# 12 The Role of Money

One conclusion keeps rising to the surface throughout our lengthy examination of the role of money in the economic trends of the past century: John Maynard Keynes "shunted the car of Economic science on to a wrong line"—to use the words that William Stanley Jevons applied to an earlier brilliant economist, David Ricardo.[1] Keynes's *General Theory* was a reaction to the circumstances of the troubled interwar years. It offered an hypothesis to explain what seemed a conflict between experience and the implications of "orthodox" monetary theory. The hypothesis was

> the right kind of theory in its simplicity, its concentration on a few key magnitudes, its potential fruitfulness. [We] have been led to reject it, not on these grounds, but because [we] believe that it has been contradicted by evidence: its predictions have not been confirmed by experience. This failure suggests that it has not isolated what are "really" the key factors in short-run economic change.
>
> *The General Theory* is profound in the wide range of problems to which Keynes applies his hypothesis, in the interpretations of the operation of modern economies and, particularly, of capital markets that are strewn throughout the book, and in the shrewd and incisive comments on the theories of his predecessors. These clothe the bare bones of his theory with an economic understanding that is the true mark of his greatness.[2]

The conclusion in this book that Keynes's hypothesis was unsuccessful is based on money and income data that were not available to Keynes himself. That is true for the pre–World War I period and for the interwar

1. *The Theory of Political Economy*, preface to the 2d ed. (London: Macmillan, 1879), p. lvii.

2. Milton Friedman, "Comments on the Critics," in *Milton Friedman's Monetary Framework* ed. Robert J. Gordon (Chicago: University of Chicago Press, 1974), p. 134.

period itself. In addition, of course, we now have evidence for more than four decades that have elapsed since the *General Theory* was published.

Our examination of this body of evidence reveals that Keynes was generalizing from an idiosyncratic episode—the interwar period in the United States. The pre–World War I period in the United States and the United Kingdom, even the interwar period in the United Kingdom, and the post–World War II period in both countries do not reveal the phenomena that Keynes regarded as contradicting "orthodox" monetary theory. On the contrary, experience during these periods and in these countries conforms to that theory very well. Indeed, one of the ironies of our examination of the evidence is that experience in Keynes's native United Kingdom conforms to the simplest version of the "orthodox" theory he attacked better than does experience in the United States.

The view that Keynes's theory, far from being general, as he labeled it, is highly special, has often been expressed but never documented as fully as we believe we have been able to do.

## 12.1 The Phillips Curve

The clearest conflict betweeen our evidence and the expectations engendered by a Keynesian vision is with respect to the relation between prices and output. Keynes's emphasis on aggregate demand as the prime mover in economic fluctuations—whether short-term movements within business cycles or the longer-term movements between cycle phases that we take as our basic unit in order to abstract from cyclical fluctuations—led to the expectation that output and prices would move together, both rising and falling together relative to longer-term trends. That view is embodied most directly in the negatively sloped Phillips curve, the idea that if output is high relative to capacity, so that unemployment is low, prices will tend to rise relative to trend (or inflation to accelerate), and if output is low relative to capacity, so that unemployment is high, prices will tend to fall relative to trend (or inflation to decelerate).

We were surprised to find that the typical relation is the reverse, that prices and output tend to be related negatively for our phase averages, not positively; that, so far as it exists at all, the Phillips curve—at least for units of time as long as our cycle phases (averaging two years for the United States, 2.8 years for the United Kingdom)—is positively, not negatively, sloping, except only for the idiosyncratic United States interwar period (chap. 9).

## 12.2 Two Extreme Theories

It has been common to contrast two supposedly extreme theories: the simple quantity theory of money and the simple Keynesian income-

expenditure theory (chap. 2). The first theory implies that the velocity of circulation of money—the ratio of nominal income to the quantity of money—is a constant aside from errors of measurement of its numerator and denominator, so that changes in income (i.e., income in dollars or pounds by contrast with "real" income) mirror changes in the nominal quantity of money—in Irving Fisher's evocative phrase, income fluctuations are a "dance of the dollar." A somewhat more sophisticated version that dates back at least two hundred years treats velocity as a stable function of a small number of variables, including the earlier behavior of the quantity of money.[3] The second theory implies that the velocity of circulation of money is a "will-o'-the-wisp" consisting of the ratio of two essentially independent magnitudes. Fluctuations in nominal income are linked to fluctuations in investment through the "consumption multiplier," so that income is a dance of investment rather than of the dollar. A somewhat more sophisticated version gives money, and velocity, some independent status via the possible effect of changes in the quantity of money on interest rates and thence on investment—both "interest rates" and "investment" being defined rather narrowly.

Neither theory, in its simple or more sophisticated version, has anything systematic to say about how a change in nominal income is divided between a change in prices and in output. The quantity theory supposes output to be determined predominantly by nonmonetary forces and supposes changes in nominal income that are produced by changes in money to be reflected ultimately entirely in prices. The Keynesian theory supposes prices to be determined by nonmonetary forces and supposes changes in nominal income that are produced by changes in investment to be reflected entirely in output (so long as employment is less than "full"). Of course, users of both theories have recognized that neither extreme is correct and have made many illuminating side comments on this issue. But neither group has succeeded in developing a satisfactory formal theory to fill this major gap.

We have offered some suggestions about how to do so in our theoretical framework (secs. 2.6–8), but these are highly tentative. Further, the chapter in which we attempt to give empirical content to these suggestions (chap. 9) largely records an unsuccessful experiment—though it does yield some highly important results to which we have already referred.

The broad survey of our basic time series with which we begin our empirical analysis (chap. 5) is sufficient to demonstrate that the simple

3. For example, see David Hume, "Of Money" (discourse 3) and "Of Interest" (discourse 4) in David Hume, *Political Discourses* (Edinburgh: Fleming, 1752), pp. 41–59, 61–78; and M. Friedman, "Discussion," at American Economic Association session "The Rediscovery of Money," *American Economic Review Papers and Proceedings* 65 (May 1975): 176–78.

Keynesian view can be rejected; the movements in the level of income and its rate of change parallel extraordinarily closely for more than a century the contemporaneous movements in the quantity of money and its rate of change, and this is equally true for the United States and the United Kingdom. Whichever is the "cause" and whichever the "effect," the two magnitudes are clearly not independently determined. Velocity varies far less than either nominal money or nominal income. Even more striking, the movements of velocity in the United States parallel those in the United Kingdom, and so do movements in the rate of change of velocity.

The one important difference between the two countries that emerges in this broad survey is that price change accounts for a larger fraction of the fluctuations in nominal income, and output change for a smaller fraction, for the United Kingdom than for the United States—a phenomenon that we encounter repeatedly in later chapters.

## 12.3 The Demand for Money

Velocity is a "real," not a "nominal" magnitude, the monetary units in numerator and denominator canceling out. Its reciprocal has the dimension of time, measuring the amount of money held in terms of the number of time units of income to which that amount of money is equal—so many weeks or months or other time units of income. In consequence, an analysis of velocity is equivalent to an analysis of the demand for money in "real" terms.

We begin our study of the demand for money (chap. 6) by examining in more detail the agreement between our data and the alternative simple theories. The conclusion is clear: both can be rejected, but the simple quantity theory comes far closer to explaining experience than does the simple Keynesian theory. Somewhat surprisingly, the simple quantity theory comes even closer to describing experience for the United Kingdom than for the United States. The later chapters show that this result too reflects the idiosyncrasy of the United States interwar years.

The simple quantity theory does not deserve the denigration that it has received in recent decades. It is a surprisingly good first approximation and clearly recommends itself as a better starting point for a more sophisticated analysis than the simple Keynesian theory. That more sophisticated analysis reveals the existence of a stable demand function for money covering the whole of the period we examine. The major variables that have affected the real quantity of money have, with one exception, apparently been the same in the two countries. The exception is the increasing financial sophistication in the United States monetary system relative to the United Kingdom system before the early 1900s. Presumably the degree of financial sophistication has continued to play a

role in the United States and has been relevant in the United Kingdom all along, but, since it seems to have been affecting both countries to roughly the same extent, we have been unable to identify it.

For the rest, the basic forces that we have identified as affecting the real quantity of money demanded throughout the period have been the level of real per capita income, the difference between the nominal yields on money and on other nominal-value assets, and the nominal yield on physical assets—which we have been able to proxy successfully by the rate of change of nominal income. In addition, two forces were operative for part of the period: (1) after both of the major wars, it took time for the quantity of money demanded to adjust to the drastic change in circumstances; (2) the Great Depression in the United States and the stagnation in the United Kingdom from the mid-1920s until World War II in both countries produced a temporary upward shift in the quantity of money demanded for given values of the other variables—a shift we interpret as reflecting a widespread perception that there was greater economic uncertainty to which a high desired level of liquidity was one response.

With one exception, the basic forces affecting demand in the United Kingdom and the United States had roughly the same quantitative impact in the two countries. The exception is the elasticity of real per capita money balances with respect to real per capita income. That elasticity was somewhat higher than unity in the United States, about 1.1, and somewhat lower than unity in the United Kingdom, about 0.9. For yields, we estimate that a one percentage point (*not* 1 percent) change in the differential yield on money would produce something more than a 9 percent (not percentage point) change in the quantity of money demanded; a one percentage point change in our proxy for the nominal yield on physical assets would produce something more than a 0.04 percent change in the quantity of money demanded. However, these estimates are less securely grounded and subject to a wider margin of error than the estimated income elasticities.

Not only are the basic forces affecting the quantity of money demanded, and their quantitative effects, largely the same in the two countries, but the single statistical equation that we estimate from data for the two countries combined leaves about the same residual variation to be explained in the United States and the United Kingdom by omitted variables or statistical error—about 5 percent for the level of money demanded, about 1.5 percentage points for the rate of change of the quantity of money demanded.

## 12.4 Common Financial System

The two countries are clearly part of a common financial system, in which monetary variables—prices, interest rates, nominal incomes,

stocks of money—are constrained to keep largely in step except as changes in exchange rates alter the number of units of one country's currency equivalent to one unit of the other country's currency. Within the unified financial system, there is much room for divergence of physical magnitudes—for example, there is only a loose link between the two countries in the movements of real per capita income (chap. 7).

Even with respect to financial magnitudes, the unification is far from complete. We have indirectly been able to examine this issue—to see how well, for example, the "law of one price" holds—by estimating year-by-year over the whole of our period the number of United States dollars that had the same purchasing power as one British pound—the purchasing-power-parity exchange rate. If the "law of one price" were perfectly satisfied, the purchasing-power-parity exchange rate would equal the market rate. In fact, it does not do so but fluctuates around the market rate. Before the early 1930s, the purchasing-power-parity rate stayed within plus or minus 10 percent of the market rate—a reasonable approximation to the law of one price. After the early 1930s, the variation was much wider, ranging from 10 percent below to 60 percent above. Government intervention in the exchange market since the 1930s has been more potent in disunifying the markets than improvements in transportation and communication have been in unifying them (sec. 6.8).

## 12.5 Dynamic Effects on Nominal Income

A stable demand function for real money balances means that an autonomous change in either nominal money or nominal income will have to be accompanied by a corresponding change in the other variable, or in variables entering into the demand function for money, in order to equate the desired quantity of money balances with the quantity available to be held. The parallelism in the temporal patterns of nominal money and nominal income means that the adjustment comes about primarily through the other nominal magnitude rather than through the yields or other variables entering into the demand function for money. Given stability of money demand, variability in conditions of money supply, and similar parallelism for the whole of the period, it is appropriate to regard the observed fluctuations in the two nominal magnitudes as reflecting primarily an influence running from money to income. The process is two-way, not unidirectional, so there undoubtedly has also been a feedback from income to money, yet the element that gives consistency to the century as a whole is the influence from money to income.[4]

4. Note that this says nothing about "endogeneity" or "exogeneity." Even for the gold standard period, when the quantity of money is an endogenous variable, the parallel fluctuations in nominal money and income can reflect primarily an influence running from money to income.

Accordingly, in chapters 8 and 9, we have examined the temporal pattern of response to changes in the quantity of money—in chapter 8, the response of nominal income; in chapter 9, of prices and output separately—with special reference to giving empirical content to the dynamic patterns in chapter 2. Though not inconsistent with the suggestions in that chapter, the empirical results do not enable us to specify those hypotheses at all precisely.

The response of nominal income to changes in the nominal quantity of money is generally cyclical, distributed over a long period, and sharply damped in amplitude. The cumulative effect is consistent with theoretical expectation: a sustained one percentage point increase in the rate of monetary growth ultimately produces a one percentage point increase in the rate of nominal income growth. However, we have been unable to pin down at all precisely the magnitude of the transitory effects to be expected en route to this long-term result.

### 12.6 Dynamic Effects on Prices and Output

The response of prices, like that of nominal income, is generally cyclical, distributed over a long period, and sharply damped in amplitude. Except only for the United States interwar period, the ultimate effect of monetary change is absorbed by prices. There is no persistent effect on output. Indeed, for the United Kingdom we have not been able to isolate even transitory effects on output. The change in United Kingdom output from cycle phase to cycle phase behaves like a strictly random series—white noise in the jargon of stochastic series. That is not true for the United States, but the output effects are smaller and less consistent over time than the price effects—again with the conspicuous exception of the interwar period.

### 12.7 Interest Rates

According to the simple Keynesian theory, changes in the quantity of money would be reflected first in changes in the opposite direction in interest rates, which in turn would affect investment and, through investment, nominal income. The implied inverse relation between changes in the quantity of money and in interest rates, widely taken for granted as recently as the mid-1960s, by now has been thoroughly discredited by the simultaneous upward trends of the past several decades in the quantity of money, nominal income, inflation, and interest rates. One result has been an explosion of economic research into the relation between prices and interest rates, research inspired largely by the seminal contributions of Irving Fisher and secondarily by some comments of Keynes in the *Treatise on Money*, which preceded the *General Theory* and out of which the *General Theory* developed.

By now there is general agreement on the theory of the relation between money and interest rates (sec. 10.1). The theoretical relation is complex, so that our empirical analysis, based on that theory, offers a rich understanding of particular episodes but does not yield any simple empirical generalization enabling an observer to predict the effects of monetary change on interest rates.

Our empirical analysis is devoted primarily to two related themes: the relation between yields on nominal and on physical assets and the effect of the level and rate of change of prices on interest rates.

For the century we cover as a whole, nominal yields on nominal assets roughly equal our proxy for nominal yields on physical assets, averaging about 4.5 percent for the United States, about 4 percent for the United Kingdom; short-term nominal assets yielded somewhat less on the average, long-term nominal assets somewhat more, in line with the widespread belief that liquidity commands a premium. The equality between yields on nominal assets and our proxy for physical assets is evidence in favor of the validity of our proxy as well as testimony to the existence of effective arbitrage between different categories of assets over long periods.

The real yield—the excess of the nominal yield over the average rate of inflation—averaged about 3 percent for the United States, about 1.25 percent for the United Kingdom. The excess of the United States yield over the United Kingdom yield is greater for real than for nominal yields because United Kingdom prices rose on the average more rapidly than United States prices, a difference that was reflected almost precisely in the average behavior of the exchange rate.

Arbitrage roughly equated yields on nominal and physical assets for the century as a whole but not for shorter periods—there is little correlation between the phase-to-phase movements of yields on nominal and physical assets. During periods of rising prices, physical assets provided higher yields than nominal assets; during periods of falling prices they provided lower yields. As between such periods, nominal yields were much stabler for nominal than for physical assets; real yields were much less stable. The conclusion is inescapable: for the greater part, inflation and deflation were not accurately anticipated and therefore were not reflected in the terms on which nominal assets were acquired: lenders did not succeed in protecting themselves against having their real returns eroded by inflation; borrowers did not succeed in protecting themselves against having their real interest payments increased by deflation.

The greater stability of the real yield on physical assets than on nominal assets does not reflect greater foresight by their holders. It reflects the absence of advance contractual arrangements. Yields on physical assets are realized mostly as a difference between receipts from the physical

assets and costs (other than the return on physical capital) of acquiring those receipts. Inflation and deflation affect both receipts and costs. A measure of automatic indexing, as it were, stabilizes the real return on physical assets.

The importance of *anticipations* as opposed to *realizations* is highlighted by one particularly instructive episode. That episode is the transition from the fear that the United States would go off the gold standard, produced by the free silver movement before 1896, to confidence in the maintenance of the gold standard following McKinley's election in 1896. From 1873 to 1896, the short-term interest rate in the United States was slightly *higher* on the average than it was from 1896 to 1914, although prices on the average were falling by nearly 2 percent a year before 1896 and rising by nearly 2 percent a year after 1896. In the United Kingdom, by contrast, the short-term interest rate was slightly *lower* on the average before than after 1896. The United States rate averaged about 2.5 percentage points higher than the United Kingdom rate before 1896, less than 1.5 percentage points higher than the United Kingdom rates after 1896.

These seemingly paradoxical results are readily explicable. Before 1896, the political strength of the free silver movement made it entirely reasonable for farsighted investors to anticipate inflation. The fact turned out to be deflation; the anticipation was inflation. The investors proved to be wrong—but there was no way anyone could know that in advance, no evidence that would have in advance contradicted a personal probability judgment that the probability of inflation was decidedly greater than 50 percent. After 1896, the *fact* was inflation; the anticipation was stability or deflation. The shift in anticipations reduced the differential between United Kingdom and United States interest rates: investors in the United Kingdom presumably shared the fears of those in the United States and hence were reluctant to lend in the United States before 1896 except at a premium that compensated them for the possibility of a devaluation of the dollar relative to the pound. After 1896 this premium was no longer necessary; the remaining excess simply reflected the premium required to compensate for investing abroad rather than at home—a differential to be expected between a capital-importing country, such as the United States then was, and a capital-exporting country, such as the United Kingdom then was. After World War I, the United States too became a capital exporter and, while the real yield in the United States remained higher than in the United Kingdom, the excess declined by about one percentage point. Nominal yields became higher in the United Kingdom than in the United States, reflecting the depreciation of the pound relative to the dollar.

## 12.8 Rational Expectations

The free silver episode is especially instructive with respect to the proper interpretation of the recently popular theory of rational expectations. In applying this idea, it is common to proceed on two assumptions: (1) that participants in whatever market is considered have "correct" estimates of the probability distribution of outcomes (itself something that is difficult or impossible to define objectively), so that on the average anticipations are correct; and (2) that errors of forecast in successive time units are uncorrelated. This episode brings out sharply the difficulty of giving a precise meaning to the first assumption, and the ambiguity of "time unit" for the second assumption. For that episode, the relevant time unit is about twenty years—so that averaging out may take a long time. Our analysis gives one example: the real yield on nominal assets matches the real yield on physical assets only for the whole century our data cover.

The formalization in the theory of rational expectations of the ancient idea that economic actors use available information intelligently in judging future possibilities is an important and valuable development. But it is not the open sesame to unraveling the riddle of dynamic change that some of its more enthusiastic proponents make it out to be.

## 12.9 Fisher and Gibson

Irving Fisher was an early and sophisticated user of the basic idea of rational expectations. His expectation (in 1896) that nominal interest rates would be relatively high during periods of rising prices and relatively low during periods of falling prices was based on the view that lenders and borrowers seek to anticipate price movements and allow for them in the interest rates they are willing to accept or to pay. His examination of empirical evidence persuaded him that there was an effect in this direction but that it was very much damped—the conclusion that we too have reached on the basis of experience for a much longer period. He suggested an explanation in terms of the slowness with which participants adapted their anticipations to experience, leading to an appearance that interest rates vary with the price *level* rather than the rate of price change.

At about the same time, Knut Wicksell was impressed by the same empirical observation that prices and interest rates apparently moved together—the observation that Keynes later termed the Gibson paradox—and suggested an explanation in terms of the slowness with which banks adapted their anticipations to changes in the productivity of physical capital. Keynes offered a variant of this explanation for the same phenomenon more than two decades later.

Our interest in the effect of monetary change on interest rates naturally led us to examine the relation between the level and rate of change of prices and interest rates, because the effect of monetary change on prices is a major channel through which monetary change affects interest rates. Our examination of the evidence confirms the doubts expressed by Frederick Macaulay more than forty years ago about the generality of the Gibson phenomenon. It does not exist over very long periods. It frequently does not exist over short periods. In particular, it is present before World War I in both the United States and the United Kingdom, though apparently, according to studies by Gerald Dwyer, not in some other countries. There are traces of it in the interwar period, almost none in the post–World War II period. The major change in price level from before to after World War I leaves no reflection at all in the level of interest rates. In short, there has at times been a Gibson phenomenon; there is no Gibson paradox.

Wicksell's and Keynes's suggested explanation of the Gibson phenomenon is clearly contradicted by the evidence and can definitely be rejected. Fisher's explanation is not, though it must be modified from his final (1930) version, in which he regarded participants as forming their anticipations of future price change as a very long weighted average of past price change. The modifications are, first, that the period of averaging is much shorter than Fisher estimated it to be, though still lengthy—something like six to nine years rather than the much longer period he estimated—and, second, that participants also take into account other relevant episodic evidence, such as the free silver movement.

A sharp break has apparently occurred in recent years in the relation between price change and interest rates in both the United States and the United Kingdom. Since about the mid-sixties, a close relation between interest rates and the rate of price change has emerged for the first time in the century we study. Gibson has been replaced by the original Fisher. Nominal returns have become more variable than real returns; nominal returns on nominal assets are as variable as nominal returns on physical assets. Lenders and borrowers apparently have been able to predict price changes more accurately and to adjust the terms of lending and borrowing accordingly.

This break has been noted by other investigators. Like Benjamin Klein, we are inclined to attribute the break to a belated and gradual recognition by market participants of the drastic change in the monetary system from a largely specie standard to a fiduciary standard. One caveat is in order: the extent of the shift may have been exaggerated by the failure to allow for the effect of taxes on the real yield from nominal assets during an inflationary period. This is a complex issue that is far from settled. In any event, whether the shift in pattern proves permanent or temporary is likely to depend on future developments in the monetary system and on the future course of inflation.

## 12.10 Long Swings

Our final substantive chapter (chap. 11) applies our findings to a fairly extensive body of research on long swings in economic activity. This research has examined swings in economic activity of a decidedly longer duration than business cycles, swings that encompass several business cycles. Investigators have documented such swings for the United States in many physical magnitudes for roughly the century our study covers. They have been asserted to exist but not comparably documented for the United Kingdom, as well as other countries.

Some investigators have maintained that these swings reflect a self-generating cyclical process of longer duration than the ordinary business cycle and have offered a series of hypotheses about the economic forces producing them.

It is a remarkable testament to the extent to which the Keynesian revolution dominated economic research during the period when the initial research on long swings was done that no investigator we have been able to uncover paid more than the most casual attention to the behavior of the quantity of money or nominal income during the asserted long swings or to the role that monetary phenomena might play. The investigators restricted themselves to "real" phenomena in the sense of physical magnitudes—an extreme example of the widespread view that "money does not matter." More recently, there has been a welcome recognition by some leading investigators of long swings of the necessity of incorporating monetary magnitudes into their analysis.

Our data demonstrate that Hamlet has been left out of most long-swing research. The swings isolated in the physical magnitudes have their counterpart in the nominal magnitudes. More important, the swings in the nominal magnitudes are wider in amplitude and more clearly marked than in the physical magnitudes. No explanation of long swings is acceptable that does not account for the monetary phenomena.

The observed swings may reflect a self-generating cyclical mechanism, as some investigators have claimed, or they may represent the smoothing, by the economic process or statistical procedures, of episodic disturbances. Our data strongly support the episodic interpretation.

Long-swing research is a clear example of a branch of economics that has been "shunted . . . on to a wrong line": masses of sophisticated statistical data and economic analysis; many useful empirical results, yet its basic theoretical generalizations null and void thanks to a shuttered vision.