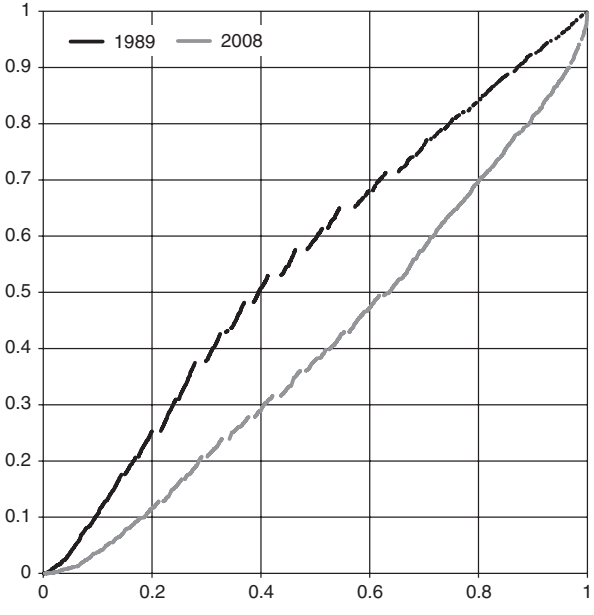


Figure 12.3 Pseudo-Lorenz curve of pensions and other transfers

12.4.3 Assessing the weights

To isolate the effect of these pseudo-Gini changes from the effect generated by changes in factor shares, in Table 12.6 we decompose the absolute variation in the Gini coefficient between one survey and the previous one according to the following formula:

$$\Delta G = \sum_{k=1}^{K} \Delta S_k C_k + \sum_{k=1}^{K} \Delta C_k S_k + \sum_{k=1}^{K} \Delta S_k \Delta C_k \qquad [12.2]$$

The first term represents the effect on the total Gini index of changes in factor shares, leaving pseudo-Gini coefficients unchanged (the factor-share effect); the second term represents the effect on total inequality of changes in pseudo-Gini coefficients, given factor shares (the inequality effect); the third term is a residual that results from the interaction between these two sources of variations.

The decomposition is of course a mechanical and in some ways arbitrary procedure. However, the factor-share effect can be interpreted as the change in total inequality resulting from a change in functional distribution that does not alter the pseudo-Gini coefficients. The inequality effect is, rather, driven by a change in the pseudo-Gini coefficients, given the share.

Let us consider decomposition [12.2] for the four sub-periods discussed above. From 1987 to 1991 the Gini index fell from 39 to 34.5, thus showing a perverse behaviour (–4.5) with respect to the fall in the wage share (–3.2). This fall in the total Gini is totally driven by the inequality effect (–4.6), a result that largely depends on the reduction of the pseudo-Gini of wage income and self-employment income (in contrast, the pure Gini coefficients of both factor shares increase). The

Table 12.6 Factor source and inequality effects

	Years			
	1987/1991	**1991/1993**	**1993/1998**	**1998/2008**
Gini index				
Factor source effect	−0.7	−0.3	0.1	−1.4
Inequality effect	−4.6	5.9	−0.4	−1.6
Residual	0.7	0.0	0.5	0.1
Delta Total Gini	−4.5	5.7	0.2	−2.9

Source: SHIW.
Divergence in Delta Total Gini due to rounding.

factor-share effect also acts positively on inequality (–0.7). This is rather surprising, given the massive fall in the wage share. At the root of this positive influence of the factor-share effect there is the following circumstance: the bulk of the fall in the wage share is absorbed by imputed rents and pensions in the well-distributed cluster. This helps us introduce a point not limited to the sub-period under scrutiny: in general, the effects of changes in factor shares are mostly small or negative, i.e., they exert an irrelevant or equalising effect on total income distribution. In other words, in isolation, changes in the distribution of income sources – dominated as they are by the fall of 10 percentage points in the wage share – have affected the total Gini only slightly. As a result, the inequality effect turns out to be the main driver of the observed variations in this index.

Over the period 1991–93, the total Gini grew by 5.7 points (from 34.5 to 40.2). The wage share fell moderately (1.1 percentage points) and the factor-share effect exerts no substantial influence over the period (–0.3). The whole increase in the Gini index depends upon the inequality effect in the wage re-ranked distribution (+5.7). The peculiar conjuncture of the period may here acquire some weight, given the recorded loss of employment over the period (more on this in Section 12.6).

The period from 1993 to 1998 covers the first phase of the substantial stability of personal inequality recorded since 1993. The total Gini index remained almost unchanged (from 40.2 in 1993 to 40.4 in 1998), with both the factor-share and inequality effects of little magnitude. Yet, the wage share recorded a fall of slightly less than 5 percentage points, while income from financial assets and actual rent grew by 1.3 percentage points. This change in both the most equally and the most unequally distributed factor sources exercised a limited effect since the former was mainly absorbed by the increase in imputed rents and pensions, while the latter was absorbed by a roughly corresponding decrease in interest income.

In the decade 1998–2008 the Gini index fell by 2.9 points. Both the share and inequality effects contributed to this fall with a similar weight (–1.4 the former and –1.5 the latter). It is interesting to note that over this period the inequality effect of all factor sources impacted distribution positively, with the exception of pensions and other transfers, which were addressed to a greater extent to the upper segment of the distribution.

Two general conclusions can be drawn from decomposition [12.2]. First, as noted above, the factor-share effect is scarcely able to drive changes in the total Gini. The limited influence of the factor-share

effect on inequality depends on the fact that, until the end of the 1990s, changes in the wage shares are absorbed by relatively well-distributed factor sources, while the modest rise in the share of income from rents and financial assets is compensated by the fall in the interest income share. Since the end of the 1990s, the factor-share effect has been weightless because shares hardly move. With almost no factor-share effect on personal distribution, the inequality effect turns out to be the substantial driver of the Gini index changes. Indeed, the index falls before the 1991–93 recession because the pseudo-Gini of wage income falls, it rises during the recession because the pseudo-Gini of wage income rises, and it is trendless thereafter mainly because the pseudo-Gini of wage income falls again, counterbalancing the increase in the pseudo-Gini of transfers. It was not, therefore, the convergence of the inner distribution of labour and capital income proper that led to the functional and the Gini measures of personal distribution in Italy de-linking over the 1990s. The combination of circumstances at the root of this disconnection can be identified as a fall in the wage share associated with a stationary or even falling non-imputed capital income share, and a wide diffusion of 'atypical' labour contracts and the consequent spreading of labour income at the bottom of the income scale. Let us now turn our attention towards these two peculiarities.

12.5 The misreporting of non-wage income

The fall in the Italian wage share of disposable income translates not into a corresponding increase in the capital share proper, but into an increase in imputed rents. If imputed rents are not taken into account, the share of labour and pension income is on the rise. In 1993 it was equal to 69.3 per cent and by 2008 it had risen to 74.3 per cent.

The huge growth in the share of disposable income allotted to imputed rent is not peculiar to Italy. What is specific to the Italian case, however, is the fact that excluding imputed rent from households' operating surplus, the modest growth in Italian value added becomes even more limited, strengthening the anomalous dismal performance of Italian productivity. An unearned income component becomes the main driver of output growth since the mid-1990s. Excluding imputed rents – which contribute to the formation of value added by increasing the operating surplus of households – the overall growth in real terms of households' disposable income over the period 1990–2008 was a mere 11.6 per cent.

The issues posed by the reliability of the recorded movements in the shares of factor sources and in their distribution among the population

are difficult to deal with. Of the two kinds of errors that can affect survey data – sampling and non-sampling errors – the literature focuses almost exclusively on the second, trying to measure and correct misreporting (see, for example, Biancotti, 2008; and Neri and Zizza, 2010).[17] The natural starting point of any attempt in this direction is a comparison between the SHIW estimates and the corresponding national account aggregates.[18] This highlights discrepancies among income sources, signalling a marked asymmetry between the well-distributed and the badly distributed cluster of income sources. For example, according to Neri and Zizza (2010, p. 29), while the ratio between post-tax income shares in the SHIW and in national accounts is equal to 88.3 per cent for payroll income and 69.1 per cent for pensions and net transfers, it is 47.5 per cent for income from self-employment in units with up to five employees, 29.5 per cent for income from self-employment in units with more than five employees, and 13.4 per cent for entrepreneur income and income from financial assets;[19] for imputed rents the ratio is 219.5 per cent. While labour and pension income is pretty well captured by the survey, imputed rent is over-reported and income from self-employment and capital income proper is consistently under-reported.[20]

In order to obtain indications on the possible effect of misreporting on personal income distribution, the bias in reported amounts has to be distributed among quintiles. The analysis operating this allocation detects the misreporting of income from self-employment as able to increase the level of the Gini index, without affecting its trend.[21] In any case, 'the order of magnitude of the discrepancy between the SHIW and national accounts is such as to demand some caution in the use of the data' (Brandolini, 1999, p. 48), in particular for interests and dividends which 'emerge as a particularly serious problem area' (*ibid.*). The seriousness of this problem is confirmed by the fact that the correction of income from financial assets, whose adjusted value is about three times the value recorded by the survey, exerts an implausible equalising influence on income distribution (the Gini index decreases due to the correction), while the correction for all the other factor shares negatively affects the inequality level.[22]

Besides the problem of the discrepancy between the SHIW and national accounts, it is very important not to lose sight of the fact that even national accounts data might suffer from some under-reporting of capital income proper. The dismal performance of Italian labour productivity could be partly the result of a measurement error. While a constant fall in the growth rate of labour productivity since the early 1980s is a feature common to almost all European countries, a zero

or slightly *negative* growth from 2000 onwards occurred in no other European country except Spain (see, for example, ECB, 2004, pp. 48–9; and Gomez-Salvador *et al.*, 2006, pp. 9–14). This decline in labour productivity has occurred in the presence of strong growth in both employment and hours worked.[23] The oddity of this combination of feeble output growth and strong employment growth has been commented on by the Governor of the Bank of Italy (see for example Banca d'Italia, 2008, p. 9), and is thus expressed by a Bank of Italy report on trends in the Italian productive system:

> Italian firms appear to be victim of some sort of economic irrationality that forces them to accumulate factors of production notwithstanding the enduring stagnation in economic activity, and while the condition of employment of such factors should grant firms much more flexibility in their use with respect to the past. (Banca d'Italia, 2009, p. 47)

In explaining this 'sort of economic irrationality' some researchers have drawn attention to an under-recording of the value added as a reasonable explanation, a circumstance that might not fail to affect factor income distribution since a revaluation of the value added – given total hours worked – would entirely translate into an enlargement of the gross operating surplus.[24]

12.6 The impact of poorly paid jobs on distribution

As we have seen in the previous sections, even if the inner distribution of wage and salary incomes declined, the pseudo-Gini of this source improved. Part of the reason why a shrinking and a more unequally distributed wage share do not affect personal income distribution negatively is that wage income increasingly moved towards the bottom of the total income ladder. As a result, the composition of disposable income has varied considerably across quintile groups of individuals arranged in an ascending order of their total disposable income.

Let us, for example, compare the data for 1989 and 2008 (Table 12.7). As we saw in Section 12.2, in 1989 the poorest quintile of the working-age population received 6.6 per cent of total disposable income. Property income constituted a minor share of that total, since the income of the poorest 20 per cent of the population mostly consists of wages and transfers (slightly less than 90 per cent). Wages and salaries, however, were only about 15 per cent of the recorded disposable income: more

Table 12.7 Factor sources by quintiles (percentage shares in disposable income)

	1st quintile		2nd quintile		3rd quintile		4th quintile		5th quintile	
	1989	2008	1989	2008	1989	2008	1989	2008	1989	2008
Wages and salaries	15.2	28.0	39.6	44.8	58.4	52.1	55.9	40.9	31.7	25.0
Income from self employment and individual businesses	6.1	7.7	11.8	6.6	12.4	7.9	15.3	12.0	33.8	28.3
Pensions and other transfers	72.1	49.5	38.9	36.8	19.4	27.2	13.5	24.7	7.1	17.0
Actual rents and income from financial assets	0.1	0.2	0.1	0.3	0.1	0.3	0.4	0.8	2.9	4.5
Interest income	0.6	0.4	1.3	0.4	1.3	0.5	2.9	0.9	7.0	1.4
Imputed rents	6.0	14.2	8.2	11.1	8.3	12.0	12.1	20.6	17.4	23.8
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Memo: share in total disposable income	6.6	5.8	12.3	12.0	16.9	16.6	22.0	22.6	42.5	42.9

Source: SHIW.

than 70 per cent was recorded in the form of pensions and transfers. In 2008 the share of total disposable income appropriated by the poorest quintile (5.8 per cent) also mostly consisted of wages and transfers. But wages and salaries grew to 28 per cent, while pensions and transfers fell to 49 per cent. While in 1989 pensions and transfers were five times the level of wages, in 2008 they amounted to less than twice that level. The same trend, albeit less pronounced, is apparent in the group from the 20th to the 40th centiles. While in 1989 wages and pensions had the same weight in disposable income of the second quintile, in 2008 pensions were 0.8 times the level of wages. This trend is instead reversed in each of the three upper quintiles of the population. The wage share falls, while the share of pensions is rising.

A different perspective on the effect of this replacement of transfers by wages at the bottom of the distribution can be gained considering the incremental effect on the total Gini of a small variation in a given factor share (see Table 12.8).

Apart from the fact that, unsurprisingly, the wage share and the transfer share are the only two sources which, if slightly increased, will lead to a reduction in the Gini coefficient, it is interesting to note that the

Table 12.8 Elasticity of the Gini index with respect to the various income sources

	Years					
	1993	**1995**	**1998**	**2002**	**2006**	**2008**
Wages and salaries	0.001	–0.017	–0.057	–0.074	–0.099	–0.103
Income from self employment and individual businesses	0.071	0.059	0.094	0.143	0.166	0.127
Pensions and other transfers	–0.183	–0.161	–0.159	–0.158	–0.146	–0.123
Actual rents and income from financial assets	0.032	0.026	0.047	0.026	0.020	0.027
Interest income	0.022	0.020	0.015	0.005	0.003	0.006
Imputed rents	0.057	0.072	0.059	0.058	0.057	0.066

Source: SHIW.

effect exerted on inequality by an increase in the wage share has been growing since 1993. In 1993, an increase of 10 per cent in the share of wages and salaries left the total Gini almost unchanged. In 1998 a 10 per cent increase in the wage share led a 0.5 per cent reduction in the Gini coefficient; in 2008, it led a 1 per cent reduction. The more wages become the prevailing income of the poorest quintiles, the more changes in functional distribution take the lead in driving changes in personal distribution. Similar considerations, with the opposite sign, apply to pensions and transfers. In 1993 a 10 per cent increase in the pension share reduced total inequality by 1.8 per cent; in 1998 by 1.6 per cent; in 2010 by only 1.2 per cent.

12.7 Conclusions

From 1993 onwards, personal income inequality in Italy as measured by the SHIW/Gini index did not show any persistent tendency to rise, although a clear tendency is detected in the behaviour of functional distribution. A retrenchment in the wage share continued throughout the 1990s. This chapter has discussed some reasons behind this recorded de-link between functional distribution and the Gini index.

The de-link between the trend in the wage share and the Gini index does not depend on a broadening of wage earners' ownership of capital which has been able to blur the very distinction between labour and

property income. No evidence of 'people's capitalism' can be detected in Italian data since there is a substantial divergence between the distribution of non-imputed property income and labour income. Some disconnection could be due, instead, to a poor representation of capital income proper among income sources. There is some evidence on this pointing towards a phenomenon of great relevance, but of limited influence on the trend in the Gini index. The fact remains, in any case, that capital income proper is a *missing* factor source, a circumstance that cannot fail to affect the Italian distributive outlook, making it difficult to draw clear-cut conclusions.

A second source of de-link between functional and personal distribution can be found in the changing economic position of wage earners. Indeed, even if the reduction in the direct and deferred wage share has been counterbalanced by the increase in the imputed rent share, the Gini index has recorded substantial oscillations over the last twenty years. These changes are mainly driven by the increasing number of poorly paid employees. In particular, when wage income has been distributed among an increasing number of receivers in the left tail of the distribution, the Gini index has shown a tendency to reduction. The point is that, at the beginning of the stage of increasing wage inequality, wage incomes were not the prevalent income source of the poorest quintile of the population. In other words, wage income, though the most equally distributed source, was not the income of the 'very poor', but constituted the prevailing income of the *middle* deciles. The very bottom of the distribution, instead, was made up of the unemployed and those in receipt of transfers. The movement of wage income towards the bottom of the distribution has partially mitigated the effect of the falling wage share and rising earning dispersion on personal income inequality.

The redistribution of a falling wage share towards low- or zero-income subjects has been of such relevance as to reverse the income position of many wage earners and pensioners. These adjustments in the relative positions of many wage and pension earners are clearly two aspects of the same phenomenon. While the distribution of wages depends on the shifts in workers' bargaining power, as well as on the considerable institutional changes in the labour market which have occurred over the last two decades, the distribution of pensions is still heavily influenced by the distribution of wages prevailing over the phase of diminishing inequality. By the same token, while a substantial proportion of wage contracts are still regulated by legislation more favourable to workers, new entrants face a much worse situation. Of course, the fact that many

receivers of transfers as well as mature workers established in the labour market enjoy a privileged status compared to the new entrants does not imply that mature workers and pensioners appropriate an unreasonably high share of production. The same issue can be looked at from another angle, noticing that the present level of wages (as established by new contracts) represent an unreasonably low demand on the part of wage earners – a situation in which they could call their own a very little slice of the cake they are producing, in the sense that the level of many wages and pensions is below the threshold that could support stable growth.

This source of de-link between functional and personal distribution should not be seen as unrelated to the problem of the under-recording of capital income proper. In fact, the move of wage income towards the bottom of the distribution can have an egalitarian effect only to the extent that it is not compounded by a persistent increase in the share of capital income proper. The recorded distributive movements prevented substantial growth in income disparities – and can thus be deemed 'horizontal' – because, with capital income proper at a standstill, changes in functional income distribution are confined among equally distributed factor sources.

The combined effect of a steady share of capital income proper in spite of a falling wage share, and of the 'wage percolation effect' that has helped stabilise Italy's Gini measure of personal distribution over the last twenty years can hardly be considered a significant long-lasting phenomenon. On the one hand, the increasing number of poorly paid jobs cannot indefinitely more than compensate the decrease in stable positions in spite of stagnant output levels; as a result, the progressive replacement of the mature and better paid workforce with young and less protected entrants could not fail to translate into a substantial reduction in the wage share. Indeed, if 'the economic irrationality that forces Italian firms to accumulate factors of production notwithstanding the enduring stagnation in economic activity' continues in the years to come, this would obviously indirectly confirm the growing inability of both the national accounts and surveys to detect capital income proper. On the other hand, the more the wage becomes the prevailing income of the poorest deciles of the distribution, the less poorer subjects will be replaced by wage earners, thus diluting the countervailing effects of the re-ranking on total inequality. For both these reasons, it is very likely that in the near future it will increasingly be possible to move readily from the functional distribution to a reliable Gini measure of personal income distribution.

Notes

1. See Torrini (2005) for an examination of the growth in the profit share in Italy since the mid-1970s.
2. On this, for example, Lilla (2005).
3. On this, see Boeri and Brandolini (2005).
4. Cfr. Brandolini (2009).
5. The year 1987 was chosen as the starting year because it is only from 1987 that interest income and dividends are recorded by the Bank of Italy's survey on household income and wealth.
6. Although it is nowadays customary to refer to disposable income net of depreciation, measuring property income net of the cost of earning it can be hardly considered a satisfactory procedure. Apart from the many statistical problems engendered by this, it must be kept in mind that labour income is instead measured gross of all its cost. Gross disposable income appears preferable in this respect since it does not involve any asymmetry between property and labour income. Unfortunately, starting from 2004, the SHIW only reports income net of depreciation. To ensure comparability with the previous years, incomes from individual businesses in 2004, 2006 and 2008 have been grossed by using the same ratio between gross and net income given by the 2002 survey.
7. The Gini index ranges from 0 (when the percentage of individuals arranged in ascending order of their incomes coincides with the percentage of total disposable income allotted to them), to 1 (when the whole disposable income is allotted to the richest). The income unit to which the calculation refers is the individual. While households appear as the relevant unit for purposes connected with the standard of living, poverty, etc., the individual income receiver seems a better choice when aiming at bridging the gap between functional and personal income distribution. All statistics are computed using sampling weight 'pesofl2', which is obtained by multiplying the sample weight 'pesofl' by a constant such that the weighted population adds to the total Italian population. Individuals aged less than 15 are not counted.
8. The year 1989 ends a 'long phase of diminishing earnings inequality …[that] is largely confirmed by the other scattered evidence available, including the information on wage differentials provided in national accounts' (Brandolini *et al.*, 2002, p. 243).
9. Moreover, some redistribution occurred away from the medium deciles toward the higher deciles (from 20/40 towards 60/80 and 80/90).
10. On the increase in the US wage dispersion see, for example, Autor *et al.* (2006).
11. Considering a group of 29 OECD countries, Italy has the highest Gini coefficient of market income inequality, except Chile. Income taxes and cash benefits play a substantial role in reducing this level of inequality; yet, the levels of disposable income inequality in Italy remain among the highest in the world, being exceeded only by the United Kingdom, Portugal, Israel, United States and Chile. On this, see OECD (2011, p. 3).
12. For an analysis of the significance and limitations of the decomposition see Shorrocks (1982), Podder (1993) and Podder and Chatterjee (2002). For a recent application of this procedure to changes in income distribution in the US see CBO (2011).

13. This index ranges from –1 to 1, and can grow for two different reasons: (a) a rise in the inner concentration of that source (the pure Gini); and (b) a higher correlation of income source k with total income distribution – a measure of the way in which the income source impacts total distribution.
14. This result partly depends on the replacement of records with negative interest income with records of zero interest income in order to obtain Gini indexes bounded in the range [0,1].
15. See Lerman and Yitzhaki (1985). See also Van Kerm (2009).
16 Plotting the cumulative proportions of income from a specific source against the cumulative proportion of households arranged in ascending order of their *total* income, the concentration curve does not have to lie below the egalitarian line.
17. The importance of the first kind of error seems to be at least equal to that of the latter, since problems of representativeness of the sample are particularly severe in the tails of the distribution. In other words, it is highly unlikely that the sample (whatever its size) could be able to include the extremely rich, a problem of great importance, given that non-sample analyses have identified the fundamental determinant of the increased inequality in the last thirty years just in the enormous increase in the share of income appropriated by the very few very rich. On this, see the essays in Atkinson and Piketty (2010).
18. Apart from differences in definitions and methods, the main difficulty in carrying out that comparison lies in the fact that while the survey records income sources net of tax, national accounts aggregates are gross of income taxes. An allotment of income taxes among different sources is thus needed.
19. These figures do not take into account undistributed corporate profits. There are good arguments for including corporate undistributed profits in households' income. It is only by consolidating households' income and income retained by corporations that a clear link can be established between functional and personal distribution. The exclusion of corporate savings, in fact, has the effect of creating a leakage in functional sources, making personal income distribution dependent on the dividend policies of corporations. Taking properly into account corporate undistributed profits it seems possible to ease the apparent contradiction between the fall in the capital share of households' disposable income recorded by SHIW over the 1990s and the fact that '[i]n Italy, according to national accounts, in the private sector the share of capital in value added at factor costs … grew steadily from the mid 1970s until 2001, reaching historically high levels. That growth, common to other industrial nations, results from increased profit margins in many sectors, not from a reallocation towards high capital intensity industries' (Banca d'Italia, 2001, p. 133).
20. This result is a permanent feature of the survey: 'Except in a few years, income from employment was satisfactorily measured until 1989, but in the early 1990s the discrepancy with the aggregate data rose to over 10 per cent. By contrast, self-employment income and net interest and dividends appear to have been very poorly captured by the survey: the shortfall of the former ranged from 50 to 60 per cent of the national account figures, while that of the latter was about three-quarters between 1987 and 1995. The gap between survey totals and national accounts tended to narrow only for transfers: in

1995 it was down by about a quarter. Rents, inclusive of imputed rents of owner-occupied dwellings, are the only item for which survey data generally overstated national accounts figures.' (Brandolini, 1999, p. 47)

21. On the other hand, the misreporting of income from financial assets is shown as statistically insignificant. See on this Brandolini *et al*. (2004), and Boeri and Brandolini (2005, pp. 14–15).
22. 'With reference to inequality, we look at two standard measures – the inter-decile ratio and the Gini index – applied to disposable income and separately to each income source… Total earnings, as well as most income sources, turn out to be less equally distributed after the adjustment; for income from financial assets only we obtain opposing indications from the two indices.' (Neri and Zizza, 2010, p. 16)
23. In the private sector, the average annual growth rate of labour productivity in 1991–95 was 2.3 per cent; in 1996–2000 it was 1.1 per cent; in the period 2001–2007 it was 0 per cent (in the industrial sector the corresponding data were 3.6 per cent, 1.3 per cent, and –0.13 per cent). Over the same period, the number of employees in the private sector increased from 12 millions (1993) to 14.3 millions (2007). Restricting our attention to the period of zero growth in labour productivity, the number of employees grew from 13 millions to 14.3 millions. More data on this in Visco (2008, pp.14–22).
24. On this controversial issue, see Banca d'Italia (2009, chapter IV). See also Bugamelli (2007) and Codogno (2009).

References

Atkinson, A.B. and Piketty, T. (2010) *Top Incomes over the Twentieth Century, Vol. II: A Global View* (Oxford: Oxford University Press).

Autor, D.H., Katz, L.F. and Kearney, M.S. (2006) 'The polarization of the U.S. labor market', *AER Papers and Proceedings*, May: 189–94.

Banca d'Italia (2001) *Relazione Annuale* (Roma: Banca d'Italia).

Banca d'Italia (2008) *Relazione Annuale* (Roma: Banca d'Italia).

Banca d'Italia (2009) *Rapporto sulle Tendenze del Sistema Produttivo Italiano* (Roma: Banca d'Italia).

Biancotti, C., D'Alessio, G. and Neri, A. (2008) 'Measurement error in the bank of Italy's Survey of Household Income and Wealth', *Review of Income and Wealth*, 54: 466–93.

Boeri, T. and Brandolini, A. (2005) 'The Age of Discontent: Italian Households at the Beginning of the Decade', *IZA Discussion Paper Series*, no. 1530.

Brandolini, A. (1999) 'The Distribution of Personal Income in Post-War Italy: Source Description, Data Quality, and the Time Pattern of Income Inequality', Banca d'Italia, *Temi di Discussione*, 350.

Brandolini, A. (2005) 'La disuguaglianza di reddito in Italia nell'ultimo decennio', *Stato e Mercato*, 74: 207–29.

Brandolini, A. (2009) *Indagine Conoscitiva sul Livello dei Redditi da Lavoro nonché sulla Distribuzione della Ricchezza in Italia nel Periodo 1993–2008* (Roma: Senato della Repubblica).

Brandolini, A., Cannari, L., D'Alessio, G. and Faiella, I. (2004) 'Household Wealth Distribution in Italy in the 1990s', Banca d'Italia, *Temi di Discussione*, no. 530.

Brandolini, A., Cipollone, P. and Sestito, P. (2002) 'Earnings Dispersion, Low Pay and Household Poverty in Italy, 1977–1998', in D. Cohen, T. Piketty and G. Saint-Paul (eds), *The Economics of Rising Inequalities* (Oxford: Oxford University Press).

Bugamelli, M. (2007) 'Prezzi delle Esportazioni, Qualità dei Prodotti e Caratteristiche di Impresa: un'Analisi su un Campione di Imprese Italiane', in A. Lanza and B. Quintieri (eds), *Eppur si Muove. Come Cambia l'Export Italiano* (Soveria Mannelli: Rubbettino).

Codogno, L. (2009) 'Two Italian Puzzles: Are Productivity Growth and Competitiveness Really So Depressed?', Ministry of Economy and Finance, *Working Papers*, no. 2, March.

Congressional Budget Office (2011) *Trends in the Distribution of Household Income between 1979 and 2007* (Washington: The Congress of the United States).

European Central Bank (2004) 'Labour productivity developments in the euro area: aggregate trends and sectoral patterns', *ECB Monthly Bulletin*, July: 47–57.

Friedman, M. (1962) *Price Theory* (New Brunswick, NJ: Transaction Publishers).

Gomez-Salvador, R., Musso, A., Stocker, M. and Turunen, J. (2006) 'Labour Productivity Developments in the Euro Area', ECB, *Occasional Paper Series*, no. 53, October.

ISTAT (2008) 'Il Valore della Moneta in Italia dal 1861 al 2007', ISTAT *Informazioni,* no. 8.

Lerman, R.I. and Yitzhaki, S. (1985) 'Income inequality effects by income sources: a new approach and applications to the United States', *Review of Economics and Statistics*, February, 67(1): 151–6.

Lilla, M. (2005) 'Disuguaglianze salariali in Italia: nuove evidenze dai microdati SHIW', *Politica Economica*, XXI(1): 71–102.

Neri, A. and Zizza, R. (2010) 'Income Reporting Behaviour in Sample Surveys', Banca d'Italia, *Temi di Discussione*, no. 777.

OECD (2011) *Divided We Stand; Why Inequality Keeps Rising, Country Notes. Italy*, http://www.oecd.org/italy/49177743.pdf, accessed: 25 January 2013.

Podder, N. (1993) 'The disaggregation of the Gini coefficient by factor components and its application to Australia', *Review of Income and Wealth*, 39: 51–61.

Podder, N. and Chatterjee, S. (2002) 'Sharing the national cake in post-reform New Zealand: income inequality trends in terms of income sources', *Journal of Public Economics,* 86: 1–27.

Shorrocks, A. F. (1982) 'Inequality decomposition by factor components', *Econometrica*, January, 50(1): 193–212.

Torrini, R. (2005) 'Quota dei profitti e redditività del capitale in Italia: un tentativo di interpretazione', *Politica Economica* XXI(1): 7–43.

Van Kerm, P. (2009) 'S-Gini – Generalized Gini and Concentration Coefficients (with Factor Decomposition) in *Stata*, v. 1.0, CEPS/INSTEAD, Differdange, Luxembourg.

Visco, I. (2008) 'Testimonianza del Vice-Direttore Generale della Banca d'Italia Ignazio Visco, 28 novembre, Indagine Conoscitiva sull'Assetto delle Relazioni Industriali e sulle Prospettive di Riforma della Contrattazione Collettiva, 11° Commissione (Lavoro pubblico e privato)' (Roma: Camera dei Deputati) ['Testimony of Deputy General Manager of the Bank of Italy Ignazio Visco, 28 November, survey on industrial relations and the prospects for reform of collective bargaining, 11° (public and private)' (Rome: Italian Chamber of Deputies)].

13

Exchange Rate Policy, Distributive Conflict and Structural Heterogeneity: The Argentinean and Brazilian Cases

*Fabián Amico and Alejandro Fiorito**

13.1 Introduction

This chapter is a comparative study of monetary policy and the inflationary process in Argentina and Brazil, based on the classical approach to the determination of prices and distribution. Our aim is to suggest some lines of inquiry in a specific framework for the analysis of inflation, and its relation to macroeconomic policies and structural heterogeneity in Argentina and Brazil.

The period under analysis is from the 1990s up to the present time. This phase is significant as it allows us to compare the different international trends (for instance, commodity prices) and their impact on domestic economies, together with several inflationary processes and the monetary policies in countries with strong structural heterogeneity features.

We will discuss the inflationary processes in developing countries that face strong external restrictions on development, in the light of the monetary theory of distribution. In the more general framework of the

* We would like to thank the organizers of the International Conference 'Sraffa's Production of Commodities by Means of Commodities 1960–2010. Critique and reconstruction of economic theory', and especially Antonella Palumbo. We also particularly thank (without attribution) Franklin Serrano for very helpful comments. Finally, we want to thank Professor Carlos Pinkusfeld for providing data about the Brazilian economy.

classical approach, we will introduce a specific topic, such as structural heterogeneity and its implications for macroeconomic policies and development. Prominent among the many factors that influence and constrain central bank policy are the conflicts that emerge from structural heterogeneity, so exchange rate policy is one of the fundamental areas where that conflict is settled.

We begin, in Section 13.2, with a brief discussion of the classical approach to price determination and inflation, examining some of the contributions made by Pivetti. Then, in Section 13.3, we will describe the inflationary processes in both countries during the period 1990–2010. We have found that the exchange rate channel to monetary policy has been of decisive importance for both the Argentinean and the Brazilian cases. In Section 13.4, in line with Diamand's (1968) analysis and some other recent contributions (Medeiros, 2010), we will discuss the relationship between exchange rate policy and unbalanced productive structures (UPSs). We also explore further the classical approach and its causes of inflation in the debate between Pivetti and Serrano. Section 13.5 contains some concluding remarks.

13.2 Monetary theory of distribution, exchange rate and inflation

In line with the classical approach to economic theory, inflation is regarded as the result of incompatible influences and claims over income distribution, like other strands in heterodox economic thought.

Within this framework, a main criticism raised by Pivetti(1991) of the Kaleckian, and other heterodox models, is about the indeterminacy of the mark-up level. Certainly Pivetti (1991), Serrano (1993; 2010) and Stirati (2001) have developed a research line focusing on the possibility that mark-up pricing, instead of being a description of pricing mechanisms, actually becomes an *exogenous* distributive variable with respect to real wage. Thus, instead of assuming a mark-up to a level consistent with a given real wage, it is the latter which depends on the former. Consequently, the share of wages in surplus (i.e., above subsistence) will depend on the level achieved by the real mark-up.

The relationship between interest rates and prices has been analysed by Brazilian and Argentinean economists, though without drawing more general conclusions in relation to distribution determination (see, for example, Frenkel, 1979; Arida and Resende, 1985). Certainly, in the monetary theory of distribution, the role of the interest rate as the determinant of the opportunity cost of capital is considered a *more*

general approach. Indeed in Pivetti's view the *whole* capital invested in the real sector is subject to the same arbitrage process between interest and profit rates, even though it takes a long time. In other words, the profitability of the productive capital cannot continue to differ significantly from the interest rate over a long period of time.

On the other hand, Pivetti (1991; 2001; 2010) has drawn attention to the relationship between the interest rate and the exchange rate. He examines the connection between the notion of endogenous money and monetary non-neutrality, 'which would become hardly disputable once the rate of interest was acknowledged as a monetary phenomenon' (Pivetti, 2001, pp. 104–5). In this approach, given money wages and production techniques, a lowering (rising) of the money rate of interest by the central bank would actually drive the price level down (up), 'owing to the adaptation of prices to normal costs caused by competition' (Pivetti, 2001, p. 112). In this case, there would not be anything 'paradoxical' (i.e., the Gibson case) in the positive correlation between interest and prices.

In this framework, the money rate of interest is conceived as an autonomous determinant of normal money production costs, which govern the ratio of prices to money wages. Thus wage bargaining and monetary policy 'come out of this analysis as the main channels through which class relations act in determining distribution' (Pivetti, 2001). The class relations are seen as tending to act primarily upon the profit rate, via the money rate of interest. Thus, the real wage in a given situation is viewed as the final result of the whole process. In other words, it is a *residual*.

In turn, it has been observed that this argument could be made to stand on its head. For instance, as Serrano said:

> One could equally argue that corresponding to the exogenously given nominal rate of interest there is always a rate of increase in money wages that would produce enough inflation to reduce the real rate of interest, and hence profits, enough to allow the workers to obtain their desired real wage. (Serrano, 1993, p. 122)

In principle, Pivetti's view that the behaviour of money wages might take the lead in the determination of real interest is correct (Pivetti, 2007, p. 245), but he considers that in fact it is generally much easier, both technically and politically, for the monetary authorities rather than the workers of money wages to establish the course of nominal interest rates. However, Serrano's observation suggests the possibility of different mechanisms through which prices and distribution could be determined,

and in the following sections we consider these mechanisms in the specific context of Argentina's and Brazil's economic structures.

13.3 A comparison between Argentina and Brazil

In the following sections we review the main factors in the evolution of the inflationary processes in Argentina and Brazil in the light of the hypothesis discussed above. Toward 1995 both countries put an end to their hyperinflation processes and in the 1980s they had followed similar paths, albeit with individual characteristics. Toward the beginning of the 1980s they both lost access to external financing and at the same time, their terms of trade worsened, causing their international interest rates to rise, and the external demand slowed down due to the world recession. The reaction of the countries in the region was to implement a recessive control on imports and to make aggressive devaluations in an attempt to increase exports. These policies led to economic stagnation and explosive inflation (Medeiros and Serrano, 2006). The objective was the repayment of external debt. The consequence was a pattern of low growth and high inflation.

13.3.1 Argentina: from the currency board to the managed floating exchange rate

The so-called high inflation regime collapsed with two hyperinflationary episodes in 1989 and 1990. In this context, the central bank aimed at keeping the nominal exchange rate relatively stable. In the meantime, the inflation rate remained very high and the resulting real exchange rate (RER) appreciation led, in early 1991, to a new round of runs against the peso and a rise in the exchange rate. Fearing that new depreciation could lead the economy towards hyperinflation again, in March 1991 the government established fixed parity between the domestic currency and the US dollar by the so-called convertibility law, fully backing the monetary base with foreign exchange reserves. This law transformed the central bank into a currency board. The system included, from early 1991, an almost complete liberalisation of trade flows and full deregulation of the capital account. This was accompanied by a strong process of market-friendly reforms, including the privatisation of state-owned firms.

Under the convertibility regime, Argentina experienced two periods of sustained capital inflows that spurred growth. The first occurred between the 1991 and the contagion of the Mexican crisis in 1995. The second was shorter; it began shortly after the 'tequila effect' and stopped in mid-1998. The economy then remained in a severe depression that

led to the dramatic collapse of the regime in 2001–02, concluding with the abandonment of the currency board, a major devaluation and default on the external public debt.

The focus on fiscal irresponsibility as the main cause of the convertibility crisis is very controversial. Several authors have shown that the authorities followed a contractionary policy from 1999 and that the increase in public expenditure was mainly due to increasing debt services (Frenkel and Rapetti, 2010). In a context of exchange rate appreciation, given that the improvement in labour productivity and other measures were insufficient to correct the lack of competitiveness in the tradable sector, a significant deflation of domestic non-tradable goods prices would have been required to correct the RER misalignment. But prices (wages) are downward inflexible.

In the peak of a cycle, RER appreciation tends to stimulate domestic demand and if there is no early correction of the RER misalignment, a persistent current account deficit may lead to an unsustainable accumulation of external debt. Then, the required RER depreciation could make foreign-indebted domestic agents (in either the private or public sector) go bankrupt.

The recovery of economic activity since 2002 has been associated with pragmatic macroeconomic policies that have offset the external and fiscal imbalances, while providing incentives for tradable activities. In both cases, the exchange rate policy was oriented towards maintaining a competitive exchange rate.[1] In 2003, the government started to intervene actively in the foreign exchange market to contain the appreciation pressure. Although the parliament passed a law revoking the currency board, the government decided to maintain the central bank's independence with a basic mandate of pursuing low inflation.

In 2004 a proposal to establish an inflation-targeting regime was rejected. Thus, going against the dominant economic opinion that the central bank has no control over the money supply, Argentina opted to follow a policy based on *quantitative monetary targets*.

To achieve this goal, the central bank relied on ´sterilisation´ operations via the issuing of central bank bonds.[2] At the same time, the Argentine government introduced capital controls in June 2005. During this period domestic interest rates were relatively low, the strong surplus obtained in the current account prior to the 2002 devaluation having enabled a reduction on the higher interest rates that were in place up until 2002 as a result of high country risk premiums or monetary policies aiming to attract capital flows (see Figure 13.1).

As a result of this strategy Argentina experienced one of the most successful growth episodes in its economic history. From mid-2002 to mid-2008,

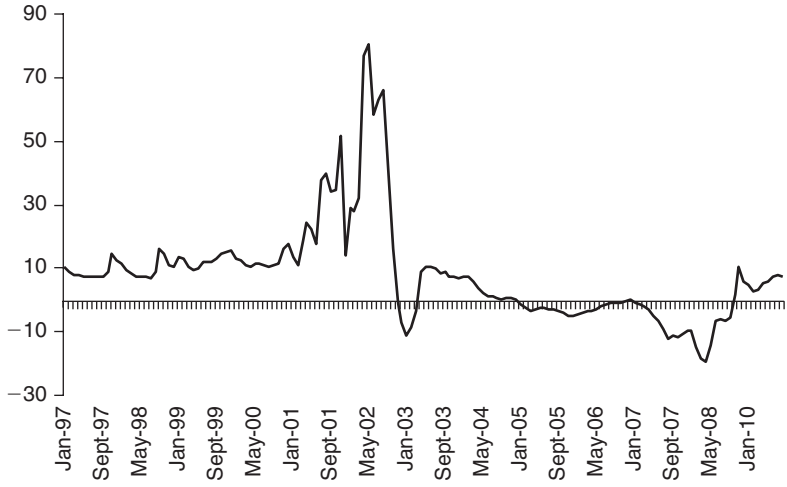


Figure 13.1 Argentina: real interest rate, 1997–2009
Sources: Banco Central de la República Argentina and Cenda.

the economy grew at an average annual rate of 8.5 per cent. The favourable external conditions were important, but they only partially explain the economic performance. It was the expansion of the tradable sector as a whole that pulled the economy up and set it on the path to rapid growth. This boost was then passed on to the domestic market, and since 2004 the growth process has been driven by domestic demand factors.

It is important to distinguish clearly between the two different roles played by exports. The expansion of exports is capable of simultaneously relaxing two different constraints on growth, namely the external financing constraint and the effective demand constraint.[3] Thus, at the beginning of the Argentinean recovery exports played an important role in *both* aspects (i.e., as a source of demand and as a source of finance for imports). But since 2003 the economic growth process has been increasingly led by domestic factors.[4]

Two crucial factors explain this unusual expansion. Firstly, the positive evolution of commodity prices in world markets, combined with a competitive exchange rate. Secondly, it was the first time in more than three decades that the government had not used demand-constraining policies and unemployment to fight inflation (more on this below). In this context, the economy expanded while maintaining a stable current account surplus.

However, it was during this period, from 2005, that inflation began to accelerate. One cause was the global rise in commodity prices,

especially during 2007 and 2008. Unlike Brazil, the effect of increasing commodity prices in Argentine inflation was not offset by the nominal appreciation of the exchange rate. Many economists in Argentina (see Frenkel, 2008) suggest that inflation is the result of demand pressures. But aggregate investment reacted very sensitively to economic growth and, with the exception of some specific sectors, capacity remained far from being fully utilised.

The evidence seems to suggest that inflation accelerated as a result of a combination of two factors: the increased bargaining power of workers (since 2003 unemployment had declined sharply and the wage bargaining position had changed) and the peak price of commodities recorded in 2007–08 (see Figures 13.2 and 13.3).

Moreover, in 2002 Argentina introduced a fixed-rate tax on the export value of commodity exports (food, soya and other crops) causing the price of wage goods to decline relative to their previous levels. But this tax couldn't prevent inflation *acceleration*. Given the fact that the Argentine export structure is based on wage goods, the rise in commodity prices pushed up nominal wages, which put pressure on costs throughout the economy. Thus, empirical evidence suggests that there are inconsistent income claims, which determine a higher and increasing level of inflation. In fact, if the workers aim at preserving their real wage level and the industrial sector pursues a policy of preserving the real mark up, this would set in motion a process of accelerating inflation, as result of inconsistence linked with real income growth.

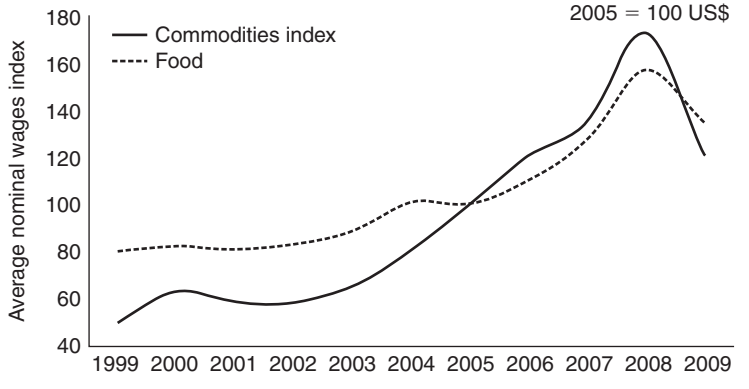


Figure 13.2 International food price index, 1999–2009
Source: Indec.

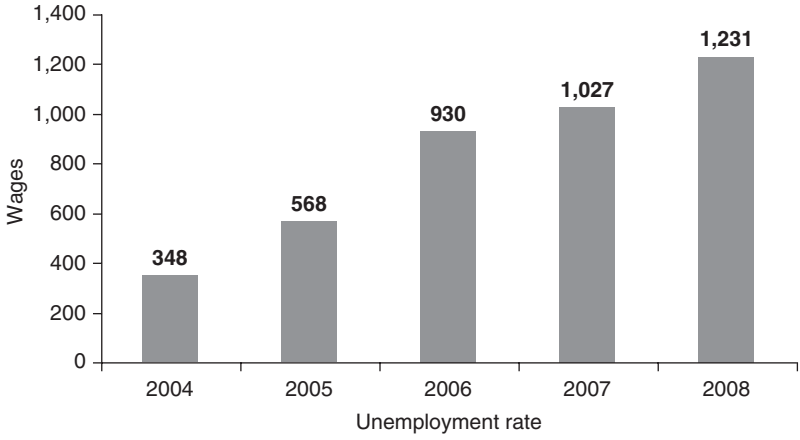


Figure 13.3 Argentina: annual wage negotiations by labour unions, 2003–08
Source: Indec.

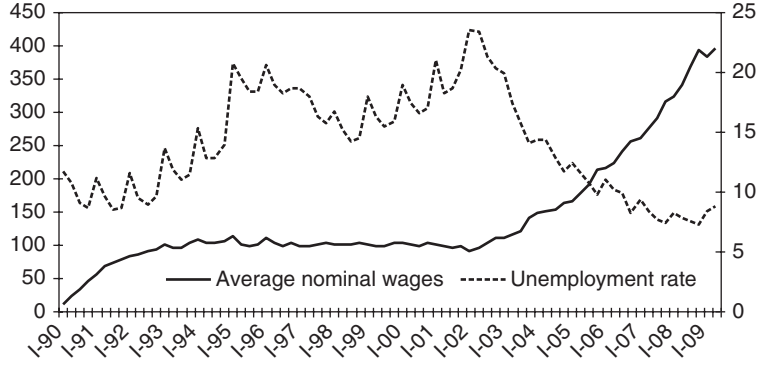


Figure 13.4 Argentina: nominal wages index (left axis) and unemployment rate, 1997–2009
Source: Indec.

In this context, a principal problem for the present regime is the lack of alternative policies to check the inflation that results from the distributive conflict. The distributive changes must translate into a rise in the price level and then it becomes difficult to maintain low inflation rates. As Figure 13.4 shows, high rates of growth after the crisis entail low unemployment and sooner or later increase the workers' bargaining power. If the government does not intervene, then the 'invisible hand

of the market' can translate these inconsistent claims on income into inflation acceleration. The exchange rate will then appreciate and will set in motion the devaluation. Finally, this process drives the economy down the traditional path of stop-and-go cycles and the so-called 'Argentinean pendulum' (see Diamand, 1986).

13.3.2 Brazil: hyperinflation, Real Plan and inflation-targeting

The Brazilian economy adopted a fixed exchange rate as part of the Real Plan in 1994. This fixed exchange rate regime led to an appreciation of its currency and greater current account deficits. After a decade of very high inflation rates, in 1994 the government reissued the real and set up a crawling peg system. The new regime stabilised inflation for the first time in decades.

But the Brazilian macroeconomic stabilisation strategy was heavily dependent on the continuous inflow of foreign capital and, as a result, the international financial position of Brazil became increasingly fragile. Thus, by the end of 1998 Brazil's current account deficit reached 4.5 per cent of GDP and the low stock of foreign reserves of the Brazilian central bank did not allow a defence of the Brazilian exchange rate, should another speculative attack hit the country (see Barbosa, 2008).

After the January 1999 crisis, Brazil adopted an inflation-targeting regime and the operational independence of central bank has *de facto* taken place. In the Brazilian case, real interest rates remained high during the inflation-targeting periods. Also, the real exchange rate tended to appreciate after the strategy was implemented. It is a fact that appreciation has benign effects on price stability and depreciation tends to be inflationary. Thus, Barbosa (2008) concluded that inflation-targeting managed to reduce inflation in Brazil after its 1999 and 2002 currency crises, with substantial help from real exchange rate appreciation. Besides, economic growth was slower under inflation-targeting than under exchange rate targeting. Likewise, the real interest rate of the Brazilian economy remained well above international standards and required a substantial increase in fiscal austerity by the government. A key has been that the high domestic real interest rates and the favourable international trade and financial conditions in the rest of the world (as we said above, mainly the rise of primary export prices and low international interest rate) allowed the Brazilian government to accumulate foreign reserves and repay most of its foreign debt.[5]

At the same time, it is important to point out that the switch to inflation-targeting preserved government's commitment to a neoliberal policy agenda.

Table 13.1 Brazil: inflation targeting, 1999–2008

Year	Target	Interval	Actual inflation	Fulfilment
1999	8.0	6.0 to 10.0	8.9	yes
2000	6.0	4.0 to 8.0	6.0	yes
2001	4.0	2.0 to 6.0	7.7	no
2002	3.5	1.5 to 5.5	12.5	no
2003	8.5	8.51	9.3	no
2004	5.5	3.0 to 8.02	7.6	yes
2005	4.5	2.0 to 7.0	5.7	yes
2006	4.5	2.5 to 6.5	3.1	yes
2007	4.5	2.5 to 6.6	4.5	yes
2008	4.5	2.5 to 6.7	5.9	yes

Source: Banco Central do Brasil.

If one looks at the inflation targets (see Table 13.1), the actual inflation rate did not achieve its target in 2001 and 2003. The use of a broad inflation index, rather than a core inflation target (i.e., excluding energy and food prices), in the Brazilian case implies that if inflation is caused by supply-side shocks, then the inflation-targeting framework, which is based on the supposition that inflation is caused by demand factors, will lead to higher interest rates in order to reduce demand pressures (Vernengo, 2008). But several authors have shown that a systematic relationship between demand and inflation acceleration does not exist (see Serrano, 2010; Summa, 2010). Thus, a question emerges: how is it possible to check inflation from interest rate in an economy in which it does not seem feasible to regulate inflation through demand control?

Serrano (2010) showed that a rise in the domestic interest rate generates positive interest differentials and produces a revaluation of domestic currency. In turn this revaluation transforms the *negative* supply shocks (i.e., increasing commodity prices in the world market) into a *positive* supply shock in terms of domestic inflation. Obviously another outcome of this interest rate policy could be less demand, slower output growth and higher unemployment. It is possible that these effects could result in a low growth of nominal wages.

Thus, in the years in which the actual inflation rate has been steered towards the target, the nominal exchange rate has appreciated. In a context of strong commodity price increases, the exchange rate was a mechanism that allowed actual inflation to fit the target. However the fundamental reason for low inflation underlying the success of the

system was low wage indexation. Since 2004, wages in Brazil have shown an invariable level of considerably slower growth than in the mid-1980s. These data are a sign of the low bargaining power of the workers (see Bastos *et al.*, 2010). Thus, the anchor of the system is the low wage resistance (Serrano, 2010, p. 68).

As a result of this, between 1994 and 2002 the exchange rate depreciation was compensated for, to some extent, by the slowdown tendency in commodity prices, reducing the depreciation impact on domestic inflation. However, between 2003 and 2008, the reverse was the case: the effect on domestic inflation from rising commodity prices was offset by an exchange rate appreciation (Serrano and Ferreira, 2010).

In the specific cases under analysis, long before the changes in interest rates could increase the price level, the rise in interest rates was set in motion by the exchange rate channel, and through this mechanism it ended up changing distribution and inflation. Moreover, the exchange rate channel affected inflation in the *opposite* direction to the straightforward effect of the interest rate on prices. In some cases, normal distribution is probably *in fact* governed by exchange rate policy, the interest rate becoming a less important variable.

Indeed, empirical analysis of Brazilian inflation in recent years shows that, since the basic interest rate sets a floor for the level of profit margins, monetary policy (based on high real interest rates) has helped to sustain high real value of profit margins in the long term (see Figure 13.5). But, in inflationary terms, this channel is more than compensated for by the effect of interest rate differentials on exchange rates. Thus, the

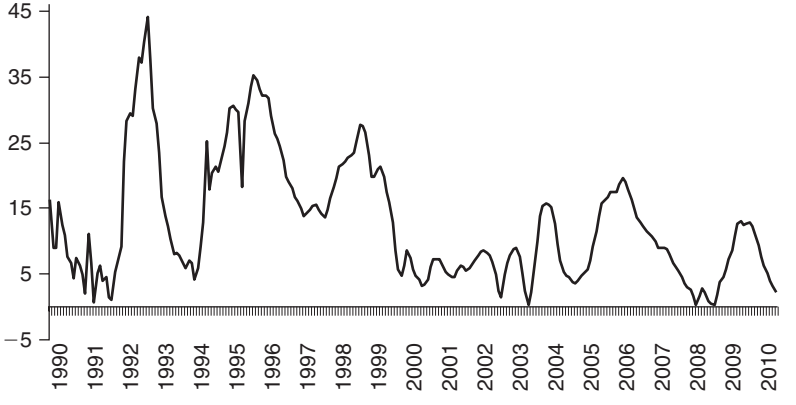


Figure 13.5 Brazil: real interest rate, 1990–2010
Source: Ipeadata.

evolution of nominal exchange rate and of the tradable goods prices in dollars were the main factors in determining Brazilian inflation for this period (see Bastos *et al*., 2010; Serrano *et al*., 2010).

More generally, in the 1990s the interest rate policy focused more on external restriction. Likewise, since the 2000s Argentina has sharply reduced the nominal and real interest rates and inflation has risen due to exchange rate devaluation. While in the 1990s, Argentina had a high real interest rate and low inflation, since 2002 there has been a relatively low real interest rate with higher inflation. In both cases, the exchange rate channel has played a decisive role in setting the path of inflation rates. Thus, the 'counterbalancing mechanism' mentioned by Pivetti appears as the *more general case* in the Latin American countries, and the exchange rate channel as a *fundamental transmission mechanism* for distribution, output and inflation.

Additionally, several research studies on Latin American countries have shown that the elasticity of aggregate demand to changes to the interest rate is weak, which rules out the new consensus prediction that increasing interest rates will fight inflation by slowing down aggregate demand (see Barbosa, 2006 and 2008, for Brazil; Galindo and Ros, 2008, for Mexico; and Chang, 2007, for several Latin American economies). Therefore, the transmission mechanism of interest rate to the inflation rate was mainly the exchange rate dynamics, and not the rate of change of aggregate demand.

In fact, inflation slowed systematically, whereas interest rates had increased and these results were related to inflation-targeting policies. All cases show at the same time sustained processes of exchange rate revaluation. Likewise, these processes have exhibited low output growth rates. Research on these countries emphasises that exchange rate appreciation is a crucial factor. It also shows a *negative* relationship between interest rate and exchange rate tendency. A tendency for the exchange rate to appreciate seems to check inflation rates.

13.4 Exchange rate policy and unbalanced productive structures (UPSs)

The inflation tendency in Argentina and Brazil, and its relationship to exchange rate policy, cannot be understood without analysing the many factors that may influence and constrain the central banks' policy. In these cases, rather than a simple relationship between workers (wage bargaining) and capitalists (central bank policy), there is a more complex social, political and economic structure based on the

determination of monetary policy, specifically exchange rate policy, as the main transmission channel to distribution, inflation and output.

The structural heterogeneity that characterises countries like Argentina and Brazil is rather different from the traditional view of ECLAC (Economic Commission for Latin America and the Caribbean).[6] According to this specific view, both Argentina and Brazil exhibit a structural imbalance between the productivity of the primary export sector and industrial productivity (see Diamand, 1986). The option to industrialise these countries implies that industrial prices will be higher than international prices.

We will now discuss some specific productive structures, the main feature of which is the coexistence of two sectors with very different productivities: the less dynamic primary sector, which works at international prices and exports, and the protected and more dynamic industrial sector, which works at higher prices than the international level and – unless it is given special industrial exchange rates for exports – produces only for domestic consumption. Marcelo Diamand (1987) called these arrangements *unbalanced productive structures* (UPSs).[7]

In this arrangement, without any special intervention from the state, the exchange rate is generally adjusted to the primary production level. This exchange rate level is inadequate for industrial exports. Rises in the exchange rate imply a fall in real wages, which societies are reluctant to accept. In this context, the imbalances in the productive structure result in an exchange rate which is not competitive in terms of industrial exports. In conditions of financial deregulation and opening capital account, these circumstances are worse if they are put together with a domestic interest rate higher than the international standard (such as in the case of Brazil). In the present situation, Argentina could partially neutralise this tendency through a tax on commodity exports, a lower interest rate and major controls over capital account.

These unbalanced productive structures mark a big difference between the development experiences of Latin American countries (Argentina and Brazil) and Asian countries. In the cases of Korea's and China's domestic currencies, depreciation strengthens and unifies the interests of productive sectors around industrial objectives. However in Argentina and Brazil things are different. In these countries the *ownership* classes and the *workers* are divided in determining the effective exchange rate, and multiple exchange rates favourable to industrial capitalists (for example through an export tax) find powerful opponents (see Medeiros, 2010).

At the same time, the determination of a higher real exchange rate favourable to export diversification *without* export tax involves a lower

real wage and creates a conflict with the workers. Thus, exchange rate appreciation is functional in increasing real wages, and then the strategy of competitive real exchange rate encounters serious political troubles. Likewise, things are even more complex given the fact that the industrial sector, although it loses competitiveness, is favoured due to increasing real wages.[8]

Basically, in these two economies, export-oriented industrialisation would require a big income sacrifice to achieve competitive international low wage costs via devaluation (Mahon, 1992, p. 242).[9] Moreover, in these countries the cyclical trend in terms of trade was a weak and inconstant motivation for a change to a policy that is sympathetic to industrial exports.

According to Diamand's view, this unbalanced productive structure is the main factor that explains the exchange rate policy and the class conflict. Due to abundant primary exports the exchange rate tends to appreciate permanently, therefore manufacturers of the country that has an unbalanced productive structure are relatively uncompetitive internationally. The pro-industrial, export-oriented reforms involve a large cost in domestic income and wages due to the huge real devaluation that is necessary to make industrial exports competitive. The 'opportunity cost' of reorienting policy toward a competitive industrial sector will be greater as a primary sector is relatively more productive. In the context of the full deregulation of the capital account, this historical fact is aggravated due to the huge capital inflows that set in motion exchange rate appreciation.

Therefore, structural heterogeneity may constitute an obstacle to the development process, strengthening primary export orientation and resulting in a growth model characterised by short-run cycles of prosperity and recession. For the same reasons import-substituting industrialisation (ISI) was supported not only by factory owners but also by workers. Unions were especially prominent in rejecting and/or postponing devaluation. Besides this, policy makers found a good excuse in the 'technical' reasons for resisting devaluations (based on 'elasticity pessimism'). These trends led to the economic and social process that has brought both countries to deadlock. Thus industrial growth was accompanied by recurrent exchange crises, and latterly, the debt crisis of the 1980s has exacerbated the original problem.

However, all this does not mean that low wages *per se* would have been a solution for structural problems. In fact, with the 1980s debt crisis real wages fell steadily in Latin American countries, which at the same time entered into the so-called 'debt-financed' deindustrialisation

process. Moreover, huge real devaluations produced hyperinflation in both Argentina and Brazil, which was part of the trend pushing real wages down.

As we saw in Section 13.2, Pivetti´s framework regards the rate of interest as a policy-determined variable, which is not subject to any general law, and is 'determined from outside the system of production' (Sraffa, 1960). The objectives and constraints that are imposed on the actions of the central bank will define the rate. These actions and constraints are able to alter at any given time within a country and it is possible that they will also differ between countries.

In the mechanism suggested by Pivetti (1991) the interaction between the real and nominal interest rates, and the inflation rate is crucial. Thus:

> given a policy-determined nominal interest rate, competition among firms within each industry should tend to cause the rate of profit to move in sympathy with the real rate of interest, rather than with [the] nominal one, because it is the former which constitutes the actual price for the use of capital in production, or its opportunity cost. (Pivetti, 1991, p. 52)

Therefore, according to Pivetti the level of the nominal interest rate determines the nominal mark-up and the nominal return, which is obtained by the firms competing at the end of the process. If the real interest rate is an opportunity cost of capital, the relative profit rate of productive capital will not permanently differ from the average real return on public bonds. Thus, firms need to have a pricing policy that is consistent with the orientation of monetary policy. When the inflation rate increases, the central bank will change the nominal interest rate and the firms will adjust their nominal margins accordingly.[10]

In this context, the rate of interest is a policy variable determined ultimately by class relations. Since the interest rate governs the ratio of prices to money wages, and given the nominal wages, a rise in interest rates by the central bank will raise the price level, because it increases the mark-ups, lowering real wages at the same time. Pivetti states:

> The rate of interest is thus regarded as an autonomous determinant of normal prices: a dearer money policy is *by itself inflationary*, through its direct impact on mark-ups. The *overall net impact* on the price level essentially depends on the effects that the policy determined interest rates will eventually exert on aggregate demand and employment, through their impact on income distribution and the other channels

> by which changes in interest rates are bound to affect activity levels, starting from the leverage they exert on net exports through *the exchange rate*. (Pivetti, 2010, p. 220 italics ours)

This suggests that a dearer money policy can (in principle) be inflationary but the final result *can be the opposite*, according to interaction with *other* effects (aggregate demand, income distribution and exchange rate).

In this sense, Stirati (2001) has proposed a kind of 'taxonomy' of conditions that can lead to different inflationary processes, focusing on the influence of labour market conditions on the dynamics of prices, and taking as *given* the alternative policies followed by the central bank on interest rates. The author does not discuss 'the many factors that may influence and constrain the central bank's policy with respect to the interest rate', but suggests that this is, in itself, 'a very important area for research' (Stirati, 2001, p. 430). This is one of the main objectives of the present chapter. Therefore, to analyse the factors that constrain and influence monetary policy, we must focus the analysis on the *specific* conditions that affect economies, such as Argentina and Brazil, with strong features of structural heterogeneity.

On this point, from the empirical side, it is clear that the exchange rate channel has actually played a decisive role in those cases in which dear money policies have succeeded in checking inflation. Therefore, in general, as Pivetti (2010) says, once *all* the transmission channels are taken into account, higher interest rates will have succeeded in checking inflation because the higher ratio of prices to money wages they bring about, through their direct impact on mark-ups, will be more than counterbalanced by:

(a) lowering the prices of tradable goods denominated in domestic currency, through the exchange rate channel; and
(b) a reduction or slower rise of money wages as a result of the likely negative impact on employment brought about by changes both in normal income distribution and in the exchange rate – i.e., by the contractionary effects on consumption spending and net exports caused by higher interest rates.

In this framework, it is important to recall that the distribution of the surplus is 'arbitrary' because it does not follow any law related to the productive structure or 'mode of production', but depends on the relative bargaining power of the parties. The price level is determined *given* the long-term rate of interest and money wages. In a *closed* economy,[11]

through manipulation of nominal interest 'it is always possible in principle … to leave distribution unaffected in the face of any increase in money wages or of any other initial agent of price increases' (Pivetti, 2007, p. 244). In this respect, Pivetti says:

> Of course, non-distributional targets – such as debt management, balance of payments or exchange rate targets – may also strongly influence, in this or that concrete situation, policy decisions concerning interest rates. Given one or another of these targets, the monetary authorities might well decide, for example, to keep nominal interest rates unchanged in the face of increases in money wages. (*Ibid.*, 244)

However, these 'non-distributional targets' have strong *distributional* implications and, therefore, also an effect on output and inflation. Obviously in Pivetti's approach the level of interest rates has distributive effects and hence affects the economy's marginal propensity to consume, fiscal expenditure and competitiveness. Therefore, a rise in interest rates would be accompanied by decreasing activity levels and a reduction in the growth rate of nominal wages. In this case, we would have an *increasing* relationship in the price level/wage ratio together with a *decreasing* inflation rate. The same mechanism underlies the exchange rate determination, which within certain limits is also an 'arbitrary' variable and hence subject to relative pressure from several social groups. As to the interest rate, it is a 'conventional' variable (see Vernengo, 2001).

In general terms and also specifically in these economies there is an inverse relationship between exchange rates and real wages. Hence, there is a clear connection between exchange rate, income distribution and inflation.

This approach in turn requires specification for a small open economy such as Argentina, where monetary policy appears for long periods to be strongly influenced by the situation of the external sector. Thus, the degree of autonomy of monetary policy depends on a greater or lesser degree of external vulnerability, and taking into consideration the degree of financial openness, the role of the interest rate depends on the sustainability of the exchange rate regime.

For example, in strategies based on openness unrestricted by capital flows and *nominal* anchor exchange rates (such as the currency board model in Argentina in the 1990s), the interest rate plays a prominent role in external sector sustainability. In this case, a low level of activity and nominal wage moderation will be essential ingredients to accommodate

the effects of the interest rate on profit margins. Therefore, the search for greater competitiveness is focused on reducing labour costs through nominal wage restraint and increased productivity in export industries. If the resistance of workers is high, unemployment is imposed by a combination of contractionary fiscal and monetary impulses.

In the case of a more expansive economic policy, with a *real* exchange rate that is more stable and has a lower degree of financial openness, the interest rate is less important in the sustainability of the external sector. Where there is external vulnerability, currency devaluations aimed at reducing wage costs in foreign currency have complex effects: on the one hand, they allow the level of employment in exporting industries to be sustained, on the other they raise the prices of wage goods, reducing real wages. If the resistance of workers to real wage loss is high, the rise in nominal wages may result in wage inflation, which is more or less virulent, depending among other factors on the extent of onward transmission to prices. Thus, if nominal wages rise more than prices, causing inflation, and monetary authorities do not correspondingly raise the nominal interest rate, the real interest rate will eventually fall and income distribution will be modified in favour of the workers.

Obviously, exchange rate appreciation reduces inflation (tradable goods) and, by slowing the growth of net exports, can generate unemployment, contributing to moderating the growth in nominal wages. But it is important to point out that no mechanical or *a priori* link can be claimed to exist between these variables (see Figures 13.6 and 13.7).

While in the 1990s both countries chose a similar path, after the 2001–02 crises Argentina and Brazil followed different paths. While Argentina's economy collapsed, Brazil was able to maintain financial stability. The government and main social actors sought to avoid the bankruptcy, devaluation and default debt that occurred in Argentina, which were inevitable consequences rather than actual decisions.

The appreciation process triggered in Brazil has made it more costly (in inflationary terms) to improve the exchange rate (i.e., to check the appreciation tendency) and has encouraged the government to pursue the same policy, strengthening the appreciation of the domestic currency. This feature is underlined by Barbosa *et al*. (2010), who analysed the impact of the real exchange rate on growth in developing countries and suggested that 'the optimal exchange rate for economic growth might not be compatible with the inflation target desired by the population' (*ibid.*, p.11). Of course, the differences between the exchange rate policies give rise to different inflation rates, especially from 2002 onwards (see Figure 13.8).

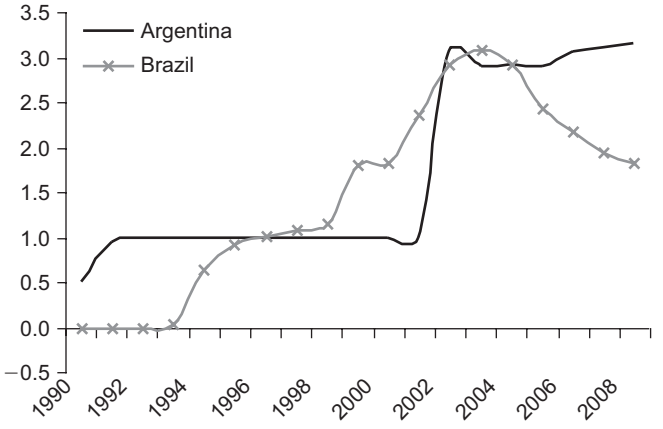


Figure 13.6 Argentina and Brazil: nominal exchange rates, 1990–2008 (domestic currency dollar)
Sources: Ipeadata and ECLAC.

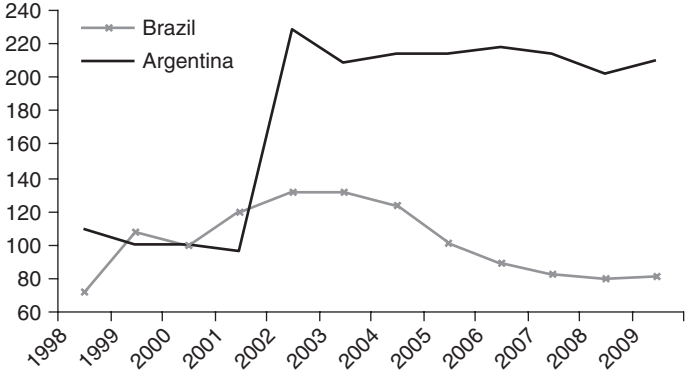


Figure 13.7 Argentina and Brazil: effective real exchange rates, 1998–2009 (index 2000 = 100)
Sources: Ipeadata and ECLAC.

A similar difficulty is pointed out by Serrano (2010), who emphasises the harsh dilemmas that emerged when the Brazilian economy sought to resume a path of development. Serrano suggests exploring a set of policies similar to Argentina post-2002. But he emphasises that the costs of policy reorientation would be lower if it had a strong expansionary effect on public investment, to improve productivity in less competitive industries, particularly in the production of wage goods.

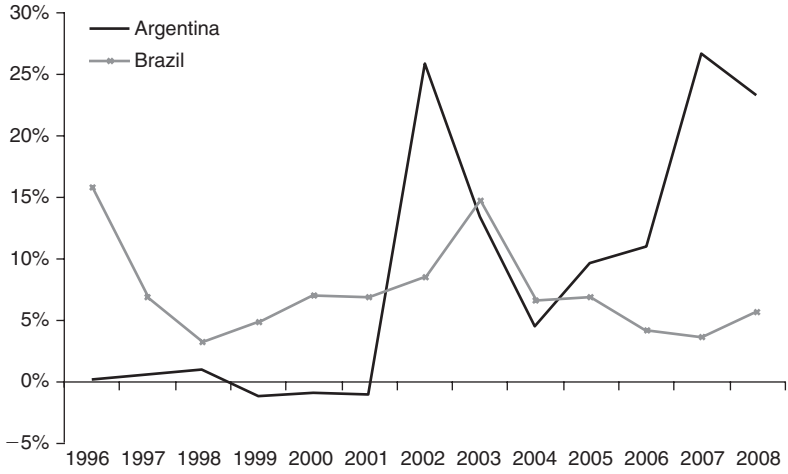


Figure 13.8 Argentina and Brazil: annual inflation rates, 1996–2009 (percentage consumer price)
Source: Prepared on the basis of ECLAC and other estimates.

Table 13.2 Argentina and Brazil: GDP growth, 2000–2009 (real growth rate %)

Year	Argentina	Brazil
2001	–4.4	1.3
2002	–10.9	2.7
2003	8.8	1.1
2004	9.0	5.7
2005	9.2	3.2
2006	8.5	4.0
2007	8.7	6.1
2008	6.8	5.1
2009	0.9	–0.2

Source: ECLAC.

In a context of high commodity prices, it may be not enough to improve the productivity in the wage goods sector and perhaps it is necessary to introduce an export tax and/or other measures to *separate* domestic and international prices of wage goods.

Although it lacks national policies for industry, Argentina would be in a stronger position than Brazil with regard to strategy development for industry. Between 2003 and 2008, Argentina's economic growth was roughly double that of Brazil (see Table 13.2). However, Argentina experienced higher rates of inflation than Brazil in the same period, while

income distribution has improved faster in Argentina than in Brazil since 2002 (see Figures 13.9 and 13.10)

On the other hand, given the fact that, unlike Brazil, Argentina does not have an industrial base, these successful results may turn out to be ephemeral if there is no marked change in the productive structure.

The macroeconomic policy in Brazil is part of a neoliberal strategy of integrating the country into the world market, where financial or rentiers' interests benefit from the current monetary regime, while the

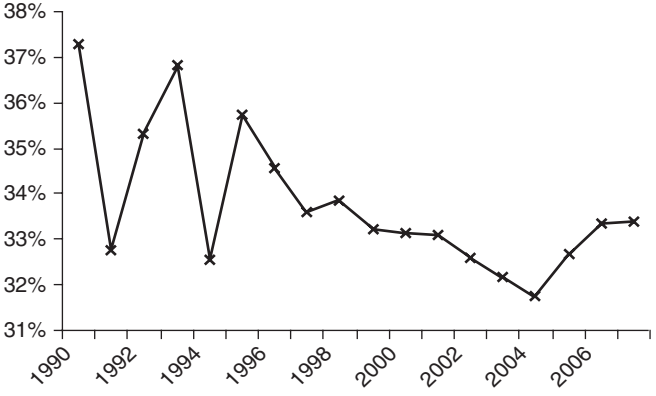


Figure 13.9 Brazil: wage share, 1990–2008, percentage of national income
Source: Ipeadata.

Figure 13.10 Argentina: wage share, 1993–2008, percentage of national income
Source: Mecon (Ministerio de Economía y Finanzas de Argentina).

manufacturing sector and workers bear the costs of this policy. However, as Bastos and Braga (2010) point out, exchange rate appreciation has a negative effect both on competitiveness and on export and import composition.

13.5 Concluding remarks

In the cases under analysis, the relationship between interest rates and price levels postulated by the monetary theory of distribution must be qualified. Thus, the exchange rate channel has actually played a decisive role in cases, such as Brazil, where dear money policies have succeeded in checking inflation. On the other hand, in the case of Argentina the exchange rate policy has been the main battlefield of the distributive conflict between workers, industrial capitalists and agribusiness-financial groups.

These cases seem to suggest that normal distribution is actually governed by monetary determination, where there is no room for a 'natural' or 'neutral money' determination. But the specific manner of governance in this determination is an exchange rate policy rather than an interest rate policy. The most important issue in this context is that no mechanical or *a priori* link can be generally claimed to exist between exchange rate, interest rate and wage rate.

Obviously, the choice of exchange rate policy depends on the relative power of each social group in the social and economic structure. Thus, after Argentina's debt default in 2002, this country chose to have a depreciated domestic currency (a high exchange rate) and to pay the cost of higher inflation in a context of higher wage resistance. Likewise, export tax (which determines a low domestic price for wage goods) entailed a conflict with agribusiness groups.

In the case of Brazil, however, there is low wage resistance and since 2003 the exchange rate policy has been used to check inflation through domestic currency appreciation. Brazil has lower inflation but prioritises its primary exports structure and has a lower growth rate.

The cases of Argentina and Brazil show that unbalanced production structures are the underlying factor influencing and constraining central bank policy, and this finds expression in the tensions over exchange rate policy. At the heart of development problems lies the conflict between several actors over redistributing the associated costs of the exchange rate policy. These are the primary structural factors governing monetary and exchange rate policies and stimulating a specific form of distributive conflict over who is going to pay for the burden of the

diverse exchange rate policy in time. Finally, these different interactions explain the differing results that we can see in the Argentine and Brazilian economies in terms of inflation, distribution and growth in the recent period.

Notes

1. By 'competitive exchange rate' is meant that a level of real exchange rate is higher than typical primary exports, so that exports are allowed to grow from the industrial sector.
2. On sterilisation process, see Lavoie and Wang (2009).
3. See Medeiros and Serrano (2003) and Palumbo (2012).
4. From 2003 to 2011 at constant prices, exports rose by 63 per cent, domestic consumption by 81 per cent and public expenditures by 72 per cent. In addition, exports' average share of GDP is only 12.4 per cent.
5. Of course, because of the endogeneity of potential output, sometimes inflation targeting cannot avoid a self-fulfilling monetary policy that locks the economy into a path of slow growth.
6. See for instance Rodríguez (2001).
7. Although in the 1960s, when industrialisation was accelerating, Brazil does not seem to fit this description, it has been the case since the 1980s.
8. The ambiguity of the effects of exchange rate appreciation on the industrial sector is only a short-term phenomenon. In the long term, *persistent* appreciation of domestic currency has an adverse impact on industrial performance.
9. 'Export-oriented industrialisation' does not mean 'export-led' growth, but a diversification in export structure.
10. The adjustment path is not necessarily a unique and simple process, either in regard to the level of the (nominal) mark-up or to the time lags involved.
11. Or in the case of the economy that issues the international currency accepted world-wide.

References

Arida, P. and Resende, L. (1985) 'Recessão e taxa de juros: o Brasil nos primórdios da década de 1980', *Revista de Economia Política*, 5(1), January–March.

Barbosa N. (2010) 'Latin America: counter-cyclical policy in Brazil: 2008–09', *Journal of Globalization and Development*, 1(1): 1–14.

Barbosa, N., Silva, J., Goto, F. and Silva, B. (2010) 'Real Exchange Rate, Capital Accumulation and Growth in Brazil', Paper presented at the Fourth Annual Conference on Development and Change, Johannesburg, South Africa, 9–11 April.

Barbosa-Filho, N. (2008) 'Inflation-targeting in Brazil: 1999–2006', *International Review of Applied Economics,* 22(2): 187–200.

Bastos, C. and Braga, J. (2010) 'Conflito Distributivo e Inflação no Brasil: Uma Aplicação ao Período Recente', *XV Encontro Nacional de Economia Política*, SEP, June 2010, São Luis, Brazil.

Bastos, C. and Salles, E. (2010) 'Adeus capitalismo dependente. Olá neo-primário exportador?', Universidad Federal Fluminense, Rio de Janeiro, Brazil, *mimeo*.

Cavallo, D. (1977) 'Los efectos recesivos e inflacionarios iniciales de las políticas monetaristas de estabilización', Banco Central de la República Argentina, *Ensayos Económicos*, no.4, December 1977, Part II: 107–48.

Chang, R. (2007) 'Inflation-Targeting, Reserves Accumulation, and Exchange Rate Management in Latin America', Rutgers University/National Bureau of Economic Research, October.

Diamand, M. (1986) 'Overcoming Argentina´s Stop and Go Economic Cycles', in J. Hartlyn and S. Morley (eds), *Latin American Political Economy: Financial Crisis and Political Change*. (Boulder, CO: Westview Press).

Dooley, M., Folkerts-Landau, D. and Garber, P. (2003) 'An Essay on the Revived Bretton Woods System', *Working Paper* no. w9971, NBER, Cambridge, MA, September.

Ferreira, S. and Serrano, F. (2010) 'Commodities, câmbio e inflação de custos no Brasil 1994–2009', *Versus Académica*, 4, April: 48–57.

Frenkel, R. (1979) 'Decisiones de precio en alta inflación', *Desarrollo Económico*, 19(75): 291–330.

Frenkel, R. (2004) '"Right" Prices for Interest and Exchange Rates', in J. Joost Teunissen and A. Ackerman (eds), *Diversity in Development Reconsidering the Washington Consensus* (The Hague: FONDAD, December), www.fondad.org.

Frenkel, R. (2008) 'The competitive real exchange-rate regime, inflation and monetary policy', *Cepal Review*, 96, December: 191–201.

Frenkel, R. and Rapetti, M. (2010) 'A Concise History of Exchange Rate Regimes in Latin America', Center for Economic and Policy Research, April.

Galindo, L. and Ros, J. (2008) 'Alternatives to inflation-targeting in Mexico', *International Review of Applied Economics*, 22(2): 201–14.

Lavoie, M. and Wang, P. (2009) 'The Compensation Thesis, as Exemplified by the Case of the Chinese Central Bank', ROBINSON, *Working Paper* no. 09-02, March.

Mahon, J. (1992) 'Was Latin America too rich to prosper? Structural and political obstacles to export-led industrial growth', *Journal of Development Studies*, 28(2): 241–63.

Medeiros, C. (2010) 'Auge e Declínio dos Estados Desenvolvimentistas. Novos Desafios', presented at the international seminar 'Estado Desenvolvimentista: Crise e Retomada?', UFRJ, Rio de Janeiro, June 2010.

Medeiros, C. and Serrano, F. (2003) 'Inserção Externa, Exportações e Crescimento no Brasil', in J.C. Ferraz, M. Crocco and L.A. Elias (eds), *Liberalização Econômica e Desenvolvimento* (São Paulo: Futura): pp. 324–49.

Medeiros, C. and Serrano, F. (2006) 'Capital Flows to Emerging Markets under the Flexible Dollar Standard: A Critical View Based on Brazilian Experience', in M. Vernengo (ed.), *Monetary Integration and Dollarization* (Northampton: Edward Elgar): pp. 218–42.

Palumbo, A. (2012) 'On the Balance-of-Payments-Constrained Theory of Growth', in R. Ciccone, C. Gehrke and G. Mongiovi (eds), *Sraffa and Modern Economics* (London: Routledge).

Pivetti, M. (1991) *An Essay on Money and Distribution* (London: Macmillan).

Pivetti, M. (1999) 'On Sraffa's cost and surplus concept of wages and its policy implications', *Rivista Italiana degli Economisti*, IV(2): 279–300.

Pivetti, M. (2001) 'Money Endogeneity and Monetary Non-Neutrality: A Sraffian Perspective', in L.-P. Rochon and M. Vernengo (eds), *Credit, Interest Rates and the Open Economy. Essays in Horizontalism* (Cheltenham: Edward Elgar).
Pivetti, M. (2007) 'Distribution, inflation and policy analysis', *Review of Political Economy*, 19(2): 243–7.
Pivetti, M. (2010) 'Interest and Inflation: Some Critical Notes on "The New Consensus Monetary Policy Model"', in A. Birolo, D. Foley, H.D. Kurz, B. Schefold and I. Steedman (eds), *Production, Distribution and Trade: Alternative Perspectives* (London: Routledge).
Rodríguez, O. (2001) 'Prebisch: actualidad de sus ideas básicas', *Revista de Cepal*, no. 75, December.
Serrano, F. (1993) 'Review of Pivetti's essay on money and distribution', *Contributions to Political Economy*: 117–24.
Serrano, F. (2010) 'Juros, câmbio e o sistema de metas de inflação no Brasil', *Revista de Economia Política*, 30(1), (117): 63–72, January–March.
Sraffa, P. (1960) *Producción de Mercancías por Medio de Mercancías* (Barcelona: Oikos Tau).
Stirati, A. (2001) 'Inflation, unemployment and hysteresis: an alternative view', *Review of Political Economy*, 13(4), October: 427–51.
Summa, R. (2010) *Um modelo alternativo ao 'novo consenso' para economia aberta*, Unpublished Ph.D. thesis, Instituto de Economía da Universidade Federal do Rio de Janeiro, March.
Vernengo, M. (2001) 'Foreign Exchange, Interest and Prices: The Conventional Exchange Rate', in L-P. Rochon and M. Vernengo (eds), *Credit, Interest Rates and the Open Economy: Essays on Horizontalism* (Cheltenham: Edward Elgar).
Vernengo, M. (2008) 'Back to the Future: Latin America's Current Development Strategy' IDEAs Network Working Paper no. 07/2008.