DALLASFED

Economic Review

Federal Reserve Bank of Dallas
September 1984

Expectations and Monetary Regimes

*By Gerald P. O'Driscoll, Jr.**

There has been an expectations revolution in economics, a revolution that has both changed economic theory and altered the way it is applied to specific issues. The theoretical and econometric innovations constituting this revolution are most noticeable in macroeconomics, but they have implications for nearly all areas of economics. This article focuses on the developments and implications for macro and monetary economics.

Conventional macro models accorded comparatively little attention to expectational issues. In Keynesian models, for instance, transactors' expectations were assumed to be unaffected by changes in the environment (including changes in monetary or fiscal policy). Monetarist analysis generally paid more attention to expectational questions but assumed that agents made *systematic* forecast errors in a wide range of circumstances. Keynesian and monetarist models alike thus violated a core postulate of economics: that individuals make the best use of available resources by attempting to maximize returns from given means. This efficient use of resources is called "rationality." In conventional macro theory, however, agents did not behave rationally. They either ignored available information or used it inefficiently. In the new macro theory, agents are rational in their exploitation of all available resources, including information—a behavioral assumption termed "rational expectations."

This article reviews the rational expectations critique of macroeconomics and presents the rational expectations hypothesis. An emerging literature is critical of that hypothesis, suggesting that conventional formulations of the hypothesis are inconsistent with the decentralized decision making characteristic of market economies. While the article supports that criticism, it also presents a recent reformulation of the rational expectations hypothesis. This reformulation links the process of expectations formation to the monetary regime or institutional environment in which individuals operate. The reformulation thereby connects the

* *Gerald P. O'Driscoll, Jr., is a senior economist at the Federal Reserve Bank of Dallas. The views expressed are those of the author and do not necessarily reflect the positions of the Federal Reserve Bank of Dallas or the Federal Reserve System.*

literature on expectations to developments in the theory of institutions.

This article concludes that the reformulated rational expectations hypothesis may obviate the recent criticism. In its reformulated version, however, the hypothesis can no longer support certain strong claims associated with it, such as the ineffectiveness of macroeconomic policy. Nonetheless, the reformulation offers promise for further theoretical development.

Expectations in a macro model

In modern economics, David Hume demonstrated the irrelevance of the price level. In the same essay,[1] however, he argued that *rising* prices facilitated trade. The idea that a "little bit of inflation" is good for economic activity reappears in economics with predictable regularity. Knut Wicksell identified the problem in the Humean argument:

> Those people who prefer a continually upward moving to a stationary price level forcibly remind one of those who purposely keep their watches a little fast so as to be more certain of catching their trains. But to achieve their purpose they must not be conscious or remain conscious of the fact that their watches are fast; otherwise they become accustomed to take the extra few minutes into account and so after all, in spite of their artfulness, arrive too late. . . .[2]

The most recent instantiation of the inflationist argument is the Phillips curve, which represents a statistical relationship between inflation and unemployment.[3] The recent controversy over expectations can be elucidated in terms of this familiar relationship.

In Figure 1, *PC* plots a series of observations (p, u) of inflation and unemployment combinations. This curve was viewed as presenting policymakers with a choice set yielding an exploitable trade-off between inflation and unemployment rates. Originally "an empirical finding in search of a theory,"[4] the Phillips curve was rationalized by a variety of theoretical stories. All related inflation functionally to unemployment by postulating an inverse relationship between the unemployment and inflation rates.

Milton Friedman provided an early critique and reformulation of the Phillips curve analysis:

> Implicitly, Phillips wrote his article for a world in which everyone anticipated that nominal prices would be stable and in which that anticipation remained unshaken and immutable whatever happened to actual prices and wages. Suppose, by contrast, that everyone anticipates that prices will rise at a rate of more than 75 per cent a year—as, for example, Brazilians did a few years ago. Then wages must rise at that rate simply to keep real wages unchanged.[5]

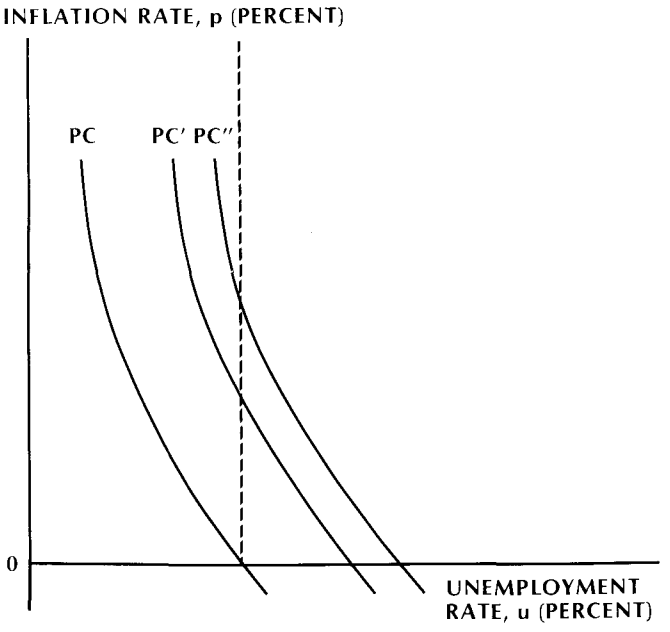
An increase in the inflation rate (that is, a higher rate of growth of prices) will have the predicted effect on unemployment only if it is unanticipated. In practice, policymakers can exploit the Phillips curve (or any other relationship between real and nominal variables) only with policies that generate *accelerating* inflation. In terms of monetary policy, money growth must accelerate. This focus on monetary acceleration or deceleration provided a name for the new view—the "accelerationist" theory.

The accelerationist thesis can be illustrated by considering a hypothetical example. Assume that

1. "Of Money," in David Hume, *Writings on Economics*, ed. Eugene Rotwein (Madison: University of Wisconsin Press, 1970), 33-46; the essay was originally published in 1752.
2. Knut Wicksell, *Interest and Prices*, trans. R. F. Kahn (1936; reprint, New York: Augustus M. Kelley, 1965), 3-4; the original was published in German in 1898.
3. Phillips found a relationship between the rate of change of money wage rates and the unemployment rate (A. W. Phillips, "The Relation Between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861-1957," *Economica* 25 [November 1958]: 283-99). Figure 1 is, strictly speaking, an adaptation of his findings substituting the rate of change in the price level for that in money wages. The adaptation is common and easily derived from Phillips' original relationship. For instance, see Paul Wonnacott, *Macroeconomics*, rev. ed. (Homewood, Ill.: Richard D. Irwin, 1978), 330-33.
4. James Tobin, "Inflation and Unemployment," *American Economic Review* 62 (March 1972): 9.
5. "The Role of Monetary Policy," in Milton Friedman, *The Optimum Quantity of Money and Other Essays* (Chicago: Aldine Publishing Company, 1969), 102. This paper constituted Friedman's 1967 presidential address to the American Economic Association. His main argument was elaborated and developed in an important volume: Edmund S. Phelps and others, *Microeconomic Foundations of Employment and Inflation Theory* (New York: W. W. Norton & Company, 1970).

Figure 1

The Inflation-Unemployment Relationship



the economy is in noninflationary macroeconomic equilibrium. Unemployment is 6 percent, which represents equilibrium in labor markets. This equilibrium rate of unemployment, designated the "natural rate" in the literature, is determined by nonmonetary factors like the structure of the labor market and information costs.[6] Assume also, however, that policymakers believe the unemployment rate is "too high" and that, for example, the monetary authority acts on this perception by fostering higher rates of money growth. Aggregate demand for goods will now grow more rapidly, bidding up wages.

Will the unemployment rate fall? The answer depends on expectations. Workers respond to changes in real, or inflation-adjusted, wages. If, influenced by past price stability, workers believe a faster rate of increase in money wages represents more rapid growth in real wages, the unemployment rate will fall temporarily.[7]

Demand-management policy might succeed for a time in reducing the unemployment rate from 6 percent to 4 percent. In doing so, however, it will eventually cause inflation to increase. The beneficial effect on unemployment will persist only so long as the resulting inflation is unanticipated.

The accelerationist theory denies that inflation illusion can persist *ad infinitum*. Though it sets no

6. Friedman, "Role of Monetary Policy," 102-3. In the literature, "natural" is synonymous with "real" or "nonmonetary." Friedman borrowed the concept of a natural rate of unemployment from Wicksell's natural rate of interest. The latter is the rate of interest that would rule without the influence of monetary disturbances (ibid., 101-2).

7. Among other things, workers will decrease the time spent searching for better wage offers. This, in turn, decreases measured unemployment. The search theory of unemployment, which evolved from the accelerationist theory, is a separate but important body of analysis to which many of the papers in the Phelps volume represent early contributions.

exact quantitative parameters on the process, the theory postulates a catch-up procedure for expectations. Expansive aggregate demand policy has moved the system northwest on the Phillips curve (*PC*). To accomplish this, policymakers must, in effect, deceive workers. They can do so temporarily if workers' inflation expectations are adaptive.

Expectations are adaptive if they are some weighted average of past rates of inflation. The weights and the number of periods averaged may vary, but the logic is the same: recent experience determines individuals' expectations of next period's inflation rate. In the current example, workers anticipate zero inflation in the future because of the experience in the recent past.

As inflation accelerates, workers may at first either not notice it or believe that rising prices are an anomaly and price stability will be restored. At some point, however, as the experience of inflation becomes undeniable, workers will realize that real wages are not rising. This realization will cause labor market conditions to return to the situation *ex ante*. Only if the authority were to increase the rate of monetary growth (that is, only if there were an acceleration in money growth) could an unemployment rate of 4 percent be maintained (remembering that a 6-percent rate represents "full employment").

In terms of Figure 1, the Phillips curve shifts from *PC* to *PC'* as inflation becomes anticipated. Higher and higher inflation rates are necessary to produce a given real effect, such as a decline in unemployment. If inflation were to accelerate further, the Phillips curve would eventually shift from *PC'* to *PC''*. Each curve in the family of Phillips curves represents an *unstable*, short-run relationship, not a long-run, exploitable trade-off. The vertical dotted line is the true long-run relationship.

A similar conclusion is reached for movements on a curve to the southeast, which represent the short-run effects of anti-inflationary policy. Unemployment will temporarily increase, falling back to its equilibrium level as workers adjust their expectations to a *lower* rate of inflation. The result is symmetric: the unemployment rate is invariant to the inflation rate. Hume's stricture about the irrelevance of the price level is extended to the rate of change of prices.

The rationalization of high inflation and high unemployment (stagflation) attracted the attention of many economists to the accelerationist thesis. Yet the thesis has a fatal flaw, which became the focal point of the rational expectations critique. The accelerationist theory assumes that economic agents use an incorrect economic model in forecasting future price changes. If inflation is accelerating, for instance, agents will make systematic (nonrandom) forecast errors. Systematic errors present unexploited profit opportunities, which alert entrepreneurs should eliminate. Persistent profit opportunities are inconsistent with the assumption of economic equilibrium. To assume that, in the face of accelerating inflation, agents expect prices to increase at a constant rate is to repeat the theoretical error of Keynesian models, which assumed static expectations.

The rational expectations hypothesis

Rational expectations theorists propose modeling economic agents as themselves modeling the economy.[8] Not only do agents have a model, but that model conforms to the actual economic process. "Expectations, since they are informed predictions of future events, are essentially the same as the predictions of the relevant economic theory."[9]

In other words, when economists theorize about individual behavior, including the formation of expectations, they must take into account the ability of individuals to theorize about their environment. Economists must also assume that individuals utilize the same ("relevant") theory being used by economists to explain individual behavior. At this broad level, the rational hypothesis constitutes a methodological principle. The hypothesis is also embedded in specific models.

Rational expectations models typically assume that economic phenomena are generated by an underlying stochastic process. Forecasts are probabilistic, but an individual's mean or expected outcome should conform to the mathematical expectation of the event being forecast.[10]

8. The classic article is John F. Muth, "Rational Expectations and the Theory of Price Movements," *Econometrica* 29 (July 1961): 315–35; reprinted in Robert E. Lucas, Jr., and Thomas J. Sargent, eds., *Rational Expectations and Econometric Practice* (Minneapolis: University of Minnesota Press, 1981), 3–22. All subsequent page references are to the Lucas-Sargent volume.

9. Muth, "Rational Expectations and Theory of Price Movements," 4.

Rational expecters make *unbiased* forecasts of future outcomes. In more narrow, technical terminology, the agent's model is a minimum-mean-squared-error generator of forecasts. Forecast errors will be neither systematic nor correlated with any information available at the time of the forecast.

Rational expectations theory has three important elements, which can be summarized as follows:

1. There is an objective probability distribution of outcomes to which subjective probability distributions will conform.
2. The agents being modeled by economists are themselves using models in forming their expectations.
3. Agents' models embody expectations that are rational in the sense outlined above, utilizing the relevant economic theory.

It was quickly recognized that, taken strictly, rational expectations would rob authorities of any ability to exploit observed relationships among economic variables. At each moment, endogenous variables incorporate the effects of expected policy changes. Policy actions may influence observed outcomes but not in any systematic way. Policymakers will not be able to predict the exact impact of their actions unless they possess an informational advantage over citizens at large. In many formulations, citizens also form expectations about the probability of a policy innovation. This assumption deprives policymakers of the advantage of being better able to predict their own actions.

Early expositions of rational expectations incorporated extreme statements of policy ineffectiveness:

> [T]here is no sense in which the authority has the option to conduct countercyclical policy. To exploit the Phillips Curve, it must somehow trick the public. But by virtue of the assumption that expectations are rational, there is no feedback rule that the authority can employ and expect to be able systematically to fool the public. *This means that the authority cannot expect to exploit the Phillips Curve even for one period.*[11]

More recently, rational expectations theorists have backed off somewhat from consistent pursuit of this line of reasoning. Models increasingly incorporate a distinction between anticipated and unanticipated policy actions. The former actions have no systematic effect on outcomes in these models, since anticipated effects are fully incorporated in plans. Unanticipated policy changes do have effects in the models. Even this distinction is not without problems, however. For one, the reformulated policy ineffectiveness proposition has failed econometric tests. Even anticipated policy changes matter for economic outcomes. Indeed, unanticipated changes in monetary and fiscal policy do not have a significantly greater impact on output and employment than do anticipated movements.[12]

The econometric battle over policy ineffectiveness (or policy *neutrality*) still rages. Important as it is, however, that empirical literature is not the primary focus of the discussion here. Rather, this article focuses on crucial issues raised in an emerging theoretical literature. This literature questions each of the three elements of rational expectations theory cited above. Overall, the critique questions the ability of agents to generate rational forecasts. It suggests that successful forecasts depend, in part, on factors that have largely remained outside models of expectations formation.

Rational expectations: a critical assessment

In this section, one basic issue will be considered. Critics argue that rational expectations models presuppose that more information is available to economic agents than is consistent with standard assumptions of economic theory. Their basic argument can be divided into three points.

1. Rational expectations models are inconsistent

10. "[E]xpectations of firms (or, more generally, the subjective probability distribution of outcomes) tend to be distributed, for the same information set, about the prediction of the theory (or the 'objective' probability distributions of outcomes)" (ibid., 4-5).

11. Thomas J. Sargent and Neil Wallace, "Rational Expectations and the Theory of Economic Policy," *Journal of Monetary Economics* 2 (April 1976): 177-78; emphasis added.

12. For a discussion of the state of the literature on this topic, see Frederic S. Mishkin, "A Rational Expectations Approach to Macroeconometrics," *NBER Reporter*, Winter 1982/3, 6-7. Also see Frederic S. Mishkin, "Does Anticipated Aggregate Demand Policy Matter? Further Econometric Results," *American Economic Review* 72 (September 1982): 788-802. That article contains references to empirical work supporting the policy ineffectiveness proposition.

with the assumption that markets operate with an economy of information.

2. Rational expectations analysis is inconsistent with economic decentralization.

3. The assumption of rational expectations ignores the problem of selecting the relevant economic theory.

Informational economy. The marginalist (neoclassical) theory of the firm has been attacked for being "unrealistic" in assuming that businessmen produce up to the point that marginal revenue is equated with marginal cost. Businessmen do not appear to know their marginal costs, much less use them in decision making. It is now generally accepted, however, that businessmen follow procedures that amount to maximizing profits. This practice results in an output level that would have resulted from equating revenues and costs at the margin. The marginalist theory is accepted because economists can identify real-world procedures that generate predicted outcomes.

In particular, the plausibility of the theory of the firm was enhanced by demonstrating how *little* businessmen (and economic agents generally) need to know in order to plan. Competitive firms need only know market data (prices, for example) and be able to react appropriately to changes in data. These firms are price takers and frequently operate in industries for which there are active futures markets. Firms need not have structural information or know market demand and supply conditions. They only react to market signals like price changes. In short, the system operates with an economy of information.

Rational expectations theory has effectively turned the economic argument around 180°. First, by emphasizing the importance of expectations, it has pointed to the speculative element in all decision making. Not only professional speculators but also savers and workers must speculate on, among other things, future inflation rates. Only by doing so can they distinguish purely nominal from real changes in returns. Agents' lives are infinitely complicated by all this, compared with the relatively simple computational problems facing them in the theory of competitive price determination.

Further, instead of emphasizing how little agents must learn, rational expectations theory assures us of how much they in fact know. Kenneth Arrow has succinctly posed the problem:

> The very concept of the market and certainly many of the arguments in favor of the market system are based on the idea that it greatly simplifies the informational problems of economic agents, that they have limited powers of information acquisition, and that prices are economic summaries of the information from the rest of the world. But in the rational expectations hypothesis, economic agents are required to be superior statisticians, capable of analyzing the future general equilibria of the economy.[13]

The rational expectations literature has thus effected an almost unnoticed but radical escalation in the information requirements of decentralized (competitive) economies. Decentralized economies have been thought to operate efficiently because of the *paucity* of information needed by each economic agent. In rational expectations models, it is *how much* agents know that ensures rapid adaptation to policy changes and other shocks.

Neoclassical general equilibrium theory has been effectively criticized for relying on price signals alone to allocate resources efficiently. The information available to individuals is insufficient to generate the allocational outcomes assumed to characterize competitive markets.[14] Yet rational expectations theory is based on general equilibrium analysis. The latter provides the theoretical structure for rational expectations models.[15] Consequently, rational expectations models can only postulate but not demonstrate that agents will acquire the knowledge assumed to be available to them in these models.[16]

13. Kenneth J. Arrow, "The Future and the Present in Economic Life," *Economic Inquiry* 16 (April 1978): 160.

14. A series of papers by Grossman and Stiglitz constitute the core of this critique. For an early statement, see Sanford J. Grossman and Joseph E. Stiglitz, "Information and Competitive Price Systems," *American Economic Review* 66 (May 1976): 246-53.

15. The dependence of the conclusions of rational expectations models on their general equilibrium underpinnings is emphasized in Kevin D. Hoover, "Two Types of Monetarism," *Journal of Economic Literature* 22 (March 1984): 58-76, especially 66-68.

16. See Gerald P. O'Driscoll, Jr., "Rational Expectations, Politics, and Stagflation," in *Time, Uncertainty, and Disequilibrium*, ed. Mario J. Rizzo (Lexington, Mass.: D. C. Heath and Company, Lexington Books, 1979), 158-60.

Rational expectations models are quite reasonable in postulating that individuals will make the best use of available information. Critics generally accept this methodological point. These models frequently assume, however, that individuals actually possess large stocks of knowledge, including structural knowledge of the economy. It is true that "information is scarce, and the economic system generally does not waste it."[17] This observation is, however, a two-edged sword. Because information is scarce, theorists cannot assume it is evenly distributed or available in abundance. It is not irrational to be poorly endowed with economic knowledge, any more than with capital or land. What is reasonable for agents to know depends on their environment.

Economic decentralization. Information specialization is a counterpart to specialized production.

> It is the essence of the decentralized economy that individuals have different information. Each individual is specialized in certain activities and has in general specialized knowledge about those activities. There is no reason therefore why his forecasts should be based only on the rather general kind of information which the econometrician can use.[18]

If market participants have specialized information that the econometrician lacks, they will, *a fortiori*, be ignorant of specialized theoretical knowledge possessed by economists and statisticians. Indeed, if specialization is to be taken seriously, the knowledge possessed by the two groups (economic agents and economic analysts) will differ systematically. Whereas theoretical knowledge is abstract and general, market information is concrete and specific. Aptly described as "knowledge of the particular circumstances of time and place,"[19] market information consists of such things as knowledge of a temporary discount on an input or of a profit opportunity resulting from a transient geographic price difference for a produced good. This is not at all the kind of abstract information yielded by any relevant economic theory.

The more specialized an activity, the more likely that factors peculiar to that market will be more important for decisionmakers than are general economic conditions.[20] This fact further diminishes the likelihood of economic agents' investing in acquiring theoretical economic information. Being based on different information sets, predictions of market participants will differ from model-theoretic forecasts. Transactors surely analyze their environment, but their "model" will differ from that employed by macro econometricians.

Individual economic agents also possess heterogeneous information. Decentralization implies strictly that marketwide information is scattered among economic agents. Heterogeneity of information and inconsistency of forecasts are predictable features of market economies.[21]

To reiterate, decentralization of information is an essential property of market economies. This feature is notably lacking in rational expectations models.[22] Its absence lends credence to the conclusions of critics that rational expectations analysis is inapplicable to decentralized economic systems.[23] Rational expectations theorists have yet to respond to this critique, which is more telling when analyzed in

17. Muth, "Rational Expectations and Theory of Price Movements," 5.

18. Arrow, "Future and Present in Economic Life," 164.

19. "The Use of Knowledge in Society," in Friedrich A. Hayek, *Individualism and Economic Order* (Chicago: University of Chicago Press, 1948), 80.

20. Arrow, "Future and Present in Economic Life," 165–66.

21. This point is developed in more detail in Gerald P. O'Driscoll, Jr., "Knowing, Expecting, and Theorizing," C. V. Starr Center for Applied Economics Research Report no. 81-22 (New York: New York University, 1981). The existence of informational differences *in* equilibrium is argued in Gerald P. O'Driscoll, Jr., and Mario J. Rizzo, *The Economics of Time and Ignorance* (Oxford: Basil Blackwell, forthcoming).

22. One conspicuous exception is Lucas' equilibrium model of the business cycle, which postulates informational differences across markets by using Phelps' island paradigm. (See Robert E. Lucas, Jr., "An Equilibrium Model of the Business Cycle," *Journal of Political Economy* 83 [December 1975]: 1113–44.) For an analysis, see O'Driscoll, "Rational Expectations, Politics, and Stagflation," 157–68.

23. Roman Frydman, "Towards an Understanding of Market Processes: Individual Expectations, Learning, and Convergence to Rational Expectations Equilibrium," *American Economic Review* 72 (September 1982): 652–68; and Roman Frydman and Edmund S. Phelps, Introduction to *Individual Forecasting and Aggregate Outcomes: "Rational Expectations" Examined* (Cambridge: Cambridge University Press, 1983), 1–30.

conjunction with the next issue.[24]

Theory conflict. Macro economists do not agree. There are an embarrassingly large number of macroeconomic theories and an indefinitely larger number of models derivable from these theories. Granted that individuals purchase economic forecasts, to which theory will they subscribe? What constitutes the "relevant" theory is itself a controversial issue among specialists. Informational disadvantages aside, the public confronts the same basic problem faced by economists in discriminating among theories. Every data observation is consistent with a large number of theories. Even sophisticated empirical tests are generally subject to alternative interpretations and analyses.

Rational expectations theory postulates an underlying objective probability distribution to which subjective probability distributions ought to conform. For conformity to occur, however, individuals must discover the parameters of that distribution. This latter point raises the question of how individuals learn the required information in a decentralized system. In other words, will there be a convergence of theories?

One author concluded that there will be no convergence to rational expectations equilibrium if the economy does not begin in such an equilibrium.[25] To forecast optimally, an individual must know the average of market forecasts (the "average opinion"). Agents can observe past and current market prices but cannot infer the average opinion from these observations. No market data provide this. Surveys and data collection could only provide ex *post* information on average opinion, while rational forecasts need ex *ante* data.[26] Without this information, no convergence to a unique forecast occurs. Hence, there is no convergence to a rational expectations equilibrium, nor will forecasts be optimal in any conventional sense. Individuals will "adopt *some subjective criteria* and make what they consider to be their own best guesses of the average opinion."[27]

This conclusion seemingly brings us full circle back to Keynes' position

> that human decisions affecting the future, whether personal or political or economic, cannot depend on strict mathematical expectation, since the basis for making such calculations does not exist; and that it is our innate urge to activity which makes the wheels go round, our rational selves choosing between the alternatives as best we are able, calculating where we can, but often falling back for our motive on whim or sentiment or chance.[28]

Rational expectations theorists endeavored to ground expectations in an objective process, thereby rendering feasible an expectational equilibrium. The attempt foundered because, in order to forecast the economic future, individuals must forecast each other's forecasts. The apparently objective process is permeated with seemingly unbounded subjectivity.

Markets are not generally chaotic, and individuals can at least sometimes form tolerably accurate expectations of future events. So long as autonomous economic agents are forming and acting on their individual forecasts, there will be no deterministic ("objective") process underlying economic activity. If, however, the range of possible or likely outcomes were sufficiently narrowed, an economic theory of expectations might be possible. In terms

24. David K. H. Begg has asserted that "Rational Expectations analysis may easily be extended to models in which individuals have access to different information because they are faced with different costs of acquiring information. Thus many criticisms refer to the simplicity of the models in which Rational Expectations has so far been embedded" (*The Rational Expectations Revolution in Macroeconomics: Theories and Evidence* [Baltimore: Johns Hopkins University Press, 1982], 255). The first statement is, of course, precisely the point of contention. As to the second point, critics can only respond to actual not hypothetical models.

25. This section draws from Frydman, "Towards an Understanding of Market Processes." See especially 652–55.

26. If individuals could revise their plans in light of average opinion, the survey forecast would no longer reflect the average of revised forecasts. If the process were to continue (that is, there were to be a new survey), transactors would face Keynes' "beauty contest" problem. Individual forecasts would continually be revised in light of the most recent revision of average opinion. There is no end to this process and no convergence to an equilibrium. See O'Driscoll and Rizzo, *Economics of Time and Ignorance*, chap. 5.

27. Frydman, "Towards an Understanding of Market Processes," 662. He calls this a "subjectively optimal forecast," but it is not a strictly defined rational expectations equilibrium.

28. John Maynard Keynes, *The General Theory of Employment, Interest and Money* (New York: Harcourt, Brace and Company, 1936), 162–63.

of conventional models, these constraints would have to be exogenous factors.

Economists are increasingly focusing on institutional constraints and rules as stabilizing forces. An emerging literature treats the analysis of expectations formation and a theory of institutions as logically related topics. Indeed, the literature argues that expectations are heavily influenced by the institutional environment.

Monetary regimes

The ability to form expectations depends crucially on the existence of some "external" constraints.[29] Institutions provide constraints on the range of possible plans and, hence, on expectations. An institution may be an organization, like a central bank, or an established relationship, like that which may exist between two firms or individuals. As such, institutions represent relatively permanent features in the economic landscape. They are by no means immutable, but they do change slowly relative to changes in economic data.

Rules determine the procedures and courses of action necessary to sustain an institution. The existence of rules implies consistency, though not necessarily rigidity, in behavior. A central bank may, for example, have a usual procedure for dealing with an outflow of international reserves (for example, "raise the discount rate"). This procedure may or may not involve a rigidly determinate course of action, such as would be dictated by a mathematical formula. In other words, rules may be discretionary or nondiscretionary, and they may be explicit or implicit.

A set of institutions and rules constitutes a monetary regime. The existence of a monetary regime is a publicly known fact. By assumption, information about a regime's characteristics is available at relatively low cost. This is precisely the kind of information, it is reasonable to suppose, that individuals will share and with which they will forecast.

Instead of acquiring structural knowledge, individuals learn about the operation of basic institutions. Institutions tend to homogenize otherwise heterogeneous information in a decentralized economic system. They do not eliminate information differences but provide a common framework within which expectations are formed and choices made.

Consistency in behavior and permanence of institutions can rationalize an expectational equilibrium. Indeed, Axel Leijonhufvud has suggested defining a monetary regime as "a system of expectations that governs the behavior of the public and that is sustained by the consistent behavior of the policy-making authorities."[30]

Different monetary institutions will, for instance, foster different sets of inflationary expectations. There is negative price change autocorrelation under a commodity standard. Thus, an increase in prices (or in the rate of change of prices) today will be followed by a decrease in prices (or in the rate of change of prices) in the future. In contrast, there has been positive price change autocorrelation under the modern fiduciary standard. Higher prices (or inflation rates) are followed by high or still higher prices (or inflation rates) in the future.[31] If the institutional environment (regime) is well established and understood, economic agents may be able to form consistent expectations. A set of consistent expectations establishes an intertemporal equilibrium.

The new institutional view does not hypothesize the same sort of tight prior equilibrium that is characteristic of rational expectations models.[32] If

29. The evolution and operation of institutions ought, nonetheless, to be explained by economic theory. Richard N. Langlois surveys the recent literature on this topic in "Economics as a Process: Notes on the 'New Institutional Economics,'" C. V. Starr Center for Applied Economics Research Report no. 82-21 (New York: New York University, 1982); also see the papers in Richard N. Langlois, ed., *The New Institutional Economics* (Cambridge: Cambridge University Press, forthcoming).

30. Axel Leijonhufvud, "Rational Expectations and Monetary Institutions," UCLA Department of Economics Working Paper no. 302 (Los Angeles: University of California, 1983), 2. He suggests this as a "loose" rational expectations concept.

31. For an important empirical analysis of the effects on expectations of the movement from the gold standard to today's fiduciary standard, see Benjamin Klein, "Our New Monetary Standard: The Measurement and Effects of Price Uncertainty, 1880–1973," *Economic Inquiry* 13 (December 1975): 461–84.

32. And before them, of the Chicago School. See Melvin W. Reder, "Chicago Economics: Permanence and Change," *Journal of Economic Literature* 20 (March 1982): 1–38 passim, especially 11–13.

it did, there would be a distinction without a difference between it and the rational expectations models generating policy ineffectiveness propositions. Rather, it "assumes that people understand the systematic components of the authorities' behavior in a general sort of way but avoids a linkage so tight as to build, for example, short-run neutrality or policy ineffectiveness assertions into the concept itself."[33]

Two corollaries of the revised equilibrium concept deserve mention. First, some monetary regimes may engender firmer expectations than others as to future movements of money prices and quantities. Information costs differ by regime. Accordingly, the scope for policymakers to exploit economic relationships, like the Phillips curve, may differ under alternative policy regimes. Unless a regime is first specified by the theorist, no definitive conclusion can be drawn on the effectiveness of such policies.

The second corollary is that changes in regime raise information costs and make the formation of appropriate expectations more difficult. This second point is especially important for understanding 20th-century monetary history. The pre-World War I era was characterized by a classic international gold standard, in which gold coin circulated domestically and international redeemability effectively fixed exchange rates. After World War I the gold standard was not completely restored internationally but was domestically. U.S. monetary policy neutralized gold flows, however, further modifying the standard.[34] In 1934, domestic convertibility was ended though international convertibility was maintained (a "gold bullion" standard). The Bretton Woods agreement further modified the international standard. Finally, in 1971 the gold window was closed in the United States, and floating exchange rates were introduced.

We have thus had four or five different monetary regimes in this century—perhaps a historical record for so short a period of time. Each change in regime altered the probability distribution of future policy actions. In other words, the "relevant" economic theory has probably changed with each change in regime.

It is by no means clear how to define the post-1934 regimes, which have consisted of a changing mixture of elements from commodity and fiduciary standards. They have been described as "vague and ambiguous."[35] Unfortunately, "a 'vague and ambiguous' standard can hardly determine a well-defined system of rational expectations."[36] Modern rational expectations theory may have developed at a time when it is least applicable to events.

The literature on monetary regimes suggests that macroeconomic theories must be less general, certainly less general than is true of rational expectations models. Macroeconomic theories must recognize the institutional dependence of expectations. In effect, distinct regimes require different macroeconomic theories.[37] If less general, however, macroeconomics may nonetheless be more applicable to the study of aggregate phenomena.

A resolution

Consideration of the role of institutions in the formation of expectations is by no means foreign to the rational expectations literature. In a series of seminal articles,[38] several rational expectations

33. Leijonhufvud, "Rational Expectations and Monetary Institutions," 2.

34. The United States was not alone in neutralizing gold flows. Ragnar Nurkse found "a striking fact" that for the interwar years (1922-38), "central banks' international and domestic assets, during most of the period under review, moved far more often in the *opposite* than in the same direction" as gold flows (League of Nations, *International Currency Experience: Lessons of the Inter-War Period* [1944.II.A.4], 68).

35. Milton Friedman and Anna Jacobson Schwartz, *A Monetary History of the United States, 1867-1960* (Princeton: Princeton University Press for National Bureau of Economic Research, 1963), 474; quoted in Leijonhufvud, "Rational Expectations and Monetary Institutions," 27.

36. Leijonhufvud, "Rational Expectations and Monetary Institutions," 27.

37. Ibid., 2-3.

38. In various issues of the *Federal Reserve Bank of Minneapolis Quarterly Review* since early 1979: Robert E. Lucas, Jr., and Thomas J. Sargent, "After Keynesian Macroeconomics," Spring 1979, 1-16; Thomas J. Sargent, "Rational Expectations and the Reconstruction of Macroeconomics," Summer 1980, 15-19; Thomas J. Sargent and Neil Wallace, "Some Unpleasant Monetarist Arithmetic," Fall 1981, 1-17; Neil Wallace, "A Legal Restrictions Theory of the Demand for 'Money' and the Role of Monetary Policy," Winter 1983, 1-7; and Neil Wallace, "Some of the Choices for Monetary Policy," Winter 1984, 15-24.

theorists have highlighted the variation in equilibrium expectations under alternative policy regimes. Consideration of the effects of the institutional environment on expectations formation has not, however, been a central feature of the rational expectations literature. Instead, the literature has been sidetracked by a fascination with questions like the policy ineffectiveness hypothesis, which is "uninstitutional."[39] Nonetheless, there are indications of a convergence of proponents and critics of rational expectations regarding the crucial importance of theoretically incorporating institutions into macroeconomic analysis.

More detailed consideration of the interplay of regimes and expectations awaits future research.[40] This research promises significant gains in our understanding of the effects of policy on economic variables. This article has analyzed only some of the effects of monetary institutions on expectations. The relationship between institutions or regimes and expectations is more general than the issues examined here. It encompasses fiscal rules and policymaking generally, as well as the whole range of economic institutions.[41]

It is not possible to specify the effects of a policy action without relating the action to the institutional environment in which the policy was adopted and the action taken. First, depending on the policy regime, given policy actions may have different effects. Second, the intelligibility and predictability of a policy action depend on the regime within which the decisionmaker operates. It is on this point that the analysis of regimes and the rational expectations hypothesis have the most in common. Rational expectations theorists have reminded us that the effects of policy are not invariant to the public's expectations. The policymaker proposes but the public disposes.

39. Two prominent proponents of rational expectations theory have observed that the implications of this theory "are sometimes caricatured, by friendly as well as unfriendly commentators, as the assertion that 'economic policy does not matter' or 'has no effect'" (Lucas and Sargent, "After Keynesian Macroeconomics," 14–15).

40. See John H. Wood, "The Search for a Monetary Policy Rule in an Uncertain World," this *Economic Review*.

41. For an examination of the current fiscal regime, see W. Michael Cox, "What Is the Rule for Financing Public Debt?" in this *Economic Review*.

The Search for a Monetary Policy Rule in an Uncertain World

*By John H. Wood**

> *"I don't think they play at all fairly," Alice began, in rather a complaining tone, "and they all quarrel so dreadfully one can't hear oneself speak—and they don't seem to have any rules in particular: at least, if there are, nobody attends to them."*
>
> —Carroll, *Alice's Adventures in Wonderland*

The price and output fluctuations of recent years have stimulated the reconsideration of a wide variety of rules by which Federal Reserve actions might be guided. Some of the proposed rules would allow the Federal Reserve System little or no latitude of choice. One of these nondiscretionary rules would require the Fed to expand the monetary base or, even more narrowly, its asset portfolio at some constant rate. Milton Friedman has argued that the Fed's control of the money stock is so complete that a constant-money-growth rule approximates his ideal of no official discretion. Other rules would permit a great deal of discretion and might better be called "general guidelines." For example, some frequently proposed rules, or guides, call for the manipulation of interest rates to achieve either price stability or an acceptable combination of inflation and "real" economic activity. A proposal that has recently gained popularity would require the Fed to respond directly to prices, specifically to conduct open market operations with the intention of moderating or reversing observed price movements.[1]

All these rules have been proposed in the past and several have been tried. And all remain controversial. The purpose of the following discussion is to give some form to the controversy over rules by summarizing the theoretical conditions under which

* *John H. Wood is a visiting scholar at the Federal Reserve Bank of Dallas, on leave from Northwestern University. The views expressed are those of the author and do not necessarily reflect the positions of the Federal Reserve Bank of Dallas or the Federal Reserve System.*

1. References to articles favoring or criticizing some of these rules may be found in R. W. Hafer, "Monetary Policy and the Price Rule: The Newest Odd Couple," *Review*, Federal Reserve Bank of St. Louis, February 1983, 5-13.

each is appropriate and describing some of the apparent consequences of the rules that have actually been applied. Much of the discussion consists of historical case studies of rules in Great Britain and the United States during the 19th and 20th centuries. Perhaps we can learn from experience. Perhaps not. In any case, there is no excuse for the sense of novelty that accompanies "new" policy proposals or for the repeated attitude of surprise at their consequences.

No unqualified conclusions will be drawn because the "right" rule is contingent both upon one's views of the structure of the economy, especially whether and how money is believed to affect real economic activity, and upon the relative intensities of one's fears of inflation and unemployment. The range of views of the economic structure has not changed significantly since the 17th century, by which time the two opposing camps had already been formed—those who favor expansionary monetary and fiscal policies to promote economic activity and the "hard money" people who believe that the benefits, if any, of inflation are greatly outweighed by its costs. Whatever theoretical contributions might be found in J. M. Keynes' *The General Theory of Employment, Interest and Money*,[2] its policy implications are not perceptibly different from those of William Petty,[3] who in 1662 urged public works for the unemployed, even if for no other reason than "to build a useless Pyramid upon Salisbury Plain" or to "bring the Stones at Stonehenge to Tower-Hill," or those of Josiah Child,[4] who in 1690 called for lower rates of interest to encourage investment.[5] The frequent and sometimes dramatic shifts between rules have been due to dissatisfaction with the results of current rules, leading to their replacement by rules that had themselves been rejected earlier, rather than to the discovery of new rules derived from new theories. The reactions of the past half century have been typical. The Great Depression heightened the fear of unemployment and increased the ranks of those advocating expansionary monetary and fiscal policies—until accelerating inflation led to the adoption of rules intended to achieve stable prices.

The following discussion considers interest-rate, price-level, monetary-base, and money rules in both their discretionary and nondiscretionary forms, principally in the context of policies that are limited to the goal of price stability. We begin with the operation of the Bank of England under the 19th-century gold standard because most of the intellectual and practical basis of central banking was formed by the conduct of the Bank of England, and conflicting opinions about that conduct, between the Napoleonic Wars and World War I. Each of the rules considered might be applied, at least for a while, under either a gold standard or the unrestricted paper standard of the present day. However, one of these systems implies limits to policy discretion whereas the other does not. The main difference between the application of rules under gold and paper standards is that in the former system rules resulting in inflation must eventually be abandoned because of the obligations of central banks and other banks to redeem their liabilities in gold.

The development and rejection of a nondiscretionary rule: the Bank of England in the 19th century

The Bank of England was formed in 1694 and remained privately owned until 1946. But it was the chief lender to the government and by the end of the 18th century had acquired all the influence on the monetary system that can be possessed by a central bank within the confines of the gold standard—although the Bank was imperfectly aware of that influence until later. By a combination of its monopoly of joint-stock banking in London, its role as the government's bank, and a conservative lending policy, the Bank of England had achieved a reputation for safety in an industry plagued by failures. The Bank's notes and deposits were considered almost "as good as gold" and the country banks held them as reserves against their own note

2. New York: Harcourt, Brace and Co., 1936.

3. *Treatise of Taxes and Contributions* (London: Brooke, 1662). Reprinted in Charles H. Hull, ed., *The Economic Writings of Sir William Petty*, vol. 1 (New York: Augustus Kelley, 1963), 31.

4. *Discourse About Trade* (London: J. Hodges, 1690).

5. Keynes recognized the similarities between his own arguments and policy prescriptions and those of Petty, Child, and other early economists in chapter 23 of *The General Theory*. In the preceding article in this *Review*, Gerald O'Driscoll discusses some of the history of the debate concerning the so-called Phillips curve trade-off between inflation and unemployment.

and deposit liabilities. This service as bank reserves meant that the Bank's liabilities, as much as gold, were high-powered money in the 18th century in the same sense that the liabilities of the Bank and the Federal Reserve are high-powered money today. An increase in the Bank's lending, and therefore in its note and deposit liabilities, had a multiplier impact on aggregate money and credit because other banks held those liabilities as reserves against expansions of their own notes and deposits.

But the Bank could not extend credit without limit. Excessive increases in money caused inflation, an adverse balance of payments, and an outflow of gold as foreigners converted their newly acquired pounds into gold. Rising prices also induced increased transactions demands for cash, including gold coin. These external and internal drains meant that the Bank eventually had to restrict its lending to protect its gold reserve.

This discipline was abandoned in 1797 when wartime inflationary finance led to gold losses and forced the suspension of convertibility. That is, banks were relieved of the obligation to redeem their liabilities for gold at the official rate of £3.89 per ounce of gold.[6] Monetary expansion and the rate of inflation were dictated by the government's war needs and by 1813 the price of gold had risen to £5.50. After the defeat of Napoleon, the government decided to return to the gold standard at the prewar rate. This meant deflation, which was achieved with a vengeance. The Bank's credit was reduced from £48 million in 1814 to £17 million in 1822, accompanied by a fall in prices of nearly 40 percent. The price deflation was accompanied by a depression in trade, high unemployment, and often violent public protests. David Ricardo argued that the severity of the deflation had been unnecessary and was caused by gross mismanagement by the Bank of England—which, however, "denied that it was in any way responsible for controlling the national currency, and indeed doubted whether it had the power even to influence it."[7]

Throughout this period the Bank sought to escape blame for the instability of prices by taking refuge in the real bills doctrine, according to which the discounting of bills for the purchase of goods ("real bills") was sufficient to cause money and credit to rise and fall with real transactions so that the price level would be stable. The Bank's directors testified in 1810 that their lending merely supplied the needs of trade. They argued that note issues result solely "from the applications made for discounts to supply the necessary want of Bank Notes, by which their issue in amount is so controlled that it can never amount to an excess." Given that the Bank discounts only for "solid persons" and "*bona fide* transactions," the public "will never call for more than is absolutely necessary for their wants."[8]

The real bills doctrine is likely to be more conducive to the instability than the stability of prices. Tying the quantity of money to the money value of goods renders both indeterminate. In any case, whatever their true beliefs, the Bank's directors discontinued avowals of the real bills doctrine after resumption of the gold standard. Gold standard central banks cannot afford the luxury of nonmonetary theories of inflation. They must look to their reserves and restrict credit when inflation causes gold losses. But monetary and price stability were not much improved after convertibility was resumed in 1821. The Bank's responses to fluctuations in its reserve were erratic and continuously a source of controversy. However, the Bank began to show signs of feeling its way toward a coherent policy in the form of the "Palmer rule,"[9] by which its security holdings were to be held constant with the purpose of causing the currency to fluctuate precisely in the manner of an unregulated metallic currency. The Bank was to pay out or redeem notes only as gold was brought to or taken from the Bank. If its note and deposit liabilities were prevented from changing

6. In fact, by the Bank Restriction Act of May 1797 the Bank of England was forbidden to make payments in gold except in a few specified circumstances. Shillings and old pence have been avoided by expressing pounds in decimals.

7. A. E. Feavearyear, *The Pound Sterling* (Oxford: Clarendon Press, 1931), 212.

8. *Report from the Select Committee on the High Price of Bullion*, House of Commons, 1810. Reprinted with an introduction by Edwin Cannan in *The Paper Pound of 1797–1821*, 2d ed. (London: P. S. King and Son, 1925), 47.

9. Named for Horsley Palmer, the Governor of the Bank who described the rule to a parliamentary committee in 1832. Palmer's testimony may be found in T. E. Gregory, *Select Statutes, Documents and Reports Relating to British Banking, 1832–1928*, vol. 1 (London: Oxford University Press, 1929), 3–6.

Table 1
THE BANK OF ENGLAND'S BALANCE SHEET ON SEPTEMBER 7, 1844

Issue Department			
Government securities	£11,015,100	£28,351,295	Notes issued
Other securities	2,984,900		
Gold coin and bullion	12,657,208		
Silver bullion	1,694,087		
Total assets	£28,351,295	£28,351,295	Total liabilities
Banking Department			
Government securities	£14,554,834	£ 3,630,809	Government deposits
Other securities	7,835,616	8,644,348	Other deposits
Notes	8,175,025	1,030,354	Seven-day and other bills
Gold and silver coin	857,765	18,117,729	Capital
Total assets	£31,423,240	£31,423,240	Total liabilities and capital

NOTE: The Bank's position (or "return") was required to be published weekly in the *London Gazette*. It was also published in *The Economist*. The above return was the first under the 1844 act and has been published in many places, including A. Andréadès, *History of the Bank of England, 1640 to 1903* (London: P. S. King, 1935), 290. The items have been rearranged and some have been renamed to conform to modern American usage.

because of changes in the Bank's lending, the Bank's liabilities would rise and fall, pound for pound, with movements of gold into and out of the Bank.

But the Palmer rule served mainly as a conversation piece, the Bank's erratic behavior continued, and so did monetary controversy. Of the many proposals for reform, one group of writers, called the Banking School, argued that no changes in the Bank's organization or powers were needed and that no strict rule of policy was desirable. All that was required was sound banking practice by the Bank of England and other banks. Specifically, the Bank should hold a sufficiently large reserve that it would not have to contract its lending upon every outflow of gold. They pointed out that the maintenance of the gold standard did not require the Bank to contract or expand its credit immediately or to the full extent of every outflow or inflow of gold. In the event of a temporary external drain due to a bad harvest, for example, public welfare would be better served by an increase in bank lending, that is, by offsetting instead of reinforcing the effects of the gold loss on money and prices.[10]

The Currency School, on the other hand, advocated the separation of the Bank into two departments. The Banking Department would be left free to behave like any other bank but the Issue Department would function strictly, without discretion, in accordance with the Palmer rule. Parliament adopted the Currency School's proposals in the Bank Act of 1844.[11] The first statement of the Bank under the new act is shown in Table 1. The monetary rule under which the Issue Department operated between 1844 and 1914 required the quantity of its notes to be kept equal to the sum of a

10. The Banking School's proposals may be found in Thomas Tooke, *An Inquiry into the Currency Principle, the Connection of the Currency with Prices and the Expediency of a Separation of Issue from Banking* (London: Longman, Brown, Green, and Longmans, 1844).

11. The Currency School's position was summarized by Samuel Jones Loyd, *Further Reflections on the State of the Currency and the Action of the Bank of England* (London: Pelham Richardson, 1837). Several interesting documents relating to the controversy leading up to the Bank Act of 1844 are contained in Gregory's *Select Statutes*.

"fiduciary" amount of £14 million and its holdings of gold and silver coin and bullion. The Bank was not permitted to issue notes in payment for additional securities but was required to issue notes in exchange for all the gold offered to it at the fixed rate of £3.89 per ounce. The Banking Department was bound by no rule and was free to borrow and lend as it chose, like other banks, with only its own safety and profits in view.

But monetary conditions during the 20 years following the act of 1844 were as volatile as in the preceding 20 years. A series of monetary expansions followed by banking crises led to the act's suspension in 1847, 1857, and 1866. On each of these occasions, in the midst of rapid declines of the Banking Department's note reserve and the Issue Department's gold reserve, the government of the day permitted the Bank to issue notes beyond the legal limit. The failure of the act during its early years has been attributed to a variety of causes—most important, the failure of both Parliament and the Bank to recognize the significance of the Bank's checking account liabilities, either as part of the money supply or as an increasingly important component of the monetary base. Fluctuations in the lending of the Banking Department continued to influence the money supply fully as much after as before 1844. Nothing had happened to diminish the significance of all claims on the Bank of England, deposits as well as notes, in the reserves of other banks. The Banking School had emphasized this deficiency of the Bank Act, and the appropriate identification of money—whether it should include bank deposits and perhaps bills of exchange and other credit market instruments as well as currency and coin—was a hotly debated issue throughout this period. The Currency School's error has seemed obvious to most later writers—although more because of their exclusion of the Bank of England's deposits from the identification of the monetary base than from the identification of money. But the correct choice was not and is not clear, and problems of this kind have always arisen from decisions to follow rules based on monetary aggregates.

The Bank eventually adjusted to its role as the nation's central bank and came to realize that its discretionary management of the Banking Department was as important to the determination of money and prices as its nondiscretionary operation of the Issue Department. But it should be emphasized that short-run price stability was never the Bank's overriding objective. Its legal commitment to the gold standard was thought (correctly, as matters turned out) to assure long-run price stability. The Bank's short-run behavior before 1914 was much like its own stop-go policies and the stop-go policies of other central banks during the era of more or less fixed exchange rates of the 1950s and 1960s. Increases in the Banking Department's reserve because of gold inflows (which were exchanged for the Issue Department's notes) permitted increases in the Bank's lending, that is, in the monetary base, which in turn led to increases in money and prices. The main difference (perhaps the only important difference) between the Bank's behavior before 1914 and in recent years is that during the earlier period gold losses eventually forced the ends of inflationary policies.

Interest-rate rules

Wicksell's variable-interest-rate (or price-level) rule. Although there was often acute dissatisfaction with the gold standard, its reputation grew with the prosperity of the late 19th century. But the gold standard probably derived most of its strength from lively memories of the even more volatile unrestricted paper-money systems of earlier periods. Most economic writers accepted the need for some kind of nondiscretionary limit on the quantity of money—and no persuasive alternative to the convertibility of money into gold at a fixed rate was forthcoming. Nevertheless, many writers still saw considerable scope for reform within this constraint. R. G. Hawtrey later argued that an important cause of price instability before 1914 was the slow response of the Bank of England to changing conditions.

> Professor Fisher in his *Purchasing Power of Money*[12] held that the fluctuations of the trade cycle were due to the adjustment of the market rate of interest to the real rate being incomplete. When prices were rising, the market rate rose, but not high enough to make the real rate normal; when prices were falling, it fell, but not low enough. . . .
>
> Professor Fisher, perhaps, laid too much stress on the tendency to overlook the effect of rising or

12. Irving Fisher, *The Purchasing Power of Money* (New York: Macmillan, 1911).

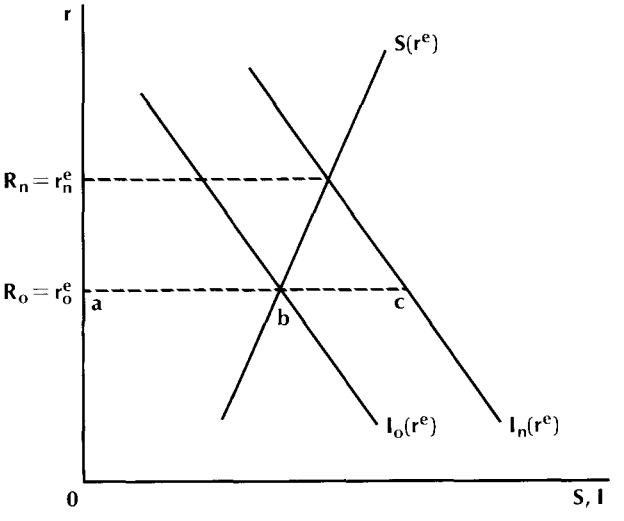
falling prices when determining the rate of interest. For Bank rate was settled by a method of trial and error, which allowed automatically for all the factors at work. The delay in adjusting it, to which Professor Fisher quite rightly attributes the fluctuations, was really due to the practice of using the Bank of England's reserve as a criterion. The reserve responded very tardily to an expansive or contractive movement, so that there were long periods in which the Bank acquiesced in such movements without attempting to counteract them. But when it did resort to the rate of interest as a corrective, the Bank simply went on raising or lowering the rate till the desired effect followed. If in raising the rate it stopped short at 7 per cent. or 9 per cent., that was because the market responded. (*A Century of Bank Rate*, 1938, pp. 213–14)

Knut Wicksell had earlier observed the same tendencies and proposed as a remedy that the central bank should respond directly to the price level without waiting for the effect of the latter on its reserve.[13] Wicksell's rule may be explained within the simplified version of his model depicted in Figure 1. Let I be real investment, that is, net additions per unit of time to plant, equipment, houses, and inventories. S is real annual saving; it is the portion of national income (output) that is not consumed. Investment and saving are shown as negative and positive functions, respectively, of the expected real rate of interest, r^e. Given the initial investment function, I_o, the market for goods and services is in equilibrium at the intersection of I_o with the saving function, S. Desired saving and investment are equal at the expected real rate of interest, r_o^e. In equilibrium the nominal rate of interest, R, is equal to the expected real rate of interest plus the expected rate of inflation, p^e. We assume no inflation initially so that nominal and expected real interest rates are the same.

Now consider an increase in the demand for investment goods to I_n, perhaps because of technological advances or the opening of new

Figure 1

The Effects of an Increase in the Demand for Investment Goods



markets that raise the expected profitabilities of new or expanded business ventures. Given an unchanged saving function, nominal and expected real interest rates will be R_n and r_n^e in the new stable-price equilibrium. But some inflation is bound to occur in the process of moving from the old to the new equilibrium. After I has increased from I_o to I_n but before R has responded, there is an excess demand for goods, represented by the distance bc. This distance also represents the real excess demand for funds required to finance the excess demand for goods. These excess demands put upward pressure on prices and interest rates, with the higher prices being financed by an increase in the velocity of money. Eventually, R rises to R_n and the new stable-price equilibrium is established at a higher price level and a higher velocity of money.

We have so far neglected two important aspects of experience: first, the universally strong positive association between money and prices, and second, the almost equally strong tendency of interest rates to lag behind price movements. The rapid rise in money and the tardy response of R may be explained in terms of bank demands for reserves. Bank

13. Wicksell's views on this subject are found mainly in *Interest and Prices*, published in Sweden in 1898, translated by R. F. Kahn (London: Royal Economic Society, 1936); and *Lectures on Political Economy*, vol. 2, *Money*, published in Sweden in 1906, translated by E. Classen (London: Routledge and Kegan Paul, 1934).

allocations of assets between loans and cash reserves depend partly on loan rates. Suppose that before the onset of an economic expansion, when interest rates are low, banks hold large reserves relative to their deposits. Under these conditions bank competition for new loans in response to the rise in investment demand from I_o to I_n might for a while inhibit increases in interest rates. Consider the extreme case in which the nominal rate of interest remains at R_o and other loan terms are not made more stringent. Borrowers obtain all the funds they desire at the interest rate R_o. The funds supplied by savers (through the capital markets, banks, and other financial intermediaries) are equivalent, at expected prices, to the quantity of investment goods indicated by the distance *ab* in Figure 1. But the excess demand *bc* is financed by banks' creation of money.

However, investors will not be able to fulfill all of their demands *ac* because only the quantity *ab* (plus possible small increments due to increases in production and the crowding out of consumption demands) is available. As much as we might hope otherwise, the demand for goods does not create its own supply, except possibly in deep depression and then not immediately. Since the excess demand is not eliminated by a rise in R to R_n as in the previous example, when the money stock did not change, it must be eliminated by price increases and delays in delivery. Would-be investors find that their loans do not command as many goods as they had expected. These price increases, and probably expectations of further increases, will be taken into account in new loan requests—so that the same excess demand for goods (*bc*) will generate greater and greater increases in money and prices in future periods as long as R does not rise. But eventually, declining bank reserves relative to deposits must, in combination with increasing loan demands caused by accelerating inflation, force interest rates to rise.

Wicksell argued that central banks at times prolonged the inflationary process by resisting increases in interest rates. The end of the process described above was hastened by the limited quantity of bank reserves. But a central bank can supply additional reserves. It can even for a time ensure that interest rates will not rise by supplying all the credit the system desires at existing rates. In terms of Figure 1, for example, a central bank has the ability to peg the rate of interest at R_o by standing ready to buy or sell unlimited quantities of securities at that rate. Of course, this requires ever-increasing purchases of securities in the presence of the real excess demand *bc* as the dollar values of loan demands rise to keep up with the prices of the goods to be financed. The gold standard eventually forced a halt to this process. But Wicksell suggested that the inflation would not have become serious if the central bank had taken early action to encourage interest rates to rise to a level consistent with price stability. Conversely, increases in central bank lending should quickly follow price declines. Wicksell stated his rule as follows:

> *So long as prices remain unaltered the banks' rate of interest is to remain unaltered. If prices rise, the rate of interest is to be raised; and if prices fall, the rate of interest is to be lowered; and the rate of interest is henceforth to be maintained at its new level until a further movement of prices calls for a further change in one direction or the other.*
>
> The more promptly these changes are undertaken the smaller is the possibility of considerable fluctuations of the general level of prices; and the smaller and less frequent will have to be the changes in the rates of interest. (*Interest and Prices*, 1898, p. 189)

A stable-interest-rate rule. Wicksell's objective was price stability and his rule that interest rates be allowed to vary without resistance from, and even with the help of, the central bank was based on the belief that economic disturbances stem principally from fluctuations in demand, such as the shift in investment in Figure 1. But the goal of price stability implies a very different central bank response to interest-rate movements when those movements are caused by fluctuations in the demand for money. For example, an increase in the demand for money relative to other assets causes interest rates to rise and, other things equal, depresses commodity demands and the price level. Prices fall until, given existing nominal money balances, real money balances conform to the public's new higher demand at the equilibrium rate of interest determined by real saving and investment. This price fall could have been avoided by a central bank determined to maintain stable interest rates. An increase in the demand for money that puts upward pressure on interest rates could have been met by an increase in central bank lending. The stock of money would rise in line with the public's demand for money, which

would be satisfied by an increase in nominal money balances instead of by a fall in prices.

So the goal of price stability may imply either a variable or a stable interest-rate rule, depending on the central bank's belief about the source of disturbances.[14] However, the high and long-lived positive correlations between private investment, prices, and interest rates, which were observed by Wicksell and many others in several countries during the 19th century and have been documented for the United States in the 20th century, strongly support the view that fluctuations in the demand for investment goods dominate those in money.[15]

A discretionary monetary-base rule: the Federal Reserve in the 1920s

The successful operation of the gold standard before 1914 was made possible by the willingness of the principal actor to play according to the "rules of the game." The Bank of England allowed gold movements to affect domestic money and prices and by its own actions reinforced those effects. Other central banks were less inclined to play this game and tended even in the 19th century to neutralize gold flows.[16] But Britain's strength in the international economy, particularly its pull on the world's gold and its willingness to allow gold to flow out again, was sufficient to make the gold standard work.[17] All this was ended by the decline of the British economy and Britain's replacement by the United States as the dominant force in world trade.

The necessities of war finance had again forced the suspension of convertibility by the Bank of England during World War I. Wartime and postwar inflation were more severe in Britain than in the United States, and by November 1920 the pound had fallen from its prewar par value of $4.86 to $3.44. (One of the indications of the dominance of the United States and the world's confidence in the gold value of the dollar was the new practice of quoting currencies in terms of the dollar.) But Britain was as determined after World War I as after Waterloo to return to gold at the prewar par. Restrictive monetary and fiscal policies were pursued and by 1925 the pound had been forced to $4.86, where it was maintained until 1931. The question of whether the pound was overvalued during this period is still controversial. In any case, the authorities were prepared to bear the costs of deflation necessary to return to and then to maintain the gold standard at the prewar parity. But they hoped that deflation would not be necessary. To a large extent British policies were based on the expectation that the accumulations of gold in the United States during and after the war would be allowed to affect American money and prices. But this hope was disappointed by the Federal Reserve's goal of price stability. Benjamin Strong, President of the Federal Reserve Bank of New York and the most influential official in the Federal Reserve System until his death in 1928, declared that the Federal Reserve would not expand its credit in response to inflows of gold. In 1923 Strong wrote to Montagu Norman, Governor of the Bank of England, that with America's "excessive gold stock we must entirely ignore any statutory or traditional percentage of reserve and give greater weight to what is taking place in prices, business activity, employment, and credit volume and turnover."[18]

There was considerable interest in Congress in legislation that would require the Federal Reserve to stabilize prices. Strong opposed such legislation because he felt that unpredictable factors outside the Fed's control also influenced prices. He testified before the House Committee on Banking and Currency that "I do not believe, and have never be-

14. It has been demonstrated that stable money growth is preferred to stable interest rates when disturbances stem from fluctuations in commodity demands but that this preference is reversed when disturbances stem from fluctuations in the demand for money. See William Poole, "Optimal Choice of Monetary Policy in a Simple Stochastic Macro Model," *Quarterly Journal of Economics*, May 1970, 197–216; and John Karaken, "The Optimum Monetary Instrument Variable," *Journal of Money, Credit, and Banking*, August 1970, 385–90.

15. See Arthur F. Burns and Wesley C. Mitchell, *Measuring Business Cycles* (New York: National Bureau of Economic Research, 1946); and *Business Conditions Digest*.

16. See Arthur I. Bloomfield, *Monetary Policy Under the International Gold Standard, 1880–1914* (New York: Federal Reserve Bank of New York, 1959).

17. See D. E. Moggridge, *British Monetary Policy, 1924–1931: The Norman Conquest of $4.86* (Cambridge: Cambridge University Press, 1972), chap. 1.

18. Lester V. Chandler, *Benjamin Strong, Central Banker* (Washington, D.C.: Brookings Institution, 1958), 188.

lieved, that any method of fixing the general level of commodity prices can be devised which would enable a monetary and credit policy by a bank of issue to accomplish that object."[19] Nevertheless, he believed that Federal Reserve actions were the principal determinant of the price level and accepted price stability as the Fed's primary goal, which he proposed to achieve by a monetary-base rule: "If I were Czar of the Federal Reserve System I'd see that the total of our earning assets did not go much above or below their last year's average, after deducting an amount equalling from time to time our total new gold imports."[20] This policy of gold neutralization was in fact applied. During 12 of the 13 years from 1921 to 1933, Federal Reserve credit moved oppositely to changes in the monetary gold stock.[21] The United States (and France) continued to accumulate gold in the late 1920s, and the inflows accelerated during 1930 and 1931 because of increasingly restrictive tariffs and the greater severity of the depression in the United States than in Great Britain and most other countries. Finally, in September 1931 the Bank of England was forced to abandon the convertibility of the pound at $4.86.

It should be pointed out that the goal of short-run price stability, if it is shared by all countries, is not inconsistent with a permanently successful gold standard. What was demonstrated by the experience of 1914–31 was the inability of a gold standard, or any fixed-rate system, to survive substantially different rates of inflation in different countries.

Friedman's rule

The stability of a free-market economy. The famous money rules of Milton Friedman and others at the University of Chicago have been based on the belief that, whatever the defects of an unregulated economy, its performance is worsened by government interference. Discretionary and therefore necessarily unpredictable monetary policies increase the uncertainties of an already uncertain world. In his paper on "Rules versus Authorities in Monetary Policy," Henry Simons argued: "An enterprise system cannot function effectively in the face of extreme uncertainty as to the action of monetary authorities or, for that matter, as to monetary legislation. We must avoid a situation where every business venture becomes largely a speculation on the future of monetary policy."[22] The only means of achieving this end was to take discretion completely away from the central bank and to make its actions fully predictable. None of the earlier rules (Simons did not regard them as rules) of central bank conduct satisfied these requirements. The Bank of England's wartime accommodative "real bills" approach placed its lending at the mercy of private and government credit demands; the gold standard left much to the Bank's discretion, even under the act of 1844, and would have done so even if Wicksell's interest-rate rule had been adopted; and Benjamin Strong valued the freedom to depart from his monetary-base rule whenever conditions required. Milton Friedman later argued that the lags of monetary policy were long and variable, so that when the central bank followed loose guidelines such as Wicksell's interest-rate rule, its actions to combat inflation became effective only after the problem had become deflation, with the result that monetary policy tended to accentuate price fluctuations.[23] "The problem with leaning against the wind is that you have to lean against the right wind, and then the problem is that you have to lean against next year's wind, not this year's wind, and unfortunately the people in Washington don't know the direction of next year's wind any better than we do."[24]

Constant money growth or constant growth of Federal Reserve credit? As the best means of achieving a monetary environment in which uncertainty would be minimized, Simons looked to the "ultimate establishment of a simple mechanical rule of monetary policy,"[25] preferably "the fixing of the

19. Ibid., 204.

20. Ibid., 191.

21. *Banking and Monetary Statistics, 1914–1941* (Washington, D.C.: Board of Governors of the Federal Reserve System, 1943), 369–71.

22. "Rules versus Authorities in Monetary Policy," *Journal of Political Economy*, February 1936, 1–30. Reprinted in Henry Simons, *Economic Policy for a Free Society* (Chicago: University of Chicago Press, 1948), 161.

23. See "The Lag in Effect of Monetary Policy," *Journal of Political Economy*, October 1961, 447–66.

24. Paraphrased from a speech in New York, November 15, 1983.

25. Simons, "Rules versus Authorities," 170.

quantity of circulating media."[26] But he did not feel that the time was ripe for this rule, partly because of "the abundance of what we may call 'near-moneys'—with the difficulty of defining money in such a manner as to give practical significance to the conception of quantity"[27] and partly because, however money might be defined, the Federal Reserve did not possess complete control of its quantity. He argued that these problems could be resolved only by a radical change in the financial system, including the elimination of all short-term borrowing (and therefore the elimination of money substitutes) and the imposition of 100-percent reserve requirements on commercial bank deposits, which following the first change would be limited to demand deposits (so that fluctuations in bank reserve behavior and the public's desired currency-deposit ratio would not affect the quantity of money).[28] Under these conditions money would be clearly identifiable and equal to the monetary base, which could be controlled by the Fed.

Friedman supported most of Simon's proposals to make monetary control effective, particularly 100-percent reserves. But he argued in 1959 that

> . . . even under present circumstances, the links between Reserve action and the money supply are sufficiently close, the effects occur sufficiently rapidly, and the connections are sufficiently well understood, so that reasonably close control over the money supply is feasible, given the will. I do not mean to say that the process would not involve much trial and some error but only that the errors need not be cumulative and could be corrected fairly promptly. (*A Program for Monetary Stability*, p. 89)

The rule by which this control of money should be exercised was described and supported at length in the last chapter of *A Program for Monetary Stability* and was later stated succinctly as follows:

> My choice at the moment would be a legislated rule instructing the monetary authority to achieve a specified rate of growth in the stock of money. For this purpose, I would define the stock of money as including currency outside commercial banks plus all deposits of commercial banks. I would specify that the Reserve System shall see to it that the total stock of money so defined rises month by month, and indeed, so far as possible, day by day, at an annual rate of X per cent, where X is some number between 3 and 5. The precise definition of money adopted, or the precise rate of growth chosen, makes far less difference than the choice of a particular definition and a particular rate of growth. (Friedman, *Capitalism and Freedom*, 1962, p. 54)

But this rule is unsatisfactory by Friedman's own criteria. In particular, it does not deprive the central bank of discretion. It is an improvement over the price-level guides of which Friedman was critical only in some unknown degree—and then only if we know what money is. Friedman has written that

> A satisfactory policy guide or rule should be connected more directly with the means available to the monetary authority than is the price level. We will, I believe, further the ultimate end of achieving a reasonably stable price level better by specifying the role of the monetary authorities in terms of magnitudes they effectively control and for whose behavior they can properly be held responsible than by instructing them solely to do the right thing at the right time when there is no clear and accepted criterion even after the event whether they have done so. (*A Program for Monetary Stability*, p. 88)

One such magnitude is the stock of Federal Reserve security holdings. A rule defined in terms of that stock would eliminate central bank discretion. It would also, by Friedman's own assumption of stable private behavior unless shocked by government interference, probably be more consistent with the stable growth of money (by a variety of definitions) than would the feedback approach to money control suggested by Friedman and quoted above. All the arguments about the proper definition of money and how best to control it, while of academic interest and crucial to some fine-tuning approaches to policy, have no relevance for a true

26. Ibid., 163.

27. Ibid., 171.

28. For example, letting m be the money multiplier, B be the monetary base, c and t be the nonbank public's currency-demand deposit ratio and time deposit-demand deposit ratio, respectively, and k_d and k_t be reserve ratios on demand and time deposits, respectively, we see that

$$M = mB = \left(\frac{1 + c}{k_d + c + tk_t}\right)B = \left(\frac{1 + c}{1 + c}\right)B = B$$

when $k_d = 1$ and $t = 0$.

nondiscretionary rule.

Still looking

In summary, those in favor of discretion might choose either an interest-rate rule or a money-supply rule. A money rule is desirable in the presence of a stable relationship between the quantity of money and the price level (which requires a stable demand for money) but an unstable relationship between the Federal Reserve's earning assets and the money supply (that is, an unstable money multiplier), so that the Fed must be prepared to adjust its actions in response to, or in anticipation of, variations in the money multiplier.

Now consider an interest-rate rule. But which one? If disturbances emanate from the demand for money, rising interest rates imply increases in central bank credit to combat deflation. If rising interest rates are caused by increases in demand, on the other hand, price stability requires the restriction of credit. These problems suggest that a better policy might be to respond directly to the price level, expanding or contracting central bank credit as price changes are less or greater than desired. However, this approach increases the chances that official actions will be too late and perhaps even exacerbate cyclical movements in prices and other variables.

So for those who desire a discretionary rule but do not claim complete understanding of the economy, the unpleasant choice lies between early actions that may be incorrect and correct actions that are too late—and the quandary prevails not only in the present flexible-exchange-rate paper-money regime but also in gold and other fixed-rate regimes. This is not a pretty picture. But a better one is not available because of the lack of a satisfactory macroeconomic model from which correct and timely actions may be inferred. The entire controversy about monetary policy rules stems from the absence of such a model. If a good model were available, complete with calculable uncertainty known to the policymakers, the controversy would be silenced and only the arithmetic necessary to the calculation of the "optimal policy" would remain. The only rules which do not require degrees of knowledge that have not been, and perhaps cannot be, even remotely attained are nondiscretionary rules of the kind logically implied by the approach of the Chicago School—such as a constant rate of change of Federal Reserve credit.

What Is the Rule for Financing Public Debt?

*By W. Michael Cox**

Of the two major types of financial paper issued by government, money and bonds, traditionally money alone has been viewed as inflationary. However, the hypothesis that bonds rolled over forever are like money, with potentially the same inflationary implications, is not new and has recently been a subject of renewed attention.

One view under which bonds are potentially inflationary focuses on net financial wealth of the private sector as the determinant of the price level.[1] Under this view, base money issued by the Federal Reserve is inflationary because it is an asset of the private sector for which there is no corresponding liability; it is net financial wealth. U.S. Treasury securities also are potentially inflationary if (1) individuals do not fully discount the future tax liabilities associated with public debt issuance or (2) public debt is not *in fact* financed by offsetting future tax liabilities.

For many years economic theory largely assumed that there were in fact equivalent offsetting tax liabilities associated with public debt issuance, and attention was focused on the question of whether individuals fully discount these tax liabilities.[2] Emphasis on this aspect is understandable because over an extended period the federal budget (including interest outlays) was roughly balanced. Over

* *W. Michael Cox is a senior economist at the Federal Reserve Bank of Dallas. The views expressed are those of the author and do not necessarily reflect the positions of the Federal Reserve Bank of Dallas or the Federal Reserve System.*

1. For a clear exposition of the hypothesis that the price level is related to net financial wealth of the private sector, see Don Patinkin, *Money, Interest, and Prices: An Integration of Monetary and Value Theory*, 2d ed. (New York: Harper & Row, 1965). The term "net financial wealth" refers to fiduciary assets less liabilities of the private sector. The focus is not on whether government bonds are net wealth in the sense of augmenting the overall consumption possibility set or utility of private individuals (they do not under this view) but on whether the stream of interest receipts from government debt actually involves an offsetting tax liability.

2. See, for example, Martin J. Bailey, *National Income and the Price Level: A Study in Macroeconomic Theory*, 2d ed. (New York: McGraw-Hill Book Company, 1971); Robert J. Barro, "Are Government Bonds Net Wealth?" *Journal of Political Economy*

the period 1948–70, for example, the average fiscal-year federal budget deficit was a comparatively small $3.7 billion. In contrast, over the period 1971–81 the average budget deficit was $39 billion, and in more recent years $110 billion to $195 billion. In view of this experience, it has become questionable whether Congress really intends for public debt to be entirely tax-financed, and attention has returned to the issue of whether government securities unbacked by taxes are like money.

This article provides some evidence on the extent to which government securities have been backed by taxes. On average, over the period 1950–81 each $1 of interest paid by the Treasury bore a tax liability of only 41 cents. The evidence also indicates that the Treasury's interest-financing behavior changed during this period. For each $1 received in the form of interest on government bonds before January 1971, there was a corresponding expected tax liability of 62 cents—hence a net receipt of interest of 38 cents by the private sector. Interest receipts accruing after that time, however, have borne no tax liability, and there has been a dollar-for-dollar expected net receipt of interest by the private sector.

Identifying a financing rule for public debt

This section models the way in which public debt is financed in the United States. As a necessary feature of that model, primary attention is paid to the consolidated budget of the Treasury and the Federal Reserve and to a mode of behavior—a "rule"—for financing the interest on public debt.

This rule is viewed as being determined by the Treasury (in the broad sense of including Congress).[3] It is not important that the Treasury has a conscious debt-financing rule per se. Whatever the behavior of the Treasury (tax policy, deficit policy, and so forth), the *result* is that a certain fraction of the interest on Treasury debt is deficit-financed. The important point is that there are a variety of policies under which the private sector may receive interest on Treasury debt without a corresponding tax liability. This article does not attempt to explore those many forms of policy or postulate a motivation for debt financing on the part of the Treasury. The approach instead is to accept policy and determine its implications for the financing rule implicitly promised by Treasury debt.

In setting out the model, it is helpful first to delineate clearly the decision-making structure. This study assumes that there are three economic decision-making units—private individuals, the Treasury, and the Federal Reserve. The Treasury is responsible for making decisions regarding the financing of government spending. This includes financing of the public debt. Specifically, the Treasury determines what part of the interest and principal on public debt it will finance with tax revenues and what part it will finance with new bond issuance. For simplicity, it is assumed that there is only one type of Treasury bond. Each bond promises to pay $1 per period to its owner, forever. It is also assumed that the Treasury views the financing of interest on a per-dollar basis. That is,

82 (November/December 1974): 1095–1117; Robert J. Barro, "On the Determination of the Public Debt," *Journal of Political Economy* 87 (October 1979, pt. 1): 940–71; and Robert J. Barro, "The Public Debt," chap. 15 in *Macroeconomics* (New York: John Wiley & Sons, 1984). For a clear discussion of the issue, see Gerald P. O'Driscoll, Jr., "The Ricardian Nonequivalence Theorem," *Journal of Political Economy* 85 (February 1977): 207–10. For empirical work on the issue of whether individuals discount the tax liabilities associated with government debt, see Levis A. Kochin, "Are Future Taxes Anticipated by Consumers?" *Journal of Money, Credit, and Banking* 6 (August 1974): 385–94; Jess B. Yawitz and Laurence H. Meyer, "An Empirical Investigation of the Extent of Tax Discounting," *Journal of Money, Credit, and Banking* 8 (May 1976): 247–54; J. Ernest Tanner, "An Empirical Investigation of Tax Discounting," *Journal of Money, Credit, and Banking* 11 (May 1979): 214–18; and John J. Seater, "Are Future Taxes Discounted?" Federal Reserve Bank of Philadelphia Research Paper no. 50 (Philadelphia, 1980). A discussion and analysis of the issue of whether bonds are like money may be found in John Bryant and Neil Wallace, "The Inefficiency of Interest-bearing National Debt," *Journal of Political Economy* 87 (April 1979): 365–81; Preston J. Miller, "Deficit Policies, Deficit Fallacies," *Federal Reserve Bank of Minneapolis Quarterly Review*, Summer 1980, 2–4; Thomas J. Sargent and Neil Wallace, "Some Unpleasant Monetarist Arithmetic," *Federal Reserve Bank of Minneapolis Quarterly Review*, Fall 1981, 1–17; Bennett T. McCallum, "Are Bond-financed Deficits Inflationary? A Ricardian Analysis," *Journal of Political Economy* 92 (February 1984): 123–35; and John Bryant and Neil Wallace, "A Price Discrimination Analysis of Monetary Policy," *Review of Economic Studies* 51 (April 1984): 279–88.

3. In using the term "Treasury" in this broad sense, it is recognized that as part of the executive branch of government, the Treasury is the fiscal agent that carries out the funding decisions of Congress.

each $1 of interest is financed by τ¢ in taxes and (δ¢ = $1 − τ¢) in bond issuance.

The Federal Reserve is responsible for making asset exchanges, commonly referred to as open market operations. Specifically, the Federal Reserve increases or decreases the monetary base through buying or selling outstanding Treasury securities. Given that the Treasury views the financing of interest on a per-dollar basis, open market operations by the Federal Reserve do not alter the tax promise associated with each unit of Treasury debt. This assumption allows the Treasury and the Federal Reserve to be treated as institutionally separate decision-making units.

Of primary importance is an examination of the Treasury's budget equation and the consolidated budget equation of the Treasury and the Federal Reserve. The Treasury periodically

1. Receives taxes from the private sector
2. Distributes government outlays to the private sector
3. Pays interest on outstanding Treasury debt
4. Increases or decreases the outstanding stock of Treasury debt.

The Federal Reserve periodically increases or decreases the stock of privately held Treasury securities through open market exchanges for base money. The Treasury's budget and the consolidated budget of the Treasury and the Federal Reserve may be illustrated as follows:

Budget of the Treasury

Expenditures	Receipts
Interest on outstanding Treasury debt	Tax receipts
	Issuance of new debt
All other government spending	

Consolidated Budget of the Treasury and the Federal Reserve

Expenditures	Receipts
Interest on privately held Treasury debt	Net tax receipts[1]
	Issuance of new debt to the private sector
All other government spending	
	Creation of base money

1. All interest earnings on Treasury securities held by the Federal Reserve are assumed to be returned to the Treasury and added to the Treasury's net tax receipts.

For simplicity, government spending will be viewed as a distribution to the private sector (a transfer), each $1 of which is valued the same as $1 of private spending.[4] This allows the analysis to focus on net taxes (taxes net of transfers) and avoids the question of whether individuals value $1 of government spending the same as they value the private spending displaced by $1 in taxes.

Adopting the notation

B = the number of privately held Treasury consols = periodical interest receipts of the private sector because each security pays $1 of interest per period
G = nominal noninterest budget outlays
T = nominal Treasury tax receipts
M = the monetary base
P_B = the price of Treasury securities
Δx = the change in x; x = M, B,

the consolidated budget may be written in equation form as

$$B + G = T + \Delta M + P_B \Delta B. \tag{1}$$

As the consolidated budget equation illustrates, from the standpoint of the private individual there is no natural reason to suppose that the interest payments on government bonds involve a corresponding tax liability. Some portion of the interest payments may be financed by net tax receipts, and the remaining portion deficit-financed.

Specifically, suppose that the Treasury follows a rule of tax-financing τ¢ and deficit-financing (δ¢ = $1 − τ¢) of every $1 of interest. That is,

$$\tau B = T - G, \tag{2}$$

which may be written in alternative form, using the consolidated budget equation, as

$$\delta B = \Delta M + P_B \Delta B = B + G - T. \tag{3}$$

Note that (τ = 1) describes an interest-financing rule whereby all interest expenditures are financed out

4. Without this assumption the focus of the analysis must be extended to include the question of whether the public sector creates net wealth by taxing the private sector and providing public goods. Public spending is here treated as neutral. Even if this assumption is violated (say, for example, that $1 of government spending is valued less than $1 of private spending), it does not follow that government bonds per se are not net financial wealth but, rather, that public spending creates a net loss effect.

of tax receipts. Such a result would obviously be obtained by a policy of balancing the budget. On the other hand, ($\tau = 0$) describes a rule of deficit-financing all interest expenditures. For ($\tau < 1$) the Treasury is not merely transferring interest to one group of private individuals through taxing another group of private individuals but is paying a part of the interest by issuing new debt.

Empirical assumptions

In empirically investigating the financing rule for public debt, one must consider not only the financing rule for interest paid but also the financing rule for repayment of principal. Over the period of this study, however, the par value of private holdings of gross federal debt increased from $212.7 billion (December 1950) to $723.6 billion (December 1980).[5] From December 1950 to December 1960, the par value of private holdings of gross federal debt increased $23.0 billion. In the 1960s this increase was $27.6 billion, and over a comparable period from December 1970 to December 1980, the increase was $460.3 billion. Over no five-year period was the par value of private holdings of gross federal debt reduced. It is, therefore, appropriate to treat the principal as rolled over and to consider only the rule for financing interest.

The second important assumption to be supported is that individuals value $1 of government spending the same as they value the private spending displaced by $1 in taxes. The major support for this assumption is that since 1950, federal outlays devoted specifically to noninterest transfer-type spending have ranged from 37 percent to 63 percent of noninterest federal outlays.

Some decision must also be made regarding the relevant length of the fiscal period. Recall that τ represents the proportion of interest that is tax-financed and δ represents the proportion that is deficit-financed. Hence, mathematically, δ is the deficit relative to interest outlays. If the federal budget is consistently in deficit, δ is greater than zero and a portion of outstanding government debt implicitly bears a claim to future currency or debt issuance.[6] Obviously, the Congress of the United States has not, on average, balanced the budget: the current magnitude of the monetary base plus the market value of privately held gross federal debt is over $1 trillion.[7] Clearly, two interesting and important variables are the magnitude of accumulated interest payments on federal debt and the accumulated federal deficit since the drafting of the first federal budget.

Unfortunately, accurate statistics on these variables are not available. However, reliable statistics on the federal deficit and interest outlays since 1948 are available from the U.S. Office of Management and Budget, and it is the 1950–81 period with which this study is concerned.[8] For that period, accumulated federal deficits total approximately $507 billion and accumulated federal interest outlays total approximately $633 billion.

By implication, the value for δ over the 32-year period is .80. The difficulty with accepting this value for δ is the underlying presumption that the fiscal period for which the federal budget is drawn up is 32 years; actually, Congress prepares the budget on an annual basis. In the study here the fiscal period is assumed to be one year.[9] Table 1 provides annual observations on the variable δ over the 1950–81 period.

Historical distribution of the deficit-interest outlay ratio

The first step in determining the financing rule is to

5. Data used in this calculation are taken from the *Federal Government Finances* publication of the U.S. Office of Management and Budget.

6. Although the discussion of δ here is for the situation where ($0 < \delta < 1$), there is no fundamental reason why δ must be limited to these bounds. Clearly, ($\delta < 0$) describes a case where the government, on average, runs a surplus, and ($\delta > 1$) describes a case where the deficit, on average, exceeds interest payments—that is, the budget exclusive of interest is in deficit. Because neither case seems particularly relevant in view of the experience of the United States over the period of this study, the analysis is limited to the situation where, on average, the government runs a deficit but the deficit is less than interest payments.

7. It is recognized that some of the base has been issued by acquiring gold and other foreign reserves.

8. Although accurate data are available on these variables since 1948, a decision was made to begin with the decade of the 1950s, during which the Treasury-Federal Reserve accord of 1951 was reached.

9. Signs of significant autocorrelation in the annual observations of δ might be interpreted as evidence that the actual budget period is longer than one year. This question is considered in the analysis that follows.

Table 1
ANNUAL INTEREST OUTLAYS AND THE FEDERAL DEFICIT

Fiscal year	Interest paid by Treasury	Federal budget deficit (+)	δ[1]
	Millions of dollars		
1950	$5,692	$3,112	.546
1951	5,557	-6,100	-1.097
1952	5,688	1,517	.266
1953	6,250	6,533	1.045
1954	6,014	1,170	.194
1955	6,032	3,041	.504
1956	6,294	-4,087	-.649
1957	6,681	-3,249	-.486
1958	6,946	2,939	.423
1959	7,073	12,855	1.817
1960	8,300	-269	-.032
1961	8,117	3,406	.419
1962	8,321	7,137	.857
1963	9,216	4,751	.515
1964	9,810	5,922	.603
1965	10,359	1,596	.154
1966	11,286	3,796	.336
1967	12,533	8,702	.694
1968	13,751	25,161	1.829
1969	15,793	-3,236	-.204
1970	18,309	2,845	.155
1971	19,602	23,033	1.175
1972	20,563	23,373	1.136
1973	22,782	14,849	.651
1974	28,032	4,688	.167
1975	30,911	45,154	1.460
1976[2]	34,511	66,413	1.924
1977	45,225	57,904	1.280
1978	43,966	48,807	1.110
1979	52,556	27,694	.526
1980	64,504	59,563	.923
1981	82,590	57,932	.701

1. The deficit divided by interest outlays.
2. Data for the third quarter of 1976 (the quarter of the transition from the June fiscal year to the September fiscal year) are included in 1977.

SOURCE OF PRIMARY DATA:
U.S. Office of Management and Budget.

examine the distribution of the ratio of the deficit to interest outlays, δ, given in Table 1. Knowledge of this distribution is important for summarizing key aspects of the Treasury's interest-financing behavior and for testing hypotheses regarding that behavior. Since many of those tests are strictly valid only for normally distributed random variables, the approach is to determine whether the observed distribution of δ resembles that which would be expected in a normal distribution.

From Table 1 the average value of δ over the entire 32-year sample is .59 and the standard deviation of δ is .69. Further characteristics regarding the distribution of the variable are obtained by several nonparametric procedures.

The first procedure is to compare the interquartile range of the observed distribution of δ to that which would be expected in a normal distribution. In a normal distribution the interquartile range is 1.35 times the standard deviation. From the historical distribution of δ given in Table 1, the first and third quartile boundaries may be calculated as Q_1 equals .18 and Q_3 equals 1.08; hence the interquartile range is .90. Dividing .90 by 1.35 gives the pseudo standard deviation of .67. Thus, the calculated historical standard deviation of .69 closely approximates the value .67 that would be expected from a normal distribution for δ.

Further evidence concerning the distribution of δ may be obtained from a stem and leaf plot. Examination of the plot indicates that the distribution of δ exhibits tails approximating those of a normal distribution.

Stem and Leaf Plot for δ	
-1.10 to -.58	| |
-.57 to -.08	| |
-.07 to .42	| | | | | | | |
.43 to .92	| | | | | | | | |
.93 to 1.42	| | | | |
1.43 to 1.92	| | | |

Another procedure used is to identify outlying values of δ and determine whether any would be considered extreme in a normal distribution. Multiplying the interquartile range by 1.5, subtracting this value from Q_1, and adding it to Q_3 give a lower bound of -1.17 and upper bound of 2.43 for outliers that would be considered mild in a normal

distribution. Similarly, the lower bound of −2.52 and upper bound of 3.83 may be established for outliers that would be considered extreme in a normal distribution. Examination of the historical distribution of δ from Table 1 reveals there are no outliers that would be considered extreme or even mild in a normal distribution.

Finally, the Kolmogorov-Smirnov goodness-of-fit test may be applied to examine the hypothesis that δ is normally distributed. Calculation of the Kolmogorov statistic at all points along the cumulative distribution of δ and comparison with its critical boundaries indicate that the hypothesis that δ is normally distributed cannot be rejected.

The individual's expected tax liability

The central hypothesis is that over the 1950–81 period, each $1 of interest received by the private sector bore an expected tax liability of 41 cents. A broad alternative hypothesis, which is considered in this section, is that the individual's expected tax liability varied in some way over the period.

The first step in investigating this hypothesis is to determine whether the sample observations on δ are independently distributed. Evidence regarding the time independence of the variable is important in deciding whether the data exhibit any pattern that may be used in forecasting future tax liabilities on government debt. Examination of the partial autocorrelation function for δ reveals no significant partial autocorrelation at any lag. The modified Box-Pierce Q statistics at lags of 6, 12, and 24 years are 7.51, 16.51, and 32.93, respectively, indicating insignificant accumulated autocorrelation at short, medium, or long lags for δ. Estimation of δ as an autoregressive moving average process related to immediately prior-period errors (MA1) reveals insignificant coefficients for the MA1 variable. In addition, the diagnostic chi-square statistics for the residual series from this regression indicate insignificant accumulated autocorrelation at lags of 6, 12, 18, and 24 years. Examination of the lag autocorrelation pattern of δ from standard Box-Jenkins time series techniques indicates only marginally significant autocorrelation at lags of 9 and 21 years. The implication of this evidence is that if δ is stationary with a mean of .59, it is best characterized as an independently, identically distributed normal random variable.

The hypothesis of nonstationarity will be examined in two forms. The first is that δ is a time-dependent variable. The second is that a mean shift occurred in the distribution of δ sometime during the 1950–81 period. Evidence regarding the stationarity of δ may be provided by a standard ordinary least squares regression of δ against time. The results of this regression are

$$(4) \qquad \delta = \underset{(.17)}{.38} + \underset{(2.79)}{.03}\ TIME,$$

$$R^2 = .21;\ DW = 2.06;\ F = 7.79.$$

where the numbers in parentheses are t statistics and the variable *TIME* takes the values 1 through 32. Note that the t statistic of 2.79 for *TIME* exceeds the critical value of 2.04 required for significance at the 95-percent level. Hence there is evidence that δ might not appropriately be regarded as a stationary random variable.

Evidence for the hypothesis that a shift occurred in the distribution of δ may be provided by maximum likelihood tests. Breaking the sample of 32 observations into two arbitrary samples of sizes n and $(32 - n)$ and calculating the standard deviation of δ over the first sample (s_1), over the remaining sample (s_2), and over the entire data set (s) allow for a test of the hypothesis that a shift occurred in the distribution of δ. The procedure used is to calculate the Chow statistic C_n for each subperiod of two years or longer ($n = 2, 3, \ldots, 30$).[10] Maximizing C_n by choice of n gives a value of ($C_{21} = 8.78$) for ($n = 21$). A critical value of 3.84 is required for significance at the 95-percent level. Hence there is substantial evidence that a shift occurred in the distribution of δ in 1971.[11]

A standard dummy variable regression may be estimated to gain further evidence on the hypothesis that a shift occurred in the distribution of δ in 1971. The results of this regression are

$$(5) \qquad \delta = \underset{(2.72)}{.38} + \underset{(2.69)}{.62}\ DUMMY,$$

$$R^2 = .19;\ DW = 1.95;\ F = 7.12.$$

10. Specifically, the Chow statistic used here is

$$C_n = -2 \cdot \ln\left(\frac{s_1^n \cdot s_n^{32-n}}{s^{32}}\right) \sim \chi^2.$$

11. Why this happened in 1971 is, in itself, an interesting question, one that is beyond the scope of this paper.

where *DUMMY* equals zero from 1950 to 1970 and 1 thereafter. Both the estimated constant term and the estimated coefficient on the dummy variable are significant at the 95-percent level. The estimated mean value for δ is .38 over the 1950–70 period but is 1.00 over the 1971–81 period. The implication is that interest receipts that accrued during the 1950–70 period bore a corresponding average tax liability of 62 cents whereas there was, on average, no tax liability over the later period.

To gain further evidence regarding the stationarity of δ, the entire 1950–81 sample is separated into the suggested subsamples 1950–70 and 1971–81, and the time dependence regressions are reestimated over each subsample. The results of these regressions are as follows:

1950–1970

(6) $$\delta = \underset{(.25)}{.08} + \underset{(1.07)}{.03}\ TIME.$$

$$R^2 = .06;\ DW = 2.19;\ F = 1.16.$$

1971–1981

(7) $$\delta = \underset{(1.12)}{1.47} - \underset{(-.04)}{.02}\ TIME.$$

$$R^2 = .01;\ DW = 1.45;\ F = .13.$$

The primary result of separating the data into these two subsamples is that the explanatory power of the variable *TIME* is substantially reduced. In each subperiod the *t* statistic for the estimated coefficient on *TIME* falls far below its critical value for significance. In addition, the estimated coefficient on *TIME* changes sign from 1950–70 to 1971–81. The R^2's, measuring the fraction of the variation in δ that is explained, fall substantially and the Durbin-Watson statistic worsens for each subperiod when compared with the time dependence regression estimated for the entire 1950–81 period. These results suggest adoption of the hypothesis that a shift occurred in the distribution of δ in 1971 over the hypothesis that δ is time-dependent.

Since there is substantial evidence to support the hypothesis that a shift occurred in the distribution of δ in 1971, the time series properties of δ must be investigated under the presumption that such a shift occurred. The observations on δ must first be made stationary by lowering the distribution of δ for the post-1971 sample to the estimated mean for the 1950–70 sample (from 1.00 to .38). This series is referred to as the adjusted series for δ (δ_A).

Calculation of the Kolmogorov-Smirnov statistic for the adjusted series indicates that the distribution of δ_A does not exhibit significant departures from normality. Examination of the lag autocorrelation pattern and partial autocorrelation pattern of δ_A reveals marginally significant autocorrelation only at a lag of 9 years. The modified Box-Pierce statistics indicate no significant accumulated autocorrelation at lags of 6, 12, or 24 years. Estimation of δ_A as an MA1 process yields a highly insignificant value of $-.356$ for the estimated MA1 coefficient. In summary, the evidence supports the hypothesis that δ is an independently, identically distributed normal random variable with a mean value of .38 from 1950 to 1970 and a mean value of 1.00 thereafter.

Conclusion

Effort directed toward identifying the interest-financing rule in the United States indicates that on average over the 1950–81 period, each $1 of interest paid by the Treasury bore a corresponding tax liability of 41 cents. There is also substantial evidence that a change in the interest-financing rule occurred in this period. Specifically, for each $1 received in the form of interest on government bonds before January 1971, there was a corresponding expected tax liability of 62 cents—hence a net receipt of interest of 38 cents by the private sector. Interest receipts accruing after that time, however, have borne no expected tax liability, and there is a dollar-for-dollar expected net receipt of interest by the private sector.

A familiar proposition in macroeconomics and monetary theory is that the price level is directly related to net financial wealth of the private sector. Treasury securities have traditionally not been viewed as net financial wealth because their stream of interest payments was presumably matched by an equivalent stream of tax liabilities. This paper considers the question of whether in fact the interest payments on government debt have involved an equivalent tax liability and finds that they have not. Evidently, Treasury securities are net wealth, at least in part. It follows that the ability of the Federal Reserve to offset the inflationary effects of deficits through open market exchanges of base money for Treasury securities is potentially limited.

The *Economic Review* is published by the Federal Reserve Bank of Dallas and will be issued six times in 1984 (January, March, May, July, September, and November). This publication is sent to the mailing list without charge, and additional copies of most issues are available from the Publications Department, Federal Reserve Bank of Dallas, Station K, Dallas, Texas 75222. Articles may be reprinted on the condition that the source is credited.