Optimum Currency Areas

Ronald I. McKinnon

The American Economic Review, Vol. 53, No.4, September 1963

In a recent note [4], Robert A. Mundell has suggested that little in the way of a systematic attempt has been made to define the characteristics of an area over which it is optimal to have a single currency regime, or—what is almost the same thing—a fixed exchange-rate system with guaranteed convertibility of currencies. The extensive literature on the relative merits of fixed versus flexible exchange rates has been rendered somewhat sterile by this omission. Existing national boundaries have been implicitly used to define the single currency area to which flexible external exchange rates would or would not be applied. However, when different possibilities for the grouping of nations in single currency areas exist, as in the EEC, or when resource mobility is low within individual countries, Mundell demonstrates that it is necessary to ask what economic characteristics determine the optimum size of the domain of a single currency. I shall develop the idea of optimality further by discussing the influence of the openness of the economy, i.e., the ratio of tradable to non-tradable goods, on the problem of reconciling external and internal balance, emphasizing the need for internal price-level stability.

"Optimum" is used here to describe a single currency area within which monetary-fiscal policy and flexible external exchange rates can be used to give the best resolution of three (sometimes conflicting) objectives: (1) the maintenance of full employment; (2) the maintenance of balanced international payments; (3) the maintenance of a stable internal average price level. Objective (3) assumes that any capitalist economy requires a stable-valued liquid currency to insure efficient resource allocation. Possible conflicts between (1) and (2) have been well discussed in the literature, especially by J. E. Meade [3], but joint consideration of all three is not usually done. For example, J. L. Stein [6] explicitly assumes internal price-level stability in his discussion of optimal flexibility in the foreign exchange rate. The inclusion of objective (3) makes the problem as much a part of monetary theory as of international trade theory. The idea of optimality, then, is complex and difficult to quantify precisely, so what follows does not presume to be a logically complete model.

"The ratio of tradable to non-tradable goods" is a simplifying concept which assumes all goods can be classified into those that could enter into foreign trade and those that do not because transportation is not feasible for them. A physical description of both tradable and non-tradable goods would correspond to that given by R. F. Harrod [1, pp. 53-56]. This overly sharp distinction between classes of tradable and non-tradable goods is an analytically simple way of taking transportation costs into account. By tradable goods we mean: (1) exportables, which are those goods produced domestically and,

in part, exported; (2) importables, which are both produced domestically and imported. The excess of exportables produced over exports will depend directly on the amount of domestic consumption, which is likely to be small when exportable production is heavily specialized in few goods. Similarly, the excess of importables consumed over imports will depend on the specialized nature of imports. Therefore, the value of exportables produced need not be the same as the value of importables consumed, even in the case of balanced trade where the values of imports and exports are equal. However, the total value of tradable goods produced will equal the value of tradable goods consumed under balanced trade. Thus, the expression "the ratio of tradable to non-tradable goods" can apply unambiguously to production or consumption.

I. *A Simple Model*

Ideally, one would like to consider a large group of countries jointly and then decide how they should be divided up into optimum currency regions. The analytical framework for such a task does not exist, so it is necessary to consider a much narrower problem and hope it throws light on the general one—besides being of interest in itself. Consider a well-defined single currency area in which we wish to determine whether or not there should be flexible exchange rates with the outside world. The outside world is itself assumed to be a single currency area which is very large.

If the area under consideration is sufficiently small, we may assume that the money prices of the tradable goods in terms of the outside currency are not influenced by domestic exchange rates or domestic currency prices.[1] In actual practice, the domestic money prices of tradable goods will be more closely tied to foreign prices through existing exchange rates than will the domestic money prices of the non-tradable goods. Under this invariance assumption, i.e., fixed foreign-currency prices, the terms of trade will necessarily be immune to domestic economic policy. Some justification is given for this in R. Hinshaw [2], even for fairly large countries. We now inquire into whether external exchange-rate flexibility or internal fiscal-monetary expansion or contraction is most suitable to maintaining external balance, i.e. shifting production and expenditures between the tradable and non-tradable goods.

[1] If we apply this assumption to the standard elasticities model, then both the elasticity of foreign demand for home exports η_f and the elasticity of foreign supply of home imports ϵ_f, are assumed infinite. Thus a devaluation, i.e., a rise in the foreign exchange rate k, would always improve the trade balance, B, by an amount proportional to the sum of the home elasticity of demand for imports and the home elasticity of supply of exports, η_h and ϵ_h respectively, i.e.,

$$\frac{dB}{dk} = Z(\epsilon_h + \eta_h)$$

where Z is the value of exports in the case of balanced trade. The trouble with this standard model is that η_h and ϵ_h depend on the amount of domestic absorption permitted in the course of devaluation as well as the openness of the economy; and it is difficult to make explicit what internal price repercussions may occur since the body of non-tradable goods does not enter explicitly in the model. Assuming both η_f and ϵ_f to be infinite is different from the usual simplification that both supply elasticities, ϵ_f and ϵ_h, are infinite, and in my opinion is more appropriate to the consideration of most small areas.

Case 1

Suppose exportables X_1 and importables X_2 together make up a large percentage of the goods consumed domestically. Suppose further a flexible exchange-rate system is used to maintain external balance. The price of the non-tradable good, X_3, is kept constant in terms of the domestic currency. Exchange-rate changes will vary the domestic prices of X_1 and X_2 directly by the amount of the change. Thus, if the domestic currency is devalued 10 per cent, the domestic money prices of X_1 and X_2 will rise by 10 per cent and thus rise 10 per cent relative to X_3. The rationale of such a policy is that the production of X_1 and X_2 should increase, and the consumption of X_1 and X_2 should decline, improving the balance of payments. Direct absorption reduction from the price rise in tradable goods may have to be supplemented by deliberate contractionary monetary-fiscal policy, if unemployment is small. Substantial theoretical justification for considering relative price changes between tradable and non-tradable goods to be more important than changes in the terms of trade for external balance is given by I. F. Pearce [5].

From Case 1, it is clear that external exchange-rate fluctuations, responding to shifts in the demand for imports or exports, are not compatible with internal price-level stability for a highly open economy, objective (3). In addition, such a policy by itself may not succeed in changing relative prices or affecting the trade balance. In a highly open economy operating close to full employment, significant improvement in the trade balance will have to be accomplished via the reduction of domestic absorption, i.e., real expenditures, which is the only possible way of keeping the price of X_3 constant in terms of the domestic currency. Thus, a substantial rise in domestic taxes may be necessary whether or not there is any exchange rate change. In the extreme case where the economy is completely open, i.e., all goods produced and consumed are tradable with prices determined in the outside world, the only way the trade balance can be improved is by lowering domestic expenditures while maintaining output levels. Changes in the exchange rate will necessarily be completely offset by internal price-level repercussions with no improvement in the trade balance.

To restate the core of the argument: if we move across the spectrum from closed to open economies, flexible exchange rates become both less effective as a control device for external balance and more damaging to internal price-level stability. In fact, if one were worried about unwanted speculative movements in a floating exchange rate in Case 1 of an open economy, a policy of completely fixed exchange rates (or common currency ties with the outside world) would be optimal. Blunt monetary and fiscal weapons which evenly reduced expenditures in all sectors could be counted on to improve immediately the trade balance by releasing goods from domestic consumption in the large tradable-goods sector. Exportables previously consumed domestically would be released for export; imports would be directly curtailed, and domestically produced importables made available for substitution with imports. The reduction of expenditures in the relatively small non-tradable-goods sector would initially only cause unemployment which, depending on the degree of interindustry resource mobility and price flexibility, might even-

tually be translated into more production in the tradable-goods sectors, and possibly improve the trade balance in the longer run. The smaller this non-tradable-goods sector, the smaller will be the immediate impact of reducing expenditures on employment and total production, and thus the more efficient this policy of expenditure reduction will be as a device for improving external balance (the surplus of production over expenditures).

Any region within a common currency area faced with a loss of demand for its products will be forced to cut its expenditures through a loss of bank reserves and regional income, thus eventually correcting the trade balance. A separate currency region with fixed exchange rates may have to carry out the cutback of expenditures more through deliberate policy if bank reserve losses are effectively sterilized. In either case, the immediate reduction in real income cannot be avoided if the trade balance is to be improved.

Case 2

Suppose the production of non-tradable goods is very large compared to importables and exportables in the given area. Here the optimal currency arrangements may be to peg the domestic currency to the body of non-tradable goods, i.e., to fix the domestic currency price of X_3 and change the domestic price of the tradable goods by altering the exchange rate to improve the trade balance. A currency devaluation of 10 per cent would cause the domestic prices of X_1 and X_2 to rise by 10 per cent, but the effect on the general domestic price index is much less than in Case 1.

The desired effect of the relative price increase in the tradable goods is to stimulate the production of tradable compared to non-tradable goods and thus improve the trade balance. On the other hand, if monetary-fiscal policy is primarily relied on to reduce domestic demand to maintain external balance, unemployment will be much higher. Much of the immediate impact of the reduction of expenditures will be in the extensive non-tradable-goods industries. If there are any rigidities in resource mobility, the trade balance will not improve much in the first instance. Through this policy, it may be actually necessary to achieve a fall in the domestic money prices of X_3, the numerous non-tradable goods, before sufficient expansion in the production of X_1 and X_2 can be obtained. Since a major component of X_3 will be labor services, it may be necessary to lower wage costs vis-a-vis the domestic money prices of X_1 and X_2, which are fixed by the inflexible external exchange-rate system. Such a policy would contain all the well-known Keynesian difficulties of getting labor to accept a cut in money wages. In addition, a successful policy of lowering prices of the numerous X_3 goods would have a large impact on the average domestic price level. Effectively, we would have permitted the tail (tradable goods) to wag the dog (non-tradable goods) in pursuing restrictive monetary and fiscal policies, with fixed exchange rates to improve the trade balance, for a small proportion of tradable goods.

Our open economy of Case 1 somewhat resembles what Stein [6] has called a "conflict" economy. In a conflict economy, export production is sufficiently large to dominate the generation of domestic income, and thus fluctuations in both are positively correlated. Therefore, with a fixed exchange

rate, periods having low income will also have unfavorable trade balances, and vice versa. For income stabilization, objective (1), Stein concludes that a floating exchange rate will be optimal for a conflict economy in a Keynesian environment. The foreign exchange rate would then rise at the top of the cycle and fall at the bottom. These exchange rate changes will stimulate domestic production and income at the bottom of the cycle and damp them at the top. But it is precisely in this case of a highly open economy that exchange-rate changes will mean great fluctuations in internal price levels—sufficiently great, that any effects of exchange rate changes on domestic production may be small. However, there may still remain a direct policy conflict between objectives (1) and (3) in the use of a floating exchange rate. Certainly, the liquidity value of the domestic currency will depend directly on the short-run fortunes of the export commodity(ies) for a floating exchange rate.

Qualifications

The sharp distinction between tradable and non-tradable goods makes the above model analytically much easier to work with; but in practice there is a continuum of goods between the tradable and non-tradable extremes. The relaxation of this sharp distinction does not invalidate the basic idea of the openness of the economy affecting optimum economic policies; but the empirical measurement of the ratio of tradable to non-tradable goods becomes more difficult. Some kind of weighting system for determining the total production in each category might be possible. Certainly, knowledge of total imports and exports would give one a good lead in determining total production of exportables and importables. In addition, the idea of openness would have to be modified when the area was large enough to affect external prices.

II. *Monetary Implications of the Model*

The above discussion has been concerned with the way by which relative price changes in tradable and non-tradable goods can be brought about, and the conditions under which monetary and fiscal policy can be used efficiently to maintain external balance. Minimizing the real cost of adjustments needed to preserve external balance hinged to a large extent on minimizing necessary fluctuations in the over-all domestic price level. Thus the argument is very much concerned with the liquidity properties of money, and it is worth while to look at some of the more general monetary implications of the model. Suppose X_1, X_2, and X_3 are classes of goods rather than single goods as in the Pearce model. One of the aims of monetary policy is to set up a stable kind of money whose value in terms of a representative bundle of economic goods remains more stable than any single physical good. Indeed, it is the maintenance of this stable value which gives money its liquidity properties. The process of saving and capital accumulation in a capitalist system is greatly hampered unless a suitable *numéraire* and store of value exists. It may be still more difficult if a more desirable money is available from another source, e.g., from a larger currency area. This latter possibility is discussed below.

If the area under consideration is sufficiently large so that the body of non-tradable goods is large, then pegging the value of the domestic currency to this body of non-tradable goods is sufficient to give money liquidity value in the eyes of the inhabitants of the area in question. It may not be sufficient from the viewpoint of potential investors in the outside world. However, if the area is large, what outside investors think need not be an overriding consideration. Efficient internal capital accumulation and full employment are more important than external capital movements. If, under these circumstances, trade patterns are so unstable that substantial relative price changes in tradable and non-tradable goods are required to maintain external balance and full employment, then flexible external exchange rates may well be optimal. Resulting internal price changes will not destroy the value of the domestic currency as money.

If the area under consideration is small so that the ratio of tradable to non-tradable goods is large and the prices of the former are fairly well fixed in the outside currency, then the monetary implications of pegging the domestic currency to the non-tradable goods are less satisfactory. Such a class of non-tradable goods may not constitute a typical bundle of economic goods. The class of importables may be more representative, and a currency pegged to maintain its value in terms of importables into a small area may have a higher liquidity value than one pegged to the domestically produced non-tradable goods. However, pegging a currency of a small area to maintain its value in terms of a representative bundle of imports from a large outside area is virtually the same thing as pegging it to the outside currency. Alternatively, if we have a number of small areas which trade extensively with each other, and each pegs its currency to a representative bundle of imports, then each currency will be pegged to the others. To maintain the liquidity value of individual currencies for small areas, a fixed exchange-rate system is necessary. In addition, capital movements among small areas are more needed to promote efficient economic specialization and growth than free capital movements among large, economically developed areas. Contractual arrangements for such movements are greatly facilitated by a common currency. These arguments give us some insight into why each of the fifty states in the United States could not efficiently issue its own currency, aside from the inconvenience of money changing.

If we have a small area whose currency is not convincingly pegged in terms of the currency of a larger area, and so on this account its liquidity value is less, then domestic nationals will attempt to accumulate foreign bank balances. This will occur even though the marginal efficiency of investment in the small area is greater than that outside. As long as the functions of savings and investment are specialized, savers will attempt to accumulate cash balances in the more liquid currency. The illiquidity of domestic currency may also reflect monetary mismanagement as well as small size. In either case, we would expect small countries with weak currencies to have a tendency to finance the balance-of-payments deficits of larger countries with more desirable currencies. Thus, we have capital outflows from countries where the need for capital may be rather high and which arise from "mone-

tary" rather than "real" considerations. Authorities in such countries are generally forced to maintain rather strict exchange controls unless the currency can be pegged in a convincing fashion to that of the larger area.

The above argument is relevant to the use of uncontrolled floating exchange rates. This device of maintaining external balance will only work well when the currency in question has liquidity value of the same order as that of the outside world—or the world's major currencies. This condition was approximately satisfied in the case of the Canadian dollar up to 1961. However, a floating exchange rate for the Korean yen may lead to less satisfactory results. If the official rate were made equal to the black-market rate and there were no further exchange restrictions, there would still be a capital flight out of Korea into currencies with superior liquidity value, aside from problems of political stability. A floating exchange rate in itself is not a sufficient control device and does not necessarily eliminate the need for exchange controls.

By contrast, short-term capital flows among currencies of approximately equal liquidity value are less likely with a floating exchange rate because of the exchange risk and the liquidity equivalence. The possibility of carrying out different degrees of easy or tight monetary policy in different countries is greater as capital flows would not be so responsive to interest-rate differentials. Once the world is divided into a number of optimal-sized currency areas permitting efficient internal capital accumulation, the desirability of short-term capital flows among areas well developed economically becomes less great, and it becomes desirable to insulate the monetary policies of the areas from each other in order that monetary policy may be used more freely to support full employment. However, it does not make any sense to advocate a floating exchange-rate system without first defining the optimal domains of individual currencies.

Suppose we look at the problem of a depressed subregion of a common currency area. Consider the case of West Virginia where non-tradable goods are largely labor services. We have an illustration of an excess supply of non-tradable goods and an excess demand for the tradable goods because of internal price rigidities. Thus, in this sense West Virginia has an *ex ante* balance-of-payments deficit even though in an *ex post* accounting sense there is a balance-of-cash flow in and out of the state. Would the adjustment of external balance and internal full employment be facilitated if West Virginia were incorporated as a country with its own currency? To the extent that the ratio of tradable to non-tradable goods was high, such a monetary system would have little chance of success. A devaluation would be associated with a large domestic price-level increase and hence money illusion would not be much help in getting labor to accept a cut in real wages [4, p. 663]. Labor unions would still continue to bargain in terms of U.S. dollars. In addition, a West Virginian currency tied to a representative bundle of non-tradable goods would not be an entirely acceptable store of value. There undoubtedly would be attempts by West Virginians to accumulate U.S. bank balances. However, if the depressed area were substantially larger, with a small proportion of production in tradable goods, a separate monetary system might be

preferable as a device for maintaining full employment and external balance in the absence of factor mobility.

III. *A Concluding Note on Factor Mobility*

The idea of factor mobility has two distinct senses: (1) geographic factor mobility among regions; (2) factor mobility among industries. I think it is fair to say that Mundell [4] had interpretation (1) primarily in mind. His discussion of optimum currency areas in large measure is aimed toward having high geographic factor mobility within each single currency area and using flexible external exchange rates to make up for the lack of factor mobility among areas. Thus, for a given amount of geographic factor mobility in the world, this method of division into currency areas would maximize the possibility of world income and employment, subject to the constraint of maintaining external balance. Of course, the currency arrangements themselves would affect factor mobility, so the extent of factor mobility has to be considered *ex post*. Once we consider problems of factor immobility among industries, it may not be feasible to consider slicing the world into currency areas along industrial groupings rather than geographical groupings. However, from our above discussion, an optimal geographic size still exists even when we are only concerned with interindustry factor immobility.

Consider the special but perhaps common case of factor immobility between regions, each with its own specialized industries, the case where it is difficult to distinguish geographical and interindustrial immobility. Suppose there is a rise in the demand for the products of region A and a decline in the demand for goods of region B. The value of the marginal products of the potentially mobile factors of production in region B in B-type industries will fall, and rise in region A in A-type industries. Now if the possibility of developing or extending A-type industries in B is feasible, then need for factor movement between A and B is not great. The existing immobility between regions can be accepted through monetary arrangements giving both regions their own currencies, thus permitting more flexibility in enabling each area to pursue monetary and fiscal policies geared to internal stability. But if B cannot easily develop A-type industries, then factor movements to A may be the only thing that will prevent a large fall in the unit incomes of potentially mobile factors of production in B. So a policy aimed directly at overcoming the immobility of factor movements between A and B may be optimal, and perhaps the two should be joined in a common currency area. This argument becomes stronger when one considers small areas trying to develop industries in which economies of scale or indivisibilities are very great instead of efficiently moving factors elsewhere.

In a world where trade patterns are not perfectly stable, there will always be the problem of changing the world pattern of resource use among various industries to preserve external balance, full employment, and efficient resource use. In the simple model given in Section I above, we considered the optimum extent of a currency area in terms of its size and structure, i.e., the ratio of tradable to non-tradable goods, in promoting shifts in resources among various industries. The model accepted the degree of internal resource

immobility among industries as an obstacle to be overcome as smoothly as possible. The arguments given there for applying flexible exchange rates to optimal-sized currency areas to efficiently overcome factor immobility hold in the main, whether the degree of internal mobility among industries is large or small. Such factor immobility among industries is a painful fact of economic life which has to be overcome as efficiently as possible. However, this criterion of size and openness of a single-currency economy in facilitating interindustry production shifts certainly has to be balanced with purely geographic factor-mobility considerations in determining the optimum extent of a currency area.

RONALD I. MCKINNON*

REFERENCES

1. R. F. HARROD, *International Economics,* 5th edition. Cambridge 1957.
2. RANDALL HINSHAW, "Currency Appreciation as an Anti-Inflationary Device," *Quart. Jour. Econ.,* Nov. 1951, *65,* 447-62.
3. J. E. MEADE, *The Theory of International Economic Policy.* Vol. I, *The Balance of Payments.* London 1951.
4. ROBERT A. MUNDELL, "A Theory of Optimum Currency Areas," *Am. Econ. Rev.,* Sept. 1961, *51,* 657-64.
5. IVOR F. PEARCE, "The Problem of the Balance of Payments," *Internat. Econ. Rev.,* Jan. 1961.
6. JEROME L. STEIN, "The Optimum Foreign Exchange Market," *Am. Econ. Rev.,* June 1963, *53,* 384-402.

* The author is assistant professor of economics at Stanford University.

Inventing and Maximizing

Cross-section studies over the period roughly 1900–50 show that patented inventions that improve capital goods tend to be distributed among U.S. industries in proportion to the industries' value-added [5]. The discoverers of this relation held that, in addition to other factors, this result was consistent with maximizing behavior on the part of inventors, since sales of capital goods tend to vary directly with the value added by the industries using those capital goods, and since the prospective net revenues from capital-goods-improving inventions would tend to vary directly with sales of the class of capital goods involved. However, it was pointed out in [2] that (a) other possible size-correlated determinants might explain the correlation, and (b) the influence, if any, of maximizing behavior on the industrial distribution of inventive activity would be more clearly revealed if the effects of industry size were suppressed. While objection (a) cannot now be met, this note attempts to follow up suggestion (b).

I

There are several ways of eliminating the effect of industry size. One way is to combine all the cross sections reported on in [5] and re-estimate the relationship, allowing each industry to have a constant term of its own.