Volume Title: A Retrospective on the Classical Gold Standard, 1821-1931

Volume Author/Editor: Michael D. Bordo and Anna J. Schwartz, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-06590-1

Volume URL: http://www.nber.org/books/bord84-1

Publication Date: 1984

Chapter Title: The Success of Purchasing-Power Parity: Historical Evidence and Its Implications for Macroeconomics

Chapter Author: Donald N. McCloskey, J. Richard Zecher

Chapter URL: http://www.nber.org/chapters/c11127

Chapter pages in book: (p. 121 - 172)

2 The Success of Purchasing Power Parity: Historical Evidence and Its Implications for Macroeconomics

Donald N. McCloskey and J. Richard Zecher

2.1 Two Views of the International Economy and Why They Matter

A model of the economic life of nations that emphasizes the mutual interdependence of the nations is easy to believe in a period of relative tranquility, such as the heyday of the gold standard, 1880–1914, or of the gold-dollar standard, 1945–71. In less tranquil periods, such as the thirty years of war and depression from 1914 to 1945 or the decade just past, one might suppose that the history is less favorable to the model. We propose to show that this supposition is misleading, and that interdependence was strong. The strength of interdependence depends on the strength of purchasing-power parity. And purchasing-power parity is stronger than it looks.

Purchasing-power parity has recently been much in the scholarly news. Some of the new interest in an old idea is attributable to the recent turbulence of international finances, giving practical reasons for wanting

Donald N. McCloskey is chairman of the department of economics and professor of history at the University of Iowa, Iowa City.

J. Richard Zecher is senior vice-president and chief economist of the Chase Manhattan Bank, New York, New York.

to know when exchange rates or prices are in equilibrium. But some interest is a result of autonomous intellectual change by itself. The chief of these changes was the monetary approach to the balance of payments developed in the early 1970s by Robert Mundell and his students, with involvement by Harry Johnson, Ronald McKinnon, Jeffrey Williamson, and various others.

The monetary approach can do without the law of one price.[1] One can approach the balance of payments as a monetary phenomenon—that is to say, not primarily the real phenomenon that the elasticities approach believes it to be—without committing oneself to any particular view of the working of international arbitrage. The monetary approach, after all, merely notes that the balance of payments is by definition a balance on monetary account (just as the current balance is a balance on commodity account and the capital account is a balance on an account of future claims), then makes the innocent-sounding suggestion that its explanation might focus on the excess supply of and demand for money. It is not essential to adopt purchasing-power parity to believe that there may be merit in this view.

Nonetheless, purchasing-power parity was in fact commonly invoked in the early theorizing in the monetary approach. Those present at the creation in Chicago in the late 1960s and early 1970s felt they were merely appropriating for use in the study of monetary affairs an assumption that was a necessary commonplace in the study of real affairs. It is hard to see how the real theory of international trade could have gotten far without postulating enough rationality on the part of economic actors to arbitrage away for each commodity any price differences outside the gold points— the term "gold points" defined to include all the risks and other costs of transportation. And, to go further, if the gold points were as wide for many commodities as is often implied in criticisms of purchasing-power parity, it is hard to see what usefulness there could have been in the standard propositions in the real theory, such as a tendency to factor price equalization or a tendency to satisfy the Heckscher-Ohlin theory of exports or indeed any tendency to equilibrium. The pioneers of the monetary approach felt they were simply bringing to international finance the intellectual habits formed in the study of the real theory, especially the intellectual habit of supposing that people exploit opportunities for profit. It is no accident that the second generation of leaders in bringing rigor to the study of international trade—the Mundells and Johnsons, following on the first generation of Samuelsons and Meades— were the inventors of the monetary approach and were inveterate users of the assumption of purchasing-power parity.[2]

We believe that the doctrine of purchasing-power parity, brought into theorizing about the monetary approach by the back door, should be the guest of honor. It is a more radical proposition than the one that the

nation's liquidity balance (or, still more generally, the asset balance) probably has much to do with the amount of money it imports or exports. It challenges in fact the whole way of doing macroeconomics.

The usual way of doing macroeconomics might be called the Martian approach. A national economy, usually the United States, is taken in its relations with the rest of the world to be similar to Mars. The price level on Mars, obviously, is determined by Martian demand and supply curves (whether for money or for aggregate goods is not important); likewise, the interest rate. An occasional spaceship might land from Earth bearing gold or Federal Reserve notes, thereby driving up the price level on Mars (or increasing speculative balances in the presence of a liquidity trap— again, the rest of one's economic ideology is irrelevant to the point at issue). The arrival of the spaceship might even be occasioned by events on Mars, but only in a very long run, since it is a long way from Earth to Mars. Mars is a closed economy that has to adjust to the money supply or aggregate demand or expectations that Martians have, period.

The Martian approach characterizes 90 percent of the articles and books on macroeconomics, written mostly by Americans. The theoreticians among them can always argue that it is unimportant to them whether or not any actual economy matches their models, for they are concerned with higher things. So much the worse for theory, one might say. A floor or two down the ivory tower the more empirically concerned theorists can argue, quite correctly, that their models might well apply to the whole world even if they are inappropriate for one part of the world. It is strange then to include institutions (such as a central bank with a national policy) in the models that have no worldwide equivalents, one might say. Empiricists, living with the computer down in the basement, can and do argue wearily that they are working on a still larger model (with 10,001 sectors) and will perhaps be able to fit the international sector into one of these. The rest of the world is to them merely another, rather small, sector of the American economy, similar in importance, say, to office equipment. They view an appeal to include the international milieu as a tiresome request to further complicate an already complex model by inserting an office-equipment sector. At the most they are willing to consider project LINK, with its plan to cure the maladies of misspecified national models by putting all the models into the same hospital.

It is notable where the other 10 percent of the books and articles on macroeconomics originate. They originate from small open economies. Cassel was Swedish. His countryman, Knut Wicksell (1918), had no difficulty believing purchasing-power parity, debating with the American Taussig (1918) on the matter.[3] The assumption of thoroughgoing arbitrage between regions and countries pervades the work of Heckscher and Ohlin. Canadians were the heirs to the Swedes in producing dispro-

portionate numbers of excellent economists. They too, living with the great bear of the United States as the Swedes lived with the great bear of the German Empire, have found it difficult to think in Martian terms about their homeland, and even about their new American home. Johnson, McKinnon, and Mundell were raised in Canada, as were many of their students.

Martian thinkers are accustomed to dismissing such cases with the remark, "precisely: they came from small economies; the United States is large." The remark is irrelevant, reflecting a common notion that recent developments in balance-of-payments theory depend somehow on the assumption that we are dealing with small economies.

The psychological disposition to recognize the existence of an international milieu may be smaller in a big country, but the milieu is still there. The United States may be so big, to be sure, that it can significantly alter the world's price level or interest rate (at least so the finance ministers of other countries believe). But the American price level and interest rate are no less the world's on this account. That General Motors is big does not put it in a different market for automobiles than British Leyland or Simca. America's money supply may well in some periods act as the world's high-powered money, with multiple effects (although the usual accounts of monetarism do not talk this way); America's policy in some periods may well affect expectations abroad (although the usual accounts of rational expectations do not talk this way). Adopting such arguments would constitute a radical break with Martianism. And in any case it would entail recognizing that prices and interest rates were world, not national, phenomena, which is quite another way of doing macroeconomics.

The other way of doing macroeconomics may be called the Iowa City approach. No one doubts that Iowa City has virtually no control over its price level and its interest rates. In very short order an attempt by Iowa City bankers to raise interest rates on loans to twice the market rate would empty the loan offices of the banks. In rather longer order (a month, say), an attempt by Iowa City grocers to raise prices to twice the market would empty the grocery stores. In still longer order (a couple of years, say), an attempt by house owners to raise rents above the cost and value of housing determined by the substitutability of housing for other goods in production and consumption would empty the city.

Likewise, no one would believe there was a useful sense in which Iowa City could have a monetary policy. It could impose tariffs or price controls, to be sure. But the more usual and subtle instruments of monetary policy would be blunt in the hands of the First National Bank. If Iowa City had its own money supply (and under the free-banking legislation before the Civil War it, like many American cities, in fact did), increasing the money supply would have no effect on the Iowa City price

level or interest rate, so long as the exchange rate was fixed. No economist would place any credence in a model of the economy of Iowa City that allowed Iowa City's prices and interest rates to be determined wholly or even largely within Iowa City by the forces of aggregate demand and supply.

The choice is one between a world in which purchasing-power parity does and does not work well enough to be a good description. Is the price level of the United States (when the exchange rate is fixed) substantially or importantly free to move independent of the price level of the rest of the world? Is the United States (or the United Kingdom or whatever) more like Mars or is it like Iowa City?

The choice between the Martian and the Iowa City approaches is an empirical one. The United States is not literally either Mars or Iowa City. The question is which approach is closer to the truth, or if you prefer, which mix of the two is true. In particular, it is not enough to remark blandly that both approaches apply to some degree and then proceed to use one or the other to buttress some conclusion on policy or history.

For much of the period 1880 to the present the major economic powers were on literally or virtually fixed exchange rates, and it is to such a case that the argument applies most easily. But it is not true, as some think, that a regime of flexible exchange rates completely unhinges an economy from the world market. With a correction for the exchange rate, purchasing-power parity might still apply (though one would expect uninsurable exchange risk to make the gold points wider). And if purchasing-power parity does apply, then the central bank can have only a neutral effect on the economy. The bank would be free to push the general price level up or down (and could just as well make the exchange rate the policy instrument as the money supply), but could not alter relative prices, pegged by world markets. Relative prices—for instance price of investment goods relative to consumer goods—are commonly objects of monetary policy. The common objectives are unattainable if purchasing-power parity works well. And, to repeat, if governments bind themselves to a fixed exchange rate, they cannot even have a neutral influence on prices.

Another red herring sometimes drawn across the trail should be avoided as well: purchasing-power parity is assured in the very long run by the price-specie-flow mechanism. Therefore, the argument goes, the monetary approach, which assumes that markets operate very quickly among nations, is merely another way of expressing conventional monetarism, which assumes that markets do not operate quickly among nations.[4]

The argument is misleading. The price-specie-flow mechanism is a disequilibrium model. It is essentially that two economies that for some reason develop a divergence in their purchasing-power parities will

generate flows of gold that will realign the parities. By contrast, the monetary approach, subspecies purchasing-power parity, is an equilibrium model. It claims, nonetheless, by virtue of an alleged quickness with which price divergences among countries are arbitraged away, to be relevant to a much shorter run than could reasonably be supposed for the other. That in one respect (namely, purchasing-power parity) the models happen to have the same outcome in the long run should not be allowed to obscure that the two exhibit radically different behavior in most other ways. In particular, monetary policy does work in the price-specie-flow model (at least in some short run and at least if the model does not belie itself by introducing a *rapid* price–specie flow) but does not work in the purchasing-power-parity model except by way of influences on the world money supply. The monetary approach takes much from monetarism but, in the end, differs importantly from the monetarist aproach to national monetary policy.

We should point out that the historical record contains little evidence that the price-specie-flow mechanism actually happened. Economists, accustomed to thinking of the Facts of History, may be surprised; it will not surprise historians, hardened to the ubiquity of the Myths of History. The intellectual status of the mechanism is similar to the kinked demand curve of oligopolists: it does not work empirically and is unreasonable besides (for instance, it would provide opportunities for speculative profit).

In an earlier paper (1976, p. 367) we reviewed the empirical anomalies in the price-specie-flow mechanism. For instance, we argued that Milton Friedman and Anna Schwartz misapplied the mechanism to an episode in American history. The United States went back on the gold standard in January 1879 at the pre–Civil War parity. The American price level was too low for the parity, allegedly setting the mechanism in motion. Over the next three years, Friedman and Schwartz argued from annual figures, gold flowed in and the price level rose just as Hume would have had it. They conclude (1963, p. 99) that "it would be hard to find a much neater example in history of the classical gold-standard mechanism in operation." On the contrary, however, we believe it seems much more like an example of purchasing-power parity and the monetary approach than of the Humean mechanism. In the monthly statistics (Friedman and Schwartz confined themselves to annual data), there is no tendency for price rises to follow inflows of gold, as they should in the price-specie-flow mechanism; if anything, there is a slight tendency for price rises to precede inflows of gold, as they would if arbitrage were shortcutting the mechanism and leaving Americans with higher prices directly and a higher demand for gold. Whether or not the episode is a good example of the monetary theory, it is a poor example of the price-specie-flow mechanism.[5]

The price-specie-flow mechanism, therefore, is not a good way to harmonize closed-economy monetarism with the fact that we live in a world of many economies. What should be clear by now is that if purchasing-power parity is found to be a useful characterization of the world, then closed-economy theorizing and empirical work in macroeconomics should be changed to allow for the direct effects of international price arbitrage. Whether monetarists or Keynesians or rational expectationists, economists should begin thinking and measuring in global terms.

2.2   The Root Definition of Purchasing-Power Parity

"Come, come," the representative Martian will say, "don't waste my time—we know that prices diverge. Purchasing-power parity fails." That prices are not identical everywhere is not an important failure of purchasing-power parity. A minor reply is that prices can be different in level but related in their changes, a distinction made in the usual statistical tests of parity. The main reply is that purchasing-power parity is a consequence of rationality in arbitraging. If all the opportunities for riskless (or insured) arbitrage among countries that are profitable at existing interest rates and other costs of arbitrage have taken place, then the price level of the world may be said to have exhausted its ability to determine the price level of one country.

Now it may have exhausted it, yet be trivial. To take the single commodity case for illustration (and only for illustration, it being a major theme below that commodity-by-commodity arguments do not suffice), the gold points might be so wide that even though they are not violated they are not useful as describing a constraint on the economy. Wheat and boomerangs in Rumania and Tasmania in 1682 (or in 1982?) may have offered no opportunities for profit by arbitrage ex ante, yet their prices were surely free to move within wide limits independent of each other. Our hypothesis is that in the modern world among the main trading nations the forces of rationality in arbitrage were powerful enough to fix any one general price level and interest rate ceteris paribus, in terms of the others.

The usual statistical question is whether or not the result is a unit elasticity of one price level with respect to another. We shall see in a moment whether the statistical question is the right one. But the price level in one country could be determined by the world in the sense that it was fixed ceteris paribus by the action of arbitrageurs—even if the elasticity were not unity. For this reason arguments that ratios of purchasing-power parities tend to drift are irrelevant.[6] They do drift, just as demand curves drift. To say that the law of demand fails because in an uncontrolled experiment the observed price does not correlate well inversely

with the quantities would be an error. So too here. The arbitrageurs could be making the American price level exogenous to, say, American monetary policy even if changing technologies of traded as against nontraded goods, changing compositions of market baskets, and changing errors in the underlying statistics caused the American price level to be poorly correlated with other price levels.

It is easy to construct examples that illustrate the point. Imagine that electrical equipment produced in the United States sells for 20 percent more than that produced in Germany, after translating the prices at the exchange rate, because of differences in the energy efficiency of the equipment. Imagine further that deviations as little as plus or minus one percentage point from the 20-percent differential would create opportunities for profitable arbitrage, exploited instantly. Now suppose the equilibrium price ratio rises (that is, the ratio dividing profitable from unprofitable arbitrage) suddenly to 30 percent from 20 percent, because of changes in energy prices (worldwide). Suppose finally that the same plus-or-minus-one-percentage-point band exists at the 30-percent differential as at the old 20-percent differential. The economist stumbling on such data might conclude that parity fails: the drift from a 20-percent to a 30-percent differential would be interpreted by him as indicating the poorness of correlation between prices in one place and in another. Yet in the sense relevant here of ceteris paribus the prices in Germany and the United States are mutually linked (plus or minus one percentage point) just as strongly at 30 percent as at 20 percent. A monetary policy in the United States that had as one intended result a rise in electrical-equipment prices in the United States relative to (exchange-adjusted) German prices would fail.

We wish, then, to appropriate for purchasing-power parity the prestige of the postulate of rationality. It surpasses belief that many opportunities to make easy money buying low and selling high persist long enough to be observed in economic data. Yet much of the opposition to purchasing-power parity seems to believe that it is so. When specialists in finance such as Richard Roll (1979) think about international markets, they assume with hardly a comment that all opportunities for arbitrage are exhausted in a matter of weeks. Roll remarks that "in the monetary approach . . . prices and exchange rates tend toward equilibrium . . . in the very long run, *say a year or more*" (italics added; p. 135).[7] His criticisms of the monetary approach come from a novel direction: instead of criticizing the approach for supposing that the long run is as *short* as a year, he criticizes it for supposing that the long run is as *long* as a year.

The view from finance is refreshing and highly relevant. We consider purchasing-power parity to be a proposition similar in more ways than one to the efficient-markets hypothesis, to be demonstrated on similar grounds, namely, on the ground of the shared belief of economists in

rationality and on grounds of whatever evidence can be adduced to confront the belief. The belief is that there must be nothing systematic or predictable in the future of prices in one country relative to present prices in another (except, indeed, instantaneous equality), or else there is money to be made in exploiting the prediction. If differences in price levels take years to be eliminated by trade, then trade right now—in one commodity or in the CPI—is profitable. The result, we believe, is a "theorem" linking the pursuit of profit and the exogeneity of the general price level. If opportunities for arbitrage are exploited (allowing fully for the cost of transport and information), then the price level of one country is fixed by the rest of the world, even in the very short run. The argument is not a test by itself, but merely a theorem, a higher-order proposition about the relation between equilibrium and the exogeneity of prices. The theorem is a curious product, to be sure, for it is not merely a logical proposition. It requires the world to be arranged in a certain way to be true. We arrive again at an empirical question.

Three points of logic nonetheless may make the fundamental theorem of the Iowa City approach more palatable. First, arbitrage does not need to occur commodity by commodity. All prices in an economy are connected to each other. Prices of bricks in New York and London are held together not only by the direct forces of arbitrage in the market for bricks itself, weak as they are, but by the indirect forces of arbitrage in related markets—the markets for brick-making labor, say, or the market for lumber for which bricks are a substitute. That nontraded goods exist is sometimes thought to be a rebuttal to purchasing-power parity. Not so, at least if the nontraded goods are provided in markets sensitive to costs (military bases, for instance, may be an exception). It is also thought to be a rebuttal to note that the law of one price need not hold for such-and-such commodity if that commodity's market is obstructed. Not so, at least if the commodity in question is related in production and consumption to other goods. Commodity-by-commodity thinking can be misleading. The simplest form of the argument is the Walrasian point that one absolute price (the numeraire) serves to set all other prices in the economy, given resources, technology, and tastes. At the extreme, then, if Mars were connected to Earth by the market in chewing gum alone, the two price levels would nonetheless be fixed in relation to each other. How much tighter, one might suppose, would be the relation between two economies connected by the prices of thousands of goods and services.

A second point of logic, seldom recognized, is that arbitrage across space down to the extent of the transport-cost wedge is reinforced, for storable commodities (housing, wheat, automobiles, cement), by arbitrage across time down to the wedge of storage costs. Suppose the price of cement rises in the United States because demand has risen. One might say that cement will not immediately flow from Canada or Spain to the

United States, that it takes time to reorganize the direction of transport. But if the cement price is expected eventually to take up its usual relation with foreign prices, then the present relation is constrained—it cannot deviate from the long-run relation by more than the cost of "transporting" cement from the long into the short run. The argument applies even to nontradable goods: if the forces of general equilibrium would eventually bring even housing prices, say, into a rough parity, then the storability of housing will enforce the parity earlier.

A third point is that a properly measured price index would be an index of characteristics, not named goods and services. The imperfections of the usual price indexes should not be used as evidence for the failure of purchasing-power parity, whether the imperfection is a sheer error in reporting or, as here, an error of concept. It is commonly argued that goods and services, especially manufactured goods, are not perfect substitutes across countries, that the category "vacuum cleaners" contains Panasonic (Matsushita Electric) model MC-881 and Sears Power-Mate model 20 A 2099 with different characteristics, and therefore that departures from the gold points are rational. The argument is that competition between the two models of vacuum cleaners is not perfect. The response must be, for one thing, that the degree of monopoly in international trade is beside the point. Matsushita Electric might well be discriminating in the prices it sets for vacuum cleaners in Japan and in the United States, yet it would still be true that there was a stable relationship between the Japanese and the American price determined, say, by relative elasticities of demand. For another thing, the characteristics making up the good may well have perfect markets. Vacuuming power, ease of use, reliability of service, and the like are separately measurable, at least in the consumer's mind, and each is perfectly substitutable across brand names. The bundle of characteristics called a Panasonic vacuum cleaner may not be exactly duplicated in any other vacuum cleaner, but the price of vacuuming power may be set on a competitive market. Another way of making the point is to think of the prices of named goods and services as being composed, speaking statistically, of many factors in principal components, shared with many other goods. This is a statistical way of stating the general equilibrium point: what must be arbitraged is a relatively small number of characteristics, not each of millions of named goods and services one by one. The degree of identity in products across countries is no more relevant than is the degree of identity of other measures of the composition of consumption.

2.3   Tests of the Efficacy of Arbitrage

Thus armed against irrelevant doubt, we turn to the empirical questions. One can cast light on the degree of thoroughness of equilibrium (or

"arbitrage") by measuring it directly. The measurement is extremely difficult, essentially because a measurement entails second-guessing people in the business. If the right amount of resources is being used for arbitrage, then prices in one place are not free to move independent of prices elsewhere. To deny purchasing-power parity in this behavioral sense is to deny that the right amount of resources is being used for arbitrage. Often unrecognized by critics of purchasing-power parity is that their conclusion that it has failed usually implies an ability to make money. Anyone who knew that purchasing-power parity was true in some long run but not yet true in the present would have a rosy financial future. The divergences from purchasing-power parity detected in the literature are so gross and the statistics purporting to show the divergences so easy to collect that the opportunities for profit are large. Go thee and prosper.

The only direct test of the rationality of arbitrageurs, then, is literal second-guessing. That is, one assembles the facts on prices and transport costs, being careful to allow for such subtleties as the cost-of-exchange risk between the time the arbitrage opportunity arises and the time it is exploited, and does the calculation that the arbitrageuse presumably did, or should have done. She herself would use list prices only after ascertaining that they reflected prices at which she could actually transact, allowing for delivery lags, credit terms, and risk of default. If she missed a profitable opportunity, one can either doubt the completeness of one's calculations or doubt her rationality, depending on the strength of one's devotion to the working hypothesis of rationality. The test is very difficult to perform, though it has the compensating merit of being most relevant to the question at issue. We urge others to attempt it.[8]

Another, more practical test is to examine whether the shipments of goods implied by the supposed opportunities for arbitrage in fact occurred. Enthusiasts of purchasing-power parity would argue that the mere threat of such flows suffices to bring prices speedily back to unprofitable levels of divergence. Doubters of purchasing-power parity would argue that only the flows themselves suffice, and these only over a considerable period. Taking the doubters' view, then, the flows themselves should be observable. A country that exhibits divergences from purchasing-power parity convincing to the doubters should also exhibit lowered exports and increased imports, a trade deficit: a place with "high" prices would have a hard time selling and an easy time buying. The reasoning is, of course, the first step in the elasticities approach to the balance of payments. In other words, the balance of trade should become more negative for a country that was above its trend of purchasing-power parity.

Nothing of the sort shows through in the U.S. statistics. One would expect excessive U.S. inflation to hurt U.S. exports. That is, one would expect a rise in the deviation from purchasing-power parity, expressed as

a rise in the difference between the U.S. price index and the U.K. or Canadian price index in U.S. dollars, to cause a fall in the trade balance. The equation for the United Kingdom meets the expectation: the sign on $\Delta (P_{US} - eP_{UK})$ is indeed negative (standard errors in parentheses):

$$\begin{bmatrix} \text{Change in U.S. trade} \\ \text{balance with U.K.} \end{bmatrix} = \underset{(19.3)}{5.3} - \underset{(2.5)}{2.8}[\Delta(P_{US} - eP_{UK})].$$

The Durbin-Watson is 1.59, in the indeterminant range, but of course much better than the result of running the levels instead of the differences. The slope coefficient, however, is insignificant at conventional levels (notice that the insignificance of the constant indicates that there is no linear trend to worry about). It is unclear what insignificance might mean with observations that are not a sample but the universe of the relevant U.S. and British variables in the period. The R^2 of the equation is a mere .02. It is clear what that means.

The equation for Canada, which in view of its contiguity with the United States should prove a better test, is worse:

$$\begin{bmatrix} \text{Change in U.S. trade} \\ \text{balance with Canada} \end{bmatrix} = \underset{(6.1)}{3.5} + \underset{(0.73)}{0.61}[\Delta(P_{US} - eP_{Can})].$$

The sign is perverse, the R^2 only .01, and again the coefficients are insignificant.[9]

One could certainly raise the R^2's here and perhaps correct the sign by embarking on a search through all possible specifications of lags and functional forms and periods. For instance, from 1880 to 1912 the United States deflates faster (down to 1896, the usual turning point for price series in the period), then inflates faster than the United Kingdom and Canada. True to expectations, the trade balance with the United Kingdom (though not with Canada) exhibits the same U, but inverted. U.K. clamor about German and U.S. competition becomes great in the 1890s and diminishes in the Edwardian boom of exports and investment abroad that followed. But such evidence would have to contend with a markedly parallel movement of net U.S. inflation and net U.S. exports—the reverse of expectations—relative to both the United Kingdom and Canada during the 1920s and 1930s.

In other words, the apparent deviations from purchasing-power parity appear not to represent unexploited opportunities for arbitrage. Another way to say it is that the deviations were only apparent. The "deviations" could be, for instance, the result of peculiarities in the price indexes— U.S. price indexes are always more volatile than the U.K. ones in the late nineteenth century, which may well say more about how the indexes were constructed than about the underlying real character of price formation. The rise and fall of net exports to the United Kingdom, then, could be a result of matters wholly internal to the countries, such as building booms

in the face of industrial maturity. Alternatively, the "deviations" could be once-for-all (or for-the-duration) changes in the cost of moving goods from one place to another, as it appears are the violent leaps up and down of parity during the Great War. Changes in tariffs or in other legal costs of movement can permit a change in the relative position of two price levels without in any way belying the assertion that the price levels are not free to move where they will. Such a tariff inflation appears to have taken place in Germany during the 1880s (although we have not as yet been able to find out how powerful the explanation is). While prices fell elsewhere, led by agricultural prices responding to competition from Russia and the United States, prices in Germany held up, as Bismarck forged a tariff politics of rye and iron. Further tests of the hypothesis that profitable opportunities for arbitrage arise when measured prices diverge might proceed commodity by commodity, but the aggregate results at least are unpromising for the hypothesis.

2.4   The Irrelevance of Price Regressions Unadorned

The representative Martian will by now be close to apoplexy. "Innumerable regressions of one country's prices on another have been performed recently, and half of them show that purchasing-power parity fails. Surely such tests are conclusive."

No, they are not. There are two points here. The first, the more fundamental, is that the regressions are not useful tests of our hypothesis, which has to do with the rationality of arbitrageurs, not with how closely one price correlates with another. Our hypothesis says that prices are linked and therefore insensitive to internal forces such as monetary policy (as we shall show presently), not that the prices are linked by some linear relationship having such-and-such a slope.

The second point is that even the hypothesis of linear relationship, which unlike ours is not based on a foundation of individual rationality, has been tested inadequately. The tests take the form of regressing, say, the U.S. GNP deflator on the U.K. GNP deflator multiplied by the exchange rate of dollars for pounds.

$$P_{US} = a + \beta[(e)(P_{UK})] + \text{error term}.$$

If the slope coefficient β is significantly different from 1.0 at conventional levels, then purchasing-power parity is said to fail. A number of criticisms can be and have been leveled at such equations, although none is the main point here. It can be argued, for instance, that allowing for simultaneity bias brings the β coefficient closer to 1.0 (Krugman 1978). It can be argued that the equation is misspecified, not allowing properly for lags or for secular trends. It can be argued that "significance" is beside the point for a nonsample. It can be argued that the prices, especially the

foreign prices on the right-hand side, are measured in error, as they are for instance if the United Kingdom is taken to stand for the world in the calculation of prices, biasing the coefficient towards zero. It can be argued that the slope coefficient could differ from 1.0 if such a difference were implied by the Walrasian general equilibrium. It can be argued that the whole procedure is misleading, because any coefficient different from *zero* bespeaks a relationship between domestic and foreign prices dangerous to ignore in explaining domestic prices. It can be argued that to judge the hypothesis of parity a failure, and to abandon it, is to flee to evils that we know not of.

The main point, however, is that the "failure" of purchasing-power parity in such an equation is not measured against a standard. How close does the slope have to be to the ideal of 1.00 to say that purchasing-power parity succeeds? The literature is silent. The standard used is the irrelevant one of statistical significance. A sample size of a million yielding a tight estimate that the slope was .9999, significantly different from 1.00000, could be produced as evidence that purchasing-power parity had "failed", at least if the logic of the usual method were to be followed consistently. Common sense, presumably, would rescue the scholar from asserting that an estimate of .9999 with a standard error of .0000001 was significantly different from unity in a significant meaning of significance. Such logic also could be applied to findings of slopes of .90 or 1.20, but usually it is not.[10]

The point is not that levels of significance are arbitrary. Of course they are. The point is that it is not known whether the range picked out by the level of significance affirms or denies the hypothesis. Nor is the point that econometric tests are to be disdained. Quite the contrary. The point is that the econometric tests have not followed their own rhetoric of hypothesis testing. For one thing, as we have said, the errors that tests of significance deal with are errors of sampling, but in many cases there is no sampling involved: we have the entire universe of observations of the general price level in the United States and the United Kingdom 1880–1940. For another, nowhere in the literature of tests of purchasing-power parity does there appear a loss function. We do not know how much it will cost in policy wrecked or analysis misapplied or reputation ruined if purchasing-power parity is said to be true when by the measure of the slope coefficient it is only 85-percent true. That is, the argument (due to Neyman and Pearson 1933) that undergirds modern econometrics has been set aside here as elsewhere in favor of a merely statistical standard, and an irrelevant one related to sampling error at that. We are told how improbable it is that a slope coefficient of .90 came from a distribution centered on 1.00 in view of the one kind of error we claim we know about (unbiased sampling error, with finite variance), but we are not told

whether it matters to the truth of purchasing-power parity where such limits of confidence are placed.

Silence on the matter is not confined to the literature of purchasing-power parity. Most texts on econometrics do not mention that the goodness or badness of a hypothesis is not ascertainable on merely statistical grounds. Statisticians themselves are more self-conscious, although the transition from principle to practice is sometimes awkward. A practical difficulty in the way of using the Neyman and Pearson theory in pure form, say Mood and Graybill (1963, p. 278), is that

> the loss function is not known at all or else it is not known accurately enough to warrant its use. If the loss function is not known, it seems that a decision function that in some sense minimizes the error probabilities will be a reasonable procedure.

Such a procedure might be reasonable for a general statistician who makes no claim to know what is a good or bad approximation to truth in fields outside statistics itself. The procedure is not reasonable for a specialist in international trade or macroeconomics. If the loss function is not known it should be discovered.

Finding the loss function amounts to finding out how close the slope has to be to 1.00 in order for it to be reckoned close enough. Every student of the matter is more or less aware of the need for some standard against which to judge the closeness, a standard beyond the probability of being misled by sampling errors (if there really are any) into gullibility or skepticism, but no one has provided them. In a superb paper that has by its sheer weight and subtlety turned the tide of battle recently against purchasing-power parity, for example, Kravis and Lipsey (1978), in reporting some calculations of parity relative to base years and some correlation coefficients (both of which "ought" to be 1.00), remark:

> Each analyst will have to decide in the light of his own purposes whether the PPP relationships fall close enough to 1.00 to satisfy the theories. As a matter of general judgment, we express our opinion that the results do not support the notion of a tightly integrated international price structure. (Pp. 214ff.)

The only guidance they provide to evaluating their "general judgment" is a footnote (p. 214) reporting that in the general judgment of Houthakker, Haberler, and Johnson deviations from parity of anything under 10 or 20 percent are acceptable to the hypothesis.[11] It happens, incidentally, that the bulk of the Kravis and Lipsey evidence passes rather than fails such a test. But accepting or rejecting one unargued standard of truth by comparing it with another unargued standard of truth does not advance the art of argument in economics very much.

To be fair, Kravis and Lipsey are in fact unusually sensitive to the case for having some standard, more sensitive than are economists working the field with more powerful statistical tools. So frequently does their paper make the point that it must be accounted one of the major ones made. Repeatedly, for instance, they draw a distinction between the statistical and the economic significance of their results: "Indeed, even high coefficients of correlation [between domestic and foreign prices] may conceal shifts in relationships that are economically important" (p. 204); "The difference [of slope coefficients from 1.00] may not be large enough to be picked up by a statistical test yet be economically significant" (p. 236); "even a high correlation does not preclude what may be economically significant variations between the two indexes" (p. 242); and so forth, *passim*. The intellectual sword is sharp.

Remarkably, however, they use the sword only against, not in favor of, purchasing-power parity, and never turn it on themselves. After stating repeatedly that they do not have a standard by which to judge the hypothesis, they nonetheless conclude:

> We think it *unlikely* that the *high* degree of national and international commodity abitrage that many versions of the monetarist theory of the balance of payments contemplate is *typical* of the real world. This is not to deny that the price structures of the advanced industrial countries *are linked* together, but it is to suggest that the links are *loose* rather than *rigid*. (Italics added; p. 243.)

Every italicized word involves a comparison against some standard of what constitutes likelihood or highness or typicality or linkage or rigidity, yet no standard is proposed. Indeed, the lone page (p. 204) in their fifty that addresses the issue, entitled "Criterion of Similarity of Price Movements," is devoted to dismissing the lone attempt in the literature to offer a criterion (namely, the Genberg-Zecher criterion: if the correlations of prices among different countries are as good or bad as they are inside one country, then the different countries act as one) and to remarking again that the merely statistical standard used elsewhere in the paper is in fact irrelevant.

The irrelevance of the merely statistical standard of fit does not bedevil only that half of the empirical literature that finds purchasing-power parity to be wrong. Towards the end of a fine article favorable to purchasing-power parity, Krugman writes (1978 p. 405):

> There are several ways in which we might try to evaluate PPP as a theory. We can ask how much it explains [that is, R-square]; we can ask how large the deviations from PPP are in some absolute sense; and we can ask whether the deviations from PPP are in some sense systematic.

The defensive usage "in some absolute sense" and "in some sense" betrays his unease, which is fully justified. There is no "absolute sense" in

which a description is good or bad. The sense must be comparative to a standard.

Similarly, Jacob Frenkel (1978 p. 175) says "if the market is efficient and if the forward exchange rate is an unbiased forecast of the future spot exchange rate, the constant term [in a regression of the spot rate today on the future rate for today quoted yesterday] . . . should not differ significantly from zero, the slope coefficient should not differ *significantly* from unity" (italics added). In a footnote on the next page, speaking of the standard errors of the estimates for such an equation in the 1920s, Frenkel argues that "while these results indicate that markets were efficient and that on average forward rates were unbiased forecasts of future spot rates, the 2–8 percent errors were *significant*" (italics added).[12] What he appears to mean is that he judges a 2- to 8-percent error to be large in some economic sense. In any event, what his results imply about their subject is unclear—purchasing-power parity, because significance in statistics, however useful it is as an input into economic significance, is not the same thing as economic significance.

2.5   The Search for Standards

Results typical of the conventional tests of purchasing-power parity are easy to replicate. Regressions for 1880 to 1940 (running through the First World War, as a more extreme test of the argument) of the U.S. GNP deflator (to avoid the usual criticism of the domination of wholesale prices by traded goods) against Canadian or U.K. prices adjusted by official exchange rates are (all variables in logarithms and standard errors in parentheses):

$$\text{Price in U.S.} = \begin{array}{c} 0.83 \\ (0.19) \end{array} + \begin{array}{c} 0.87 \\ (0.04) \end{array} \left[\begin{array}{c} \text{Exchange-adjusted} \\ \text{Canadian price} \end{array} \right].$$
$$R^2 = .87, \text{D.W.} = 0.29$$
$$\text{Price in U.S.} = \begin{array}{c} 0.26 \\ (0.14) \end{array} + \begin{array}{c} 0.93 \\ (0.03) \end{array} \left[\begin{array}{c} \text{Exchange-adjusted} \\ \text{U.K. price} \end{array} \right].$$
$$R^2 = .94, \text{D.W.} = 0.46$$

Additional regressions adjusting for autocorrelation and for the simultaneity of prices and exchange rates (during the floating-rate period) yielded β's for the United Kingdom that ranged from .91 to 1.02, and for Canada from .35 to 1.00. Most of these regressions have in common the result that at conventional levels of significance, the slope coefficients would be said to be different from 1.00, and about half of the trends would exhibit "drift." In other words, by the usual standards, these regression results would lead to the conclusion that purchasing-power parity "fails."

But the regressions also have something else in common—a uniformly
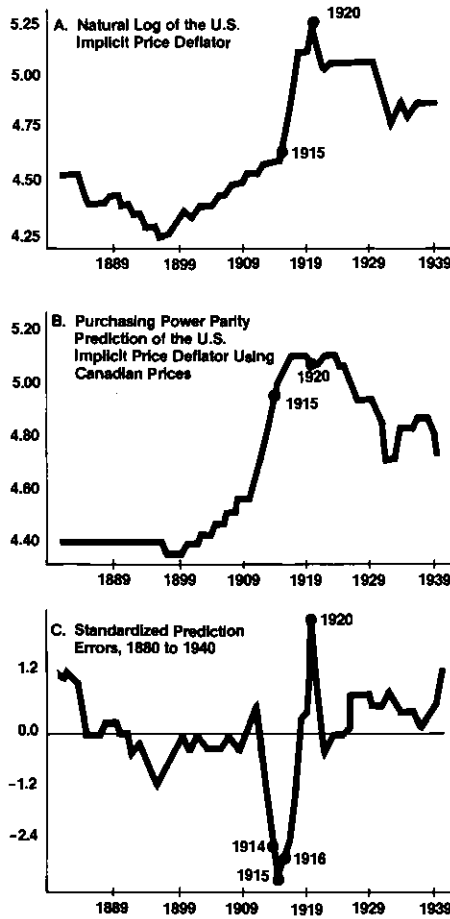
Fig. 2.1   U.S. prices predicted by Canadian prices.

"high" correlation as measured by R^2. Consider the plots in figures 2.1 and 2.2 of the actual U.S. price, the price estimated from foreign prices, and their difference, for Canada and the United Kingdom. Although the period includes the Great War and the Great Depression, the foreign prices do predict the gross outlines of the U.S. price. Such eyeballing is another way of saying that the R^2's are high, as they are. A fuller treatment would make the sensitivity analysis explicit, introducing the cost of being wrong in adopting a model in which a was zero and β was 1.00.

A second standard is suggested by the bottom graphs, which plot the forecasting error in units of standard deviation. The period 1880–1914 is a standard for relative tranquility. This is the standard, a relative one over
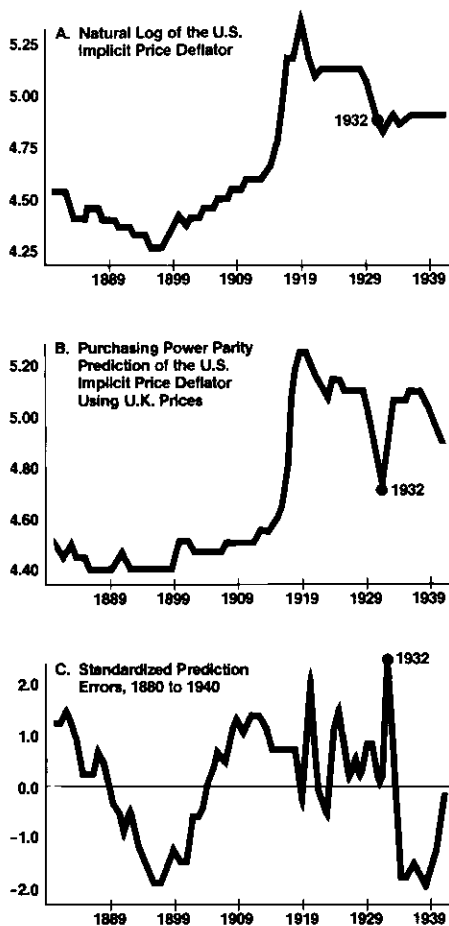
Fig. 2.2   U.S. prices predicted by U.K. prices.

time. If one is willing to think of 1880–1914 as tranquil (one may not, of course), then it lends meaning to the "success" or "failure" of purchasing-power parity. Comparing 1880–1914 with 1921–40 reveals no difference in the average deviation from purchasing-power parity. The comparison involving the United Kingdom, indeed, can even include the Great War and its immediate aftermath with no change in result. The turbulence of the 1920s and 1930s, which is said to have loosened the economic ties among nations, appears not to have done so. The price relations are equally close, which suggests that the economic behavior causing prices to move in parallel was uniform. It is no surprise that turbulence would offer high rewards to arbitrageurs and that arbitrageurs would take them.

A third standard (our favorite) is the Genberg-Zecher criterion (see section 2.4 above). Our paper, and papers by Genberg (1976, 1978), expanded on this theme in the early 1970s.[13] It will suffice for the present to recall the conclusion of our work, namely, that the matrix of the United States, the United Kingdom, Germany, and Sweden (with special attention to the United Kingdom) had correlations of prices in the 1880–1914 period similar to those between different parts of the United States. Genberg performed similar tests with similar results for recent times. The few attempts since to undermine the conclusion leave us unmoved. Arthur Gandolfi and James Lothian (1982) have written an interesting paper on the subject, but demand that the hypothesis pass a test of a lower correlation between very close states of the United States than between far nations of the world.

A common reaction to the standard is, "Well, suppose we do not even accept the premise that the United States is a unified market?" If one does not accept such a premise, then of course one will refrain from speaking of "the" U.S. price level. Nor will one talk of "the" U.S. GNP. If the United States itself is a mere collection of wholly local markets, there is no more use for the talk than for talk about macroeconomic variables for a federation of Fiji-Botswana-Iceland-Saskatchewan, that is, for a random assortment of places.

Fourth and finally, purchasing-power parity can be tested against the standard of explaining major events or illuminating puzzles in history. We have mentioned the event of the United States returning to gold in 1879 at the wrong parity. Consider another episode, puzzling to all who have noticed it, the inflation of 1933–34. It was a spectacular outlier of the Phillips curve, an example of stagflation forty years before that hideous neologism was coined. Unemployment was very high, yet prices rose, whatever notion of "the" price level one uses: a 14 percent rise in wholesale prices would have caused alarm even in 1981; a 3 percent rise in consumer (retail) prices would have caused alarm at least in 1960 (table 2.1). Why did it happen?

While recognizing the indirect force of the world price level, Friedman and Schwartz (1963, pp. 498–99), argued that:

> Another [factor] was almost surely the explicit measures to raise prices and wages undertaken with government encouragement and assist-ance, notably, NIRA [the National Industrial Recovery Act, leading to the National Recovery Administration, or NRA], the Guffey Coal Act, the agricultural price-support program, and the National Labor Relations Act. . . . We have grave doubts that autonomous changes in wages and prices played an important role [after World War II. But] there seems to us a much stronger case for a wage-price or price-wage spiral interpretation of 1933–37.[14]

Table 2.1   **U.S. Price Indexes 1932–36 (1933 = 100)**

Definition of Prices	1931	1932	1933	1934	1935	1936
1. GNP deflator, 1929 weights	113	101	100	106	105	109
2. GNP deflator, 1954 weights	113	102	100	106	107	108
3. Personal consumption, 1954 weights	117	104	100	106	108	109
4. Consumer price index	118	106	100	103	106	107
5. Retail food	124	103	100	112	119	120
6. Wholesale prices, 1926 base	111	98.3	100	114	121	123
7. Average annual earnings per full-time employee, all industries	122	107	100	104	108	113

Source: U.S. Bureau of the Census 1960. Row 1, series F 1/2; row 2, F 67, 87; row 3, F 68, 88; row 4, E 113; row 5, E 114; row 6, E 13; row 7, D 696.

This is an unusual line of reasoning for such crusaders against mixing up the determination of relative and absolute prices. It also seems to us to square poorly with the evidence.

The chief factual difficulties with the notion that the official cartels sanctioned by the NRA codes caused a rise in the general price level is that most of the NRA codes were not enacted until *after* the price rise. *Ante hoc ergo non propter hoc*. Look at the plot of wholesale prices of 1933 in figure 2.3 (retail prices, including such nontradables as housing, show a similar pattern). Most of the rise occurs in May, June, and July of 1933, but the NIRA was not even passed until June. A law passed, furthermore, is not a law enforced. However eager most businessmen must have been to cooperate with a government intent on forming monopolies, the formation took time. As the Bureau of Labor Statistics described it:

> The monthly load of code approvals reached its peak in the period from October 1933 to March 1934; thereafter there was a rapid decrease. Many of the large employing industries were codified in the latter part of 1933 [mentioning cotton textiles, petroleum, bituminous coal, retail trade, fabricated metal products, retail food]. The National Recovery Administration estimated in a report issued in February 1934 that codification of American industry under the industrial self-government program contemplated by the act was 90 percent completed.[15]

Exchange Rate
($ per £)

$5.00 —

$4.50 —

$4.00 —

$3.50 —

$3.00 —

Jan | Feb | March | April | May | June | July | Aug | Sept | Oct | Nov | Dec

2nd Week, FDR Inaugurated, Bank Holiday, Foreign Exchange Markets Closed

April 20: FDR Announces Intent to Devalue

Scale →

Exchange Rates
($ per £)

U.S.A.
Wholesale
Price Index

Scale →

July 2 or 3, FDR Depreciation Message to
London, World Monetary and Economic
Conference

July 1

April 15

1934
Range

— 72
— 71
— 70
— 69
— 68
— 67
— 66
— 65
— 64
— 63
— 62
— 61
— 60

February 1, 1934: $ Price of Gold Set at $35 per Ounce

Sources:
Friday Spot Exchange Rates in London from Paul
Einzig, *The Theory of Forward Exchange* (1937).
Wholesale Price Index, *BLS Handbook of Labor
Statistics 1936 Edition*, pp. 684–85. Averages for
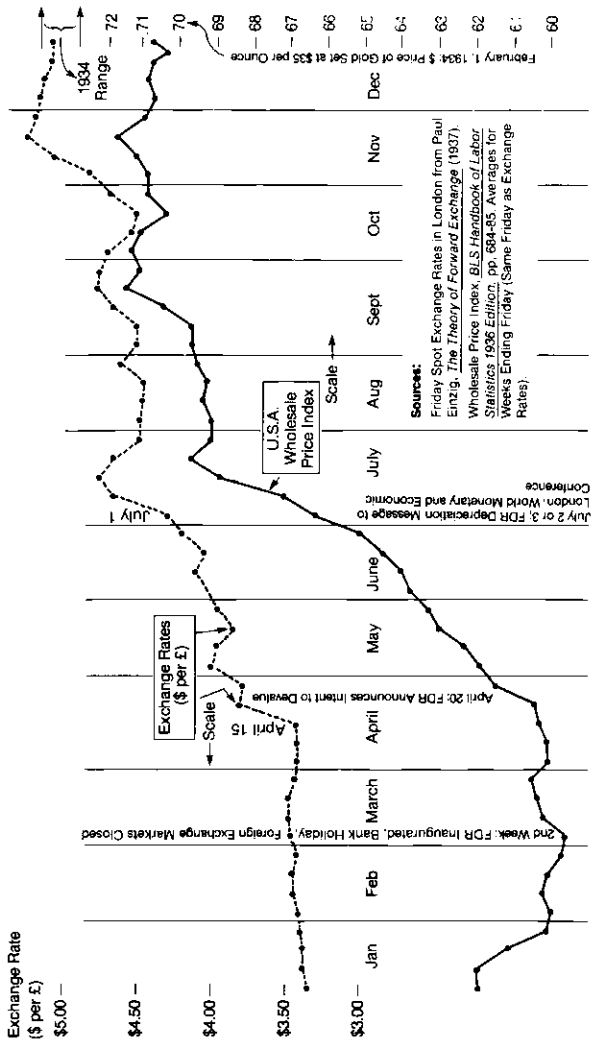Weeks Ending Friday (Same Friday as Exchange
Rates).

Fig. 2.3 Close weekly relationship between the exchange rate and the wholesale index, U.S.A., 1933. *Source:* Friday spot exchange rates in London, from Einzig 1937; wholesale price index, *BLS Handbook of Labor Statistics, 1936 Edition,* pp. 684–85, averages for weeks ending Friday (same Friday as exchange rates).

By September 1933, apparently before the approval of most NRA codes—and, judging from the late coming of compulsion, before the effective approval of agricultural codes—three-quarters of the total rise in wholesale prices and more of the total rise in retail food prices from March 1933 to the average of 1934 was complete. On the face of it, at least, the NRA is a poor candidate for a cause of the price rise. It came too late.

What came in time was the depreciation of the dollar, a conscious policy of the Roosevelt administration from the beginning, although not usually believed to have taken effect until the fall of 1933. There was certainly no contemporaneous price rise abroad to explain the 28-percent rise in American wholesale prices (and in retail food prices) between April 1933 and the high point in September 1934. In fact, in twenty-five countries the average rise was only 2.2 percent, with the American rise far and away the largest. The close link between exchange rates and U.S. wholesale prices is clear in the weekly series graphed in figure 2.3 above. Note especially the two sharp jumps around 20 April and 2 July in response to explicit announcements by Roosevelt of the intent of his administration to devalue the dollar. Wholesale prices move up simultaneously.

It would appear, in short, that the economic history of 1933 cannot be understood with a model closed to direct arbitrage. The inflation was no gradual working out of price–specie flow; less was it an inflation of aggregate demand. It happened quickly, well before most other New Deal policies (and in particular the NRA) could take effect, and it happened about when and to the extent that the dollar was devalued. By the standard of success in explaining major events, parity here works.

2.6 Purchasing-Power Parity and Monetary Policy

In the style of the doubts expressed above about price–specie flow, the success of parity can be judged by the failure of the alternatives. A common view in much of purchasing-power-parity literature is that while international trade places limits on what exchange-adjusted domestic prices can be, there is nevertheless considerable flexibility for prices to move up and down. It is argued in particular that a country can raise its price level relative to the exchange-adjusted price abroad by expanding the domestic money supply. If the activities of traders and arbitrageurs fix prices within very narrow ranges, however, such policies would not work. Consider, then, the relationship between monetary policy and the "errors" in purchasing-power-parity forecasts.

The model is monetarist, postulating that excessively rapid money growth will put upward pressure on prices, leading to domestic prices that are systematically over the purchasing-power-parity predictions. This is a

test, then, to what degree a country, the United States in this case, can through its monetary policies affect its price level (P_{US}) relative to prices (adjusted for exchange rates) in the rest of the world $(\hat{P}_{US})$.

In general equilibrium terms, the hypothesis concerns a state of disequilibrium in the goods markets that is matched by an offsetting disequilibrium in the domestic money market. The regression model below represents the goods-markets disequilibrium by the difference between actual U.S. prices and the purchasing-power-parity prediction of U.S. prices using Canadian and U.K. prices and exchange rates.

Disequilibrium in the U.S. money market is the difference between growth in money supply Δ (M^S) and money demand Δ (M^D). Two measures of money supply are used: (1) M2, which reflects the Martian view that the United States could control its total money supply over the 1880–1940 period, and (2) domestic credit, omitting the effects of specie flows on money supply, which reflects the view that for most of this period the United States could only affect domestic credit, not total money supply.

Growth in money demand is represented by the sum of growth in real income and in prices, on the assumptions that money demand is unit elastic with respect to both of these variables: i.e., $\Delta M^D = \Delta y + \Delta P$, where y is real income and P is the price level. Thus the regression equation becomes:

$$\Delta (P_{US} - \hat{P}_{US}) = \alpha + \beta\Delta (M^S - M^D)_t$$
$$+ \gamma\Delta (M^S - M^D)_{t-1}.$$

The regression results are reported in tables 2.2 and 2.3. Of the thirty-six estimated coefficients relating excess money growth to changes in the purchasing-power-parity forecast error, twenty-two are negative and sixteen are positive. Two of these coefficients are significant at the 5 percent level; both are negative and both are for regressions using M2 as the measure of money. Only two of the twenty-four regressions have R^2s above .08. There is little support here for the notion that the errors in purchasing-power parity are related to domestic monetary conditions. The alternative to the international determination of prices appears to work poorly.

2.7   Conclusion

The argument and evidence presented here make a pronouncement of the failure of purchasing-power parity impossible, and without a pronouncement of failure much of modern macroeconomics is badly damaged. The failure of purchasing-power parity must be large indeed to leave the Martian models and their empirical implementations unscathed, whether these are Keynesian, monetarist, or rationally expecta-

tionist. Under either the cleanly fixed exchange rates that typify the historical periods used to test the models or under the dirty float that typifies the years in which the conclusions thus tested have been used for policy, it is hard to believe that foreign prices or interest rates did not matter. Yet the silence of most American macroeconomists on the role of

Table 2.2   **Purchasing-Power-Parity Forecast Errors For the U.S. as a Function of Measures of Excess Money Growth, 1880–1940**

Calculated Forecast Errors	Constant	Current Excess Money Growth	Lagged Excess Money Growth	R^2	D.W.
Implicit Deflator					
Canadian	0.018	0.109		0.007	1.905
	(0.038)	(0.166)			
	0.019	0.125	0.163	0.023	1.893
	(0.037)	(0.166)	(0.166)		
U.K.	0.087	0.040		0.000	2.160
	(0.098)	(0.429)			
	−0.089	−0.080	0.417	0.016	2.108
	(0.098)	(0.431)	(0.431)		
CPI					
Canadian	0.026	0.077		0.001	2.208
	(0.060)	(0.263)			
	0.028	0.112	0.360	0.033	2.327
	(0.060)	(0.262)	(0.262)		
U.K.	−0.117	0.154		0.001	2.481
	(0.117)	(0.512)			
	−0.117	−0.167	0.132	0.003	2.484
	(0.118)	(0.519)	(0.519)		
WPI					
Canadian	0.489	−1.114		0.106	1.609
	(0.096)	(0.421)			
	−0.488	−1.126	−0.123	0.107	1.621
	(0.097)	(0.427)	(0.427)		
U.K.	−0.404	−0.343		0.005	2.100
	(0.139)	(0.610)			
	−0.402	−0.304	−0.406	0.013	2.124
	(0.140)	(0.615)	(0.615)		

Sources: Canada: Deflator, Firestone 1958; CPI, Urquhart and Buckley 1965, tables J165, 139; WPI, ibid., table J34. U.K.: Deflator, Feinstein 1971; CPI, McCloskey and Zecher 1976; Mitchell 1975, series 12; WPI, Mitchell 1975 series 11; Board of Trade series spliced at 1919–20 and 1929–30. U.S.: Deflator, U.S. Bureau of the Census 1975, series F5; CPI, ibid., series E 135; WPI, ibid., series E 23; money supply and gold flows, Friedman and Schwartz 1963, pp. 704–7; real GNP, McCloskey and Zecher 1976; U.S. Bureau of the Census 1975, series F 3.

Note: Values in parentheses are the standard errors of the coefficients.

Table 2.3   Purchasing-Power-Parity Forecast Errors For the U.S. as a
                     Function of Measures of Excess Domestic-Credit Growth, 1880–1940

Calculated Forecast Errors	Constant	Current Excess Domestic-Credit Growth	Lagged Excess Domestic-Credit Growth	R^2	D.W.
Implicit Deflator					
Canadian	0.004	−0.002		0.001	1.719
	(1.175)	(0.010)			
	−0.410	−0.004	0.144	0.026	1.726
	(1.223)	(0.010)	(0.124)		
U.K.	−0.674	−0.008		0.015	2.564
	(1.069)	(0.009)			
	−1.014	−0.010	0.119	0.035	2.632
	(1.116)	(0.009)	(0.113)		
CPI					
Canadian	0.146	−0.003		0.003	1.912
	(0.756)	(0.006)			
	0.012	−0.003	0.047	0.010	1.926
	(0.795)	(0.006)	(0.080)		
U.K.	−0.094	−0.010		0.016	1.694
	(1.270)	(0.010)			
	−0.233	−0.010	0.048	0.019	1.726
	(1.338)	(0.011)	(0.135)		
WPI					
Canadian	−0.405	−0.006		0.023	2.307
	(0.609)	(0.005)			
	−0.088	−0.004	−0.110	−0.076	2.246
	(0.264)	(0.005)	(0.063)		
U.K.	−0.012	−0.008		0.026	1.937
	(0.835)	(0.007)			
	0.098	−0.008	−0.038	0.029	1.898
	(0.880)	(0.007)	(0.089)		

Sources: See table 2.2.

Note: Values in parentheses are the standard errors of the coefficients.

the rest of the world in their models implies such a belief. It is hard to believe that American prices and interest rates are not at all constrained directly by the forces of arbitrage. Yet the journals are filled with work embodying this belief. The failure of arbitrage necessary to validate a Martian model must be gross, not a matter of the fourth digit of accuracy but of the first. To the first digit of accuracy, and even to the second, the hypothesis of parity succeeds. One wonders what would happen to estimates of wage and price equations, or of the effects of domestic monetary policy on prices and interest rates, or of optimal economic policy under

rational expectations, if they were each asked to embody the international milieu of the U.S. economy to the second digit of accuracy.

The hypothesis of parity survives the test for the reasons usual in economic arguments. Economists are embarrassed to assert in print that they possess the economic equivalent of a perpetual-motion machine, and the gross violations of rationality that opponents of purchasing-power parity believe they see entail such a machine. General equilibrium makes the hypothesis still more robust; even nontraded goods are substitutes in consumption and production with traded goods; a few characteristics are tradable in markets even when named goods are not; and the tendency for parity to hold in the long run gives opportunities for speculative profits if it does not hold in the short.

The standards by which parity "fails" empirically are unclear, as many opponents of the hypothesis readily admit. The literature contains no articulation of standards. The introduction of standards casts into doubt all the recent attacks on parity. By standards that make intellectual sense the hypothesis succeeds. It succeeds in explaining the U.S. price level from 1880 to 1940 to a standard of accuracy demanded of such explanations.[16] It succeeds in explaining the price level in turbulent periods by the standard of tranquil periods. It succeeds in explaining the difference in prices among countries by the standard of the difference in prices among places in a single country. One of its competitors, closed-economy monetarism, fails to explain the residual deviations from parity by the standard of statistical fit. Another of its competitors, the elasticity approach, fails to explain the balance of trade by the same standard, although again we express our doubt that much can be inferred from the uncontrolled experiments in curve fitting that characterize the literature and that we have dutifully followed here. And by the standard of good storytelling that underlies all economics, applied to episodes from their beginning to their end, the hypothesis of parity explains what happened under the gold standard. Purchasing-power parity is not a failure. On the contrary, by the standards we have examined, it is a great success. And at the least, speaking to the most skeptical reader, it is not so great a failure that macroeconomics can go on ignoring the rest of the world.

Notes

1. See R. Dornbusch and D. Jaffee's (1978) introduction to the special issue of the *Journal of International Economics* on purchasing-power parity. They remark that the Kravis and Lipsey paper in the issue leaves purchasing-power parity "rather in a shambles" (p. 159).

2. Samuelson, the chief theoretician of the real theory, however, is an implacable foe of purchasing-power parity.

3. Gottfried Haberler once described Cassel's work as "one part Wicksell and nine parts water." One can also read the *General Theory* as taking prices to be given (not constant), because given by international factors such as purchasing-power parity.

4. Friedman and Schwartz (1982, p. 318) assume no quick operation of price–specie flow. Were it as quick as in the monetary approach, many of their other conclusions would be wrong, especially the effect of domestic money on prices.

5. Prices might have risen in anticipation of gold flows, though such rationality runs counter to the usual price-specie-flow argument.

6. See Darby 1982. Kravis and Lipsey (1978) make a similar point, arguing that the drift of purchasing power from parity over periods of a decade or so shows that parity is false.

7. Roll says further that "In an efficient market, something so easy to detect and so intoxicating to the arbitrager as a relative price difference in two locations would presumably display an immeasurably short half-life" (p. 136).

8. One of us (McCloskey 1981, chaps. 4, 6) has in fact done similar work with the notion that English businessmen failed to adopt profitable novelties in the late nineteenth century. The task is to see whether the second-guesser could have done better, recognizing the limitations of resources they faced (including a limitation on prescience). It is worth knowing for the present context that the second-guessing showed that the businessmen knew what they were doing.

9. The source for the trade statistics is U.S. Bureau of the Census 1960 based on declarations to American customs. The period covered is 1880–1940.

10. A good—or bad—example is the paper by J. D. Richardson (1978). Richardson regresses Canadian on U.S. prices (multiplied by the exchange rate) for a number of industries and concludes: "it is notable that the 'law of one price' fails uniformly. The hypothesis of *perfect* commodity arbitrage is rejected with 95 percent confidence for *every* commodity group" (p. 347).

11. To which may be added an authority overlooked by Kravis and Lipsey, Leland B. Yeager (1958). Yeager reckoned that if twenty-six out of thirty-five countries were within 25 percent of their 1937 parity by 1957, the hypothesis is confirmed.

Richard Caves and Ronald Jones (1973) may be added to the affirmation. They remark on Yeager's results that "this performance seems rather good—just how good it is hard to say" (p. 338). The rueful remark, "just how good is hard to say," illustrates well how urgent it is to develop some standard. The state of play, largely favorable to purchasing-power parity, is well described in a comprehensive review by Lawrence H. Officer (1976).

12. In another article, Frenkel (1981), concludes that the "collapse" of PPP during the 1970s was due to a fundamental difference between exchange rates and prices. "Exchange rates reflect expectations about *future* circumstances while prices reflect more *present* and *past* circumstances" (p. 162). In this view, arbitrageurs in commodity markets look backwards in time while arbitrageurs in financial markets look forwards. For a different view, see Nattress and Zecher 1982, where a theory of the arbitrageur is developed.

13. Genberg writes (1976, p. 302): "it is evident . . . that the differences between OECD countries are no greater on the average than those between cities within the United States. Thus, if we believe that the whole of the U.S. can be treated as a single market in a macroeconomic context, then the area composed of the above countries can be treated likewise."

14. Friedman and Schwartz speak of 1933–37 because of their commitment to analysis by cycles, 1933–37 being an upswing. Notice that by all measures except the consumer price index most of the price rise had occurred before 1935. The discussion of indirect influence of devaluation occurs at any rate on pp. 465ff, at the beginning of which a direct influence on "most farm products and raw materials exported by the United States" is mentioned.

15. *BLS Handbook of Labor Statistics, 1937*, pp. 512ff. The report referred to was "Report on the Operations of the National Recovery Act," p. 7.

16. Another related standard that we shall explore in later work is that of the relative

convergence of prices over time. Spooner's magnificent graphs (Braudel and Spooner 1967, pp. 470–71) show ranges of wheat prices expressed in silver in Europe and its offshoots of 6.66 to 1 around 1400, falling steadily to 1.88 to 1 around 1750. The divergences of the late nineteenth century, not to speak of the twentieth, look trivial beside these. Correspondingly, fixation on the "failure" of the unity of world markets for the period 1880–1939 looks odd indeed.

References

Braudel, F. P., and Frank C. Spooner. 1967. Prices in Europe from 1450 to 1750. In *The Cambridge Economic History of Europe*, ed. E. E. Rich and C. H. Wilson, vol. 4. Cambridge: Cambridge University Press.

Caves, Richard, and Ronald Jones. 1973. *World trade and payments.* Boston: Little Brown.

Darby, Michael R. 1982. Does purchasing power parity work? In *Proceedings of Fifth West Coast Academic/Federal Reserve Economic Research Seminar*, December 1981. San Francisco: Federal Reserve Bank of San Francisco.

Dornbusch, Rudiger, and Dwight Jaffee. 1978. Purchasing power parity and exchange rate problems: Introduction. *Journal of International Economics* 8 (May): 157–61.

Einzig, Paul. 1937. *The theory of forward exchange.* London: Macmillan.

Feinstein, Charles H. 1971. *National income, expenditure, and output of the United Kingdom, 1855–1965.* Cambridge: Cambridge University Press.

Firestone, O. J. 1958. *Canada's economic development, 1867–1913.* Income and Wealth series V11. London: Bowes and Bowes.

Frenkel, Jacob A. 1978. Purchasing power parity: Doctrinal perspective and evidence from the 1920s. *Journal of International Economics* 8 (May): 169–91.

————. 1981. The collapse of purchasing power parities during the 1970s. *European Economic Review* 16 (May): 145–65.

Friedman, Milton, and Anna J. Schwartz. 1963. *A monetary history of the United States, 1867–1960.* Princeton: Princeton University Press.

————. 1982. *Monetary trends in the United States and the United Kingdom, 1867–1975.* Chicago: University of Chicago Press.

Gandolfi, Arthur E., and James R. Lothian. 1982. A comparison of price movements among countries and states. Citibank. Typescript.

Genberg, A. Hans. 1976. Aspects of the monetary approach to balance-of-payments theory: An empirical study of Sweden. In *The monetary approach to the balance of payments*, ed. J. A. Frenkel and H. G. Johnson. Toronto: University of Toronto Press.

————. 1978. Purchasing power parity under fixed and flexible exchange rates. *Journal of International Economics* 8 (May): 247–76.

Kravis, Irving B., and Robert E. Lipsey. 1978. Price behavior in the light of balance of payments theories. *Journal of International Economics* 8 (May): 193–246.

Krugman, Paul R. 1978. Purchasing power parity and exchange rates: Another look at the evidence. *Journal of International Economics* 8 (Aug.): 397–407.

McCloskey, Donald N. 1981. *Enterprise and trade in Victorian Britain: Essays in historical economics.* London: Allen and Unwin.

McCloskey, Donald N., and J. Richard Zecher. 1976. How the gold standard worked, 1880–1913. In *The monetary approach to the balance of payments. See* Genberg 1976.

Mitchell. B. R. 1975. *European historical statistics.* New York: Columbia University Press.

Mood, Alexander M., and Franklin A. Graybill. 1963. *Introduction to the theory of statistics.* 2d ed. New York: McGraw-Hill.

Nattress, Dayle, and J. Richard Zecher. 1982. The theory of the arbitrageur and purchasing-power parity. Chase Manhattan Bank. Typescript.

Neyman, Jerzy, and E. S. Pearson. 1933. On the problem of the most efficient tests of statistical hypothesis. Royal Society of London. *Philosophical Transactions* 231 (series A): 289–337.

Officer, Lawrence H. 1976. The purchasing-power-parity theory of exchange rates: A review article. *IMF Staff Papers* 23 (Mar.): 1–60.

Richardson, J. David. 1978. Some empirical evidence on commodity arbitrage and the law of one price. *Journal of International Economics* 8 (May): 341–51.

Roll, Richard. 1979. Violations of purchasing power parity and their implications for efficient international commodity markets. In *International finance and trade,* ed. M. Sarnat and G. Szego. Cambridge: Ballinger.

Taussig, Frank W. 1918. International freights and prices. *Quarterly Journal of Economics* 32 (Feb.): 410–14.

U.S. Bureau of the Census. 1960. *Historical statistics of the United States, colonial times to 1957.* Washington, D.C.: Government Printing Office.

————. 1975. *Historical statistics of the United States, colonial times to 1970.* Part 1. Washington, D.C.: Government Printing Office.

Urquhart, M. C., and K. A. H. Buckley. 1965. *Historical statistics of Canada.* Toronto: Macmillan of Canada.

Wicksell, Knut. 1918. International freights and prices. *Quarterly Journal of Economics* 32 (Feb.): 404–10.

Yeager, Leland B. 1958. A rehabilitation of purchasing-power parity. *Journal of Political Economy* 66 (Dec.): 516–30.

Comment   Robert E. Lipsey

How to define purchasing-power parity and how to test its "success" is the central issue of McCloskey and Zecher's paper. The authors come to two conclusions about purchasing-power parity, both favorable. One conclusion is derived from their reasoning and one from their tests. The first conclusion is that purchasing-power parity works instantaneously and perfectly, immediately eliminating any genuine differences in prices or divergences in relative price movements. The second is that the theory works sufficiently well that foreign influences on prices cannot be ignored. I shall argue that the first conclusion, as stated, depends on a definition of purchasing-power parity that makes it a tautology, not susceptible to proof or disproof, and that the second conclusion is valid but would not be disputed by many "critics" of the theory. I shall argue in addition that the strongest test they propose has been performed, with results that are unfavorable to the theory by their own criterion.

The Definition of Purchasing-Power Parity

In discussing the meaning of the concept, the authors say that they view purchasing-power parity as resulting from arbitrage in commodity markets. To me that means they identify the concept with the operation of the law of one price rather than with vaguer notions of aggregate country price levels or price changes. For example, "purchasing-power parity is a consequence of rationality in arbitraging. If all the opportunities for riskless (or insured) arbitrage among countries that are profitable at existing interest rates and other costs of arbitrage have taken place, then the price level of the world may be said to have exhausted its ability to determine the price level of one country" (p. 127), and "If opportunities for arbitrage are exploited (allowing fully for the cost of transport and information), then the price level of one country is fixed by the rest of the world, even in the very short run" (p. 129).

Two things can be said about this definition or "theorem" or "higher-order proposition" as the authors refer to the latter version. One is that it is not what most people mean by purchasing-power parity. A second is that, stated this way, it is only a definition, not a theory, and is not susceptible to testing. The authors make no effort to apply it empirically. If one includes and takes seriously all the qualifying phrases about costs of arbitrage over space and time, most of which are not measured and probably cannot be measured, it is impossible to show that purchasing-power parity has been violated. The price of product x in country a plus *all* costs of arbitrage (including information costs, advantages of continuity of supplier relationships, costs of adaptation of existing machinery and

Robert E. Lipsey is professor of economics at Queens College of the City University of New York and research associate of the National Bureau of Economic Research.

work habits, etc.) equals the price of product x in country b. That statement is an identity; it tells us nothing about how the world works. By this standard, prices on Mars and Venus satisfy purchasing-power parity even though there is no communication between them. The information and arbitrage costs are so high that any prices or price changes satisfy the authors' criterion: no one can make money by arbitrage between the two planets.

Critics of purchasing-power parity do not deny this tautology. What they argue is that adjustment costs are large in some cases, that it may take a long time to overcome them, that prices can move relative to each other in the meantime, and that consumers and producers react to these price differences and relative price movements.

The authors illustrate their point about arbitrage with examples of changing price differences for electrical equipment, cement, and vacuum cleaners. It is clear from these examples that they picture the mechanism often called the law of one price, referring to prices of individual commodities, as enforcing purchasing-power parity. They do not give vivid examples of arbitrageurs buying the U.S. CPI or WPI or GNP deflator and selling that of the U.K. Of course, if the law of one price operated exactly and instantaneously, that is, if prices of carefully defined individual products were identical everywhere or moved identically in different countries, the levels of prices in general and their movements would be similar or would move similarly. However, aggregate price levels and price changes would not be the same. Given the differences among countries in the composition of consumption and production and what are probably even greater differences in the way that aggregate price measures are constructed, one must be careful in moving from one kind of statement to the other. The authors take little note of this point and speak about the movements of vaguely defined and badly measured price aggregates in different countries and the responses to them as if they were useful for testing the opportunities for, or existence of, arbitrage. I do not think they are.

Testing Purchasing-Power Parity

The first test the authors propose is that a critic of purchasing-power parity show that he has gotten rich on arbitrage profits. "It surpasses belief that many opportunities to make easy money buying low and selling high persist long enough to be observed in economic data" (p. 128). "Often unrecognized by critics of purchasing-power parity is that their conclusion that it has failed usually implies an ability to make money. Anyone who knew that purchasing-power parity was true in some long run but not yet true in the present would have a rosy financial future. The divergences from purchasing-power parity detected in the literature are so gross and the statistics purporting to show the divergences so easy

to collect that the opportunities for profit are large. Go thee and prosper" (p. 131). "Economists are embarrassed to assert in print that they possess the economic equivalent of a perpetual-motion machine, and the gross violations of rationality that opponents of purchasing-power parity believe they see entail such a machine" (p. 147).

The number of times the point appears shows that the authors take it seriously, but I have never seen a criticism of purchasing-power parity that implied irrationality on the part of purchasers. There is strong evidence that in the capital goods and other complex products that form a large part of the trade of developed countries, price differences and divergences in price movements exist. They are not arbitraged away immediately, but they do eventually bring about shifts in trade that tend to remove them (an indication that they are not simply illusory or due to differences in specifications). The reasons are implicit in some of the authors' own discussion: information is costly and the risks of purchasing unknown types of machinery and dealing with unfamiliar suppliers are high and uninsurable.

A fall in the price of a Japanese machine might at first produce no shift in purchases because buyers were unaware of the change, uncertain about its permanence, or skeptical about the quality of the machine or the availability of spare parts. There would be no violation of the law of one price by the authors' definition, because the price difference was insufficient to offset information or risk or insurance costs. After the lower price had been in effect for a year, information would become more widely and cheaply available, risk and insurance costs would decrease, and some buyers would switch. After another year more information would be available and still more buyers would switch. At each point, the price in Japan plus information, risk, and insurance would have been equal to the price in the United States. Therefore this sequence of events, which would suggest a violation of the law of one price to most observers, would be in conformity with it by the authors' definition.

In view of the impracticality of calculating in retrospect all the costs of arbitrage, the authors go on to suggest a "more practical" test. That test is "whether the shipments of goods implied by the supposed opportunities for arbitrage in fact occurred. . . . A country that exhibits divergences from purchasing-power parity convincing to the doubters should also exhibit lower exports and increased imports . . . a place with 'high' prices would have a hard time selling and an easy time buying" (p. 131).

One might expect that the authors would perform their "practical" test on the data for commodities defined as narrowly as possible to observe the action or inaction of arbitrageurs. Instead, they perform their tests of responses to deviations from purchasing-power parity between the United States and the United Kingdom and the United States and Canada using aggregate price indexes. If they seriously believe that profit

seeking by arbitrageurs is the force at work, and they wish to observe that action, why measure purchasing-power parity from price aggregates? The arbitrageur cannot easily buy the U.S. wholesale price index or GNP deflator. A more appropriate test of the workings of arbitrage, and one the authors themselves suggest, would be to compare price and trade changes for identical or related goods such as the electrical equipment they refer to. As they say, "Further tests of the hypothesis that profitable opportunities for arbitrage arise when measured prices diverge might proceed commodity by commodity."

In fact, Irving Kravis and I, using price indexes that were constructed so that the same goods with the same weights were represented in two countries' price measures did perform many versions of the "practical" test they suggest, although not with the intention of testing purchasing-power parity. That is, we investigated whether price changes and differences in price levels did lead to shifts in trade. We found strong evidence that they did and that these shifts took years before they were completed. We explained the reasons for these lags in our book on price competitiveness (Kravis and Lipsey 1971) and in a series of later articles.[1]

The reasons we gave for the price and price-change divergences we found did not imply that there were overlooked opportunities for above-average profit, given the costs of information, costs of adjustment, and uncertainties about the permanence of price changes. It was true that the first U.S. buyers of foreign electrical generating equipment and foreign steel paid less than their competitors who hesitated. We cannot say what the rational policy was for a buyer of generating equipment given the possible uncertainties at the time. The first buyers were public systems which may have faced less danger from mistakes than private utilities. The private utilities may have had little incentive to lower costs, given the way their prices were regulated and the lack of incentives for managers to break cozy relationships with domestic suppliers. The important point is that whatever the reason, gaps in prices persisted for a long time and produced not sudden but gradual shifts in trade; but they did produce the shifts that McCloskey and Zecher imply would refute purchasing-power parity.

In looking back at these past episodes in which price differences gradually gave rise to shifts in trade, we do not know whether the first U.S. buyers of foreign electrical generators and transformers or the first U.S. buyers of foreign steel reaped exceptional profits for a time. Even if they did, the profits, as usually measured, may have done no more than compensate for the risks, given the uncertainties about the quality of foreign products, the reaction of the U.S. government and regulatory agencies, the commitment of foreign suppliers to technical assistance, service, and continuity of supply, and many other factors. Thus these events may not represent a violation of purchasing-power parity by their

definition (it would be difficult to do that), but fail their test of conformance with the theory by the "practical" test.

If, as stated in the text (p. 127), the hypothesis has to do with the rationality of arbitrageurs, the tests they offer are irrelevant. I know of no writer skeptical of purchasing-power parity who has stated or implied or assumed that the reason for deviations from it is that traders ignore profit opportunities. They believe there have been deviations from purchasing-power parity because it was *not* profitable to exploit all opportunities or to exploit them immediately, given the costs and uncertainties involved. One does not add to understanding of the international economy by assuming away all these deviations from purchasing-power parity.

The paper includes several of the authors' own tests, but these are tests of what I called at the beginning of my comments their second conclusion, or the second version of purchasing-power parity. That second, or weak, version is that "it is hard to believe that foreign prices . . . did not matter" or "that American prices are not at all constrained directly by the forces of arbitrage" (p. 146). They test this version by regressions of the U.S. GNP deflator against aggregate Canadian and British price indexes adjusted by exchange rates, and by plots of the U.S. price indexes predicted by Canadian and British prices and of the prediction errors from the equations. The correlations are "high," they conclude, and "the foreign prices do predict the gross outlines of the American price." Furthermore, they report, the equations predict as well in the turbulent period 1921–40 as in the "tranquil" period 1880–1914. Aside from the point that these equations represent the "uncontrolled experiments in curve fitting" the authors are so critical of in others, the equations are unconvincing in other ways. For example, for the first twenty years of the chart the Canadian price predicts nothing of U.S. price movements. We are then told that foreign prices did no worse as predictors of U.S. prices in turbulent times than in "tranquil" times. But for half of the "tranquil" period foreign price did not predict at all—not a very exacting standard for judging the predictive power of Canada's prices to estimate later U.S. prices. A more serious objection to these tests is that there is no consideration of the possibility that the high correlation stemmed from the effects of common factors, such as World War I and the Great Depression on U.S., Canadian, and British prices at the same time, rather than from the dependence of one country's prices on another's.

Judging the Results of Tests of Purchasing-Power Parity

A substantial part of the paper is devoted to standards by which to judge the results of tests of purchasing-power parity. The authors take many of their colleagues to task for using vague and ill-defined standards. I cannot see that they escape the same problem. Their paper is filled with the same undefined terms they deplore in others. One reason for their

vagueness and that of everyone else is that no one standard can be used for all purposes. The precision required by a speculator might be greater than that needed by a company making a long-term investment or by a government or international agency trying to set an exchange rate, and all of these standards may be far above that required for satisfaction by devotees of purchasing-power parity.

The authors' judgments of the results reported by others are clouded by the fact that they seem to conceive of only two possible conclusions. Either the theory is a "success" and is graded A + , or it is a failure and is graded F. They seem to be determined to come out with a grade of A + and to think that all analyses that find deviations from purchasing-power parity imply a grade of F. In fact, most of the studies they cite seem to imply judgments of B or C rather than total success or failure.

In summary, my reaction to the two tests of purchasing-power parity proposed by the authors is that one test cannot be failed and that the other test, conducted with the right type of data, usually is failed.

Fortunately, there is another theme to the paper, although it is obscured by the extravagant claims made for purchasing-power parity. It is the important and reasonable one that "it is hard to believe that foreign prices or interest rates *did not matter*. It is hard to believe that American prices and interest rates are *not at all constrained* directly by the forces of arbitrage," and "it [purchasing-power parity] is not so great a failure that macroeconomics can go on ignoring the rest of the world." If that is the point the authors really want to make, even many skeptics about purchasing-power parity could agree.

Note

1. Some of these were referred to in the article (Kravis and Lipsey 1978) quoted by McCloskey and Zecher; we might add to the list Kravis and Lipsey 1982. Of course, similar studies by many others of price elasticities in trade have yielded similar results.

References

Kravis, Irving B., and Robert E. Lipsey. 1971. *Price competitiveness in world trade.* New York: National Bureau of Economic Research.
————. 1978. Price behavior in the light of balance of payments theories. *Journal of International Economics* 8 (May): 193–246.
————. 1982. Prices and market shares in the international machinery trade. *Review of Economics and Statistics* 64 (Feb.): 110–16.

Comment   Milton Friedman

I believe that a fundamental confusion runs through the McCloskey-
Zecher treatment, both in this paper and in their earlier paper—a confu-
sion between two very different propositions. One proposition is whether
the quantity of money in a country is an endogenous or an exogenous
variable—an important and very interesting question. There is no doubt
that in a world of fixed exchange rates and stable barriers to trade, the
quantity of money in each country separately is ultimately an endogenous
variable. That proposition is perfectly clear and everybody has accepted
it for a long time. It's the proposition that Keynes presented so well in the
appendix to his *Tract on Monetary Reform* ([1923] 1971), that there are
three things of which a country can choose any two: stable internal prices,
stable exchange rates, and free trade. You can't have all three; you can
have only two. That proposition is critical.

It should be noted, however, that while the quantity of money is
ultimately an endogenous variable, there can be and is much leeway in
the short run, before the external forces overwhelm the independent
internal effects. And we have repeatedly been surprised in our studies by
how much leeway there is and for how long—frequently a number of
years.

There's a very different proposition that is easily confused with the
endogeneity or exogeneity of the quantity of money, namely, if money is
endogenous, there is no causal relation between money and prices. That
is a whole different proposition. Whatever may determine the quantity of
money within a country, that quantity of money may still largely deter-
mine—or at least, be the conduit through which other forces determine—
prices within a country. The confusion between these two wholly differ-
ent propositions is apparent in the statement by McCloskey and Zecher
that "if purchasing-power parity is found to be a useful characterization
of the world"—and they should have added "and fixed exchange rates
characterize the world"—"then closed-economy theorizing and empiri-
cal work in macroeconomics should be changed to allow for the direct
effects of international price arbitrage. Whether monetarists or Keynes-
ians or rational expectationists, economists should begin thinking and
measuring in global terms" (p. 127). Economists have consistently
thought and measured in global terms in examining the determinants of
the quantity of money in a country, both for periods of fixed exchange
rates and of dirty floating—and McCloskey and Zecher cite no examples
to the contrary. In respect of the second proposition, the money supply

Milton Friedman, Nobel laureate, is senior research fellow at the Hoover Institution,
Stanford University, and Paul Snowden Russell Distinguished Service Emeritus professor
of economics at the University of Chicago.

may be endogenous after a sufficient interval, yet not in the short run. And whether endogenous or exogenous, the domestic money supply is absorbed primarily within the domestic economy; it is highly fruitful to examine the process whereby that occurs and the relation within a country between changes in the money supply and changes in other variables. There is no such sharp dichotomy between "closed economy" and "global" thinking and measuring as the straw man set up by McCloskey and Zecher.

Some findings from Anna Schwartz's and my book *Monetary Trends* are relevant to this subject. We calculated correlations between the United States and the United Kingdom in the rates of change of various variables between cycle phases for almost a century. We pointed out, and this is strictly in accord with the McCloskey-Zecher view, that the correlation between prices in Britain and the United States is closer than the correlation between any other two magnitudes. It is closer than the correlation between money supplies in the two countries, closer than the correlation between income in the two countries, real or nominal. We stressed that that result is strictly consistent with what they call our Martian view of the economy, and indeed demanded by it. Because, we said, in a world of fixed exchange rates, the money supplies in the different countries have to accommodate themselves in such a way as to be consistent with equality of prices and goods among countries. And therefore, the relation between the quantity of money in the two countries would be expected to be less close than between prices because changes in the demand curve for money within an individual country must be accommodated by changes in money rather than in prices (Friedman and Schwartz 1982, pp. 310–15).

Let me turn to a couple of other points. First, McCloskey and Zecher assert that "the turbulence of the 1920s and 1930s, which is said to have loosened the economic ties among nations, appears not to have done so" (p. 139).

One comparison in our book supports a very different conclusion. We started with the Kravis, Heston, and Summers estimate for the purchasing-power-parity exchange rate between the United States and the United Kingdom in 1970. We used price indexes in the two countries to extrapolate the purchasing-power exchange rate annually from 1870 to 1970. We then calculated the ratio of the purchasing-power-parity exchange rate to the market exchange rate.

The resulting chart is fascinating (Friedman and Schwartz 1982, chart 6.5, p. 291). Before about 1932, the ratio of the purchasing-power-parity exchange rate to the market exchange rate varied within a range of plus or minus 10 percent. All of us would say that is a fairly close relationship to purchasing-power parity. After 1931, the range is between minus one-quarter and plus one-third—an enormously wider range. There is no

doubt that a comment by Jacob Frenkel is right: Hume traveling in the 1950s would have found more deviations from purchasing-power parity than he would have found in the 1780s. Exchange controls, tariffs, other impediments to trade were far more important in dividing the world than improvements in communication and other technologies were in uniting it. I don't see that notion is any contradiction to the purchasing-power-parity theory of exchange rates. It is simply a consequence of the fact that there has been an enormous increase in barriers to trade among countries since 1931.

I want now to discuss two particular episodes for which McCloskey and Zecher take Anna Schwartz and me to task for our analysis in *A Monetary History* (1963). The first example is resumption in 1879. They quote our statement that "it would be hard to find a much neater example in history of the classical gold-standard mechanism in operation" (p. 99). Their look at that episode on the basis of monthly data is interesting and most welcome, but on closer examination it does not, contrary to their claims, contradict our interpretation of the episode.

McCloskey and Zecher compare price rises to inflows of gold, concluding, "In the monthly statistics . . . there is no tendency for price rises to follow inflows of gold . . . ; if anything, there is a slight tendency for price rises to precede inflows of gold, as they would if arbitrage were shortcutting the mechanism" (p. 126).

Their comparison is the wrong one for determining whether prices were reacting to arbitrage rather than reflecting changes in the quantity of money. For that purpose the relevant comparison is with the quantity of money. Gold flows are relevant only as a proxy for the quantity of money. If we compare price rises with changes in the quantity of money directly, a very different picture emerges than McCloskey and Zecher draw (see table C2.1). Our basic estimates of the quantity of money for this period are for semiannual dates, February and August. Resumption took effect on 1 January 1879. From August 1878 to February 1879, the money supply declined a trifle, continuing a decline that had begun in 1875 in final preparation for resumption. From February 1879 to August 1879, the money supply rose sharply, according to our estimates, by 15 percent. The Warren-Pearson monthly wholesale price index fell in the first half of 1879, reflecting the earlier decline in the money stock. It started its sharp rise in September 1879, or at least seven months later than the money supply.[1]

As to gold, the total stock of gold, as well as gold held by the Treasury, had been rising since 1877 as part of the preparation for resumption. But it had been rising at the expense of other components of high-powered money, which actually fell slightly. However, the decline in the money stock before 1879 had been due primarily to a decline in the deposit-currency ratio and the deposit-reserve ratio. After successful resumption,

both ratios rose, which enabled the stock of money to rise despite no initial increase in gold flows. The large step-up in gold inflows in the fall of 1879, to which McCloskey and Zecher call attention, was mostly absorbed in raising the fraction of high-powered money in the form of gold rather than in speeding up monetary growth.

Table C2.1 shows these developments in more detail.

On the basis of this reexamination, I am inclined to repeat the statement in our *Monetary History*, "It *would* be hard to find a much neater example in history of the classical gold standard mechanism in operation" (italics added), at least on a sophisticated interpretation of both the gold standard mechanism and the historical data.

At any rate, the gold standard mechanism, as I understand it, has always incorporated a variety of channels of adjustment, subject to different lags. Any result one wants can be gotten, depending on the relative speed of adjustment of the various channels. It is an important scientific question to try to identify and isolate these relative speeds of adjustment. I believe that McCloskey and Zecher make a real contribution in examining aspects of that issue.

The second episode is the behavior of prices after the U.S. went off gold in 1933. The figure 2.3 shows a close parallelism between the weekly movements of wholesale prices and the exchange rate. However, the different scales used for the price index and the exchange rate in the chart give a misleading impression.[2] For example, from April 1933 to July 1933, wholesale prices rose less than a sixth, the exchange rate by nearly a half, yet the total impression from their chart is that prices rose more sharply. Logarithmic scales would give a more accurate picture and make clear

Table C2.1	Relations between U.S. Prices, Money, High-Powered Money, and Gold, 1878–80					
	Whole-sale Prices (P) (1)	Money Stock (M) (2)	High-powered Money (HPM) (3)	Gold Stock (G) (4)	M/HPM (5)	G/HPM (6)
Aug. 1878	90	1.57	0.767	0.182	2.05	0.24
Feb. 1879	88	1.55	0.752	0.198	2.06	0.26
Aug. 1879	86	1.78	0.815	0.219	2.18	0.27
Feb. 1880	105	1.94	0.897	0.302	2.16	0.34
Aug. 1880	97	2.05	0.972	0.378	2.11	0.34

Sources: Col. (1) U.S. Bureau of the Census 1949, app. 24, p. 344; col. (2) Friedman and Schwartz 1970, p. 5; col. (3) Friedman and Schwartz 1963, p. 799; col. (4) ibid., notes to table A-1, p. 723 and table A-3, p. 765.

Notes: Col. (1) Warren-Pearson index of wholesale prices (1910–14 = 100); col. (2) currency held by the public plus adjusted deposits of commercial banks (billions of dollars); col. (3) in billions of dollars; col. (4) in billions of dollars.

how much narrower the relative movement in the wholesale price index was than in the exchange rate. In any event, as McCloskey and Zecher note, we pointed out in *A Monetary History* that there was a direct effect of devaluation on prices. However, the existence of a direct effect on wholesale prices is not incompatible with the existence of many other prices, as Moe Abramovitz has remarked, such as non-tradable-goods prices, that did not respond immediately or responded to different forces. An index of rents paid plotted against the exchange rate would not give the same result. An index of wages would not give the same result.

It may be worth quoting what we actually said on the issue, especially in view of the McCloskey-Zecher comment on a quotation from the *Monetary History* that "this is an unusual line of reasoning for such crusaders against mixing up the determination of relative and absolute prices" (p. 141).[3]

Here is what we said: "The aim of the gold policy was to raise the price level of commodities, particularly farm products and raw materials. . . . Most farm products and raw materials exported by the United States had a world market in which this country . . . was seldom dominant. . . . Hence, the decline in the foreign exchange value of the dollar meant a roughly proportional rise in the dollar price of such commodities, which is, of course, what did happen to the dollar prices of cotton, petroleum products, leaf tobacco, wheat and similar items" (Friedman and Schwartz 1963, pp. 465–66).

Thirty-odd pages later, after noting that the rise in the implicit price index from 1933 to 1937 was of roughly the same order of magnitude as in 1879–82 and 1896– or 1897–99, but in wholesale prices decidedly larger: "What accounts for the greater rise in wholesale prices in 1933–37, despite a probably higher fraction of the labor force unemployed and of physical capacity unutilized than in the two earlier expansions? One factor, already mentioned, was devaluation with its differential effect on wholesale prices" (p. 498). This was followed by the passage McCloskey and Zecher quote in which we referred to "the implicit measures to raise prices and wages undertaken with government encouragement and assistance" (p. 498).

Contrary to the impression McCloskey and Zecher give, we did not try to assess the relative importance of various factors in explaining the rise in prices from 1933 to 1934—the period to which they limit their chart and discussion. On the contrary, we explicitly cited these measures as helping to explain the "rise in wholesale prices in 1933–37." The wholesale price index continued to rise after its initial sharp rise in 1933 and did not reach its peak until mid-1937 when it was 47 percent above its low point in February 1933 and 28 percent above its level in July 1933. Hence there was ample time for the factors we referred to to play their part after the enactment of the legislation we listed.

Note finally that McCloskey and Zecher have faced up in this paper to the problem of floating or flexible exchange rates to only a very minimal extent. This paper is written primarily for a world of fixed exchange rates, and indeed, fixed exchange rates with nonchangeable barriers to trade. To be applicable to the current world, those elements must be added.

Notes

1. This paragraph and the next two were added after the conference in revising my comments for publication.

2. This sentence and the next two were added in revising these comments for publication.

3. This paragraph and the next three were added in revising these comments for publication.

References

Friedman, Milton, and Anna J. Schwartz. 1963. *A monetary history of the United States, 1867–1960*. Princeton: Princeton University Press.

———. 1970. *Monetary statistics of the United States*. New York: National Bureau of Economic Research.

———. 1982. *Monetary trends in the United States and the United Kingdom, 1867–1975*. Chicago: University of Chicago Press.

Keynes, J. M. [1923] 1971. *The applied theory of money: A treatise on money*. Vol. 4 of *The collected writings of John Maynard Keynes*. London: Macmillan and New York: Cambridge University Press for the Royal Economic Society.

General Discussion

Abramovitz suggested another way of posing the question that lies at the heart of the McCloskey-Zecher paper: How can tradable-goods prices remain equal to one another or move in similar ways in different countries without destabilizing the gold standard, in the face of differing national and sectoral rates of productivity growth? Abramovitz pointed out that McCloskey and Zecher offer one possible adjustment mechanism. In contrast, the traditional specie-flow mechanism, involving changes in prices and nominal wages in different countries, offers a rather different mechanism.

McCauley asked McCloskey and Zecher to justify the leap from purchasing-power parity, however defined, to the assertion that monetary policy cannot alter prices. This assertion appears in weak form (p. 146)—that American prices are not at all constrained by the forces of

arbitrage—and in rather stronger form (p. 128)—that American price levels are exogenous to American monetary policy. Surely as the share of world output produced by a single country approaches unity, an individual country's monetary policy becomes capable of raising prices.

McCauley argued that these questions imply technical objections to McCloskey and Zecher's analysis. While McCloskey and Zecher analyze the residuals from their purchasing-power-parity equation, their procedure presumes that U.S. monetary policy is incapable of influencing price levels in the rest of the world.

ZECHER responded that the Zecher-McCloskey paper effectively incorporates McCauley's point. In response to Friedman, Zecher disagreed with his statement that McCloskey and Zecher fail to take into account the difference between periods of flexible and fixed exchange rates. Under flexible exchange rates, McCloskey and Zecher's assertion is *not* that a country cannot affect its own price level or rate of inflation, but rather that relative prices, or the deviation from purchasing-power parity, is constrained by arbitrage in commodity and other markets. Thus, their analysis is capable of dealing with flexible-exchange-rate periods.

FRIEDMAN restated the central point of his argument: suppose a country's money supply is endogenous, determined by the outside world. One can still examine the relationship between the quantity of money in that country (call it Illinois) and the price level and nominal income in Illinois. The change in the quantity of money, however produced, has effects internal to Illinois.

McCLOSKEY responded that all monetarists share a belief in a stable demand for money. But to go from the presumption that money demand is stable to the assertion that money supply in Illinois determines prices and interest rates in Illinois is a large jump.

FRENKEL raised the question of what exactly McCloskey and Zecher mean when they speak of purchasing-power parity? He suggested that McCloskey and Zecher essentially mean the law of one price. Purchasing-power parity is enforced by the mechanism of commodity and asset arbitrage, converting the whole discussion of purchasing-power parity into a discussion of financial flows and profit opportunities.

Frenkel suggested posing a very different question, which was the original question underlying the development of purchasing-power parity: How can one determine an appropriate exchange rate for the period following a serious market dislocation? How much information can be obtained from aggregate price indexes? One issue is which "aggregate" to look at. Needless to say, this question is based on the presumption that aggregates provide useful information for determining equilibrium exchange rates; changes in relative prices call this view into question. The crucial question, therefore, is under what conditions it is likely that aggregate price indexes will provide useful information about equilib-

rium exchange rates? It may be very important to know whether the shocks to the system originate from the real side or from the monetary side.

In the wake of monetary shocks, Frenkel argued, it is advisable to focus on the aggregate that best represents monetary conditions. This view suggests looking at the price indexes that cover the broader domain of goods and services, which is what Cassel had in mind. In his view, purchasing-power parity was not a theory of individual prices but of price indexes because it was intended as a measure of the monetary conditions. His view is fundamentally different from the view that foreign exchange rates have nothing to do with aggregate price levels but only with individual commodity prices.

Frenkel pointed out that the original view of purchasing-power parity refers to the ratio of rates of inflation of purchasing power, where inflation is to be understood as inflation of the quantity of money. These concepts are completely divorced from individual commodity prices per se.

Frenkel made a number of points concerning McCloskey and Zecher's econometric results. One interesting exercise in their paper asks whether deviations from purchasing-power parity have real effects? McCloskey and Zecher choose to concentrate on the trade balance; they ask if there is a visible, statistically significant relationship between apparent deviations from purchasing-power parity and the trade balance. They find no such evidence. Frenkel hesitated to infer from these results any particular conclusion about purchasing-power parity, since it is not clear that changes in relative prices should always have a particular trade-balance effect. If one thinks of the current account as the difference between income and spending, then there is a determinate theory that links changes in the relative prices of commodities to the aggregate saving ratio and hence to the current account. There is a determinate link between the terms of trade and the current account—the so-called Laursen-Metzler effect. However, under a variety of plausible circumstances, the Laursen-Metzler effect might not hold.

ZECHER responded to Frenkel by emphasizing the importance of questioning the extent to which national markets are integrated. The problem with many recent empirical studies, he suggested, is that after concluding that purchasing-power parity fails, subsequent theorizing simply neglects the rest of the world. It is important to attempt to define criteria that permit one to label markets as more or less integrated.

PIPPENGER pointed out that several economists have analyzed the residuals from purchasing-power parity calculations. He himself had examined deviations from purchasing-power parity using annual data, going back in some cases to the 1870s. The evidence indicates that deviations follow a random walk. There appears to be no tendency for

relative prices to return to any normal long-run level. This random-walk property holds even for the regional price indexes within the United States, which is indeed curious. Of course, this result is inconsistent with the continuous maintenance of purchasing-power parity, because the result permits exchange rates to drift in any direction. It also raises some interesting questions about the proper way of interpreting time-series data. One possible interpretation is that purchasing-power parity simply fails to hold. Another, suggested by Richard Roll, is that the random-walk property is evidence of efficient commodity markets. This latter interpretation would suggest that Zecher and McCloskey are right, but for a different reason. What we may be seeing is simply the fact that price indexes for different countries are made up of different commodity bundles. Many problems must be sorted out before we can distinguish between the traditional view and the Zecher-McCloskey view.

BRUNNER argued that one may wish to distinguish a shorter run, perhaps up to one-and-a-half or two years, over which the money stock is exogenous, and a longer run over which it is endogenous. Support for this distinction can be found in the history of the Italian monetary affairs in the 1960s. At that time there were a number of one-and-a-half- or two-year periods when macroeconomic accelerations were fueled essentially by the domestic-credit component of the monetary base. Such credit creation was able to alter the money supply for one-and-a-half or two years with accompanying adjustments in prices. Only thereafter was the balance of payments affected. Such lagged responses are quite consistent with some of the adjustment mechanisms sketched here.

DORNBUSCH suggested that McCloskey and Zecher had provided insufficient room in their framework for the considerations emphasized by Abramovitz. When an economy is growing and the composition of activity is changing over time, simple tests of purchasing-power parity will be biased. For example, anyone who tests the purchasing-power-parity hypothesis for the last twenty years will find that real exchange rates in manufacturing, of the United States as well as of any other industrialized country, are well explained by differentials in sectoral growth rates but not by national rates of price inflation. Therefore, serious tests of purchasing-power parity must incorporate time trends or other variables designed to account for differentials in sectoral growth rates and other real changes.

Dornbusch also drew attention to the financial research of the last three years, which demonstrates that one cannot reject the hypothesis that the stock market deviates for long periods from market fundamentals. On purely statistical grounds, even fifty years of Dow-Jones data are incapable of rejecting, at a .99, a .95 or even a .90 level of confidence, the hypothesis that the stock market is driven for long periods by fads and fashions.

MCGOULDRICK asked whether empirical tests of purchasing-power parity should include in the price indexes not only the prices of goods but the prices of securities.

MCCLOSKEY agreed that in principle the capital market should be considered in studies of purchasing-power parity. He pointed out that Lipsey takes the McCloskey-Zecher analysis of purchasing-power parity as an empirical test and argues that it is not in fact properly interpreted as a statistical test of a hypothesis. Rather, the analysis is better thought of as a way of looking at particular episodes which might persuade people of the plausibility of a particular view of how markets function.

McCloskey suggested that Friedman was in substantial agreement with the authors' main point. Friedman concedes that under fixed exchange rates the money supply of Illinois does not determine prices and interest rates in Illinois. That was McCloskey and Zecher's main point.

FRIEDMAN referred back to some of the work in his recent book written with Anna Schwartz. Friedman and Schwartz discovered an appreciable difference in the relation between interest rates in the United States and United Kingdom before 1896 and after 1896. Interest rates in the United States before 1896 are much higher relative to the British interest rates than after 1896. In other words, before-1896 interest rates act as if there was widespread anticipation of a depreciation of the U.S. dollar. After 1896, they act as if there was widespread expectation of an appreciation of the U.S. dollar. This behavior bears on the question of whether there can be significant deviations in prices and interest rates in various countries over substantial periods of time.

FRATIANNI suggested parallels with the Italian experience. In Italy, he argued, systematic deviations from purchasing-power parity are matched by deviations from interest-rate parity. Also, the way in which governments finance budget deficits should be accounted for in regressions.

GREGORY reported an experiment conducted by himself and colleagues (Baltagi and Sailors) at the University of Houston. They estimated a three-equation model, pooling thirty-four cross-section time-series observations for seven countries (France, Germany, Japan, Russia, Sweden, the United Kingdom, and the United States), to investigate the working of the gold standard. The model included a money-demand equation, a balance-of-payments equation, and a money-supply equation, with all variables expressed as first differences. The pooled regression results suggest that the classical gold standard was a fairly simple system, not the complex one that Bloomfield reported. According to the model, if supply shocks were to set off domestic inflation, domestic inflation would then cause a worsening of the external balance, and a worsening of the external balance would cause the domestic money supply to drop. While the pooled model yielded these statistically significant and plausible results, the individual-country time-series regressions

yielded generally insignificant coefficients. Gregory suggested that pooled data might serve as an avenue for expanding the empirical data set in exploring the working of the classical gold standard.

THOMAS raised a major question with respect to the model that Gregory described. An outstanding characteristic of the period 1880–1913 was the unique role of Great Britain in the Atlantic economy, as major capital exporter and center of what was virtually a sterling standard; her status was different in kind from that of the borrowing countries of the periphery. In 1880 the United Kingdom was responsible for 41 percent of world exports of manufactured goods as against 3 percent for the United States; even as late as 1889 the U.K. share of world exports of capital goods was 44 percent as compared with the U.S. proportion of 23 percent. A unique feature of the growth process was the fact that the long swings in capital formation, productivity, and real income in the center country, Great Britain, were inverse to the corresponding long swings in the borrowing countries of the periphery. The voluminous evidence confirming the validity of these inverse long swings was summarized by Arthur Bloomfield in his well-known *Patterns of Fluctuation in International Investment before 1914* (1968). No account of the working of the international gold standard can afford to neglect these special features of the pre-1913 period.

Five of the seven countries on which the pooled model was tested— Germany, France, Sweden, Russia, and Japan—were not Atlantic-oriented either in trade or foreign lending, so that only two—Great Britain and the United States—reflected the special characteristics of the pre-1913 international economy, namely, the center-periphery interaction. However, the way the model was specified made it impossible to pick up this interaction. The pooled results have drowned the peculiarity of Britain's interaction with the United States in general averages for the seven countries.

The working of the international gold standard between the "regions" of the nineteenth-century Atlantic economy has a close resemblance to the working of the internal gold standard between the regions of the United States. Under this internal standard the ease of adjustment was greatly facilitated by the existence of two fundamental conditions—free interregional migration of labor and the transfer of Treasury funds into weak regions. Among the most important reasons why the international gold standard worked fairly smoothly were, first, the high degree of international mobility of labor and, second, the fact that Britain, the dominant creditor, with a high propensity to import, was always putting money back into international circulation, either through a substantial upswing in imports or through a substantial upswing in foreign lending.

Thomas also commented on the McCloskey and Zecher paper. Proceeding from the analogy between the international and the national

economy, McCloskey and Zecher emphasize what they regard as the necessary corollaries—unified prices of products, assets, and labor across national boundaries. Hence the monetary theory, which is an equilibrium model; but the theory overlooks some awkward facts. One has to recognize the nature of the growth process in the pre-1913 Atlantic economy, particularly the inverse relation between investment upsurges in the center country and in the overseas country of new settlement. Export capacity in a given phase of the long swing was a function of the infrastructure investment in that country in the previous swing. There was a long-run symbiotic relationship, but it necessarily entailed opposite movements at the center and the periphery, and serious disequilibria when the peaks of the long swing were reached and the Bank of England had to protect its reserve. This process would occur particularly when under-effected transfer was experienced. The U.S. trade balance determined the gold flow, and the gold flow determined the rate of growth of the money supply. There is no basis for the notion that investment upswings, by generating excess demands, attracted net capital inflows that more than offset the unfavorable trade balance, thereby inducing gold inflows. Gold inflow, and as a consequence the money supply, rose most rapidly in the phases of the long swing when U.S. exports were surging upwards and infrastructure expenditure and imports were in a downswing. Simultaneously Britain was having an upswing in home investment, her exports as a proportion of imports were falling, and there was an external flow of gold from the Bank of England. When it was the turn of the United States to have its upswing in investment, her trade balance deteriorated, gold flowed out, and the rate of growth of the money stock fell.

See what happened during the 1890s. Between June 1892 and June 1896 there was an absolute fall in the U.S. money stock, the first such decline since the 1870s, whereas the Bank of England's reserve increased spectacularly from £15 million to no less than £49 million. When the United States was struggling desperately to stay on the gold standard, Britain was enjoying such a surfeit of liquidity that the market rate of discount was below 1 percent. In the second half of the 1890s the reverse happened as a result of the massive upsurge in United States exports in relation to imports coinciding with the opposite in Britain. The Bank of England reserve as a proportion of liabilities fell almost as fast as it had risen, while the money stock of the United States went up by 52 percent.

The Old Lady of Threadneedle Street was not managing the international gold standard: she was just minding her own business and doing it on an investment in gold stocks inexcusably small in relation to her responsibilities. Her status as central bank of the center country endowed her with clout. McCloskey and Zecher are scornful of Keynes's description of the Old Lady as "conductor of the international orchestra"; they

regard her as "no more than the second violinist, not to say the triangle player, in the world orchestra" ("How the Gold Standard Worked, 1880–1913," in *The Monetary Approach to the Balance of Payments*, ed. J. Frenkel and H. G. Johnson. [London: Allen and Unwin, 1976], pp. 358–59). As a superb understatement that must constitute something of a record.

EICHENGREEN elaborated upon one of Brinley Thomas's points. The purpose of pooling national time series and of attempting international comparisons is to extract as much information as possible from historical data. This approach is predicated upon the assumption that the structural relationships under consideration are identical across countries. In the case of monetary relations under the classical gold standard, there is considerable historical evidence of the existence of important structural asymmetries that would call into question the validity of this assumption. Eichengreen drew attention to the work of Triffin and others that pointed to the unique degree of market power exercised by the Bank of England under the classical gold standard and to asymmetries in the impact that changes in the monetary conditions in different countries had on the balance of payments of the countries participating in the system. For example, changes in monetary conditions in Britain appear to have had a much more powerful impact on short-term capital flows than did comparable changes in monetary conditions abroad. As Triffin suggests, the Bank of England had an ability to influence international gold flows unrivaled by other central banks. Asymmetries of this sort are not taken into account in Gregory's analysis.

ABRAMOVITZ pointed out that there is no obvious connection between long swings, such as fifteen-to-twenty-year Kuznets cycles, and international gold movements. This is not surprising, since many factors can substitute for the actual movements of gold: the growth of high-powered money from domestic sources, changes in high-powered money multipliers, and changes in the income velocity of money. What, then, produced the long swings in high-powered money that parelleled so closely the long swings in the growth rate of real output? Abramovitz suggested focusing on the growth rate of the sum of exports and capital imports. That sum traces long swings that parallel the long swings in the growth rate of nominal and real incomes in the United States.

This line of inquiry suggests a further question: are exogenous changes in the growth rate of the sum of exports plus capital imports driving the growth rate of the money supply and of nominal income? Or is the growth rate of nominal income determined by independent changes in real output and world prices to which the money supply of a country must adjust?

The latter is recognizably the view of the monetary approach to the balance of payments, which is arguably the right view when one has in

mind very long periods of time. In shorter periods, however, and even in the long swings that run across ordinary business cycles, the answer is much less clear. In the long swings of income in the United States, there appears to have been an interaction between the sum of commodity exports and capital imports which together constitute the positive elements underlying the balance of payments and the real-income changes with which they are associated. Independent movements of commodity exports or capital imports stimulated change in nominal income and in real income as well. The changes in real income, in turn, generated further increases in capital imports. Within limits, there was a self-sustaining cumulative process.

Abramovitz noted that the central point, raised by Brinley Thomas, concerns the inverse pattern of long swings in Britain and the United States. Abramovitz maintained that there were no inverse swings in real output or in nominal income in the aggregate in the United States and United Kingdom. In point of fact, the long upswings in the United States were matched by surges of capital export from Britain and capital import into the United States. Similarly, surges of exports from Britain were matched by surges of imports into the United States, and declines in home investment in Britain were matched by rises in home investment on the other side of the Atlantic. In Britain fluctuations in commodity and capital export offset one another and left the British economy growing smoothly over the business cycle, in contrast to the United States where capital imports and exports were not matched so closely by surges and declines in commodity imports and exports. These observations lead to a further question: Why did the long swings come to an end? Why did they result in the United States in serious depressions that culminated each of these episodes, while in Great Britain there were no comparable breakdowns that would have produced, had they occurred, the appearance of long swings in aggregate output as in the United States? One reason was the difference between the banking systems of the two countries. At intervals, the United States suffered banking and financial panics far more violent than those to which Britain was subject. A more severe impact on money and real income was felt in the United States. The two countries, however, differed in other respects as well, and the matter deserves a lot more study.

II. Technical Procedures: Rules of the Game