

IMF WORKING PAPER

© 1991 International Monetary Fund

This is a working paper and the author would welcome any comments on the present text. Citations should refer to an unpublished manuscript, mentioning the author and the date of issuance by the International Monetary Fund. The views expressed are those of the author and do not necessarily represent those of the Fund

WP/91/37

INTERNATIONAL MONETARY FUND

Research Department

Do Exchange Rates Work? Another View

Prepared by Robert A. Mundell*

Authorized for Distribution by Jacob Frenkel

April 1991

Abstract

This paper investigates the role of exchange rates in balance of payments theories. It explores the sixteen approaches to the balance of payments, the concept of an "equilibrium" trade balance and sequential "stages" of the current account. It examines fiscal and demographic influences on the U.S. deficit. The final section considers the breakdown of the international monetary system after World Wars I and II; an evaluation of alternative proposals to correct the defects of the system; and an examination of the extent to which deficits of reserve countries have their origins in systemic problems.

JEL Classification Number

F30

* Professor of Economics at Columbia University and Visiting Scholar in the IMF Research Department. This paper was prepared for presentation at the Conference on International Adjustment at the Institute for International Economics, Washington, D.C., October 27-28, 1990.

<u>Contents</u>	<u>Page</u>
Summary	iii
I. Introduction	1
II. Classical Theory	1
1. The Gold Standard	2
2. The Transfer Debate	3
3. Transfers, Gold Flow and the Exchange Rate	7
4. Bickerdike's Breakthrough	10
III. Alternative Deficit Theories	13
1. Sixteen Approaches to the Balance of Trade	13
2. Is There an Equilibrium Trade Balance?	18
3. The "Stages" Theory	20
4. How Demography Creates Trade Deficits	22
5. The Fiscal Approach to the Deficit	23
6. Exchange Rates and the Trade Balance	24
IV. Trade Balances as System Imbalances	27
1. The Gold Exchange Standard of the 1920's	29
2. The Bretton Woods Problem	31
3. The Collapse of the Bretton Woods System	35
4. Flexible Exchange Rates	38
5. Currency Areas	41
6. Instability of the Real Exchange Rate	44
V. Conclusions	47
Chart	48a
References	49

Summary

This paper discusses the role of exchange rates in economic theory and policy. The theoretical problem in integrating exchange rates in economic theory arises from the fact that the exchange rate is the ratio of two nominal quantities whereas most targets of economic policy are real variables. The school-specific nature of macroeconomics therefore invades theorizing about exchange rates.

After a historical review of the theoretical literature, the paper describes alternative approaches--sixteen in all--to the balance of payments, made up of four static and four inter-temporal approaches for each country and its mirror-image counterpart in the rest of the world. Thus, the elasticity (or direct) approach, the absorption approach, the external financing approach, and the net internal savings approaches have both international and inter-temporal counterparts. The paper also discusses the meaning of equilibrium trade balance, the stages theory of balance of payments development, the demographic theory of net savings and the trade balance, the fiscal approach to the deficit, and exchange rate theories of the deficit. It argues that depreciation of a currency in a country with rigid money wages may worsen a country's trade balance if it raises the marginal efficiency of capital and attracts capital imports.

The paper discusses the under-valuations of gold after World Wars I and II and the consequences for the international monetary system of the different ways adopted for coping with it. It considers and evaluates the consequences of alternative reforms of the Bretton Woods system--an increase in the price of gold-retaining reserve currencies, the gold standard with elimination of the reserve-currency system, a world central bank, and new instruments of international reserves, such as the SDR and flexible exchange rates. The final sections deal with systemic effects on current account balances of the asymmetric use of national currencies as international assets, the defects of the system of flexible exchange rates from 1973 to 1985, and the rise of large currency areas, the instability of the real exchange rate, and the movement to a dollar-management system of coordinated exchange rates after 1985.

I. Introduction

The appropriate role of exchange rates in economic policy is one of the most complicated issues in international economics and public policy. The difficulty stems from the fact that the exchange rate is a ratio of two monetary quantities and most targets of economic policy are real quantities. Solution of the problem requires specification of how nominal variables affect real variables. The same difficulty arises in relating nominal to real variables in closed-economy macroeconomic theory.

Much of the controversy in the literature has turned on differences relating to (1) the mechanism for changing exchange rates in a flexible exchange rate system; (2) the effect of devaluation on absolute and relative prices; (3) the relation between the exchange rate and the trade balance; (4) the relation between changes in the trade balance and relative prices; (5) the role of exchange rates in general equilibrium theory; (6) the link between the international monetary system and exchange rate changes; and (7) the relation of national currencies to the international standard (if any).

Models of international economic behavior are school-specific. This seems more true today than in the past because of the proliferation of models in closed-economy macroeconomic theory. Edmund Phelps, for example, identifies "Seven Schools of Macroeconomic Thought" in a recent book of that title. These include (1) the macroeconomics of Keynes; (2) monetarism; (3) the new classical school; (4) the new Keynesian school; (5) supply-side economics; (6) real business cycle theory; and (7) the structuralists. The schools are distinguished by different assumptions about information, expectations, price flexibility or disequilibrium-equilibrium methodology. It is possible to derive a distinct exchange rate theory corresponding to each of these schools.

This paper sketches some important developments in exchange-rate doctrine and analyze some necessary links between theory and the international monetary system. Part I considers the theory of exchange rates and trade balances in the context of classical theory. Part II outlines the sixteen approaches to the balance of trade and considers alternative theories of the deficit. Part III analyzes "systemic deficits"--those deficits that arise out of the working of the international monetary system itself. Concluding remarks about the current account imbalances of today complete the paper.

II. Classical Theory

Alfred Marshall could write--in 1912--that the most important thing that can be said about currency is that it is totally unimportant. In this way he summarized the attitude of a century of the classical school acclimatized to the monetary stability of the gold standard. Currency theory, incorporating the effects of exchange rate changes on real variables, had no place in the classical framework except as an offshoot of the quantity theory of money. Unlike the preceding age of mercantilism,

and the succeeding age of Keynesianism, exchange rate changes were relegated to the realm of economic pathology.

1. The Gold Standard

The gold standard, like earlier bi-metallic and silver standards, encouraged a fundamental conservatism in monetary and fiscal policy. It imposed financial discipline. The convertibility requirement pinned down the money supply to the level compatible with balance of payments equilibrium. Long-standing adherence to the gold parity ensured stabilizing inelastic expectations. Future monetary policy was predictable from knowledge of demand and supply conditions for the monetary metal.

A government would tamper with the monetary standard only at its peril. Changing the standard would have a shock effect with reverberations on expectations for generations to come. Devaluation of a paper currency in terms of gold would result in a proportionate increase in the price level and a proportionate scaling down of the real value of debts; once tried, interest rates would adjust upward to allow for the cost of a possible repetition. On the other hand, belief in the standard--and in its restoration if a major, but infrequent, disturbance (such as a major war) had undermined it, was itself a resource that kept government credit cheap; stabilizing inelastic expectations, resting on the assumption of early restoration of a broken standard, prevented the collapse of the bond market. 1/

1/ Examples of this stabilizing element in expectations is afforded by the currency history of Britain and other countries. Instead of devaluation in the 1690s to finance the great recoinage, the government harked back to Elizabeth I's recoinage over 130 years earlier that paid for the coinage out of taxes rather than depreciation. Britain was able to compound its public debt to finance the several wars of the 18th century without devaluation; and when the Bank of England succumbed to inconvertibility in 1797, interest rates for two decades reflected the strong probability that Britain would restore the sacred parity after the war (which happened in 1819). Again, the uninterrupted history of the gold standard in Britain in the nineteenth century, despite periodic crises, contributed to the belief, when payments were suspended during World War I, interest rates again reflected the expectation that, after the war, convertibility would be resumed at the old parity. This occurred in 1925, but it proved abortive when, for systemic reasons, the undervaluation of gold that had been a consequence of world war inflation led to a liquidity crisis, tight money and the great deflation. In the face of this external stability Britain correctly opted out of convertibility. In a recent paper, Bordo and Kydland (1990) analyze the gold standard as a contingent rule, meaning that the authorities could "temporarily abandon the fixed price of gold during a wartime emergency on the understanding that convertibility at the original price of gold would be restored when the emergency passed." See also Bordo and White (1990) for a recent discussion of British and French finance during the Napoleonic Wars.

The gold standard has been variously referred to as the Golden Rudder, the Golden Brake and the Golden Umpire. As rudder, it gave direction to monetary policy; as brake, it inhibited excess monetary expansion (beyond that provided by fluctuations in supply of the precious metals); as umpire, it allocated the burden of adjustment fairly between surplus and deficit countries. 1/ The classical economists did not believe that these functions could be adequately performed by any monetary system that was not tied to one of the precious metals.

The consensus of the classical school rested on five pillars: the homogeneity postulate, Walras' 2/ Law, the quantity theory of money, purchasing power parity and the assumption of flexible wages and prices. Together these assumptions made up the classical dichotomy. Changes in exchange rates engineered by monetary policy would result in equilibrium changes in wages and prices with no change in the underlying real equilibrium of real wages, the real rate of interest or the terms of trade. Devaluation--which in the classical world meant an increase in the national-currency price of gold--would increase prices and wages, improve the balance of payments and increase the money supply until the initial real equilibrium was restored. Money was neutral and there was no place for a theory of the exchange rate related to effects on the balance of trade or real variables.

2. The Transfer Debate

The classical consensus on exchange rates and the neutrality of money, however, did not imply unanimity on the nature of the international adjustment process. Specifically, are changes in relative prices necessarily associated with change in the trade balance induced by unilateral transfers? The subject can most easily be addressed in two parts, first that applying

1/ If fiduciary monetary issues were permitted, they had to be kept within the limits of safeguarding the nation's central gold reserves. Gold also ensured that the domestic price level would remain, in the long run, basically constant relative to the world price level. The stability of the world price level would itself depend on the relation between the world demand and supply of gold. When gold supply increased more rapidly than gold demand, the value of gold would fall and the world price level would rise; when gold supply increased less rapidly than gold demand, gold would appreciate and the world price level would fall. Although there were disturbing trends in the value of gold over the long period due to changes in supply, new gold production in any given year was a small proportion of the outstanding stock of gold so that annual changes in the price level were comparatively small. Over the long run, relative changes in the real value of gold would affect supply and demand in a stabilizing way, with the result that periods of rising prices tended to be offset by falling prices.

2/ I have argued elsewhere (Mundell 1968a) that John Stuart Mill at least, among the classical economists, did not confuse Walras' Law and Say's Law.

to barter, and second, that applying to the international monetary economy. The first problem is to see how changes in transfers and the balance of trade affect changes in the terms of trade and the real exchange rate. The second problem is to determine how transfer affects the balance of payments (under fixed exchange rates) and the exchange rate (when the money supply is constant).

Analysis of the barter theory reveals two camps, the difference between which crystallized only during the Keynes-Ohlin-Rueff debate on the transfer problem in the 1920s. The debate that emerged therein pitted the income-expenditure approach against the relative prices approach, with Wheatley, Ricardo, Longfield, Bastable, Nicholson, Wicksell, Ohlin on one side and Mill, Taussig and Keynes on the other. ^{1/} The Mill school argued that the generation of a balance-of-trade surplus required a fall in relative prices of the transferring country, whereas the income-expenditure school denied that changes in relative prices were an essential ingredient in the adjustment process. As more sophisticated analysis by Pigou, Haberler, Meade, Samuelson and others subsequently showed, the income-expenditure school was vindicated.

The income-expenditure school got the better of the argument as far as the effect of transfer on the terms of trade is concerned; Viner (but not Keynes ^{2/}) conceded the point. A revision of the literature became necessary to give credit to those classical writers who had fully incorporated expenditure effects into the analysis. Within the framework of the theory of exchange a transfer of purchasing power produces expenditure effects that worsen the balance of trade of the receiving country and improve that of the paying country. The redistribution of world expenditure from the paying to the receiving country could leave global excess demand unchanged (if the expenditure effects cancel) or it could result in an excess demand for one country's goods, leaving room for changes in relative prices. The essential factor in the adjustment mechanism is the shift in expenditures relative to incomes in the two countries, the relative price changes being incidental to the process.

Mill and Marshall had investigated the stability question of whether a change in relative prices (a change in the terms of trade) would shift excess demand onto the good whose relative price has fallen. Marshall deduced that it depended on whether an elasticity criterion--the sum of the

^{1/} See Mundell (1989a and especially 1989b) for a recent review of the literature on this subject and for detailed references. An excellent review of the literature up to 1937 is offered by Viner (1937).

^{2/} Keynes did not admit Ohlin's point in his reply to Ohlin's article in the *Economic Journal* but later in correspondence with Ohlin he wrote: "As to your point that reparations cause a shift in the demand curve of the receiving country irrespective of any rise in the price level of that country, I do not think I disagree with you." This is quoted in Dimand (1988, p. 124) from Patinkin and Leith (1978 p.162).

elasticities of demand ^{1/} for imports minus unity--was positive or negative. There is no *a priori* way to determine that such a criterion will be satisfied. Economists in the post-war period came to be divided into two camps--elasticity optimists and elasticity pessimists, with the former stressing the advantages of relying upon the price mechanism, the latter the need for direct controls. Nevertheless, it is usually assumed in economic analysis that the stability conditions are met.

Assuming that the underlying stability condition is satisfied, the direction of change in the terms of trade is determined by the expenditure effects, whereas the extent to which the terms of trade must move in order to correct any residual excess demand depends on the magnitude of the price elasticities. Because the elasticities fully incorporate supply effects, the latter play a role in determining the extent of any necessary changes in relative prices. The higher are the elasticities, the more effective any given change in relative prices will be in eliminating a given excess demand. It should therefore be noted that even if the re-arrangement of world demand creates excess demand, supplies might adjust to prevent changes in relative prices if one country were incompletely specialized and produced both goods at constant costs. ^{2/}

It is necessary now to consider the role of non-traded goods. The discussions of the 1920s had been couched in the framework of a two-good two-country model with primary focus on the terms of trade. But writers of both schools, going as far back as Cantillon and Hume, and including Ricardo, Taussig and Keynes, had given some consideration to domestic (non-traded) goods. The existence of domestic goods brought into consideration not only the terms of trade, but also the ratios of the prices of domestic and import goods and the ratios of the prices of domestic goods and export prices in each country. One convenient definition of the real exchange rate is the ratio of the prices of home-produced goods at home and abroad. For a small country, in which the terms of trade are constant, the real exchange rate is sometimes defined as the ratio of the price of domestic (non-traded) and international (traded) goods.

^{1/} This assumes trade is initially balanced; a slight adjustment is required where this is not the case.

^{2/} Constant costs and incomplete specialization imply that the elasticity of demand for imports is infinite. Meade (1950) produced the first algebraic formula for the change in the terms of trade explicitly integrating expenditure and price effects. Note that because expenditure propensities and price elasticities are related by the Johnson-Slutsky condition, import propensities high enough (greater than one-half on the average) to shift demand onto the goods of the paying country (the anti-Mill direction) are sufficient to ensure the elasticities of demand are large enough (greater than one-half on the average) to ensure stability of exchange equilibrium. See Mundell (1960 and 1968).

The introduction of domestic goods into the transfer debate seemed to qualify the victory of the income-expenditure school somewhat, giving the Mill-Taussig school a second line of defense. Changes in relative price levels can take place even if the terms of trade are constant. In his little treatise, *Capital Imports and the Terms of Trade*, Roland Wilson (1931) argued that an income transfer would shift demand onto domestic goods in the receiving country and away from domestic goods in the paying country. The effect on the relative prices would then depend on cost conditions. Under the assumption of increasing (opportunity) costs, the price level will tend to rise in the receiving country and fall in the paying country.

Increasing costs, of course, represent only one possibility. If opportunity costs between domestic and international goods are constant, no change in the real exchange rate will result. And if there are decreasing costs, the relative price of domestic goods would tend to move in the opposite direction from that postulated by the Mill-Taussig school. The result will depend upon the conditions of supply.

An interesting case involving domestic goods which lends some support to Mill's conclusion that the terms of trade of the paying country worsen as a result of transfer, was recently analyzed by John Chipman. ^{1/} If each country produces export and domestic goods (but no import-competing goods), there will be a tendency for the terms of trade to change in the Millian direction. This conclusion follows because resources are shifted from export to domestic industries in the receiving country, and in the opposite direction in the paying country, restricting the supply of exports in the receiving country and augmenting the supply of exports in the paying country. Note that in this case it is the relative price of domestic and export goods that will be constant in both countries, but the relative price of domestic and import goods will rise in the receiving country and fall in the paying country; the real exchange rate will thus also improve along with the terms of trade in the receiving country and fall in the paying country.

Chipman's result depends partly on the supposition that there are no import-competing goods in the two countries. To complete the analysis it is worth considering the results from a two-factor three-commodity Heckscher-Ohlin model under conditions that permit factor price equalization in the two countries. ^{2/} In this case it can be shown that a transfer may have no effect on relative prices despite the increase in the output of domestic goods in the receiving country and the decrease in the output of domestic goods in the paying country. Let us consider first the condition of a receiving country which faces fixed terms of trade, given by the rest of the world. This fixes factor prices in the import and export industries of the receiving country and therefore costs of production in the domestic goods

^{1/} Chipman's analysis was presented at the 1988 meeting of the Eastern Economic Association in Baltimore.

^{2/} This paragraph was part of my comment on Chipman's Baltimore paper.

industries. The receiving country's increase in expenditure (financed by the financial transfer) on import and export goods can be supplied by a decrease in its trade balance, but the increase in spending on domestic goods (assuming no inferior goods) must be supplied internally, by a shift of labor and capital out of the export and import competing goods industries. It is possible to reduce the productions of the two international goods in exactly that proportion needed to release the needed factors--along Rybczynski-theorem lines--required to produce the additional domestic goods at constant factor prices. The conclusion does not, however, depend on the small country assumption, provided that expenditure effects cancel. If marginal tastes for commodities are the same in the two countries, there will be no change in the global demand for factors and therefore no change in relative factor or commodity prices. But even if expenditure effects cancel, there is no presumption (given as always incomplete specialization) that the real exchange rate will change in the Millian direction.

3. Transfers, Gold Flow and the Exchange Rate

Now consider the equilibrium after unilateral transfers in a monetary economy. Consider first the case where there is a single money--call it gold--in the world economy so that changes in monetary conditions will be reflected in the balance of payments rather than exchange rates. In the absence of transfers, equilibrium will prevail when the trade balances and the excess demands for gold are zero. Now let the equilibrium be disturbed by a unilateral transfer. How will this affect the balance of payments of the two countries? Will gold move to the receiving or the paying country? 1/

1/ Jacob Viner addressed this question in his *Studies* (1937), analyzing the issue on the assumption that what he called the "final-purchases velocity" of money was constant. Viner's analysis resulted in an exchange with Dennis Robertson (1938) in the *Economic Journal*, Robertson arguing that it was more natural to assume that the income velocity of money was constant. In the ensuing debate a compromise position was arrived at, with Viner conceding that the money balances required to service the increase in expenditure in the receiving country would be, while still positive, less than that required to service expenditure from income produced at home, with the result that gold would still flow in the direction of the receiving country--changes in prices being abstracted from--but by a smaller amount than he had initially concluded. Robertson, however, held to his view that the receiving country would not require more money to service the additional expenditure made possible by tax reductions than was previously required to service the payment of taxes that the inward transfer made unnecessary.

A correct analysis of the issue requires splitting the demand for money balances into sectoral demands by consumers, producers and governments. Because production in both countries is assumed to be constant, there would not ordinarily be a change in producer's money balances. But the money requirements of consumers would rise in the receiving country and fall in the paying country on account of the expenditure shifts. Government

The answer to this question derives from the monetary approach to the balance of payments: money will flow to the country in which the transfer creates an excess demand for money. Of course the demands for money in the two countries will depend partly on how demand shifts affect price levels. Money will normally flow to the country in which the price level increases. But there are also important independent effects that do not depend on changes in price levels. To isolate these effects let us suppose that expenditure effects cancel.

The direction of change in the excess demand for money will depend on the arguments--apart from prices--in the demand-for-money function. To the extent that liquidity requirements depend on transactions, there will be an increase in demand for money in both countries to satisfy the additional transactions involved in arranging the transfer. 1/ The extra transactions will therefore imply that domestic expenditure in the receiving country will rise by less, and in the paying country fall by more, than the transfer itself until the additional monetary requirements are met. If the stock of monetary gold in the world is constant, the increased world demand for money will result in a once-for-all decline in the world price level in order to raise the real value of the world gold stock to that required by the higher level of transactions.

Over and above this once-for-all transactions effect affecting both countries, there will be a differential effect on the two countries in view of their different positions after the transfer. The movement of gold between the two countries will depend on how the demand for money is affected by the increase in expenditure (absorption) in the receiving country and the decrease in the paying country. 2/ If, as seems natural,

1/ (Cont'd from page 7) spending, however, would rise in both the receiving and the paying countries in view of the transactions associated with the disbursement of the proceeds through (say) an income subsidy in the payee and the raising of the proceeds by (say) taxation in the payer.

1/ I have noted this transactions effect in Mundell (1988). As noted in the preceding footnote, a precise treatment of the increased transactions demands would require splitting national demands into household, producer and government sectors.

2/ An alternative formulation would make the demand for money a function of wealth. Insofar as we are restricting our analysis to unilateral transfers, wealth, and therefore the demand for money, is lower in the paying country and higher in the receiving country, resulting in a balance-of-payments surplus in, and a flow of gold to the receiving country.

This formulation would have to be modified in the case of capital movements. Unlike unilateral transfers, capital movements do not imply a change in wealth. They represent a geographical redistribution of wealth. The purchasing power (absorption) of the borrowing country is increased, and of the paying country is reduced, by the capital transfer, but the GNPs, which take into account international interest payments, are only slightly affected (by the extra rent associated with differences in

the demand for money in the two countries depends partly on expenditure, the receiving country will require more, and the paying country less gold. To accumulate the additional liquidity, inhabitants (including the government) in the receiving country will increase spending by less than the transfer, thus improving the balance of payments and attracting gold; and similarly, the paying country will for a time decrease its expenditure by less than the transfer, generating a balance-of-payments deficit until the lower equilibrium quantity of money is established.

The three individual types of effects of a transfer on the demand for money must be summed to determine the direction of the gold flow. The changes in the demand for money in the two countries arising from expenditure shifts must be super-imposed on the pure transactions effect of the transfer discussed earlier; both effects operate to increase the demand for money in the receiving country, but they have opposing tendencies in the paying country. To these two effects must then be added those additional effects resulting from any altered patterns of demand that induces changes in price levels. To the extent that there is a presumption that the real exchange rate of the receiving country rises as a result of the transfer, the price effect will combine with the other two effects to increase unambiguously the demand for money and therefore the balance of payments of the receiving country. The transfer will in this case result in a flow of gold from the paying country to the receiving country in addition to some deflationary effect in the world as a whole in the absence of any increase in the global money supply.

An alternative monetary assumption is that exchange rates are fixed but that the central bank has some flexibility in its credit policy. Consider, therefore, two countries with gold-convertible currencies, but in which central bank assets are composed not only of gold but of government or private-sector assets. Because it raises the level of transactions, the transfer will now, as before increase the demand for money. Now, however, the central banks in the two countries could prevent world deflation by open market operations in government or private securities. Now consider the changes in the demand for money induced by the increase in the receiving country and the decrease in the paying country of spending. To preserve balance of payments equilibrium at fixed exchange rates, making a gold flow unnecessary, the central bank in the receiving country could expand credit, and the central bank in the paying country could contract credit. In these cases, however, the ratio of gold to total monetary liabilities of the central bank would be correspondingly altered.

2/ (Cont'd from page 8) rates of return). On these grounds, therefore, the inter-temporal pattern of consumption would not theoretically be affected by borrowing; most of borrowed money should therefore be devoted to the formation of physical or human capital. This implies that the product mix of marginal expenditure effects will be influenced by the proportion of the transfer that is unilateral and the proportion that is merely a loan.

Let us now consider the alternative monetary assumptions where the authorities hold the money supply constant and allow flexibility of the exchange rates. According to the monetary approach to the exchange rate, the transfer will create an appreciation or depreciation of the exchange rate in those circumstances in which, under a common money or fixed exchange rates, the balance of payments would move into surplus or deficit. Given the requirement of purchasing power parity as an equilibrium condition, an appreciation of a country's currency implies either a fall in the domestic price of tradable goods in the appreciating country or an increase in the price of tradable goods in the other currency.

As a result of a transfer--with fixed money supplies and flexible exchange rates--there will be, as before, a transactions effect that increases the demand for money in both countries, leading, to some degree, to world-wide deflation. Superimposed on this tendency will be an excess demand for money in the receiving country and an excess supply of money in the paying country due to the redistribution of world expenditure; this leads to appreciation of the currency of the receiving country and a depreciation of the currency of the paying country. Both effects therefore work to create some deflation in the receiving country, but there are opposing tendencies in the paying country. Changes in relative prices induced by the rearrangement of world demand will have additional effects. To the extent that the prices of home-produced goods rise in the receiving country and fall in the paying country, the appreciation of the currency of the receiving country will be reinforced.

4. Bickerdike's Breakthrough

Classical reasoning, as we have just seen, emphasized an equilibrium methodology, price flexibility, the neutrality of money and the long run. This was in contrast to the methodology of the mercantilists, who stressed disequilibrium transitional effects, rigidities and short-run solutions. The classical school's rejection of mercantilist theory seemed much too harsh after the monetary chaos and short-run problems, including mass unemployment, thrown up by the breakdown of the gold standard during World War I. The exchange rate theories developed in the inter-war period--and they survived long after World War II--had closer affinity to mercantilist than to classical doctrine.

Mercantilist writings have a modern flavor in their account of lags in price adjustments, giving rise to real effects of devaluation. Early in the 17th century, mercantilist writers recommended devaluation as a means of stimulating the economy. Schumpeter (1954, p. 365, n.8) credits Simon Clement (1695) with the best early statement of the "description of the sequence of events that devaluation will produce as long as domestic prices do not respond to it: bullion will flow, exports increase, imports decrease. He was not the first to see that, but his is the first compact statement, so far as I know, of the particular piece of the mechanism, made with a full sense of its importance." In the following century, many writers, not least

being Hume, emphasized the importance of employment effects in the transition from a lower to a higher price level.

Even during the classical period, there was important dissent from conventional prescriptions. Dissent (as usual) emerged especially in times of stress, particularly after the dislocations of big wars. Thomas Attwood and other members of the Birmingham school, for example, advocated either devaluation (a higher price of gold) or flexible exchange rates in the deflationary period following the Battle of Waterloo. A century later these ideas found their echo in the works of Fisher and Keynes, openly criticizing the attachment to external rather than internal stability.

A breakthrough in theory of exchange rate analysis occurred with the publication of C.F. Bickerdike's analysis (1920) of the "instability of foreign exchange." In a few short pages this landmark contribution 1/ established the "elasticity approach" to the balance of trade--defects and all! Bickerdike starts out with an additional payment of Z dollars "conceived of as an old loan falling due." He then differentiates the balance-of-payments equilibrium equation (inclusive of the transfer) to arrive at the elasticity expression made famous later by Joan Robinson and Lloyd Metzler." 2/ He arrives at the effect of a transfer payment on the exchange rate, thus answering the following question: By how much must the exchange rate change in order to equilibrate the balance of payments after a transfer. Bickerdike also develops the special case--usually identified with Lerner--for the change in the exchange rate when supply elasticities are infinite. 3/

It is ironic that the elasticity condition came to be called the "Marshall-Lerner" condition. Marshall repudiated the notion that his real equilibrium could be applied to the analysis of devaluation, while Lerner added nothing new to Bickerdike's formulation. It would be more appropriate to call it the Edgeworth-Bickerdike condition.

There are, however, defects in Bickerdike's analysis, not usually corrected in subsequent expositions. It is not in general valid to assume, as Bickerdike did, that "the shape of the demand and supply curves for commodities remains unaltered." Bickerdike's approach is not valid for

1/ Among a number of other important contributions to theory and policy, Bickerdike (1907) originated the theory of the optimum tariff. Edgeworth (1908), following along the lines of Bickerdike's optimum tariff work, had explicitly developed the criterion for stability of the real exchange equilibrium. Bickerdike was the first, however, to relate the elasticity formula to changes in exchange rates, using the device of two inconvertible currencies.

2/ Bickerdike uses the inverse of elasticities, what he calls "inelasticities."

3/ Bickerdike's "inelasticities" are zero.

comparing the *equilibrium* positions before and after an exogenous change such as "an old loan falling due." A payment cannot be made unless it is first financed, whether by money creation, taxation or a voluntary act of saving. If there is an excess demand for foreign exchange there must be an equal excess supply of the quid pro quo offered in exchange.

Although Bickerdike's analysis is not correct for analysis of a transfer, starting from and ending with an equilibrium position, it is valid for the study of the transition from a *disequilibrium* position to an *equilibrium* position. Starting from a position in which the expenditure effects of "an old loan falling due" have already been expended, Bickerdike's analysis gives us the answer to the question "by how much must the exchange rate fall to restore equilibrium."

Bickerdike is, at least in the short run, an elasticity pessimist. He worries that "if the Americans have not purchased British goods or stocks, or services to an extent sufficient to create a supply of dollars payable at the rate of exchange previously existing, the situation is not necessarily remedied by a further fall in the value of pounds in dollars. Prima facie, indeed, that fall makes the gap wider. If four million dollars are due to be paid, and Americans are due to pay one million pounds, if the exchange rate were four dollars to the pound, that would clear the market. If the exchange rate is three and a half dollars to the pound, however, there are not enough dollars obtainable with the million pounds, and competition does not tend to drive the rate up to four dollars, but, on the contrary, forces it lower and makes the impasse worse. It is like a "bear squeeze" on a stock market, when bulls have been caught short."

There are fortunately mitigating factors "which may generally be reckoned upon to prevent the rate falling away indefinitely; but it is important to note that the mitigations do not act as promptly as might be desired, and the fundamental instability is the dominating fact." He then goes on to discuss the factors that help, including the cheapening of stocks in terms of dollars, loans that the British can make, and the cancellation of some new business. But he correctly notes that insofar as stocks and shares pay interest in pounds, they must fall in dollars more than the exchange rate "unless the lower rate is expected to be temporary."

Bickerdike thus emphasizes capital market transactions and financial payments as dominating factors in the short run, and reserves his mathematical analysis for consideration of the adjustment of the trade account in the long run. Thus "In the longer period, such as a year or two years, the influence of relative demands for goods and services will tell, but still it is comparatively short-period demands and supplies with which one is concerned. Large changes in the sources of supply can be effected over periods of many years, and the degrees of inelasticity may be considerable over periods of one or two years." All that is needed for the J-curve is a graph!

Bickerdike's contribution founded the elasticity approach that would occupy such a prominent place in subsequent literature. His framework, however, had the defect that it analyzed the nominal exchange rate outside a framework of full general equilibrium. Its omission of macroeconomic relationships created the opening for the "absorption approach" and its omission of conditions of money supply and demand left it vulnerable to the monetary approach.

III. Alternative Deficit Theories

A correct analysis of the relation between the balance of trade of a country and other variables must take into account all the factors that enter into the balance of trade. The literature has concentrated on three approaches to the balance of payments or trade: (a) the elasticities approach; (b) the absorption approach; and (c) the monetary approach. The usual assumption made in the literature is that the capital account is either zero or exogenous, unchanged by the effects of exchange rates so that the analysis applies indifferently to the balance of trade or the balance of payments. But in a complete analysis where capital flows are allowed for, the elasticity and absorption approaches are ways of analyzing the balance of trade, and only the monetary approach is a way of analyzing the balance of payments.

1. Sixteen Approaches to the Balance of Trade

There are exactly sixteen ways of looking at the balance of trade (or balance of payments). ^{1/} Eight of the approaches can be deduced from the application of Walras' Law in the national and international economy; the other eight approaches derive from the application of inter-temporal budget constraints.

General equilibrium requires analysis of the conditions of equilibrium in three broad aggregates: goods, securities and money. Each sector, of course, could be disaggregated into domestic and foreign components, and subdivided further into the innumerable categories relevant in the real world. To illustrate the sixteen approaches to the balance of trade or payments, however, is sufficient to maintain the degree of aggregation implied by the three-fold classification. For the sake of simplicity, I shall also bypass the difference between GDP and GNP, and the corresponding difference between the trade balance and the current account. But a big warning is in order: this difference cannot be ignored when we deal with the eight intertemporal approaches to the balance of payments!

^{1/} The next few paragraphs draw on and summarize the analysis of a paper delivered first at the Caracas Conference on Financial Markets in January 1979 and published in Spanish in the Conference volume. See also Mundell (1989, 1990).

The first equation we write down is the identity utilized by practitioners of the elasticity approach, namely that the balance of trade is equal to exports minus imports, $B = X - M$. By specifying the arguments of the functions relations determining supplies of exports and demands for imports we arrive at approach (1) to the balance of payments. Approach (2) is in the same family; it results from specifying the functions determining the supply of exports and the demand for imports in the foreign country. From "Cournot's Law", the trade surplus of one country is the trade deficit of the rest of the world, due allowance being made for asymmetries in the treatment of reserve suppliers and gold producers.

Analysis of these factors owe much, as we have seen, to Edgeworth and Bickerdike, subsequently elaborated by Robinson, Lerner, Haberler, Meade, Machlup and Metzler. A more sophisticated development of this approach would incorporate into the demand functions for exports and imports not just exchange rates, prices and output, but domestic expenditure, capital movements and wealth; it should then not be called the elasticity approach (why not income propensities approach?) but rather the "direct approach" because it examines directly the forces determining exports and imports. The lesson from the direct approach is that no policy variable, including the exchange rate, can improve the balance of trade without increasing exports relative to imports. The direct approach is no more and no less valid than the other seven pairs.

The next pair of approaches makes use of the absorption relations: $B = Y - E$ where Y is income and E is expenditure (absorption). This relationship can be subdivided into separate sectors to yield the equation $S - I + T - G = B$, where S , I , T and G , refer, respectively, to savings, investment, tax revenues and government spending. There are two equations of this type, of course, one for the country under analysis and the other for aggregate of economies in the rest of the world. These approaches owe much to Gervaise, Metzler, Machlup, Chipman, Meade, Alexander (especially) and Harry Johnson. Alexander called "absorption" what Meade had called "domestic expenditure," and this inspired the name usually given to the approach. The lesson from this approach is that no policy variable, including the exchange rate, can alter the equilibrium balance of trade (current account balance) unless it increases output (income) more than expenditure.

The next pair of approaches relies on the balance-of-payments identity. The accounting balance of payments is the external counterpart to Walras' Law. The current account surplus is equal to net capital exports plus the balance of payments deficit. Put another way, the current account surplus finances an import of securities (which is a capital export) or an inflow of money (balance-of payments surplus). We can write $B = K + P$, where K is the net capital outward flow and P is the balance of payments surplus. Similarly, any current account surplus of the rest of the world must be matched by net capital exports of the rest of the world or an increase in

its net reserve assets. ^{1/} The lesson from this approach is that no policy variable can improve the balance of trade without generating a greater net capital outflow and/or increase in international reserve assets that is acceptable in the rest of the world.

It is worth noting here a long-standing dispute about the direction of causation in the balance-of-payments identity. There are those who argue that trade balance has an independent existence requiring external finance; on the other hand, there are those who argue that capital exports and other transfers induce changes in the balance of trade through the mechanism of financing changes in expenditures. It should be emphasized, however, that the identity itself says nothing about causation; there are merely two sides to the same coin. It cannot be inferred from the balance of payments identity whether the financial transfer determines the trade balance or the trade balance determines the financial transfer.

The classical school, up until Mill, put the capital flow in the driver's seat. It emphasized the fact that the mere act of making a financial transfer automatically implies the additional saving to finance it, so that the mere existence of an outward capital flow implies an excess of income over expenditure. But the Mill-Taussig-Keynes omission of expenditure effects in the adjustment mechanism opened the way for Keynes's idiosyncratic view of the trade balance as a given entity independent of financial transfers. Although this involves a false dichotomy--independence of spending from finance--Keynes' view has strongly influenced the modern literature (in my opinion to its detriment). It is easy to see why economists who write down equations specifying exports as a function of output and prices fall prey to the single-entry book-keeping approach that ignores the mechanism through which capital inflows or transfers stimulate an increase in expenditures.

Causation can in principle go in either direction. But it is unlikely that enduring changes in the balance of trade can be motivated other than by basic financial transfers. It is true that a shift of domestic expenditure from home to foreign goods will worsen the balance of trade and immediately force accommodating finance, a loss of reserves or a change in the exchange rate. But accommodating finance is purely temporary and short-run. Major trade deficits and surpluses are nearly always not only accompanied by but induced by voluntary financial transfers of one kind or another.

Because there are two balance-of-payments equations--one for the home country and one for the rest of the world--we now have six approaches to the balance of trade; there are only ten remaining.

^{1/} This abstracts from asymmetries in the treatment of the reserve assets that result in a recorded non-zero balance of trade for the world as a whole.

Approaches (7) and (8) make use of Walras' Law directly. The sum of all excess demands is zero, so that any excess supply of commodities must be matched by an equivalent excess demand for securities and money. Because the trade balance is equal to the excess supply of goods it must also be equal to the excess demand for securities and money. Formulated in a different way, the excess supply of goods is equal to the sum of the budget surplus and the excess of saving over investment. Any budget surplus must be expended on money or securities; it implies an excess demand for money or government bonds on the part of the government. Similarly, an excess of saving over investment equals an excess demand for securities and/or money. We can therefore deduce the current account deficit solely by examining the markets for securities and money. This approach applies to both countries so that we now have eight approaches to analysis of the balance of trade.

It will be convenient at this point to pause to summarize the eight approaches in the form of a system of equations. These equations can be looked upon as identities after the fact or as equilibrium conditions when they are defined in terms of intentions and plans.

$$\begin{array}{rclcl}
 (1) & Y - E & = & B = -B^* = E^* - Y^* & \text{(Goods)} \\
 & + & & + & \\
 (2) & A - Z & = & -K = K^* = Z^* - A^* & \text{(Securities)} \\
 & + & & + & \\
 (3) & G - H & = & -P = P^* = H^* - G^* & \text{(Money)}.
 \end{array}$$

All the variables in the above equations are flows. A and Z refer to the supply and demand for securities, respectively, and G and H refer to money creation and hoarding, respectively, and P is the balance-of-payments surplus.

The plus signs between the lines are inserted to indicate that the rows can be added vertically to equal zero.

Now for the inter-temporal dimension, the other eight approaches. Inter-temporal conditions are implied when we say that a deficit is "unsustainable." If a deficit is unsustainable the worry is not that it will endure but that it will disappear in an unpleasant way. The above equations do not reflect this fact because they have not as yet taken into account the inter-temporal version of Walras' Law. Trade deficits integrated over time yield net debt; and debt has to be paid back. We owe to David Hume--the theorem that Charles Kindleberger called Hume's Law--that exports equal imports. Hume meant that the balance-of-payments adjustment mechanism would eventually bring imports into equality with exports. This is emphatically not true in the short run; the U.S. has had a trade deficit for nearly a generation, as it had a surplus for two prior generations. But there is a sense in which Hume is right after all. Exports equal imports over long periods of time.

A country's net indebtedness is limited by its perceived export capacity. The U.S. can sustain a trade deficit because it has the productive power to service and ultimately repay its debts. If, on the contrary, it were perceived that the United States could not generate the future trade surpluses to service a growing debt, external finance would dry up and expenditure would be brought back in line with income, eliminating the deficit. No country, of course, can pile up debts indefinitely unless they are at the same time generating increased means for repaying them. A country borrowing to expand its capital plants and productive power has more credibility than a country borrowing to finance pure consumption spending. It is sometimes relevant to note which sectors are doing the borrowing or lending. If the private sectors do the borrowing, the responsibility for making repayments rest with the individuals; households or firms might legitimately borrow to achieve a better inter-temporal pattern of consumption along life-cycle lines. ^{1/} Budget deficits, however, raise other questions, creating the potential for mistaken shifts in the inter-generation distribution of income.

At some point diminishing returns will set in and a country will reach its debt limit, i.e., the limit beyond which it can no longer finance current account deficits. Higher debts require larger interest payments which at some stage would require unattainable trade surpluses to maintain debt service. At this stage risk of default soars. An example is afforded by many of the less developed countries. The LDC's borrowed excessively in the negative real interest days of the late seventies and early eighties. When real interest rates rose and their future export capacity seemed insufficient to service their debts, the flow of capital dried up and they were left with a painful adjustment process that has abated only moderately.

And so we come, finally to the additional eight approaches to the balance of trade, that each one of the above-mentioned eight approaches has its reverse counterpart in the inter-temporal dimension when surpluses and deficits have to be reversed. In the long run, therefore, we need to pay special attention to the difference between the GDP and GNP, and to the equivalent difference between the trade balance and the current account balance. Expected future net outpayments to service the debt, or in payments of debt service, like repayments, lower or raise disposable income and domestic expenditure. The larger the debt the larger the gap between GDP and GNP, and between the trade balance and the current account.

We must also pay attention to the pall that heavy indebtedness casts over expectations about the exchange rate and bond prices signalling decreased credit-worthiness. In the long run debts involve a subtraction from wealth that should affect our current propensities to save.

^{1/} Even where the inter-temporal borrowing is legitimate, however, there might be externalities associated with changes in the terms of trade that may create a divergence between private and social benefits from external borrowing.

Individuals pay attention to those negative factors pertaining to personal debts, but only super-rational individuals correctly subtract from personal wealth the share of the national debt that implies future tax obligations.

2. Is There an Equilibrium Trade Balance?

The major difference between a creditor and a debtor nation is that the former is enabled to consume more, while a debtor must consume less than its income; corresponding differences develop (given balance of payments equilibrium) between the current account and the trade balance, and between Gross Domestic and Gross National Product.

How can disequilibrium be distinguished from equilibrium? Is the balance of payments the criterion of equilibrium or is it the balance of trade? When is a current account deficit or surplus a symptom of disequilibrium?

Attitudes on this subject have changed with the international monetary system. Under fixed exchange rates, the balance of payments occupied center stage. There were, to be sure, alternative definitions of the balance of payments and some debate about which was the most useful definition for the United States. One important, and widely used, measure as target variable was the "basic balance," the difference between the trade balance, repetitive transfers and "normal" capital exports. Another measure was the "official settlements balance" which measured the change in net liabilities of the monetary authorities. Whichever of these measures was employed little attention was given to the trade balance or the current account balance. The most important problem was to ensure that the balance of trade and the net capital account offset one another, not that any particular level of the balance of trade or current account was achieved. In the Tinbergenian language of targets and instruments, the exchange rate and the balance of payments (not the balance of trade) were targets, not instruments, of economic policy. 1/

1/ Some attention was, it is true, paid to the impact of the trade balance on the level of employment, theorizing that was at least partly based on errors in economic theory. Superficial mercantilist reasoning, arguing from the national income accounts, Keynesian equations, and the fact that exports are source of demand while imports are not (except for repercussion effects), seems to imply that an improvement in the balance of trade increases employment. This conclusion may be correct if the disturbance to the balance of trade originates from an increase in export demand on the part of the rest of the world; but it is false if the improvement arises from a reduction in domestic expenditure. It is a fallacy to argue that measures that improve the balance of trade *ipso facto* increase employment.

In late 1971, after the breakdown of the Bretton Woods system, an improvement in the trade balance became, for a brief period, an explicit target of government policy. Secretary-of-the-Treasury John Connally made

To the extent that the level of the current account was analyzed at all, capital movements, combined with a corresponding trade balance, were thought to be a desirable way of allowing countries to optimize the inter-temporal distribution of spending. By borrowing and running a trade deficit, a high interest-rate country could accelerate its economic development and service the loan out of the augmented income; and similarly, a low interest-rate country could find an outlet for high savings by investing abroad and earning a higher rate of return than that available at home. Under fixed exchange rates the current account was expected to adjust itself to the optimum extent provided capital moved from low-interest to high-interest rate countries. The function of capital imports was to enable a country to attain economic objectives at an earlier stage than would otherwise be possible.

What came to be called "the composition problem" 1/ of determining the correct structure of the balance of payments was defined as follows:

If the balance-of-payments statement is divided into a current account, a capital account and a reserve account, and if adjustment is defined as the correction of the reserve account, should adjustment be achieved by improving the capital account or the current account? What criteria should determine the appropriate composition of adjustment? 2/

Put another way, what are the costs and benefits--the welfare implications--of a larger or smaller current account surplus fully financed by corresponding changes in capital exports? Alternatively, at what rate should a country build up or run down its international equity? What is the optimum level of external equity or debt and what is the optimum speed of adjustment of the equity or debt, i.e., the optimum current account surplus

1/ (Cont'd from page 18) his famous assertion that the United States wanted a turnaround in its trade balance of \$13 billion. Treasury officials naively hoped that the devaluation of the dollar negotiated at the Smithsonian Institution would bring about that turn-around in the trade balance.

The devaluation failed to make a dint in the trade balance. The trade balance deficit actually severely worsened in 1972 to over \$6 billion, the largest deficit up to that time in US history. Although the balance became positive in 1973, it worsened again in 1974, only to become positive again in the recession year 1975.

The year 1975 marks a watershed in the history of the international accounts. It was the last year the United States had a trade surplus. Deficits rose after 1976 to over \$30 billion and after 1983 to over \$100 billion.

1/ At the 1966 Chicago Conference on International Monetary Problems. See Mundell and Swoboda (1969).

2/ See Mundell and Swoboda (1969, p. 38).

or deficit. Unfortunately, beyond the general recognition that international capital movements were means of altering the inter-temporal distribution of expenditure relative to income, there was no explicit criterion for determining the rate a country should import or export capital.

Concern over a country's long run creditor position partly reflects concern over its power position. In an age of war, a country's creditor position gave it a potential source of finance for war-time imports, supplementing the "gold chest" to which every country sought access. Other things equal, a debtor country's power position is weaker than a creditor country's. What Hume called the jealousy of the balance of trade may just as aptly be applied to the jealousy of a country's creditor position.

The distribution of power is a function of the size and wealth structure of countries in the world economy. It has often been remarked that the world economy is an oligopoly; each country reacts directly to the actions of other countries. The United States and Japan have reversed positions as the world's largest creditor. ^{1/} If the rest of the world were composed of small economies that, singly or in combination, could never be a political or military threat to the United States, there would be little concern over the US trade deficits beyond the ordinary economic problem of servicing the rising level of debt.

3. The "Stages" Theory

Historically, current account deficits can persist for long periods of time. The United States had a chronic deficit in the nineteenth century as it made use of capital imports to accelerate its development. From World War I to about 1975, the United States had substantial surpluses. In every year since 1975 the United States has had substantial deficits, rising sharply to over 3 percent of GNP in the 1980s. Other countries have experienced similar long trends, giving rise to theorizing about a sequence of "stages" of a country's balance of payments over time. According to this theory, the stages of the balance of payments can sometimes be related to the level of development of a country. Because capital exports (according to classical theory) are determined by relative rates of return and international savings rates, with international capital flowing from low profit to high profit countries, the stages approach implies systematic developments in net saving rates or rates of return on capital at different levels of development.

^{1/} There was a short period in which the decline in the US creditor position made Saudi Arabia the largest creditor country, before Saudi Arabia was overtaken by Japan and Germany.

The simplest framework suggests four stages of the balance of payments. ^{1/} A young country passes from the positions of (1) immature debtor to (2) mature debtor to (3) immature creditor to (4) mature creditor. A debtor pays interest on borrowed capital; a creditor receives interest on capital invested abroad. A debtor is immature when its trade balance is less than its debt service (implying that it is still borrowing); a debtor is mature when its trade balance exceeds debt service so that the country is no repaying its debts. A creditor is immature when its trade balance deficit is less than its interest receipts (implying it is still lending) and mature when its trade deficit is at least equal to its interest receipts.

The stages approach is an optimistic one in that it suggests a reversibility of debt positions. It seems natural that some prominent European and Pacific countries--notably Germany and Japan--moved quickly through several stages since the end of World War II, passing, in a matter of only three decades, from an immature debtor to an immature mature creditor position. One would predict from the theory gradually increasing absorption until trade surpluses were exchanged for trade deficits, arriving eventually at the position of mature creditor.

The United States, on the other hand, has passed from an immature creditor at the end of World War II quickly through the position of mature creditor to that of immature debtor, commencing the cycle again. It could be argued that the supply-side tax revolution in the United States in the 1980s rejuvenated the US economy, raising its marginal efficiency of capital and inviting capital from all over the world to partake in the largesse of reborn efficiency. This transformed the United States from the world's largest creditor to the world's largest debtor. According to the stages approach, the position will reverse only when either the rate of return on investment in the United States falls to the levels abroad, discouraging further capital inflow, or the US savings rate rises by enough to restore the export surplus of the earlier decades. With a reduction in absorption induced by the need to finance outward payments of interest and dividends, and a larger capital stock that lifts Gross Domestic Product above Gross National Product, the trade balance will move again into surplus.

There is a grain of tautology in the stages approach: what is borrowed has to be repaid (if it is not defaulted). The agents contracting the loan are committed to repayments. Apart from differential rates of return it is necessary to seek answers to chronic trade balances in differential

^{1/} See Cairnes (1874) for the development of an pioneering version of the stages approach; see also Taussig (1927). Crowther (1957) elaborates six stages: immature debtor-borrowers; mature debtor-borrowers; debtor-repayers and debtor-lenders; immature creditor-lenders; mature creditor-lenders; and creditor-drawers and borrowers. Halevi (1971) develops a twelve-fold classification (allowing for borderline cases). Other early work in the modern analysis of stages includes Onitsuka (1970, 1974), Neher (1970), Fischer and Frenkel (1974a, 1974b).

savings rates. One possibility arises from demographic changes that create systematic biases in the levels of borrowing and lending.

4. How Demography Creates Trade Deficits

If the various generations have different propensities to save and invest, as in the *life-cycle* (with or without bequest) hypotheses, blips in the age distribution of the population will alter net national lending and, via the transfer mechanism, the balance of trade. In Mundell (1990a, 1990b) I developed a four-generation model designed to elucidate the trade balance effects of demographic blips. The four generations included not only the very young and very old dependent generations, but two working generations. The cutting edge of the model lay not so much in the ratio of dependent to working generations--although that produced some of the usual effects--but in the different characteristics of the two working generations. While the working generations had similar earning and saving potential, the junior working generation invested--mainly in consumer durables--more than it saved, borrowing the difference; whereas the senior working generation saved more than it invested, lending the difference. The balance of trade will thus be a positive function of the proportion of senior to junior members of the working generations even though, as I assumed, each individual obeys its inter-temporal budget constraint on life-cycle principles.

Empirically, the model would predict big deficits in the 1980s, coupled with big surpluses in the late 1990s, taking into account two facts: (1) a small cohort of births in the 1930s (depression babies); and (2) a large cohort of post-war births (the baby-boomers). Thus the small numbers of depression babies, followed by the large number of baby-boomers, result in a small number of senior workers to do the lending in the 1980s and early 1990s, combined with a large number of junior-workers who are chronic borrowers. The small number of grown-up depression baby lenders is no match for the hordes of yuppie borrowers. The excess of borrowers over lenders lead to high real interest rates and a capital inflow that imposes a current account deficit on the country as a whole.

The demographically-inspired deficit is benign because it is reversible. In the succeeding period, when the baby-boomers mature into the senior working-lending generation and there is a relatively small cohort of junior borrowing-workers, there will be an excess of lenders that will swamp credit markets, create a substantial capital outflow and produce a sharp turnaround in the current account.

The effect of this cycle will be more important, the more asymmetrical is US experience compared to surplus countries like Japan and Germany. ^{1/} If the net saving rate theory, based on reversible age-distribution trends,

^{1/} Note, however, that the model would predict a turnaround again after the first decade of the 21st century as the baby-boomers retire and are replaced as senior workers by another small cohort.

is of sufficient importance, it would not be necessary to alter the policy of benign neglect of the current account deficit. If left alone, the deficits will reverse by structural-demographic adaptations in private saving.

5. The Fiscal Approach to the Deficit

Yet another possibility--not incompatible with the under-saving demographic theory--is that the US trade deficit has been dominated by the budget deficit and that a reduction of the budget deficit is a precondition for elimination of the trade deficit. The budget and trade deficits are of course connected by the national income identities of the absorption approach.

The budget deficit had been ranging from 0-2 percent of GNP for many years up to 1974, but it suddenly jumped up to 3.4 and 4.2 in the recession years of 1975 and 1976. From 1977 to 1981 it was consistently (except for 1979) above 2.5 percent of GNP: 2.6 in 1977, 2.6 in 1968, 1.4 in 1979, 2.8 in 1980, 2.6 in 1981. In this period, it will be remembered, the trade deficit lurched upward.

In the recession year of 1982, the budget deficit jumped to 3.97 percent of GNP. It then rose to 5.95 in 1983, 4.72 in 1984, 5.28 in 1985 and 5.01 in 1986. These were the peak years of the budget deficit. A remarkable turnaround came in the next three years with the budget deficit percentage falling to 3.26 in 1987, 3.17 in 1988 and 2.71 in 1989. The improvement in the budget deficit coincided with an improvement in the trade deficit, both however, remaining well over 2 percent of GNP.

More significantly, the changes in the budget deficit have been followed, usually with a one-year lag, by changes in the trade deficit. Figure I shows the budget and trade deficits as a percentage of GNP with the budget deficit advanced one year. The correlation is striking.

The future of the current account deficit will depend critically on net capital flows. Elimination of the trade deficit will depend on the United States resuming its historic position, since 1915, of being a net capital exporter. There are several reasons, however, to expect the United States to be the recipient of continuing substantial gross capital and money inflows in the 1990s. Increasing global financial integration, continued use of the dollar as a major international currency and continuing advantages from the tax reforms in the United States will be factors favorable to gross capital inflows. On the other hand, rising savings due to demographic factors will act to lower interest rates and induce larger capital outflows. As the decade of the 1990s unfolds, the capital outflow factor will get larger relative to the capital inflow factor and the current account deficit should gradually fall.

6. Exchange Rates and the Trade Balance

Exchange rate changes present another possible explanation of the trade deficit. It is widely believed that dollar depreciation can correct the trade balance, and that depreciation since 1985 has started an adjustment process that is gradually reducing the trade deficit.

At the outset, however, a difficulty sets in. Under the flexible exchange rate system, the exchange rate is an endogenous variable. Variations in the exchange rate represent the outcome of a solution of a system of general equilibrium equations. The exchange rate is an endogenous variable in a flexible exchange rate system. The effect of a change in the exchange rate on the trade balance can only be determined in the context of the changes in the exogenous variables that brought about the change in the exchange rate.

In the case of wage and price flexibility, depreciation of a currency brought about by an increase in the money supply would probably not affect the trade balance at all. An increase in the money supply would bring about a proportionate increase in wages and prices, including the price of foreign exchange, and no change in competitiveness or relative prices. Depreciation brought about by inflation clearly cannot be relied on to improve the trade balance. 1/

Another possibility is that the depreciation is brought about by monetary changes in the rest of the world. But if the dollar depreciates because of deflationary policies abroad, the deflation, or lower rates of inflation, will in the long run cancel the competitive advantages of the exchange rate changes. It can hardly be said that depreciation in a world of flexible wages and prices--whether brought about by inflationary policies at home or deflationary policies abroad--would have an ascertainable and predictable effect on the trade balance.

An alternative and widely-adopted assumption, drawn from the Keynesian model, is based on rigidities. Monetary changes in an environment of price and wage rigidities have real effects. In the Keynesian theory employment, it is usually assumed that wage rates are rigid. This assumption works best

1/ It is, of course, possible that inflation can shift the distribution of wealth, resulting in a rearrangement of world demand and a change in relative prices. Inflation also has fiscal effects insofar as it reduces the real fiscal burden of the public debt. If saving increases (because of the wealth-saving relation) and the budget deficit is reduced (because of reduced real interest payments), the trade balance will improve unless these changes are offset by increases in investment financed by increased capital inflows.

if there is unemployment. If the monetary authorities expand the money supply, the exchange rate will fall and prices will rise, lowering real wages and raising employment and output. Such a policy might well be in a country's interest regardless of its effect on the balance of trade. But that is not the issue. The issue under consideration is whether the exchange rate will improve the balance of trade.

In a flexible exchange rate system, the equilibrium exchange rate is that rate which clears the foreign exchange market. If there is no official intervention in the foreign exchange market--as in the case of a freely floating rate--the exchange rate will equilibrate the current account and the capital account. Any reduction in the trade deficit must be exactly matched by an equal reduction in capital imports; and similarly, any worsening of the trade deficit must be accompanied by an increase in capital imports. It is possible, therefore, to determine the direction of change in the trade balance by predicting the direction of change in capital imports.

The pure theory of international trade was mainly developed on the assumptions of either no capital movements or exogenously-determined capital movements. But if there are no capital movements (and no change in reserves) exports must equal imports both before and after the disturbance that changed the exchange rate. In the absence of capital movements, even under Keynesian assumptions of wage rigidity, the exchange rate--however it is brought about--cannot bring about a change in the balance of trade. ^{1/} In the absence of capital movements domestic expenditure and the gross national product must be equal.

But let us now relax the assumption that there are no capital movements and that the latter are unaffected by exchange rate policies. Assume that initial balance-of-payments equilibrium prevails with both a trade deficit and an equal capital inflow. We must now ask in what direction capital will move as a result of the policy change. Those who argue that a change in policy that brings about a depreciation of the dollar will improve the trade balance must also argue that the same change in policy will reduce the inflow of capital. If there is a decreased capital inflow as a result of the policy change, the trade deficit will be lower; but if there is an increased capital inflow, the trade deficit would worsen.

^{1/} In offer curve analysis of the Marshall-Meade type there is no place for a balance-of-payments deficit. A price vector that differs from the equilibrium produces an excess demand for the foreign country's good, which can only constitute a transactions configuration if there is intervention: Domestic commodity authorities must be selling foreign goods out of their own stockpiles, and simultaneously buying stocks of the domestic good. When official transactions are added to the private transactions, there is no deficit or surplus.

To answer this question we need to know how capital inflows are affected by the exchange rate. 1/ There is a difference between classical and Keynesian assumptions. Under classical assumptions, the level of the exchange rate would not have an important effect on capital flows. It is true that financial assets are now cheaper under the depreciated dollar; but so is the income from the financial assets. The ratio of income to capital value, which is the rate of interest, is unchanged by the exchange rate. Under classical assumptions of price and wage flexibility, therefore, neither capital movements nor the balance of trade would be affected by changes in the exchange rate. 2/

In the Keynesian case of rigid wages, however, a different conclusion applies. Monetary expansion that brings about depreciation lowers real wage rates. The resulting increase in actual and expected profits amounts to an increase in the marginal efficiency of capital. Foreign companies now find it profitable to shift production facilities to the United States to take account of the lower wage costs, amounting to a capital import. Even if exports increase due to the lower selling prices in terms of foreign currencies, imports must increase by an even larger amount in order to create an increased gap between imports and exports equal to the increased inward financial transfers. Paradoxically, the very factor that is often used to justify the effectiveness of depreciation in improving the trade balance--rigid wages--works against an improvement in the trade balance. 3/

Another possible rigidity is land prices. Unless the depreciation has been brought about by an increase in the US money supply, the price of land will not increase *pari passu* with the price of foreign exchange. Fluctuations in the dollar over the past few years have not been matched

1/ We might note that it is also necessary to determine the effect on the net capital outflow of the policy that creates the change in the exchange rate. If, for example, the central bank buys foreign exchange with newly-printed money and invests the foreign exchange in a foreign security, the trade balance must improve by the amount of the capital outflow.

2/ As noted earlier, this was pointed out by Bickerdike (1920, p.118-119). After discussing the fall in stock prices in terms of dollars as one of the mitigating factors that would encourage greater American demand for pound assets, he writes: "...it has to be noted that if the stocks and shares pay interests in pounds, a fall in value arising in this way must exceed the fall in the exchange rate before purchases by Americans becomes attractive, unless the lower rate is expected to be temporary."

3/ A stronger case can be made for changes in relative prices in the manufacturing sectors of large countries than in small open economies producing primary products for which there is a world market. The relative price of basic commodities are determined by real variables and it is unlikely that relative price changes play an important role in the process of adjustment.

by equivalent fluctuations in the price of land. The depreciation of the dollar since 1985 has made the yen and mark prices of American land much cheaper relative to the corresponding prices of Japanese and European land. To the extent that real assets are only slowly affected by the exchange rate, they become a target for bargain-hunting on the part of both investors and speculators. To the extent that this occurs the capital inflow will finance a larger excess of expenditure over income and worsen the trade deficit.

IV. Trade Balances As System Imbalances

It frequently happens that monetary problems have their origin, not in domestic policy mistakes, but in circumstances arising in the rest of the world over which the domestic authorities have no control. A typical example is the deterioration in a country's terms of trade because of outside factors. Oil shocks are a typical example, but most raw material producing countries have faced markets in which staple exports can only be sold at disastrous prices.

We are not concerned here with problems of this type, important as they are in real life. We are rather concerned with problems that have their origin in the international monetary system itself. An outstanding example of this type of shock is the great depression, which had its origin in the breakdown of the gold standard in World War I and its unwise restoration at an undervalued price of gold in the 1920s. A second example is the balance of payments deficit of the United States in the 1960s, which developed as an inevitable outgrowth of the post-war gold exchange standard in the face of a recurrent gold shortage. A third example lay in the great inflation that occurred in the wake of the collapse of the Bretton Woods system and the severing of the gold link of currencies to the commodity world on August 15, 1971. Yet another example was the renewed outbreak of wild inflation in the late 1970s that produced negative real interest rates and a level of third world indebtedness that was not viable after the disinflation of the 1980s.

Supply conditions have always played a key role in the selection of monetary metals. When countries fixed the legal prices of both gold and silver, an increase in the supply of one metal would result in a change in the standard as the abundant metal was exchanged at the mint for the scarcer metal. In the middle of the 16th century, silver became dominant after the great discoveries in Potosi, a dominance that was to last over two centuries. In the 18th century, new supplies of gold from Brazil enabled Britain to move toward a gold standard. In the middle of the nineteenth century, France, the bi-metallic power of the first half of the century, exchanged her silver currency for gold with the great discoveries of that

metal in Australia and California. 1/ The great gold discoveries in South Africa in the 1880s ensured an easy trend to gold by most of the rest of the world in the 1890s. After World War I, countries could choose among gold, silver (rarely), the pound or the dollar. 2/ In 1936 a tripartite

1/ Commodity standards were subject to instability arising from two main sources: fluctuations in the supply of the commodity; and shifts of demand. Although changes in supplies of the commodity were disturbing enough, it was possible to anticipate them, in view of long production lags. More disturbing were sudden and dramatic shifts of demand due to countries going from one standard to another.

From the 1820s to the late 1840s, when bi-metallism in France and the United States gave the world a monetary unity, expanding silver supplies from Mexico were insufficient to make up for the shortfall in gold production. This position was reversed in mid-century when the gold discoveries in Australia and California were sufficient to double in a decade world monetary gold stocks; gold drove silver out of circulation in France and world prices began to rise. The United States currency became inconvertible during the Civil War, leaving France alone to bear the brunt of bi-metallism; but France suspended convertibility during the Franco-Prussian War. By that time silver production had rapidly increased, forcing France to safeguard her gold currency by abandoning bi-metallism for a limping gold standard; even earlier the new German Empire had dumped silver for gold, further lowering the price of silver.

The 1870s witnessed an explosion of silver production, leading to the abandonment of bi-metallism and the march to the gold standard. This shift of standards created an excess demand for gold and an excess supply of silver which led to deflation in the gold bloc and inflation in the remaining silver countries. Criticism at this time was directed at the shift away from bi-metallism, which aggravated deflation in the gold bloc and inflation in the remaining silver countries. Objections to deflation led to agitation to restore bi-metallism, which, however, failed when South African gold arrived to prevent further deflation, and, indeed, induce a mild inflation for almost two decades before the outbreak of World War I.

2/ Gold had become unstable after the outbreak of World War I when the belligerent countries engaged in inflationary finance, exporting gold to the few countries (including the United States and Japan) remaining on the gold standard. The commodity value of gold fell in half as US prices doubled. The immediate post-war deflation in the United States was insufficient to restore the pre-war gold prices. The stock of gold was sufficient to maintain the US on gold, but inadequate for an international gold standard of the pre-war type. Nevertheless, the international gold standard was restored in the midst of a state of monetary uncertainty occasioned by wide fluctuations in exchange rates. Despite feeble attempts at international monetary coordination, the gold scarcity led to deflationary policies that inaugurated the great deflation, made worse by mass unemployment. Belatedly, some countries left the gold standard or devalued, leading to a revised system.

arrangement was adopted among three gold standard countries. ^{1/} At the Bretton Woods conference in 1944, the choice was between gold and the dollar (except for a brief flurry of interest in silver). At the Smithsonian meeting in 1971, countries could choose among gold, the dollar or the SDR. When (or if) a European currency is created, the choice of international standards will be further widened.

The choice among monetary systems in recent history is elaborated below. The purpose is to investigate *systemic* problems that have a bearing on the relation between exchange rates and trade balances. Because of the major role that superpowers have played in the international monetary system, the solution of some their national problems lies in a change in the international monetary system.

1. The Gold Exchange Standard of the 1920s

A century of comparative stability under the bi-metallic or gold standards was shaken by World War I, when the belligerent countries--except the United States and Japan--left the gold standard. Keynes (1923) warned against restoring gold at the old parity, arguing that at current prices relative to the United States, the pound would be overvalued and that resumption would bring on tight money and deflation in Britain. More to the point, from the standpoint of the system as a whole, Cassel warned of an approaching deflation if the gold standard were restored at current price levels.

Britain did restore the pound at the historic parity, overvaluing it in terms of inflated post-war prices. This action has been universally condemned, not least by Winston Churchill himself. Many of the subsequent ills of the world have been blamed on this unfortunate event. But the modern economist must be skeptical that such an apparently minor event as a mere overvaluation of (according to Keynes) 10 percent--could have such catastrophic consequences on the world as a whole. Notwithstanding that Britain was still the center of an empire, it seems merely silly to observers in an age accustomed to wild gyrations of major currencies, to blame so many of the ills of the world on this little event.

^{1/} The new system relied heavily on the U.S. dollar after its devaluation in 1934. The devaluation more than corrected the gold shortage (given the prohibition on privately-owned gold in the United States), leading to an initially under-valued dollar (relative to gold). The dollar shortage, however, lasted only from 1934 to 1950. By 1950 wartime and post-war inflation had raised prices, creating a gold scarcity, concealed somewhat by the prohibition of gold for US citizens and the disproportionately-large U.S. gold holdings. Attracted by interest returns and confidence in the dollar, the rest of the world was initially content to use dollars in lieu of gold. But gold losses to European central banks, revealed the true nature of the global excess demand for gold.

The difficulties of sterling represented a symptom rather than a cause of the breakdown of the inter-war gold standard. Whatever the inconveniences of adjustment experienced by Britain because of the overvaluation of the pound against the dollar, they could have been managed if the system itself had been robust. Had the fundamentals of the international monetary system been sound, Britain's actions would have turned out to be right, gaining long-run advantages for the short-run trouble. The fetish with Britain's short-run dilemma diverted attention from the fundamental systemic problem of the system.

The fundamental problem of the inter-war gold standard was the undervaluation of gold. This had occurred in World War I when European countries left the gold standard, resulting in massive gold imports into the United States that were immediately monetized, doubling the dollar price level even though the United States continued to adhere to the gold standard. ^{1/} This overvaluation of the dollar against gold was only partly corrected by the post-war deflation, leaving gold undervalued by about 35 percent. In the 1920s, all currencies (including the franc) were overvalued against gold; the pound was overvalued relative to the dollar; and the dollar (after 1927) was overvalued relative to the franc.

As already noted, Gustav Cassel (1925) was one of the few economists who spotted the systemic problem in the early 1920s; he warned that restoration of the international gold standard would bring on deflation (appreciation of gold). Even if Britain had restored the pound at an equilibrium parity vis-à-vis the dollar, world deflation would have been resulted from the systemic problem of the undervaluation of gold.

Alternative policy steps could have prevented it. The descent into deflation could have been averted as late as 1930. The solutions, however, were outside the intellectual framework of the times. One possible solution was an increase in the price of gold in terms of all currencies. Another possibility would have been flexible exchange rates in 1930, before the onset of deflation instead of after it. A third solution would have been the creation of an alternative form of liquidity through a world central bank.

Despite some earlier indications of awareness of gold scarcity in the 1920s--*vide the Genoa Agreement of 1922*--the problem was not generally recognized as a systemic problem. None of the needed measures were taken,

^{1/} I have discussed this effect--what I have called the "Thornton effect"--in Mundell (1989, 1990). Hume and Smith had recognized that an increase in credit or paper notes would not cause prices to rise provided convertibility of the currency was maintained; gold or silver money would be exported to the same extent that it was replaced by soft money. Thornton, however, took account of the effects of the export of specie on the price level in the rest of the world, which would have a corresponding, if small, effect in raising the domestic price level.

not even discussed. The closest approximation to a solution--at least pointing in the right direction--was the creation of the Bank of International Settlements, an institution which, however, while creating a forum for discussion of international monetary matters, focussed most of its attention on reparations.

2. The Bretton Woods Problem

History repeated itself in a corollary mistake in the decades after World War II. From 1934 to 1945, gold appeared to be slightly overvalued due to (1) dollar devaluation; (2) the prohibition of gold for U.S. citizens; and (3) the international concentration of monetary gold in the United States. The United States had ample gold for the backing of Federal Reserve notes after the reserve requirement was lowered from 40 percent to 25 percent in 1945. The apparent gold redundancy, however, concealed an incipient scarcity in the international monetary system. There was sufficient gold for US purposes, but not enough to permit a general redistribution of US gold to Europe. Although European countries pegged currencies to the dollar, they nevertheless wanted to hold an important part of their reserves in gold.

The 1950s paradox of a strong dollar (measured against other currencies) coupled with a balance-of-payments deficit--as measured by gold losses and the increase in liquid liabilities to foreigners--can only be understood in the context of the new role for the dollar in the international monetary system. 1/ Concurrent with the massive redistribution of monetary gold to Europe, the dollar was becoming the world's money. The acquisition of dollar liabilities abroad, combined with gold losses, was interpreted as a deficit in the U.S. balance of payments. Even in the hey-day of what was called the dollar shortage, the United States developed a substantial balance-of-payments deficit. 2/

Triffin posed the problem of the international monetary system as a dilemma: If the United States corrected its balance-of-payments deficit, the world would suffer a liquidity shortage and potential deflation; but if the

1/ One of the problems associated with the confusion between systemic and national issues is the concept of the balance of payments deficit, conventionally defined as a loss of gold or increase in liquid liabilities. This definition is inappropriate if deficits and surpluses are thought of as error signals because it makes no distinction between desired and undesired changes in reserves. In Mundell (1965) I advanced a definition of the balance of payments restricted to *undesired* changes in reserve assets or foreign liquid liabilities; I still believe this is the appropriate concept, despite its operational difficulties.

2/ A dollar shortage is not, however, incompatible with a gold shortage. In the I.M.F., the dollar was the currency that was "needed to be drawn" even while countries with balance of payments surpluses were exchanging dollars for gold.

United States failed to eliminate its deficit, gold losses would bring on a crisis of confidence in the dollar. Triffin's solution to the liquidity problem was to create a new source of liquidity through a world central bank reminiscent of the plan proposed by Keynes two decades earlier. 1/

The Triffin plan was an idea whose time had not yet come. The political ingredients of a solution along these lines did not exist. There was no political constituency for an international central bank. 2/

An alternative to the Triffin solution was an increase in the price of gold, as recommended by Sir Roy Harrod and Jacques Rueff. 3/ But, as already noted, an increase in the price of gold could eliminate the undervaluation of gold but not the US balance of payments deficit. With an increase in the price of gold high enough to restore confidence in the dollar, the rest of the world would not only continue to accumulate dollars as reserve assets, aggravating the US balance of payments problem, but might even cash in their gold hoards for interest-bearing dollar assets. An increase in the price of gold without a reform of the system might have made the US balance-of-payments deficit permanent--that is, until the next crisis a couple of decades down the road. 4/

It is necessary, however, to distinguish between Harrod's and Rueff's proposals for increasing the price of gold. Whereas Harrod--always an advocate of elastic money--would condone the continued use of reserve currencies, Rueff advocated abolition of the gold exchange standard in favor of a full gold standard. Rueff's proposal thus took full account of the need to resolve the liquidity, adjustment and confidence problems, albeit by means that would impose the discipline of the automatic monetary adjustments of the gold standard. The necessary doubling of the price of gold, Rueff

1/ Keynes had earlier suggested what he called an "international clearing union" in 1943. I presented a plan for a world currency in testimony to the Joint Economic Committee in Mundell (1968).

2/ Certainly the Viet-Nam War did much to sour the trans-Atlantic climate. More important, however, the dominant power--always the country with, in the short run, the most to lose by powerful supranationalism--was not yet willing to demote itself.

3/ There was a provision in the Articles of Agreement of the International Monetary Fund for a change in the par values (specified in gold) of all currencies, so the founding fathers of the IMF anticipated the possibility. Note that this proposal would not necessarily involve any changes in exchange rates.

4/ Other (often mutually contradictory) objections were that an increase in the price of gold would (1) be inflationary; (2) create expectations of future increases and lead to a gold shortage in the future; (3) tend to reinstate the gold standard; (4) unfairly redistribute wealth toward gold-holding countries; (5) penalize countries that had accumulated dollars rather than gold; (6) lower the gold value of existing contracts; (7) help South Africa and/or the Soviet Union.

argued, would not be inflationary because the additional dollar value of gold liquidity would be used by the United States to extinguish reserve liabilities of the United States.

Many economists, arguing along traditional lines, believed that devaluation of the dollar was the correct remedy for the US balance of payments deficit. But this approach failed to distinguish between national and systemic problems. It failed to recognize the unique position of the US dollar in the world payments system. The economists who recommended devaluation of the dollar (as opposed to a general increase in the price of gold) were thinking in terms of devaluation against other currencies. 1/ Apart from the fact that it would not be easy to bring this about under the existing exchange system, it was not the right remedy for the problem. 2/

There was no indication that exchange rate changes would make US products more competitive or improve the US balance of trade. There had been, it is true, a gradual decline in the US share of world exports since the early post-war recovery period; European and Japanese goods were becoming much more important as their economies recovered from the devastations of war. The United States had been worried about its growth rate in the post-Sputnik years, but slow growth was not an exchange rate problem. Throughout the post-war period, the United States had both current account and trade balance surpluses. (The United States had a trade surplus in each of the six decades from 1914 to 1974!) No systematic empirical study in economics had ever demonstrated convincingly that devaluation would improve the trade balance. The gap between theory and practice was even more apparent when devaluation of a key currency that constituted a major reserve asset was in question. The basic problem of the 1950s and 1960s was systemic not national.

Devaluation of the dollar would not have been a solution to the problems of the Bretton Woods system. Only a substantial increase in the

1/ The difficulties attendant upon official devaluation of the dollar became clear in the debates at the Smithsonian meeting when a political compromise was reached at which the United States devalued the dollar against gold and some other countries revalued their currencies against gold.

2/ The devaluation of the dollar against other currencies could have been effected in either of two ways: (1) An increase in the dollar price of gold, the par values expressed in gold of other currencies remaining constant; (2) The raising of the gold values of other currencies, the par value of the dollar remaining constant at .888671 grams = 1/35 of an ounce. Either method would have required other countries to raise the price at which they bought dollars in the exchange markets. These difficulties were by no means insurmountable--they were resolved when exchange rates were changed at the 1971 Smithsonian meeting--but they indicated the great complications associated with the asymmetry of the dollar's position in the Bretton Woods system.

price of gold would have eliminated its undervaluation. On the other hand, an adequate increase in the price of gold would have restored confidence in the dollar and, in the absence of a change in the system, resulted in a reflux of gold to the United States at the higher price. The international monetary system based on the new price of gold would have lasted only until the ensuing great expansion of liquidity and inflation once again threatened the dollar. 1/

The dark horse of the Bretton Woods solutions was a proposal for a system of flexible exchange rates. That proposal, initially advanced after the Napoleonic Wars, and advocated in modified form by Fisher and Keynes, had been supported by Frank Graham and others in the 1940s. It received extensive attention after James Meade and Milton Friedman advocated it for the world economy in the 1950s. It was one of the four proposals discussed in the G-32 meetings 2/ in 1964. In 1966 sixteen prominent economists endorsed a proposal for a system of flexible exchange rates.

This solution was expected to give each country monetary independence. Fixing the national rate of monetary expansion would give each country the rate of inflation it desired and international differences in inflation rates would be managed by changes in exchange rates in the free market. The liquidity problem would be solved under flexible exchange rates by dispensing with the need for international reserves; the adjustment problem would be solved by using exchange rate changes to offset rigidities in money wage rates; and the confidence problem would be resolved by a free market in international reserve assets. 3/

Another proposal, stopping short of a world central bank, was to create a new international reserve asset to supplement gold. Several plans were made along these lines, including those by Sir Maxwell Stamp,

1/ As always, systems have to be compared against the alternatives. Despite the arguments against such an arrangement, it might have been better than the alternatives, including that actually adopted. There are worse systems than one which required a change in the price of the international monetary asset every generation or so.

2/ See Machlup and Malkiel (1964). The other three proposals coming in for detailed discussion were the gold standard, a world central bank and system of currency areas or large floating blocs.

3/ The facts proved otherwise. National demands for international reserves would not abate with flexible exchange rates. On the contrary, as Harrod pointed out even in the 1960s, countries would need more rather than less liquidity under flexible exchange rates because of the increase in uncertainty. Moreover, as I argued, a flexible-exchange-rate world without an official standard would naturally use the dollar as the most important reserve asset, creating a normal current account deficit equal to the secular demand for dollar reserve assets.

Edward M. Bernstein and Valery Giscard d'Estaing. But there was both doubt about what the right solution was and an inability to negotiate one if it had been found. A solution along these lines would have required a sufficiently large splash of paper gold needed to eliminate its undervaluation. The amount of paper gold required would depend on whether or not paper gold would be used to replace dollar reserves. If so, the creation of perhaps \$35 billion of paper gold (SDRs) would have been necessary. If dollars were to remain in reserves--and some dollars would have been required for working balances--perhaps a once-for-all increase of SDRs 10 billion would have been sufficient, followed by a continued growth along the lines that were actually enacted (about SDR 3 billion per year).

A combination of measures would have yielded a more balanced solution with a better chance of success. The dollar could not, in practice, have been completely replaced by either gold or the SDRs. But limits could have been placed at the rate at which dollar holdings were increasing. Disincentives or penalties could, at least theoretically, be imposed on countries whose deficits or surpluses exceeded pre-determined limits. ^{1/} Combined with both a modest increase in the price of gold and the annual creation of paper-gold SDRs, a restricted form of dollar-gold standard could have preserved the essentials of the Bretton Woods system. The solution actually adopted--the SDR--did nothing to prevent the demise of the Bretton Woods system.

3. The Collapse of the Bretton Woods System ^{2/}

The gold scarcity, misread as a balance-of-payments deficit of the United States, generated expectations that the dollar price of gold might be raised; gold was naturally preferred to dollars. The U.S. Treasury sold hundreds of millions of ounces of gold before finally restricting further dollar conversions in the middle 1960s. Even though the United States was formally committed to buying and selling gold freely under a 1949 agreement between the Fund and the United States (which absolved the U.S. from intervening in the exchange market to preserve exchange rates), other countries

^{1/} Penalties in the form of asset conversions would have to be negotiated in advance, rather than imposed. At the Copenhagen meetings of the IMF in 1970, IMF Managing Director Pierre-Paul Schweitzer proposed that the major reserve country accept some gold losses in view of its balance-of-payments deficit worsened, a proposal which, at the time, was resented by the U.S. Treasury.

^{2/} The collapse of Bretton Woods was stretched out over three episodes: 1968, when the market price of gold was severed from the official price and the members of the I.M.F. withdrew from the private gold market; 1971, when the dollar became inconvertible; and 1973, when flexible exchange rates was adopted.

accepted dollars without converting them into gold, rather than risk U.S. action to dismantle the system. The gold pool, formed in 1961 to allocate scarce gold 1/ in the private market, was abandoned when, in 1968, market demand in the private market overtook supply.

With the famous Communique of March 1968, the market price of gold was detached from the official price, creating what came to be called a "two-tier" system. 2/ Gold became immobilized: no country wanted to sell gold for dollars when the official price was much below the market price. Countries now needed dollars to fill the vacuum created by the immobilization of gold. After an initial stint of tight money in 1969, dollars became plentiful when the recession of 1969-71 unfolded. A so-called "overhang" of dollars developed in Europe 3/ bringing on the mark crisis in the spring and the dollar crisis in the summer of 1971. After the Joint Economic Committee of the US Congress explicitly recommended devaluation of the dollar, foreign central banks requested gold conversions, and the United States responded by closing the gold window.

Gold now became officially, as it had been already *de facto*, inconvertible after August 15, 1971. At the Smithsonian meeting in December of that year, the dollar was devalued, raising the price of an ounce of gold to \$38. Because the price of gold was much higher in the market, no transactions would take place at that price. President Richard Nixon called the Smithsonian agreement the most important monetary agreement in the history of the world. Ironically, it was the first general

1/ The gold pool, organized at the initiative of the United States, with the central banks of Belgium, France, Italy, the Netherlands, Switzerland, West Germany, and the United Kingdom. The pool was actually a gentleman's agreement to divide the burden of stabilizing the price of gold in the London gold market, with the United States having a 50 per cent share. The Bank of England acted as agent for the pool in its market operations.

2/ The communique of March 17, 1968, stated that the seven members of the gold pool (France had dropped out in 1967) "decided no longer to supply gold to the London gold market or any other gold market" and further asserted that in view of the prospective establishment of the Special Drawing Rights, the existing stock of monetary gold is sufficient.

3/ The accumulation of dollars by the rest of the world was partly desired to compensate for the immobility of gold; but it is probable that US monetary policy was more expansive over the period 1965-73 than the countries on the European continent desired. The other countries temporarily had an excess supply of dollars which, however, disappeared with the rising prices level in the early 1970s.

monetary agreement establishing an international monetary system in which no currency would be convertible into one of the monetary metals. 1/

The economic theory that treated the US balance of payments deficit as a national rather than a systemic problem prevailed at the Smithsonian meeting. 'Plausible' elasticities based on back-of-the-envelope intuition were substituted into the "Marshall-Lerner" condition to suggest an exchange rate change of about 10-15 percent against other currencies, accomplished by a rise in the price of gold in the United States and a fall in the price of gold in other countries. 2/

The full implications of the Smithsonian agreement, in establishing a fixed exchange rate system without asset convertibility, were by no means fully realized at the time. With a continuing US balance-of-payments deficit--wanted or unwanted--other countries had to accumulate dollars with no prospect for asset settlement. Confidence in the dollar, previously the only currency with some pretence to gold convertibility, was shaken. More important, the link to gold--fictitious though it may have been in terms of operational convertibility--had been an important background factor maintaining confidence in money and the international monetary system. The breakdown of the system undermined confidence in money and abetted inflationary expectations.

Every country in the world could now inflate--provided they did it at the same rate--without experiencing balance of payments deficits. The absence of an anchor--asset convertibility--undermined the effectiveness of balance-of-payments discipline and led to monetary policies that accommodated inflationary wage settlements and price increases. No sooner was the ink dry on the Smithsonian papers devaluing the dollar than the OPEC

1/ Gold was still, however, legally used as a numeraire. A few days before the Smithsonian agreement of December 17-18, 1971, Presidents Nixon and Pompidou met at the Azores and agreed that the United States would devalue in terms of gold. The dollar was devalued by 7.89 per cent, raising the price of gold to \$38 an ounce. The DM was appreciated by 4.61 per cent against gold (13.58 per cent of the old dollar parity); the yen was raised 7.66 per cent against gold (16.88 per cent against the old dollar parity). The other Group of Ten countries, with the exception of Canada, whose currency was left floating, revalued against the dollar, but some devalued against gold.

2/ Exchange rate changes could not, as we have argued, remedy the systemic problem of the undervaluation of gold against all currencies. The Smithsonian agreement failed to correct the basic problem of the undervaluation of gold and in 1973 a second dollar crisis emerged with another equally futile, devaluation of the dollar. Only a few months after this devaluation (June 1973), the system broke down into flexible exchange rates, removing what little monetary discipline remained. The subsequent increase in oil prices was quickly ratified by inflationary finance and the explosion of credit in the Eurodollar market.

countries raised the dollar price 1/ of oil, a foretaste of more important events two years later.

The Smithsonian agreement was an exercise in futility, enacted by politicians with at best a shallow knowledge of its full implications. It set in motion a train of events that would lead the world into the greatest monetary inflation in its history. Far from the tiny devaluation improving the US balance of payments, it worsened it. Hardly more than a year later, the dollar was confronted with another crisis of confidence that led, in February 1973, to an equally futile second devaluation of the dollar. With the official dollar price of an ounce of gold put at \$42.22, still far below the market price, confidence in the dollar did not recover. The United States and her trading partners were still using national weapons to deal with a systemic problem.

The failure of the second devaluation of the dollar to resolve the US balance-of-payments problems stirred up interest in a joint European float. The attempt at a joint European float in the spring of 1973 failed, however, because Britain would not join the float unless the surplus continental countries agreed to unlimited support of sterling. Failure to organize a joint European float in the face of persisting US deficits in the spring of 1973 led to the abandonment of the fixed exchange rate system and the advent of flexible exchange rates.

4. Flexible Exchange Rates

As already noted, flexible exchange rate had been more popular among academics than among officials. 2/ In the early 1970s, however, it received the patronage of George Schultz, US Secretary of the Treasury. 3/ After the Committee of Twenty failed to find a solution through international monetary reform, it approved a regime of flexible exchange rates, leaving responsibility for containing inflation to the individual countries.

1/ The connection between the two events was brought home to me in January 1972, before the Smithsonian agreement had been fully ratified. Participants at two independent meetings--one of OPEC ministers and the other of the Bellagio group of academics and officials--were lodged at the Intercontinental Hotel in Geneva debating simultaneously the same Smithsonian agreement from entirely different directions. The OPEC group then and there decided to raise the dollar price of oil in reaction to the increase in the dollar price of gold.

2/ It had been, like devaluation, an "unmentionable" at US Treasury Consultants meetings in the 1960s.

3/ In the 1960s, George Schultz, as Dean of the Graduate School of Business overlapped with Milton Friedman as Professor of Economics at the University of Chicago.

Flexible exchange rates represented the antithesis of an international monetary system. Its advocates had argued that the exchange rate was a price "like any other price" and that any attempt to fix it would lead to disequilibrium and the need for controls. If, instead of the monetary coordination that is necessary in a currency area, countries fix monetary targets and let the exchange rate float, the balance of payments will be maintained in equilibrium. If the rate of monetary expansion is fixed at about the rate of growth (after due allowances for secular changes in the demand for money) monetary stability, price stability and balance of payments equilibrium will be simultaneously achieved. It was believed that surpluses would result in appreciation, deficits in depreciation of a currency, until equilibrium was restored. Speculative capital movements would tend to be stabilizing. Its advocates argued that balance of payments problems would be solved by floating exchange rates.

Events did not turn out that way. There was a large gap between the theory and the practice of flexible exchange rates. At the end of 1973 (the year floating began) international reserves were \$102.9 billion. By the end of 1989 foreign exchange reserves were \$542 billion, an increase of 426.7 percent. Over the same period, imports rose from \$542 billion to \$2,974.6 billion, an increase of 443.2 percent. In other words international foreign exchange reserves rose about the same as world imports. The ratio of reserves to imports was 18.8 percent in 1973 and 18.22 percent in 1989, about the same. World reserves were just as high under flexible exchange rates as under the system of fixed exchange rates.

If the increase in reserves had been evenly spread among countries around the world, there would be no implications for current account balances. But this is not the case. The bulk of the increase in reserves represented an increase in liabilities of the United States, financing a corresponding deficit in the US balance of payments. ^{1/} The prediction that flexible exchange rates would eliminate balance of payments problems has not been fulfilled. Most countries have found it necessary to intervene in foreign exchange markets and accumulate huge levels of reserves.

Important relevant features of the real world were left out of the models relating exchange rates to the balance of payments. International trade theory had developed largely on the assumption either of a symmetrical gold standard, a system in which national money is only used inside the nation; or currency areas where satellite countries hold the mother country's currency. Little attention had been paid, up to the 1960s, to the

^{1/} The gross US deficit, measured by the increase in liquid liabilities of the United States to foreign countries, has amounted to over a trillion dollars since 1973; total liabilities rose from over the period from \$92.5 billion to \$1,101.7 billion. Over the same period, liquid external claims rose from \$26.6 to \$658 billion. On an official settlements basis alone, US liquid liabilities to foreign central banks and governments has increased from \$66 billion to over \$300 billion.

explicit introduction into theoretical models of massive movements of portfolio capital. Exchange rate theory was largely based on long run real models with emphasis on the balance of trade and the elasticities of demand for imports and exports.

The elasticity approach to adjustment under flexible exchange rates was based mainly on a theory that did not take account of capital movements. 1/ In the absence of capital movements, the theories of the balance of trade and the balance of payments are essentially equivalent. If devaluation improves the balance of payments it must do so by improving the balance of trade. Now it is easily shown in monetary models that devaluation tends to improve the balance of payments (e.g., through real balance effects); it would therefore improve the balance of trade if there were no capital movements. However, the conclusion does not follow if there is capital mobility. All that can be proved by the monetary theory of the exchange rate when there is capital mobility is the tendency for devaluation to prove the *sum* of the trade balance and the capital account.

If devaluation induces a capital inflow in excess of the improvement in the balance of payments, it will also induce a worsening of the current account of the country of the depreciating currency. To the extent that exchange rate depreciation is not offset by changes in wage rates, the marginal efficiency of capital rises and the resulting capital inflow increases absorption and worsens the trade balance. Depreciation can also temporarily underprice capital assets and land relative to the products produced from them, another factor raising the marginal efficiency of capital, attracting capital imports and worsening the trade balance. Instead of increasing the export of ordinary goods and services, depreciation can lead to the sale of financial and real assets, a worse trade balance and an increase in net international indebtedness.

Market exchange rates have been dominated, in the short run, by capital transactions rather than trade. Global imports and exports amounted to \$5.8 trillion in 1989. But transactions in the foreign exchange market amount to thirty times that amount! 2/ Instead of reflecting trade accounts and inflation differentials, the exchange rate is dominated by sustained swings based on speculative capital market transactions. This has led to gyrations of exchange rates that for substantial periods of time result in sustained variations from purchasing power parity. From DM 3.2 at the end of 1972, the dollar fell to DM 1.73 at the end of 1979, then rose to DM 3.15 at the end of 1984, only to fall to DM 1.58 at the end of 1987, and then rise to DM 1.78 the following year, and then fall to DM 1.5 in October 1990.

1/ Although Bickerdike pays close attention to capital movements, his elasticity formula is restricted only to the trade account.

2/ Estimates of foreign exchange market put it over \$600 billion per working day, which implies total yearly transactions over \$150 trillion.

These swings, which were of course considerably greater when measured from peak to trough, were equally large against other major currencies.

The demand for dollar reserves continues steadily, both on the part of central banks and governments and on the part of private financial institutions. In every year from 1960 to 1990 liquid dollar liabilities has increased. This would in itself be sufficient to create a trade deficit in the United States if it were not offset by income transfers or capital movements. Neither devaluation nor the fall of the dollar under flexible exchange rates in the 1970s eliminated the US balance of payments problem. The US dollar was still the major world currency asset for which neither the SDR or any other currency were adequate substitutes. While the United States continues in that role, it will be saddled with a balance of payments deficit whatever the exchange system.

Flexible exchange rates can, it is true, eliminate excess demand in the foreign exchange market; and zero excess demand in the foreign exchange market is an alternative definition of the balance of payments. But rather than solving the basic problems of equilibrium for the United States in its national role and in the international monetary system, it transformed the problem into another form. Instead of worrying about the level of reserves government officials worried about the path of the exchange rate; instead of concern over the balance of payments, they worried about the balance of trade or the current account balance.

Flexible exchange rates have probably produced a less efficient international trading system, in view of the wide departures from purchasing power parity and the great gyrations of the nominal and real exchange rates. It has also resulted in a breakdown of the monetary discipline that prevailed under the gold exchange standard and even under the Bretton Woods System. It has not, however, resulted in the "chaos" that some of its critics had predicted.

5. Currency Areas

One reason why the movement to flexible exchange rates did not result in chaos is that countries protected themselves by joining currency areas. With the adoption of flexible exchange rates, it soon became apparent that it was not practical for all countries to adopt floating exchange rates. Currencies differ considerably from one another in liquidity, marketability, stability and market power. Market power is an increasing function of the size of the transactions domain of a currency, relative to typical market shocks. Because transactions domains are related to national income, the hierarchy of currencies reflects economic size. Other things equal--especially policy credibility--the currencies with the greatest monetary

properties are the largest economies. 1/ The countries in the economic hinterland of large and stable countries can improve the monetary properties of their own currencies by attaching them at fixed exchange rates to the hegemon. In turn additional currencies attached to the lead currency reinforce the monetary properties of the dominant country. 2/

Another important factor is the monetary will of a country, and the record and reputation of its central bank. As Otto Poehl (1987) has 3/ said, "...credibility is the capital stock of any central bank." Over the centuries, of course, Britain has had the longest record of a stable currency without a currency conversion; the reputation of the pound has been damaged by inflationary policies and depreciation in the post-war period, however, and the transactions area of the pound has declined below 5 percent of world transactions.

The dollar has stood out among currencies since 1914, both with respect to stability and transactions area. In recent decades, however, Germany, Japan and Switzerland have established reputations for dynamic economies, stable monetary policies and strong currencies, with the consequent benefit of relatively low interest rates on currency-denominated financial assets. Taking both transactions area and reputation into account, three monetary leaders emerge in the world community: the United States, Germany and Japan. In 1975, almost 80 percent of official holdings of foreign exchange was in dollars, with less than 7 percent in marks and less than 4 percent in pounds sterling. By 1988, however, dollars accounted for only 63.3 percent, while the mark share had risen to 16 percent and the yen share had risen to over 7 percent; the pound sterling was a not-so-distant fourth with 3.1 percent. The possibility of large currency areas centered on the dollar, mark and yen

1/ The comparative monetary properties of a currency are related to the slopes of the liquidity preference schedules in each country, the absolute value of which is normally proportionate to size. See Mundell (1974).

2/ The cumulative improvement in the monetary properties of a currency area as it expands leads to the theoretical proposition that the optimum number of currencies in the world is one. Flexible exchange rates has usually been presented, by its major proponents, as a second best arrangement, as an adaptation necessary in the real world to overcome nominal wage rigidities arising from the existence of labor unions. There is, however, no reason to expect the zones of labor bargaining power to overlap with the domain of currencies.

3/ Quoted in Frenkel, Goldstein and Masson (1989, p. 194).

emerged as a more realistic alternative to a universal system of flexible exchange rates. 1/

The planned movement to flexible exchange rates did not turn out that way. The international system did not become a truly flexible exchange rate system. There are at least three reasons why this is so. In the first place few countries abandoned their foreign exchange markets to the winds of speculation that would have blown freely under a clean float. As we have seen countries built up foreign exchange reserves to offset private fluctuations judged to be against the interests of the community.

Second, and more important, most countries formed themselves into one of a few great currency areas. The International Monetary Fund makes monthly lists of currency areas or exchange rate arrangements, among which the most important are, of course, the dollar area and the ERM (exchange rate mechanism) of the Economic Community. Smaller countries have found it desirable to stabilize their currency into optimum or feasible (Corden) currency areas. Currency areas give members a chance to coordinate monetary policies and share the same inflation rate at the expense of ceding their monetary independence to a group solution. From an abstract point of view, currency areas are second-best in relation to a world system, but they may be the best that can be achieved in the absence of a coordinated currency area for the whole world. Currency areas are stimulated by the gains involved in exploiting, by internalizing, an externality: the larger the currency area, the less vulnerable a single currency will be to outside shocks. 2/

1/ See Tavlas (1990). The rise of the mark is accounted for by the growth in activity of the EMS and the use of European currencies for intervention purposes. Tavlas supplies, among others, the following figures: Dollars accounted for 71.5 per cent of EMS intervention in 1979-82, but only 26.3 per cent in 1986-87; whereas the mark accounted for only 23.7 per cent in the early period but 59 per cent in the later period. As far as US intervention is concerned, almost 90 per cent was in marks in 1979-82, with virtually the rest in yen; but in 1986-87, the mark share fell to 57.5 per cent and the yen share rose to 42.5 per cent.

2/ There is nothing automatic about the advantages of a large currency area to a single country. It depends, among other things, on whether the prospective partners have a more stable monetary policy than the members; or more exactly whether the new currency area will have a more stable monetary policy than individual country. The ERM system offers a good example where some currencies, previously more inflationary than the DM, have been able to use the system as a political excuse for a more stable monetary policy. Germany may have been induced to follow a slightly more inflationary monetary policy as a result of the ERM, but offsetting this cost is the beneficial effects of European monetary leadership and a more important currency.

Third, and perhaps most important, the existence of the dominant super-power currency has imparted a natural stability to the currency organization of the world economy. Despite the weakness of the dollar at different times, the huge deficits and the build-up of a huge international debtor position, the dollar remains at the center of the system. That position--especially now that the United Kingdom has opted for ERM--could be dislodged when or if the European countries form a continental currency, a possibility neither remote nor highly probable. But until, or unless, that event comes to pass the dollar will play the leading role in the foreign exchange markets.

The history of the international monetary system from 1973 to the present has not been dominated by a system of flexible exchange rates; its central theme has been the dollar against the mark (and its currency associates) and the yen, by far the most important exchange rates in the world. The next most important exchange rate has been the dollar price of gold, frequently referred to, in lieu of a better measure, as an index of inflation in the dollar area.

The era of flexible exchange rates came to an end in 1985 when the United States and other G-5 members decided to coordinate exchange rates and allow, or induce, a depreciation of the dollar against the major currencies. This shift to coordinated rates at the G-7 Louvre meetings reinforced the shift away from flexible exchange rates toward a system in which exchange rates were coordinated according to the goals of G-7 policy. It is interesting that, twenty years after the breakdown of the Bretton Woods system, the concern at G-7 meetings is still the relation between the dollar and the other currencies, with little attention paid to the other cross rates. In this respect the G-7 discussions have turned into dollar-management sessions.

6. Instability of the Real Exchange Rate

Under the comparatively fixed exchange rates of the Bretton Woods system, real exchange rate fluctuations were minor. Under the flexible rate system, *real* exchange rates fluctuated a great deal. Substantial evidence suggests that changes in nominal exchange rates induced substantial changes in real exchange rates. To the extent that nominal exchange rates fluctuated excessively, instability was transmitted to real exchange rates.

The dollar depreciated too much in the late 1970s, falling from DM 3.2 in 1972 to DM 1.73 at the end of 1979, rising to DM 3.4 in early 1985, then falling to DM 1.58 at the end of 1987, only to rise to DM 1.9 in 1989 and fall recently below DM 1.5. Overshooting is followed by undershooting with disequilibrium prices and efficiency losses in between. Taking 1985 = 100, the US real exchange rate, as measured by the IMF, was 79 in 1976, 68.1 in 1979, 100 in 1985, 64.1 in 1988, rising again to 67.8 in 1989 (III), falling to 62.3 in July 1990. These fluctuations in competitiveness have had the effect of alternating profitability between domestic and international

industries due to the whims of international capital movements that have little or nothing to do with international trade or even the balance of trade. 1/

Three observations can be made about the depreciation of the dollar, its effect on the real exchange rate and the change in the balance of trade 2/ since 1985: (1) the real measured exchange rate of the dollar has closely tracked the nominal exchange rate; (2) the real exchange rates of other countries have not as closely followed their nominal rates; and (3) interest rates seem to have played a role in the direction of the discrepancy between real and nominal rates. Countries like Canada and the United Kingdom, with interest rates higher than the United States, seem to have worsened their competitive position, judged by the real exchange rate, relative to countries like Germany and Japan, with lower interest rates.

With the exchange rate dramatically falling in the late 1970s, then rising in the first half of the 1980s, and then falling in the second half of the 1980s, the trade balance seemed to have been governed by forces other than exchange rates. In the early 1980s it was tempting to associate the rising dollar with the worsening trade balance, and to expect a reversal of the exchange rate to correct the trade balance. That, however, has not occurred. Sharp-eyed economists claim to see an important reduction in the trade deficit vindicating, they hoped, the elasticity school. But a far more interesting observation is that the deficit has been remarkably stubborn.

Secular forces are at work. The trade balance has been in deficit since 1975. Three stylized facts stand out:

1/ The real exchange rate seems to have been an innocent bystander--a victim--of the speculation and erratic monetary policy. After the 1979 oil price increase, an easy monetary policy was adopted to try to prevent the economy from falling into a recession in the 1980 election year; the result was too much inflation and an undervalued dollar. The rise of the nominal and real exchange rate in the first Reagan term was motivated by five factors: (1) a corrective of preceding inflationary policies; (2) reflux into the dollar from foreign currencies as confidence was restored; (3) an increase in capital imports due to the effect of the tax cuts in raising the marginal efficiency of capital in the United States; and (4) overkill by an over-zealous Federal Reserve Chairman (Paul Volcker) as a reaction against the presumed effect of the Kemp-Roth tax cut on the budget deficit; and (5) an underestimation by the influential Friedman wing of the Republican party of the increase in demand for money at home (reduction in velocity) due to a reduction in the rate of expected inflation. There was little or no attention to the role of the real exchange rate in these deliberations and little thought to the harmful consequences of arbitrary fluctuations in the real exchange rate.

2/ The first two of these observations were made to me by Herbert Grubel.

(1) The trade balance deficit jumped up from a period of average balance in the period 1970-75 to a plateau ranging between \$25-\$33 billion, in the period 1977-82. As a percentage of GNP the balance was +.56 in 1975, -.53 in 1976, but -1.56 in 1977, -1.51 in 1978, -1.10 in 1979, -0.92 in 1980, and -1.15 in 1982.

(2) The trade balance deficit jumped up again after 1982. As a percentage of GNP it stood at 1.97 in 1983, 2.98 in 1984, 3.04 in 1985, 3.42 in 1986, 3.52 in 1987, 2.61 in 1988 and 2.16 in 1989. The figures for 1990 are of course not out, but the deficit is not likely to be below 2.5 percent of GNP.

(3) The increase in the deficit, as a percentage of GNP, from the period 1977-1982 to the period 1983-1989 was slightly larger than the increase in the deficit from the period 1970-76 to the period 1977-82. It rose by 1 percentage points in the period 1977-82 from the early 1970s, but by 1.8 percentage points in the period 1983-89 from the period 1977-82.

As already noted, the deficit has stubbornly resisted changes in both the nominal and real exchange rate. The fluctuations in the dollar did not occur as a result of conscious policy aimed at changing the real exchange rate. The turnaround in the dollar came in February 1985 as a consequence of an earlier reversal of monetary policy, aimed at the domestic economy. In the summer and fall of the election year 1984, the recovery was petering out and a recession was widely predicted. As growth in the economy fell, the Federal Reserve eased, first halting and then reversing the soaring dollar. The dollar had already fallen substantially by the time of the Plaza meeting in September 1985, from DM 3.4 in February to DM 2.5. The Plaza Accord, designed to pre-empt growing protectionist pressure in the Congress, did not involve any major changes in policy in the United States beyond the comparatively easy money stance of the Federal Reserve enacted earlier. Most of the burden of lowering the yen-dollar rate fell on Japanese monetary policy. In the subsequent year, the fall in oil prices probably had greater impact on the yen-dollar rate than any further changes in policies.

If there is to be correction of the US current account deficit, there must be a decline in surpluses, or an increase in deficits in the rest of the world. In a competition for reduced deficits or increased surpluses, which countries are likely candidates to give way by enough to match a substantial reduction in the US deficit?

It is not easy to find candidates for substantially increased deficits or reduced surpluses in the rest of the world. Countries must have assets to sell that other countries are willing to buy. From a welfare standpoint, it might be desirable for US deficits to be shifted to the LDCs. But the LDCs lack sufficient credit-worthy debt instruments that potential surplus countries are inclined to accumulate. Equity transfers have proved to have at best limited prospects.

There will probably be a continuing surplus of export capital from the Asian "tigers" with North America as the main target. An improvement in the US balance would probably require some reduction in the export surpluses of these countries.

Another possibility is for needy Eastern Europe countries to "take over" US deficits. But again, the Eastern European countries run into a similar position, with respect to credit-worthiness, as that of the LDCs. There is, of course, considerable scope for the surplus countries, and especially Germany, to shift surpluses to Eastern Europe, particularly in what was formerly East Germany. Some reduction in the German trade surplus can be expected on this account.

A correction of the US trade deficit therefore depends primarily on a reduction in Japanese and German surpluses. What are the prospects? The Japanese surpluses were \$56 b. in 1985, \$92.8 b. in 1986, \$96.4 b. in 1987, \$95 b. in 1988 and \$77.1 b. in 1989; because of the increase in the price of oil, it is probable that the Japanese surplus will be reduced in 1990. In Germany, the surpluses were \$28.5 b. in 1985, \$55.7 b. in 1986, \$69.9 b. in 1987, \$79.4 b. in 1988 and \$76.7 b. in 1989. Because of the expenses of absorbing East Germany, the German surplus, as already noted, will drop further in 1990 and the following years. There are therefore some grounds for optimism that a reduction in German and Japanese surpluses will be reduced, making room for a further reduction in the US current account deficit.

Two other factors will play an influential role. The budget deficit, influenced negatively by both an increase in government spending as a result of the Iraqi crisis and a decline in revenue due to sagging economic growth, will be a factor worsening the deficit. On the other hand, as the decade advances, the demographic factor will result in an increase in the net savings rate and a potential resumption of net capital exports sufficient to make overt policies designed to improve the trade balance unnecessary.

V. Conclusions

The current account of the balance of payments is the outcome of a complicated system of sectoral and market interdependences in the world economy; taking into account intersectoral, international and intertemporal budget constraints, there are sixteen different perspectives on it. The current account balance represents the rate at which a country is building up equity (debt if negative) in the world economy and should always be measured against the sacrifice or accumulation of resources that constitute its counterpart.

There is no scientific body of evidence--theoretical or empirical--that supports a consistent relation between the exchange rate and the current

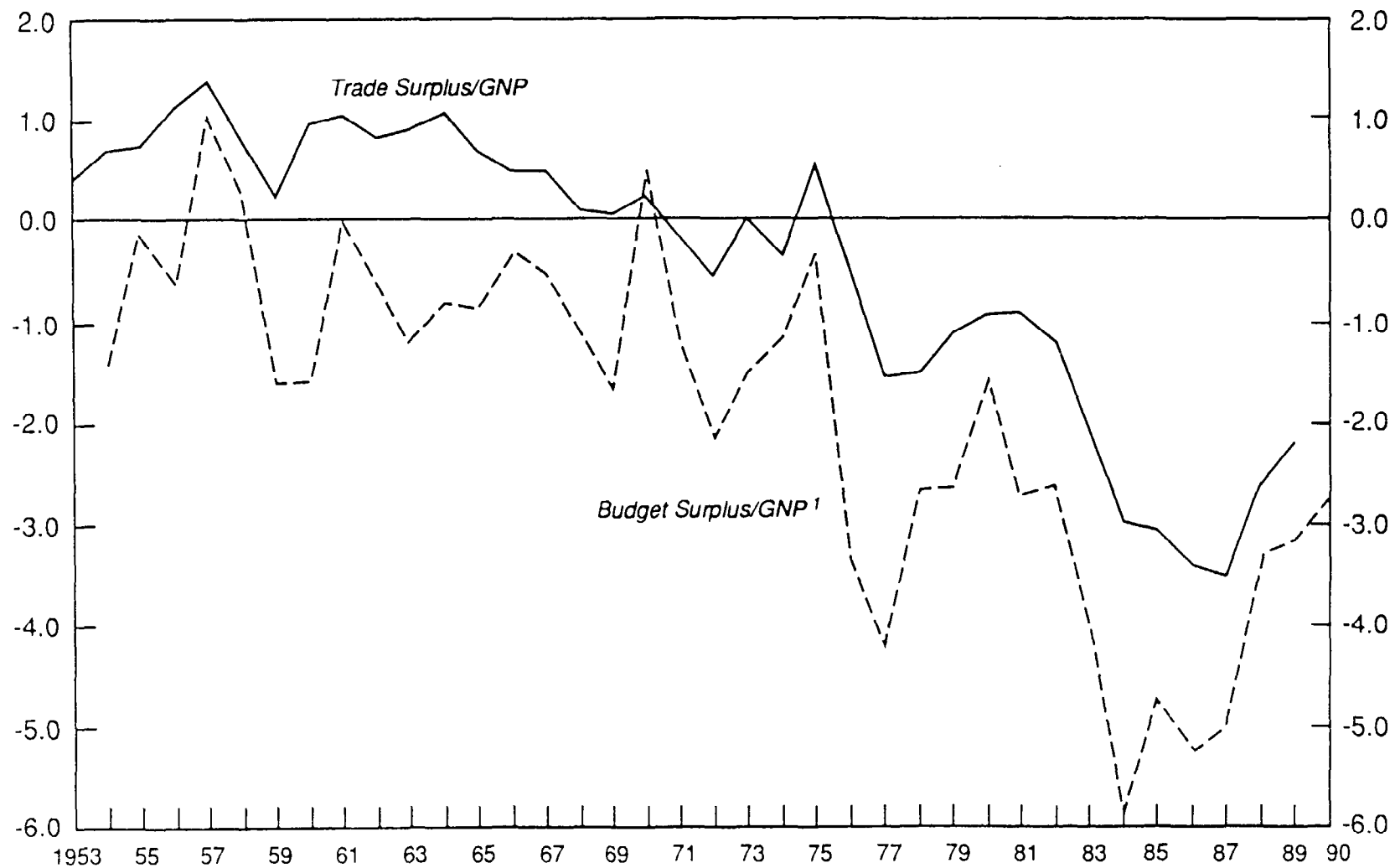
account. In the smaller countries, depreciation brought about by monetary expansion brings on inflation, not greater competitiveness; it is at best an inevitable accommodation to past inflation generated by monetized budget deficits. In the larger countries, sluggish wage reactions in the face of monetary acceleration may lead to greater short-run competitiveness, but the overdepreciation in the exchange market also increases profits that invites exports of capital assets rather than goods.

The large current account deficits in the United States that began in the late 1970s and soared in the 1980s have been misconceived as an exchange rate problem. Two factors have been responsible for both the trade and the budget deficits, one cyclical and the other systemic. The cyclical factor is the demographic shift that led to a collapse of the saving rate, starting in the late 1970s, continuing into the 1990s. These same demographic factors will be reversed in the 1990s, transforming the deficits of the 1980s into large surpluses in the late 1990s.

The systemic factor relates to the absence, in the international monetary system, of a world currency. The dollar has filled the vacuum. The world economy generates a steady demand for U.S. assets that finances a chronic current account deficit; a deficit would be imposed on the United States even in the absence of its low savings rate. The dependence on the dollar that was so marked in the heyday of the Bretton Woods system was aggravated, not eliminated, by the move to floating exchange rates in 1973. Long after the demographic savings problem has been reversed, the systemic problem will survive until a dollar substitute has been discovered.

CHART 1

U.S. Trade and Budget Surpluses as a Percentage of GNP, 1953-1990



Source: IMF, *International Financial Statistics Yearbook 1990*.

¹ Budget Surplus/GNP of previous year.

References

- Angell, James W. (1926), The Theory of International Prices.
- Bergsten, C. Fred (1975), The Dilemmas of the Dollar. New York: New York University Press.
- Bickerdike, C. F (1906), Theory of Incipient Taxes, Economic Journal (December).
- ____ (1907), Protective and Preferential Import Duties, Economic Journal (March).
- ____ (1920), The Instability of Foreign Exchange, Economic Journal (March), pp. 118-122.
- Bordo, Michael D. and Finn E. Kydland (1990), "The Gold Standard as a Rule," NBER Working Paper 3367 (May).
- Bordo, Michael D. and Eugene N. White, A Tale of Two Currencies: British and French Finance during the Napoleonic Wars (August). Manuscript.
- Brock, P., M. Connolly and C. Gonzalez-Vega (1989), Latin American Debt and Adjustment: External Shocks and Macroeconomic Policies, Praeger: New York.
- Cassel, Gustav (1925), The Restoration of Gold as a Universal Monetary Standard. In Young (1925), pp. 205-6.
- Clement, Simon (1695), A Discourse of the General Notions of Money, Trade, and Exchanges...
- Crowther, G. (1957), Balances and Imbalances of Payments, Boston.
- Dimand, Robert W. (1988), The Origins of the Keynesian Revolution. Stanford: Stanford University Press.
- Edgeworth, F. Y. (1908), Appreciations of Mathematical Theories, Economic Journal (December), pp. 541-56.
- Fischer, Stanley, and Jacob A. Frenkel (1974a), Economic Growth and Stages of the Balance of Payments. In Horwich and Samuelson, pp. 503-21.
- ____ (1974b) Interest Rate Equalization and Patterns of Production, Trade and Consumption in a Two-Country Growth Model. The Economic Record (December), pp. 555-79.

- Frenkel, Jacob A., Morris Goldstein, and Paul A. Masson (1989), International Dimensions of Monetary Policy: Coordination versus Autonomy. In Monetary Policy Issues in the 1990s. Federal Reserve Bank of Kansas, pp. 183-233. A Symposium Sponsored by Fed. Reserve Bank of Kansas City.
- Halevi, Nadav (1971), An Empirical Test of the "Balance of Payments Stages Hypothesis", Journal of International Economics 1: pp. 103-117.
- Horwich, George and Paul A. Samuelson (1974), Trade, Stability and Macroeconomics: Essays in Honor of Lloyd A. Metzler. New York and London: Academic Press.
- Host-Madsen, Paol (1962), "Assymetries Between Balance of Payments Surpluses and Deficits," IMF Staff Papers, Vol. 9, pp. 182-99.
- Kindleberger, Charles P. (1985), Keynesianism vs. Monetarism and Other Essays in Financial History, London: Allen & Unwin.
- Kojima, Kiyoshi (1990), "Proposal for a Multiple Key Currency Gold-Exchange Standard," Journal of Asian Economics, Vol. 1, No. 1, pp. 19-33.
- Marshall, Alfred (1925), Memorials of Alfred Marshall.
- Mundell, Robert A. (1968a), Man and Economics, New York: McGraw-Hill.
- _____ (1968b), International Economics, New York: MacMillan.
- _____ (1968c), "A Plan for a World Currency," In Next Steps in International Monetary Reform, Reuss Subcommittee (September 9), pp. 25-28.
- _____ (1974), "The Optimum Balance of Payments Deficit and the Monetary Theory of Empires," In Salin and Claassen, pp. 69-86.
- _____ (1988), "Latin American Debt and the Transfer Problem," In Brock, Connolly and Gonzagez-Vega.
- _____ (1989a), "Trade Balance Patterns as Global General Equilibrium: The Seventeenth Approach to the Balance of Payments," Rivista di Politica Economica (June).
- _____ (1989b), "The Global Adjustment System", Rivista di Politica Economica (December), pp. 351-466.
- _____ (1990a), "The International Distribution of Saving," Rivista di Politica Economica. Paper presented at the second Mondragon Conference in July 1990 (forthcoming).

- Mundell, Robert A. (1990b), "Fiscal Policy and the Theory of International Trade," forthcoming in Giersch (ed.), Proceedings of the Sohm Memorial Conference: Tegernsee, Bavaria.
- Mundell, Robert A. and Alexander Swoboda (1969), Monetary Problems of the International Economy, Chicago: University of Chicago Press.
- Mundell, Robert A. and John McCallum (eds.) (1990), Global Disequilibrium in the World Economy. Oxford University Press (forthcoming).
- Neher, Philip A. (1970), International Capital Movements and Balanced Growth Paths. Economic Record 46, pp. 393-401.
- Onitsuka, Yusuke (1970), International Capital Movements and Patterns of Economic Growth, unpublished Doctoral Dissertation, University of Chicago.
- ____ (1974), International Capital Movements and the Patterns of Economic Growth, American Economic Review 64 (1) (March), pp. 24-36.
- Patinkin, Don and J. Clark Leith (eds.) (1978), Keynes, Cambridge and the General Theory, Toronto: University of Toronto Press.
- Poehl, Karl Otto (1987), Cooperation--A Keystone for the Stability of the International Monetary System. First Arthur Burns Memorial Lecture at the American Council on Germany, New York (November 2).
- Robertson, Dennis H. (1938), Professor Viner Again, Economic Journal.
- Salin, P. and E. Claassen (1974), Stabilization Policies in Independent Economies, Amsterdam: North Holland.
- Schumpeter, Joseph A. (1954), History of Economic Analysis, New York: Oxford University Press.
- Swoboda, Alexander (1989), Global Disequilibrium and International Monetary Reform. Rivista di Politica Economica (December). Reprinted in Mundell and McCallum (1990).
- Taussig, Frank W. (1928), International Trade, New York.
- George S. Tavlas (1990), "On the International Use of Currencies: The Case of the Deutsche Mark," IMF Working Paper, 90/3. (Washington, D.C., International Monetary Fund (January)). Forthcoming revised version in Essays in International Finance, Princeton University (January 1991). See also an unpublished paper by Tavlas and Yusuru Ozeki (December 1990). The Japanese Yen as an International Currency.

Jacob Viner (1937), Studies in the Theory of International Trade, London.

_____ (1938), Reply to Robertson: Rejoinder, Economic Journal.

Young, John P. (1925), European Currency and Finance, Government Printing Office, Washington, D.C.