

CHRISTOPHER A. SIMS

*University of Minnesota*

## *Policy Analysis with Econometric Models*

RECENTLY the rational expectations school has mounted an attack on the conventional use of simultaneous equations models for policy analysis. One might go further and say that among academic macroeconomists the conventional methods have not just been attacked, they have been discredited. The practice of using econometric models to project the likely effects of different policy choices, then choosing the best from among the projected outcomes, is widely believed to be unjustifiable or even the primary source of recent problems of combined high inflation and low economic activity. Instead, it is claimed, policy analysis should be formulated as choice among rules of behavior for the policy authorities and estimates should be made of the stochastic properties of the economy under each proposed rule to choose the best.

This point of view has gained such wide acceptance in part because of its association with Lucas's theoretical demonstration that a Phillips curve could emerge in an economy in which such an association between inflation and real activity was not a usable menu for policy choice. Because users of conventional simultaneous equations models sometimes presented the Phillips curve as just such a menu, and because it became apparent in the 1970s that this menu was not helpful, an analysis that provided a cogent explanation for why the menu was chimerical had great appeal.

As in most revolutions, the old regime toppled by the rational expectations revolution was corrupt and in some sense deserved its fate. However, as is often the case, the revolution itself has had its excesses, destroying or discarding much that was valuable in the name of utopian

ideology. This paper tries to assess where the revolution itself could use revision.<sup>1</sup>

In this paper I argue that it is a mistake to think that decisions about policy can only be described, or even often be described, as choice among permanent rules of behavior for the policy authorities. A policy action is better portrayed as implementation of a fixed or slowly changing rule. I also argue that explicit identification of expectation-formation mechanisms is not necessary for policy analysis, concluding that the rational expectations critique of econometric policy analysis is a cautionary footnote to such analysis rather than a deep objection to its foundations. From this perspective, the conventional use of econometric models to aid in policy choice is neither self-contradictory nor meaningless.

### **Applying Decision Theory to Economic Policymaking**

Formal quantitative analysis of policy choice must begin with a model of the effects of policy. The model must describe the “outcome,” usually in the form of a probability distribution over future events in the economy, for each possible “setting of policy.” To choose policy optimally, one evaluates the outcomes according to some objective function and chooses the best. Although this description appears static, in the sense that it refers to a single decision rather than to a time sequence of them, it can also be applied to dynamic problems. In the conventional description of a dynamic problem there is at each date,  $t$ , a measure,  $y_t$ , of the current state of the economy (which may include information on the history of the economy and of policy), and a value must be chosen for the policy action at  $t$ ,  $c_t$ . The model gives a conditional distribution for  $y_{t+1}$ , the state of the economy in the next period, as a function of  $y_t$  and  $c_t$ . If the objective function included only  $y_{t+1}$ , one would simply choose  $c_t$  so as

1. The senses in which the “old regime” was in my view “corrupt” are discussed in more detail in my “Macroeconomics and Reality,” *Econometrica*, vol. 48 (January 1980), pp. 1–48. There I argue that the identifying restrictions used to obtain equation-by-equation interpretations of these models are incredible. They are mainly simplifications, chosen empirically so that they do not conflict with the data. Such simplifications may be a useful part of estimating a forecasting model, but they do not represent a priori knowledge and therefore cannot help in identifying the model.

to obtain the best probability distribution of  $y_{t+1}$ . In general, however, the objective function will depend on  $y_{t+s}$  (and perhaps also  $c_{t+s}$ ) at many or all dates beyond  $t + 1$ . Then, in order to set  $c_t$  optimally in the current period, one must determine a set of optimal contingency plans for all future dates, describing how each  $c_{t+s}$  will be chosen as a function of information available at  $t + s$ . Defining the outcome as the conditional distribution of all future  $y_{t+s}$  and  $c_{t+s}$  given  $y_t$ ,  $c_t$  and the optimal contingency plans, the dynamic problem becomes analogous to the static problem. A policy setting is a complete set of contingency plans for the future course of policy, and one chooses among such settings to achieve the best outcome.

However, this abstract description of the problem of policy choice appears at first glance not to match the problems policymakers actually face. Decision theory portrays optimal policy choice in a dynamic setting as a single analytical exercise. One has to produce contingency plans specifying policy actions at all future dates under all conceivable circumstances in order to find the best current policy action. Once this has been done, there should be no need for new analysis at future dates. Yet in practice macroeconomic policymaking does not seem to be this sort of once-for-all analysis and decision. Policymakers ordinarily consider what actions to take in the next few quarters or years, reconsider their plans every few months, and repeatedly use econometric models to project the likely effects of alternative actions. Furthermore, optimal policy should be a deterministic function of information available to the policymaker, but actual policy seems to include a large component that is unpredictable even to observers with the same information set as the policymaker.

On closer examination, these difficulties with decision theory as a description of actual policy choice are probably not important. It is likely that policymakers' objective functions discount the future, so that actions and consequences far enough ahead receive little weight. Precise contingency plans for dates far in the future will therefore be relatively unimportant in solving the current problem. A good approximate solution will thus concentrate on determination of the current policy action,  $c_t$ , and the contingency plans for policy actions in the near future, but will tolerate cruder approximations for contingency plans in the increasingly distant future. This is more likely because the state of the economy in the distant future is relatively unpredictable, so that the range of

contingencies that would have to be considered in formulating precise contingency plans for that time period is great. The cost of precise analysis of such wide ranges of contingencies is high, while the value of such accurate analysis of the distant future in improving the choice of  $c_t$  is relatively low. It is perfectly reasonable, therefore, for even a nearly optimal procedure for choosing policy to involve reconsidering policy plans at regular intervals.<sup>2</sup>

Policy is not made by a single maximizing policymaker, but through the political interaction of a number of institutions and individuals. The people involved in the process, the nature of the institutions, and the views and values of the public shift over time in imperfectly predictable ways. Thus even if each individual who has some power over the choice of policy has a coherent view of his or her objectives and of a dynamic model of the economy, leading that policymaker to believe in the appropriateness of a particular value for  $c_t$  as a function of the state, the actual setting of  $c_t$  will be partly random from the point of view of the public. The observed unpredictability of policy need not, therefore, imply that any individual policymaker is randomizing his or her actions.

### **Rational Expectations**

Control theory's dynamic version of the problem of optimal decision theory rests on the following apparently reasonable assumptions:

At each date  $t$  there is a list of numbers, the vector describing the state of the economy,  $y_t$ , which summarizes all aspects of the history of the system being controlled that are relevant to its future evolution.

At each  $t$  there is a function,  $f_t$ , determining the probability distribution of the next period's state  $y_{t+1}$ , from the current state,  $y_t$ , and the current policy action,  $c_t$ , where  $f_t$  is taken as given when policy choice is optimized.

Policy actions,  $c_t$  (the vector of values for the control variables at  $t$  in

2. In textbook dynamic control problems, the form of the function relating optimal  $c_t$  to the current state often tends to be a fixed form over time. If economic policy choice were like these textbook problems, the need for reanalysis would steadily diminish through time as one refined the knowledge of the function describing the optimal choice of  $c_t$ . However, in economics the form of the model relating the state and policy to the next period's state inevitably drifts in time so that there is no reliable tendency for  $c_t$  to become a fixed function of the state of the economy.

control theory), can depend on information available at  $t$ , but not on information available only later than  $t$ .

In engineering applications the  $f_t$  functions and the state of the economy and control vectors,  $y_t$  and  $c_t$ , are determined by the physical causal structure of the system being controlled. In economics it is not so clear what vector of variables summarizes the state of the economy or what vector of variables summarizes the effects on the economy of all policy actions. People trying to make a direct link between econometrics and control theory have assumed that the  $f_t$  functions are the equations of an econometric model and that the two vectors,  $c_t$  and  $y_t$ , are obtained by dividing the variables in the econometric model into two categories, policy variables (those in  $c_t$ ) and others.

The rational expectations critique challenges this way of bringing control theory to bear on econometric policy analysis. It points out that econometric models describe the behavior of people, not an inanimate system. People's current economic choices depend not only on the actual values of variables entering an econometric model, but on their expectations about future values of those variables. For a given setting of  $c_t$  and  $y_t$ , people may behave differently according to what their views are of how  $c_s$  will be chosen in the future. The critique argues, therefore, that it is a mistake to treat econometric models and the variables they contain as playing the roles of  $f$ ,  $c$ , and  $y$  in control theory. There is no well-defined function determining the distribution of  $y_{t+1}$  from  $y_t$  and  $c_t$  alone if  $y_t$  and  $c_t$  are simply a list of variables drawn from an econometric model. The views of the public about the future will always be an omitted argument when one attempts to construct such a function. In particular, the econometric model itself has a form that changes as people's expectations of the future change.

Control theory teaches us, though, that in choosing  $c_t$  a policymaker will in general have to form contingency plans for choices of  $c_{t+s}$  for all positive  $s$  also. To the extent that the public finds out about these contingency plans for future policy, its behavior is likely to be affected by them. This leads to the conclusion by rational expectations analysts that econometric models of the type in wide current use, which make no provision for examining the effects of the public's views about plans for future policy choices, are useless for policy analysis.

Having reached this destructive conclusion, what way out does rational expectations provide? One possibility is the use of the rational expectations assumption in policy analysis. The rational expectations

assumption (which is stronger than what is needed to justify the rational expectations critique) is that the public behaves optimally, given its own objectives and the information available to it, and further that the public understands precisely what contingency plans have been chosen for future policy. Under this assumption there is no sharp distinction between actions taken now and contingency plans for future action. Both are chosen now and both have effects now. Policy analysis under this assumption is not a special case of control theory; it becomes a special case of game theory. As in game theory, what is being chosen is a complete set of contingency plans, a strategy. Also as in game theory, the other player (the public in this case) is assumed to understand one's own strategy and to play optimally.

The rational expectations critique, by using examples generated with the rational expectations assumption, shows what serious errors can be made in econometric policy analysis if the response of expectation-formation mechanisms to policy is ignored. As a warning about potential pitfalls, the critique is certainly correct in principle. There can be little doubt, for example, that the consequences of anticipated and unanticipated inflation are different, and that policy analysis using a dynamic model which ignores this fact can be seriously misleading. But the resolution of this difficulty by simply accepting the rational expectations assumption is of mainly academic interest for a number of reasons.

In the theory of zero-sum games the assumption that other players know one's own strategy and use this knowledge optimally leads to a solution with appealing properties, even if we expect that most other players will either not understand our strategy or use their knowledge suboptimally. The result of a strategy chosen on this assumption will be that other players will lose at least as much, under incomplete knowledge or suboptimal play, as they would if they played optimally. If their losses are our gain, such a strategy is then appealingly conservative; its results can be no worse than what they are calculated to be. But the losses of the public are not the gains of the economic policymaker—in fact almost precisely the opposite. In the problem of economic policy choice, the assumption that the public knows our strategy and reacts optimally cannot be justified as in the theory of zero-sum games, that is, as a conservative assumption; if it is incorrect, policy chosen on this assumption could produce results much worse than expected. The only way to justify the assumption is to claim that it is a close approximation to reality.

Macroeconomic policymakers cannot in fact guarantee that the public understands the policymakers' choices of contingency plans for future policy or that the public acts optimally on such understanding as it has. Besides the usual objection that the public may not have the computational capacity to absorb and quickly act on new information about future policy, there is the problem that, historically, announcements about the future course of policy have often later turned out to be misleading. Even if it were to announce its plans publicly, therefore, the policy authority could not realistically assume that the public takes the announcement seriously.

Supporters of the rational expectations assumption have not ignored these difficulties with it. Kydland and Prescott have observed that, under rational expectations, optimal policy choice is time-inconsistent.<sup>3</sup> That is, if a set of contingency plans for current and future  $c$  is chosen optimally at  $t$ , the contingency plan chosen for  $c_{t+s}$  will generally not yield the same value for  $c_{t+s}$  as the one that would be chosen if the problem were solved afresh at  $t + s$ . This occurs because under the rational expectations assumption one may gain benefits at  $t$  from the effects of an announced policy plan for period  $t + s$ . It may be best to forgo some benefit at  $t + s$  for the benefits at  $t$ . Yet if the problem is solved anew at  $t + s$ , those benefits at  $t$  generated by people's expectations at  $t$  of how policy would be chosen at  $t + s$  have already been formed and will not be lost if those expectations turn out to be false. There is no trade-off of benefits at  $t$  for benefits at  $t + s$  any longer, so the optimal choice will ignore the now-obsolete trade-off.

Whenever the effects of current policy choices depend on expectations of how future policy will be chosen (and whether or not expectations are rational) there is likely to be an incentive for policymakers to renege on previous commitments about the nature of future policy plans. This may explain why, historically, announcements of future policy plans have frequently not proved correct and why policymakers encounter public skepticism over their announced intentions for future policy.

Advocates of the rational expectations approach have recognized the strength of the objections that the public is not likely to absorb and act quickly on new information about plans for future policy and that, due to the inconsistency of optimal policy under rational expectations, the

3. Finn E. Kydland and Edward C. Prescott, "Rules Rather than Discretion: The Inconsistency of Optimal Plans, *Journal of Political Economy*, vol. 85 (June 1977), pp. 473-91.

public is not likely even to believe announced plans for future policy. Their response has usually been to suggest that policy ought to follow fixed “rules.” By a rule is meant a function,  $h$ , determining  $c_t$  as a function of  $y_t$ . That is, rather than finding the optimum  $c_t$  at  $t$ , a policy authority following a fixed rule finds the best  $h$  on the assumption that it will set  $c_t = h(y_t)$  now and at all dates in the future. The truly optimum choice of  $c_t$  on the rational expectations assumption would involve contingency plans in which a different function,  $h_s$ , was used to determine  $c_s$  from  $y_s$  at each  $s$  in the future.

The rational expectations assumption is less unrealistic when applied to analysis of the long-run effects of fixed rules. If the rule is implemented and adhered to for some time, people will eventually come to believe that there is a high probability that future policy will be set using the rule. The difficulty of getting the public to believe in announced plans for future policy is thus eliminated. Also, in these circumstances the public should eventually learn to adapt its behavioral patterns so that they are not far from optimum given the announced rule—even if only by trial and error. This removes the objection that the public may not have the computational capacity to adapt optimally to announced plans for future policy.

But these arguments only suggest limitations, stringent ones at that, to the range of applicability of rational expectations analysis of policy. They should not be interpreted, as they sometimes have been, as prescriptions that policy should always take the form of fixed rules and that one should be restricted to analysis of the effects of permanent shifts in fixed rules. As I will argue further below, most analysis of macroeconomic policy is not properly treated as a problem of choosing an optimal fixed rule. This means that the rational expectations assumption is a treacherous tool in analyzing most problems of macroeconomic policy choice. As shown below, fortunately, the rational expectations assumption is not the only logically consistent way to proceed with econometric policy analysis in the face of the rational expectations critique.

### **Valid Reduced Forms for Policy Analysis**

There is a distinction between actions policymakers can take now, at  $t$ , and their plans for future actions. When this distinction is made properly, there is no doubt that a model capable of making accurate



conditional projections of the effects of various potential policy actions is possible. Indeed, no reasonable decision about policy action can be made without such a model.

The rational expectations critique argues that what happens to the economy following a policy action depends on the public's expectations of the future as well as on the current action. But policy can affect the public's expectations only through policy actions. While an incorrect model is certainly capable of ignoring these effects, there is no reason why a model necessarily fails to include indirect effects of policy actions through their influence on expectations. When a model correctly includes these effects, it will make valid conditional projections of the effects of policy and will be useful in guiding policy choice. Control theory could even properly be used in conjunction with such a model to guide policy choice.

One kind of policy action is an announcement of plans for future policy; one part of the state of the economy might consist of previous announcements about plans for policy. Rational expectations analysis may help us understand the effect of announcements as policy actions. Kydland and Prescott might be regarded as having explained, using rational expectations, why announcement effects are ordinarily of minor importance. Certainly policy actions other than announcements are at least as important as announcements themselves in influencing the public's beliefs about future policy.

The fact that some effects of a policy action occur through effects on expectations does not necessarily imply that one must explicitly identify the parameters of expectation-formation mechanisms to obtain models that correctly project the effects of the action. An analogy may be helpful. Suppose one wanted to project the effects on the peanut market of an excise tax on peanuts. The effects will depend on parameters both of the supply curve and the demand curve. But this is only because the reduced form of the model, which relates the excise tax to its consequences for peanut prices and quantities, depends on the parameters of supply and demand. All one needs to know to discover the effects of the excise tax are the parameters of this reduced form. These may be discovered without explicitly determining the parameters of demand and supply, and their usefulness for projecting the effects of policy does not depend on how they are discovered. Correspondingly, effects of policy actions that affect expectation-formation mechanisms can be correctly evaluated with models that are reduced form in the sense that

they do not explicitly display the parameters of the expectation-formation mechanism.

Kydland and Prescott display an example which purports to show that attempts to use reduced-form models to make conditional projections and then to choose policy actions on the basis of those projections is likely to lead to bad policy choices—precisely the opposite of what is claimed here.<sup>4</sup> In their example they assume that the economy is characterized by a simple rational expectations Phillips curve,

$$(1) \quad U_t = a - b[P_t - E_{t-1}(P_t)],$$

where  $U_t$  is the unemployment rate and  $P_t$  is the price level. They assume further that the policy authority has an objective function that puts negative weight on the level of unemployment and on the absolute value of the rate of inflation and that  $P_t$  is directly controlled by policy. If the economy has long been in a situation such that  $E_{t-1}(P_t) = P_{t-1}$ , that is, price follows a random walk and expected inflation is zero, econometricians might correctly discover that

$$(2) \quad U_t = a - b(P_t - P_{t-1}).$$

If econometricians took  $b$  in equation 2 to be structural, in the sense of invariant under policy intervention, they might be misled into suggesting that the mean of  $U_t$  could be lowered by increasing the mean rate of inflation.<sup>5</sup> But obviously increasing the mean price change from zero to, say,  $Q$  would change the public's price expectations. Assuming the public adapts its expectations to the new policy, eventually  $E_{t-1}(P_t) = P_{t-1} + Q$ , which from 1 we can see would change the constant term in 2 from  $a$  to  $a + bQ$ . The result would be a higher mean level of inflation, but no change in the mean level of unemployment. If econometricians mistakenly supposed that there had been an exogenous shift in  $a$  so that the new version of 2 with  $a$  increased by  $bQ$  was taken to be structural, they could repeat their initial mistaken analysis and advice. The contin-

4. Ibid.

5. This use of "structural" to mean invariant under intervention follows Leonid Hurwicz, "On the Structural Form of Interdependent Systems," in Ernest Nagel and others, eds., *Logic, Methodology and Philosophy of Science* (Stanford University Press, 1962), pp. 232–39; and Tjalling C. Koopmans and Augustus F. Bausch, "Selected Topics in Economics Involving Mathematical Reasoning," *SIAM Review*, vol. 1 (July 1959), pp. 138–48.

ued revision of 2 and change in the mean rate of inflation that would come with repeated attempts to affect unemployment by manipulating inflation could eventually result in a stable equilibrium in which the ultimate version of 2 implied a trade-off between increased inflation and decreased unemployment such that no further change in the inflation rate appeared desirable.

What this example shows is only that a false reduced form can lead to bad policy. If there are no announcement effects, or if the public does not pay attention to such announcements as they are made, the public's expectations about future prices can be affected only by history.

The public's expectations depend on the mechanism they use to learn about policy formation, but in the absence of direct observation of the psychology and politics of policymaking, their price expectations will be based only on the history of prices and unemployment. This is because prices and unemployment are the only variables of interest to the policymakers and because setting the price level is the only action available to them.

There are no other influences on the economy to which policymakers might be responding. But expectations will depend in a complicated way on the whole history of prices and unemployment up to the current date. The correct reduced form, therefore, must make the current level of unemployment depend on the current price and many lagged values of prices, probably in a nonlinear way. The mistaken course of policy Kydland and Prescott describe depends on econometricians persisting in believing in the incorrect reduced-form 2, even as policy produces a historical sample in which it should become quite clear that 2 is not correct. If instead econometricians experimented with distributed lag versions of 2 during the policy of deliberate inflation, they would soon discover that the correct specification makes the unemployment rate depend on something closer to the difference between the current inflation rate and its recent average level. Even an approximate reduced form of this nature would lead them to suggest a policy not far from the optimal one, and it would not be necessary that they realize that the underlying behavioral model was 1.

Not only are valid reduced forms possible, they are essential. Economists may confine themselves to rational expectations analysis of the effects of permanent changes in policy rule, renouncing any claim to be able to produce estimates of the immediate effects of policy actions.

But before they can recommend that a change in rule be implemented, they must somehow decide whether it is likely that there will be short-run negative consequences to making the change large enough to offset the permanent long-run gains. Such an assessment necessarily is based on a probability distribution for the immediate consequences of the change in rule, that is, on a reduced-form model.

### **Regime Shifts versus Normal Policymaking**

From the above discussion it can be seen that for policy analysis one needs a valid reduced-form model and that one can know the form of such a model without knowing how expectation-formation mechanisms are embedded in it. It is also apparent that a valid reduced form will make relatively precise conditional projections for the effects of policy actions or sequences of actions that are close in form to what has been observed historically. The role of a priori knowledge, or subjective guesswork, in the projections one makes with the valid reduced-form model will be increased as one makes projections conditional on sequences of policy actions more remote from what has been observed historically.

Rational expectations analysis has focused attention on a type of action that can be analyzed only with strong a priori assumptions about the behavior of the private sector—a one-time shift in regime or policy rule. The change in rule will result in a sequence of policy actions that is completely without precedent, so there is no hope of simply looking at the historical data for similar sequences of actions to determine the likely consequences.

But permanent shifts in policy regime are by definition rare events. If they occurred often they would not be permanent. Political rhetoric that sounds as if it is concerned with a permanent change in the policy rule is not rare. This is only natural when groups with different interests and ideals contend for control of policy. Each would change the rule if given exclusive and permanent control of policy. In a system in which this does not happen policy is generated by a steady, though not entirely predictable, swinging of the political pendulum. These swings are reversible, and the public understands them to be so. They are not changes in the policy rule of the type for which rational expectations analysis is suited.

Normally policy actions are generated by a mechanism that, from the point of view of the public, forms a more or less stable stochastic process. Treated as an abstract problem in control theory, this situation may seem to present no interesting analytical difficulties; it may seem to arise only when policymakers are satisfied with a preexisting solution to the control problem. But as I argue above, solutions to dynamic decision problems are not in fact computed once for all. They are regularly recomputed with increased precision in the analysis of the consequences of current and immediate future policy actions. These recurrent reevaluations of policy options are, by construction, frequent. Although any one of them may be of less social importance than a single permanent shift of policy regime, they are nonetheless cumulatively important. Use of valid reduced forms to project the likely consequences of various possible policy actions has been and will continue to be a useful part of this continuing process of normal policymaking.

Even in normal policymaking, the question of evaluating the effects of unprecedented policy actions arises. As one solves the problem of optimal choice of  $c_t$  at  $t$ , one must in principle project the effects of every possible set of contingency plans for policy action and choose the best. Some of those possible contingency plans will be far out of the range of historical experience. But in a world of shifting political fortunes and imperfect analysis by policymakers, the historical record will show a considerable range of policy actions, and the strongest disputes are likely to be between people advocating policy actions of types that have been tried before. The historical record is likely to be especially valuable in projecting the effects of such actions. Radically new types of policy are likely to be, and ought to be, discounted because it is recognized that their effects are uncertain. For a formal decision-theoretic approach to do this discounting properly, the reduced-form model used would have to recognize explicitly the uncertainty about the a priori restrictions embodied in the model's specification. Informally the same effect is obtained when policymakers treat suggestions for radical shifts in policy skeptically, even when the point estimates from econometric models suggest that the results would be good.

There are cases that seem intermediate between the case of a pure permanent shift in regime and the case of normal policymaking.

Suppose, for example, that policy is concerned only with setting the money stock and that there are just two views on how this should be done, view A and view B. There is an unobservable variable,  $s_t$ ,

measuring the political strength of A relative to B, and the actual money stock is determined by a well-defined function with  $s_t$  as an argument. For example,  $s_t$  might range between zero and 1 and in each period the actual  $m_t$  might be  $m_t^A s_t + m_t^B (1 - s_t)$ , where  $m^A$  and  $m^B$  are the values favored under views A and B. If the shift back and forth between views A and B occurred persistently, it would eventually generate historical data covering all values of  $s_t$ . From the point of view of the public, the two views and the stochastic process for  $s_t$  jointly generate actual policy choice according to a single probability mechanism. One can therefore expect, from analysis of a sufficiently long stretch of data, to obtain by direct observation useful information about how the state of the economy responds to policy actions, just as in any situation in which policy has been generated by a single probability mechanism in the historical data. Of course, a fully accurate model would be quite nonlinear. One might hope, though, that it could be well approximated by a linear model with unknown stochastically varying coefficients. Since there is no particular reason to suppose the economy is well characterized by a linear model with fixed coefficients even in the absence of complicated probability mechanisms for policy formation, it is not clear that the problem of estimating the response of the economy to policy is fundamentally more difficult in the presence of such persistent oscillation in the policy mechanism than in its absence.

This argument is not meant to deny that explicit modeling of expectation formation and its dependence on policy could be useful. If in the example  $s_t$  changed slowly with time, there might be long periods in which it was nearly 1 or nearly zero. A rational expectations analysis of data drawn entirely from a period with  $s_t$  near zero might, from knowledge of the policy mechanism favored by view B and use of a priori identifying restrictions, succeed in predicting how the economy would behave in a period with  $s_t$  near 1. Thus explicit modeling of the connection of expectation-formation mechanisms to policy in an accurately identified model would allow better use of the data. The resulting probability model for the effects of policy actions would, with a given amount of data, be sharper in the sense of providing less uncertain predictions. The model estimated without using so much previous knowledge would, however, still be a true probability model, giving an accurate picture of the uncertainty in its own predictions. And its validity would not depend on the validity of the extra previous knowledge.

### **The Case for Empirical Policy Analysis**

Returning to the example of the excise tax on peanuts, how can one estimate the reduced form needed? If there had never before been an excise tax on peanuts, one would have to rely on hunch and intuition to generate identifying restrictions. With these one could estimate supply and demand curves. In effect, one would be applying *a priori* restrictions to use historically observed data on variation in things other than an excise tax (supply and demand shifts) to estimate the effects of an excise tax.

At the opposite extreme, there might have been a great deal of historically observed variation in excise taxes, all of it generated by considerations (fiscal needs of the government, for example) unrelated to random disturbances to supply and demand for peanuts. Then one could estimate the required reduced form by regression (possibly nonlinear, possibly dynamic) of price and quantity of peanuts on the tax rate in the historical sample. Even in this case one would be using some *a priori* restrictions, choosing some functional form for the regression, and assuming that form to be stable in time.

Intermediate cases are possible also. The tax rate might be strongly related to employment in the whole economy, for example. It would then not be predetermined, but with some knowledge about determinants of employment and their relation to the peanut market, one might find identifying assumptions that permitted estimation of the effect of the tax on peanut quantity and price (for instance, by finding an instrumental variable for the tax rate). In this case one would be using some *a priori* restrictions on the private sector in the estimation, but the supply and demand of peanuts would not need to be estimated.

As another example, if a change in the money stock were contemplated and the money stock had never changed in recorded history, the task of using historical data to estimate a valid reduced form to predict the effects of a change in the money stock would be difficult, though not impossible. One would have to invoke hypotheses about how variation in other things could be used to deduce the effect of variation in money supply. If there had been some variation in money supply historically, it would be possible to obtain more accurate estimates of the effects of money-supply variation, and one would probably abandon some of the more doubtful hypotheses invoked in the absence of historical variation

in the money supply. The smaller that variation is relative to the contemplated change in the money supply, the less reliable would be conditional forecasts based on the historical data. This is simply an example of the general principle that extrapolation of a statistical model far beyond the range of history to which the model was fit yields unreliable results. This unreliability arises because the model's extrapolations lean more heavily on aspects of the specification dependent on one's inevitably uncertain *a priori* beliefs, less on evidence from the data, as the conditions from which one extrapolates move further away from historical experience.

All the points in the preceding paragraph are commonplace to anyone who has used statistical models. They have nothing to do with rational expectations in principle. The rational expectations critique of econometric policy analysis has been taken by some to show that policy analysis based on attempts to extrapolate the effects of alternative policy actions using econometric models is logically unsound. A careful look into the matter, however, reveals that this is not so. The rational expectations critique is only a special case of the more general cautionary note—statistical models are likely to become unreliable when extrapolated to make predictions for conditions far outside the range experienced in the sample.

One of the main contributions of the rational expectations assumption to macroeconomics has been to provide examples showing how even a sequence of policy actions,  $c_t$ , whose size at each date,  $t$ , is within the historically normal range could be far outside the range of past experience in the relevant sense if its time pattern were historically unusual. The same considerations that would lead to caution in predicting the effects of an increase in  $c_t$  of 25 percent in one quarter when the historical standard error of quarterly changes in  $c_t$  is 3 percent should also signal caution about predicting the effects of twelve successive quarters of positive 3 percent change in  $c_t$  when historically  $c_t$  has behaved like a serially uncorrelated stochastic process. Both types of action are outside the historically normal range; both would have effects on the public's prediction methods and uncertainty levels.

### **Are Existing Large-Scale Models Useful?**

I have argued elsewhere that existing large-scale models embody identifying restrictions that are not in fact believed, even as approxima-



tions, by most economists.<sup>6</sup> I also argued, though, that they represent a valuable summary of a great deal of historical experience and that forecasts from them are useful. The reason is that the identifying restrictions are pragmatically adjusted to avoid obvious conflicts with the data, so that they can be regarded as convenient simplifications, not as a priori knowledge imposed on the data. According to this interpretation, the restrictions cannot in fact be of value in identifying the model, though they are useful in improving its forecasting performance.

Though this is harsh criticism, it does not lead to the iconoclastic conclusion of Lucas and Sargent: “that modern macroeconomic models are of *no* value in guiding policy and that this condition will not be remedied by modifications along any line which is currently being pursued.”<sup>7</sup> They seem to base this conclusion on the view that no aspect of the structure of such models is likely to remain fixed under policy interventions because of the rational expectations critique. It should be clear by now that if this paper’s argument is accepted, the only aspect of a model’s identification that is crucial to its yielding useful projections of the effects of policy is its distinction between policy-setting behavior and the behavior of the public given policy acts. False “identifying” restrictions that do not distort the conditional forecasting properties of the model’s sector displaying the reaction of the public to policy acts will not prevent the model from making useful conditional forecasts of policy effects. There is a real question as to how accurately these models identify the reaction of the public to policy, as will be shown in the next section. However, these reasons for doubt apply just as strongly to existing empirical implementations of rational expectations concepts.

Existing large-scale models are useful as they stand, and there are plenty of currently suggested lines of work, mostly not based on rational expectations, which are likely to improve such models.

## Identification

While the valid reduced forms discussed above are reduced in the sense that they do not require explicit estimation of all the behavioral

6. “Macroeconomics and Reality.”

7. “After Keynesian Macroeconomics,” in Federal Reserve Bank of Boston, *After the Phillips Curve: Persistence of High Inflation and High Unemployment*, Conference Series 19 (FRBB 1978), p. 50, reprinted in Robert E. Lucas, Jr., and Thomas J. Sargent, eds., *Rational Expectations and Econometric Practice*, vol. 1 (University of Minnesota Press, 1981), p. 296.

parameters of the private sector, they are not the standard reduced form of econometric theory except in special cases. For one thing, predetermined variables that are not policy variables are still part of the state-of-the-economy vector. Because the models discussed here give a conditional distribution for the full state-of-the-economy vector next period, they must include equations for every nonpolicy variable in that vector. More importantly, every policy variable is an explanatory variable in these models. A standard reduced form contains only predetermined variables as explanatory variables. Policy variables need not be predetermined. When they are not, estimation of the type of reduced form under discussion raises the usual problems of identifying parameters in a simultaneous equations model.

While the rational expectations critique of econometric policy analysis is not as fatal as it might seem, imposing rational expectations does create difficulties in distinguishing between policy behavior and private sector response that have not been confronted in most macroeconomic work, even in work that assumes rational expectations. The Lucas critique explains that one must rethink the definition of structure, and hence of identification, when choosing a fixed policy rule.<sup>8</sup> It does not directly say anything about the problem of identifying and distinguishing between policy actions and public reaction. Indeed, the simple models commonly used to illustrate the Lucas critique handle this kind of identification problem by assuming that policy variables are predetermined, which is the same assumption as that underlying the usual approach to estimating and simulating standard macro models.

For example, Sargent and Wallace use a model with reduced form:

$$(3) \quad y_t = a_0 + a_1 y_{t-1} + u_t + a_3 e_t,$$

where  $u_t$  is the disturbance to an aggregate supply equation,  $y_t$  is output, and  $e_t$  is the disturbance to the money-supply equation.<sup>9</sup> Their point is that equation 3 does not contain the parameters of the money-supply rule, and hence the choice of those parameters cannot affect the stochastic behavior of  $y_t$ . But they introduce their paper with a discussion

8. Robert E. Lucas, Jr., "Econometric Policy Evaluation: A Critique," in Karl Brunner and Allan H. Meltzer, eds., *The Phillips Curve and Labor Markets*, Carnegie-Rochester Conference Series on Public Policy, vol. 1 (Amsterdam: North-Holland, 1976), pp. 19–46.

9. Thomas J. Sargent and Neil Wallace, "Rational Expectations and the Theory of Economic Policy," *Journal of Monetary Economics*, vol. 2 (April 1976), pp. 169–83.

of Milton Friedman's proposed fixed percentage growth rate rule for the money stock, making it clear that they are presenting grounds for believing such a rule would not be suboptimal. In their model if such a rule did not eliminate or reduce variance in  $e_t$ , it would have no effect at all and there would be little point in discussing it. There is thus an implicit assumption that 3 would remain fixed not only for changes in the systematic part of the policy rule but also for changes in the stochastic behavior of  $e_t$ . In particular, a change in rule that set the variance in  $e_t$  to zero would leave 3 invariant. Furthermore, the main reason for paying attention to 3 is that there are strong, relatively well-behaved, dynamic statistical relations between the money stock and  $y_t$  in historical data.

Equation 3 explains this observation as arising from a connection between random disturbances in money supply and movements in output, but without further assumptions one can neither interpret this connection as causal nor draw policy conclusions. Part of the money-supply disturbance may not be generated by policy; and at least this part may be affected by contemporaneous movements in disturbances to other equations in the system. If these possibilities are important in practice, the monetarist implications of 3 would be substantially altered, even though the formal conclusion that changes in the parameters of the money-supply rule do not affect the behavior of  $y_t$  would remain valid. In the extreme, if no part of the disturbance to money supply could be affected by policy, then no monetary rule could affect  $y_t$  and discussion of the Friedman rule would be pointless. In making conditional forecasts of the effects of policy, it is necessary to identify the policy-induced component of  $e_t$  and to make an assumption about how  $u_t$  responds to this component.

The assumptions that Sargent and Wallace make, that  $e_t$  is uncorrelated with other disturbances and that the money-supply equation involves no current values of variables other than money, amount to the standard assumption of simultaneous equations econometrics—that the money stock is predetermined. The assertion that money stock is predetermined is not testable by itself because, as is well known, any simultaneous equations model can be transformed to a Wold causal chain form, in which an arbitrarily chosen variable becomes predetermined in all but one of the equations. It is, however, testable jointly with other restrictions. In a BPEA paper ten years ago Goldfeld and Blinder tested for the importance of policy endogeneity by using instrumental

variables for policy variables in estimating the system. Instruments were available because of maintained hypotheses on the form of the nonpolicy equations of the system.<sup>10</sup> It is often—perhaps even nearly always—the case that an assertion that policy variables are predetermined can be plausible only because of a belief that policy can respond to nonpolicy variables only with a delay. For the delay to be exactly one period in the time units in which the data happen to be measured would be a bizarre coincidence in most applications. If the delay is more than one period, and one imposes this delay in the form of the policy-formation equations, overidentifying restrictions are generated that allow testing for the predeterminedness of policy variables without restrictions on the nonpolicy equations of the system. In the limit as the delay becomes infinite, so that there is no feedback from nonpolicy variables to policy, policy variable predeterminedness is equivalent to exogeneity, which is testable by methods I have described elsewhere.<sup>11</sup>

Economists are accustomed to testing their specifications, formally and informally, for the presence of substantial feedback over time periods of a quarter or a year. In the case of the money stock, I and others have tested a strict exogeneity assumption and found it well supported by the U.S. data in small systems not including an interest rate. Although his statistical work is less formal, Friedman's empirical work on the relation of money and income is most naturally interpreted as building evidence for a predeterminedness assumption. His work acknowledges the possibility that feedback exists and that some of the observed cyclical variation in money stock represents response of the stock to changes in business conditions rather than deliberate policy-induced shifts, but he brings forward evidence that most of the large variations in the stock did not result from feedback. His use of evidence on the relative timing of turning points in the money stock and in business activity is one category of such evidence. But he also tries to show that the time from a turning point in the money stock to a subsequent turning point in business activity is less variable than the corresponding gap when the turning point in business activity precedes that in the money stock. And he

10. Stephen M. Goldfeld and Alan S. Blinder, "Some Implications of Endogenous Stabilization Policy," *BPEA*, 3:1972, pp. 585–640.

11. "Exogeneity and Causal Ordering in Macroeconometric Models," in Federal Reserve Bank of Minneapolis, *New Methods in Business Cycle Research* (FRBM 1977), pp. 23–43, and "Money, Income, and Causality," *American Economic Review*, vol. 62 (September 1972), pp. 540–52.

displays evidence that the variability of the money stock is related to subsequent variability in business activity.<sup>12</sup> I interpret Friedman as trying to convince us that most of the observed correlation of money with income stems from the correlation of unpredictable disturbances in the money stock with subsequent movements in income. This amounts to convincing us that money can appropriately be treated as predetermined.

But evidence, of whatever sort, that money stock is predetermined can only show that a reasonable statistical model can be developed in which one equation is a regression of money stock on lagged values of other variables and in which this equation's disturbance is independent of other equations' disturbances. Such evidence can never prove that the equation in question, or its disturbances, is actually what we think it is—a reflection of monetary policy choices. The application of rational expectations to financial market modeling has generated a wide range of examples in which speculative arbitrage gives a misleading impression of exogeneity and causality. These examples give good reason for concern that usual econometric analysis, including especially monetarist models, may be badly biased by an identification problem. Variables determined in financial markets, such as interest rates, asset prices, and nominal stocks of assets, are likely to appear predetermined and to have substantial explanatory power for other variables in the model, even when they are only passive reflections of real economic activity.

That this can occur with, say, a stock price, is not hard to see. Obviously for small time-intervals, to a reasonable approximation, a stock price must follow:

$$(4) \quad E_t(P_{t+1}) = P_t,$$

where  $P_t$  is the stock price and  $E_t$  is expectation conditional on information available at  $t$ . If 4 were far from true, there would be a predictable component to price changes, creating room for arbitrage profits. Even for a small company, there is likely to be a relation between earnings and business cycle conditions, not because the company's performance

12. These remarks refer to Milton Friedman, "The Demand for Money: some Theoretical and Empirical Results," *Journal of Political Economy*, vol. 67 (August 1959), pp. 327–51; "The Lag in the Effect of Monetary Policy," *Journal of Political Economy*, vol. 69 (October 1961), pp. 447–66; and Milton Friedman and Anna J. Schwartz, "Money and Business Cycles," *Review of Economics and Statistics*, vol. 45, no. 1, pt. 2 (February 1963), pp. S32–S64.

affects GNP, but because general economic conditions affect the company's performance. Yet 4 implies that  $P_{t+1} - P_t$  is uncorrelated with GNP occurring at  $t$  or earlier. The level of  $P_t$  has already captured any predictable effect on  $P_{t+1}$  of GNP occurring at  $t$  or earlier. Thus  $P_{t+1} - P_t$  is correlated only with information about GNP that is new at  $t + 1$ .

If the public has ways of anticipating movements in GNP that are better than extrapolating past movements in GNP itself,  $P_t$  will have predictive value for GNP beyond that contained in past values of GNP itself. Yet 4 implies that GNP does not help predict  $P$  in this sense. As is by now well understood, this implies that there is a distributed lag regression of GNP on  $P$  with exogenous  $P$ .<sup>13</sup>

More intuitively, historical investigators will find that major movements in GNP are *preceded* by *unpredicted* movements in  $P$ . If they were already disposed to think that movements in  $P$  were causing movements in GNP, this finding might appear to confirm their belief. Nonetheless, it is clear that for an asset price like this such a finding does not indicate that the correlations of  $P$  with GNP represent a causal effect of  $P$  on GNP. Arbitrage and forecasting creates a pattern of correlations that mimics what would reasonably be expected of a causal relation.

No one is likely to be misled into thinking a single company's equity prices determine GNP; but when an asset price is an overall stock market index, or a long-term interest rate, or Tobin's  $q$ , the risk may be greater. In fact the risk even exists for the stock of money.

The stock of money is, of course, not an asset price. It is not even the product of a market price with a quantity of an asset. One is accustomed to thinking of it as in effect set by the monetary authority. As I have argued elsewhere, it is therefore not apparent that worries about statistical exogeneity of money being induced by speculative activity are of any importance.<sup>14</sup> But it is in fact plausible that much movement in the money stock, particularly in the short run, is responsive to demand. It is possible to construct a rational expectations, equilibrium model in which the monetary authority is completely passive, in which the stochastic process generating output could not be affected by the monetary authorities even if they attempted to do so, yet in which money stock has predictive power for output and is statistically exogenous.

One way to construct such a model is to introduce money into Hall's

13. See my "Money, Income, and Causality."

14. "Exogeneity and Causal Ordering."

permanent income consumption model.<sup>15</sup> Hall assumes that consumers maximize expected lifetime utility of the form,

$$(5) \quad E \left[ \sum_{s=0}^{\infty} (1 + \delta)^{-s} U(C_s) \right],$$

subject to the budget constraints,

$$(6) \quad C_t + W_t = Y_t + (1 + r_t)W_{t+1}, \quad t = 0, \dots,$$

where  $C_t$  is consumption,  $W_t$  is net worth,  $r_t$  is the interest rate,  $\delta$  is the subjective discount rate, and  $Y_t$  is income. The budget constraints imply that borrowing or lending at the same rate,  $r_t$ , is possible for each consumer, regardless of that consumer's current net worth.

If  $r_t$  is fixed forever at  $r$ , and if  $Y_t$  is given exogenously, Hall shows that a necessary condition for a solution to the consumer's optimization problem is

$$(7) \quad E_t(DU_{t+1}) = [(1 + \delta)/(1 + r)]DU_t,$$

where  $DU_t = U'(C_t)$ . As in equation 4 and the stock price example, this implies that the marginal utility of consumption will behave like an exogenous variable in any model including it and any other variables observable by the public at  $t$ . Because in this problem  $DU_t$  is a monotonic function of  $C_t$ , the latter will itself behave to a close approximation as exogenous. It will be exactly exogenous if  $U$  is quadratic.

Now suppose there is a transactions demand for noninterest-bearing money,  $M_t$ , which depends on the volume of consumption. This idea can be represented by including money as an argument in the utility function. Then  $U(C_s)$  in 5 can be replaced by  $U(C_s, M_s)$  and the constraints 6 by

$$(8) \quad C_t + W_t + M_t = Y_t + M_{t-1} + (1 + r_t)W_{t-1}, \quad t = 0, \dots$$

The first-order condition, equation 7, is unaffected by adding  $M_t$  to the problem, except that in it  $DU_t$  is replaced by  $D_1U_t = \partial U(C_t, M_t)/\partial C_t$ . It still can be interpreted as "saying that marginal utility of income is best predicted by its own current value," but now, in general, marginal utility of consumption depends on  $M$  as well as  $C$ .

15. Robert E. Hall, "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, vol. 86 (December 1978), pp. 971-1007.

One additional first-order condition can be obtained for the extra choice variable,  $M$ :

$$(9) \quad D_2 U_t = D_1 U_t - E_t(D_1 U_{t+1})/(1 + \delta).$$

But from 7 this is

$$(10) \quad D_2 U_t = [r/(1 + r)]D_1 U_t.$$

Equation 10 provides an exact contemporaneous relation between  $M_t$  and  $C_t$ . Because of this exact dependence, the marginal utility of consumption can be written as a function of either  $C_t$  or  $M_t$  alone. Hence the fact that marginal utility is its own best predictor means that this is approximately true also for both  $M_t$  and  $C_t$ . Both would be exactly their own best predictors if  $U$  were quadratic.

Hall's is a version of the permanent income theory of consumption and implies the familiar distributed lag relation of income with consumption—though with consumption, not income, exogenous. With  $M_t$  a function of current  $C_t$ , clearly  $M_t$  is also strongly related to income and predetermined.

The example treats the price level as fixed and hence implies that monetary policy ratifies the shifts in money demand that occur as permanent income shifts, keeping the price level fixed by open market operations in the single commodity. In this sense monetary policy is completely passive. The statistical exogeneity of money in distributed lag regressions of income on money in this model has nothing to do with predeterminedness of policy decisions.

The most serious objection to the model, which applies also to the original Hall model, is that no reasonable general equilibrium framework could leave real interest rates fixed forever. Equation 7 implies that marginal utility cannot be a stationary stochastic process with positive mean. This is because with fixed  $r_t$  optimal accumulation leads to consumption drifting off to zero or infinity. Thus the model must be taken as an approximation. If  $r_t$  is a continuous function of time, 7 will be true to an arbitrarily close approximation with data measured at sufficiently small time intervals.<sup>16</sup>

16. The approximation involved here is really the same as that needed to justify 4, the equation describing the unpredictability of stock prices. That equation ignores interest rate effects and cannot hold exactly in the face of a random interest rate. But both 4 and 7 will work arbitrarily well as econometric specifications, in a certain sense, if the time unit is small. The argument for this is given in my "Martingale-Like Behavior of Prices," Discussion Paper 489 (National Bureau of Economic Research, June 1980).



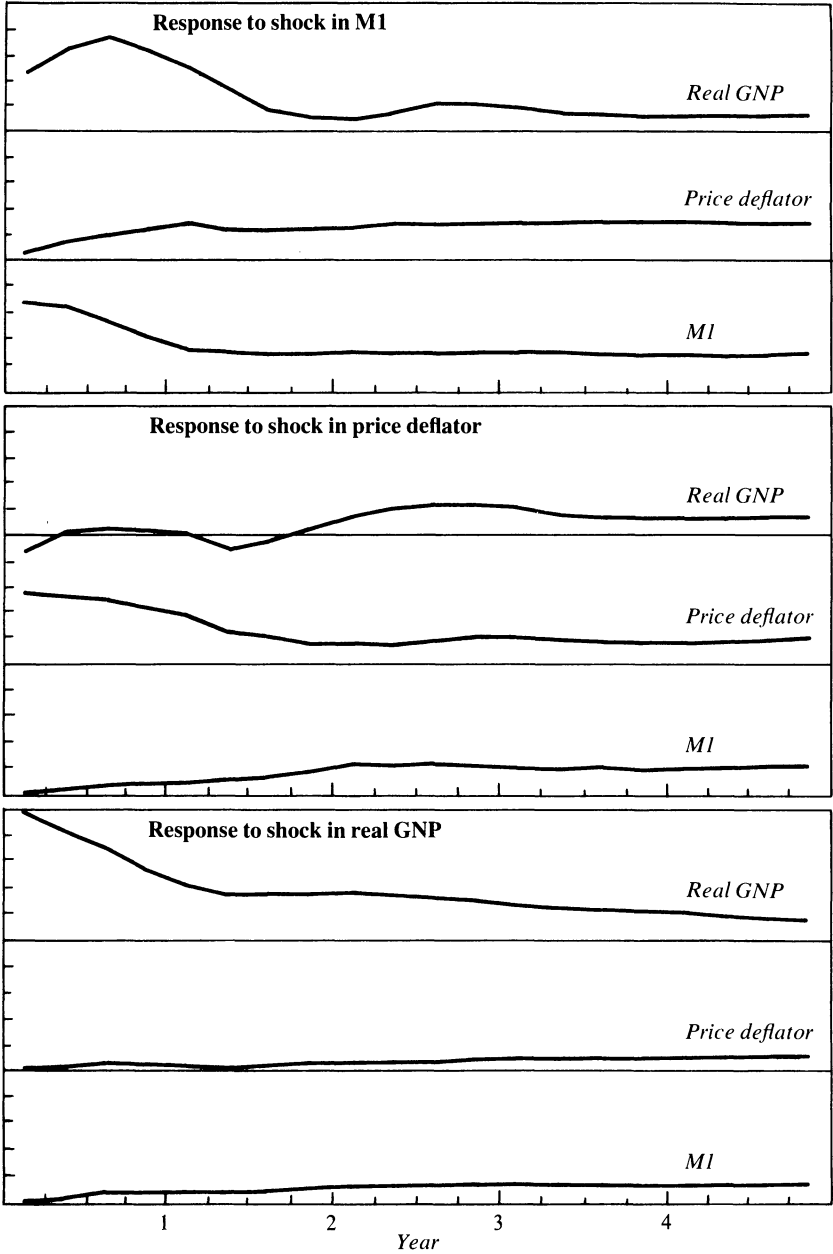
This example is not meant to be taken as an assertion about the way the world works. It is an illustration showing that we cannot take a demonstration that money stock is statistically predetermined to be good evidence that dynamic regressions of real variables on the money stock display the effect of monetary policy disturbances on those real variables. Although the example includes some arbitrary assumptions, it is in this respect like rational expectations models of the Phillips curve, which illustrate that one cannot suppose that a Phillips curve will remain stable under a change in monetary policy rule. It also functions like Tobin's model, which, while holding money passive, generates timing relations among money stock and real variables like those found by Friedman.<sup>17</sup> The present example generates a stochastic process for money and income in which movements in output are preceded by unpredictable movements in money stock, even though monetary policy has nothing to do with generating the movements in output.

Such an example can be considerably elaborated—to include a role for money as a buffer against temporary fluctuations in income, to relate transactions demand for money to current income as well as consumption, to allow for the possibility of inflation. But even a much more realistic example could preserve the basic idea that changes in money demand, accommodated by the monetary authority, reflect anticipations of future fluctuations in real activity.

Granger causal priority of money—the condition that the best forecast of money be formed from lagged values of money alone—remains a characteristic of U.S. data through the current period for systems including money, prices, and a measure of real activity. A natural measure of the degree to which Granger causal priority holds is the percentage of forecast error variance accounted for by a variable's own future disturbances in a multivariate linear autoregressive model. In such a system the  $k$ -step ahead forecast error for each variable is a linear combination of forecast errors 1 through  $k$  steps ahead in the variable itself and in other variables in the system. A variable that is optimally forecast from its own lagged values will have all its forecast error variance accounted for by its own disturbances. Table 1 displays results, shown in parentheses, for a three-variable system formed from M1, real GNP, and the GNP deflator and estimated from data through the third

17. James Tobin, "Money and Income: Post Hoc Ergo Propter Hoc?" *Quarterly Journal of Economics*, vol. 84 (May 1970), pp. 301–17.

Figure 1. Dynamics of the Three-Variable System<sup>a</sup>



a. See the discussion of table 1 in the text and the appendix for details. The variables are defined in the notes to table 2. The vertical scale makes the distance from one horizontal axis to the one above it correspond to 1 percent of the original level of the variables before the shock.

**Table 1. Percentage Variance Decompositions, Six and Fourteen Quarters Ahead, in Three- and Six-Variable Autoregressive Systems, 1948:3 through 1981:3<sup>a</sup>**

Innovation variable	Forecast horizon (quarters)	Error variance					
		M1	Three-month Treasury bill rate	GNP deflator	Real GNP	Federal expenditures	Federal revenues
M1	6	(96.0)	...	(19.1)	(41.0)	...	...
	6	57.5	25.4	29.7	15.1	4.7	27.2
	14	(81.0)	...	(38.0)	(36.0)	...	...
	14	39.1	24.0	33.7	13.9	7.5	24.4
Three-month Treasury bill rate	6	35.4	72.6	0.3	21.4	1.3	4.5
	14	31.5	72.8	2.7	18.5	1.3	5.3
GNP deflator	6	(1.6)	...	(80.3)	(0.3)	...	...
	6	2.1	0.5	65.7	0.4	20.1	16.7
	14	(11.7)	...	(59.7)	(4.5)	...	...
	14	9.6	1.1	47.7	3.7	22.4	16.7
Real GNP	6	(2.4)	...	(0.6)	(58.8)	...	...
	6	1.5	0.2	0.9	56.5	8.1	18.2
	14	(7.3)	...	(2.4)	(59.5)	...	...
	14	8.2	1.1	1.5	47.4	12.7	19.8
Federal expenditures	6	1.5	1.1	2.8	0.8	60.1	1.1
	14	1.6	1.2	2.0	1.6	47.7	1.5
Federal revenues	6	1.9	0.2	0.5	5.8	4.6	32.3
	14	10.1	0.4	7.3	15.0	8.4	32.3

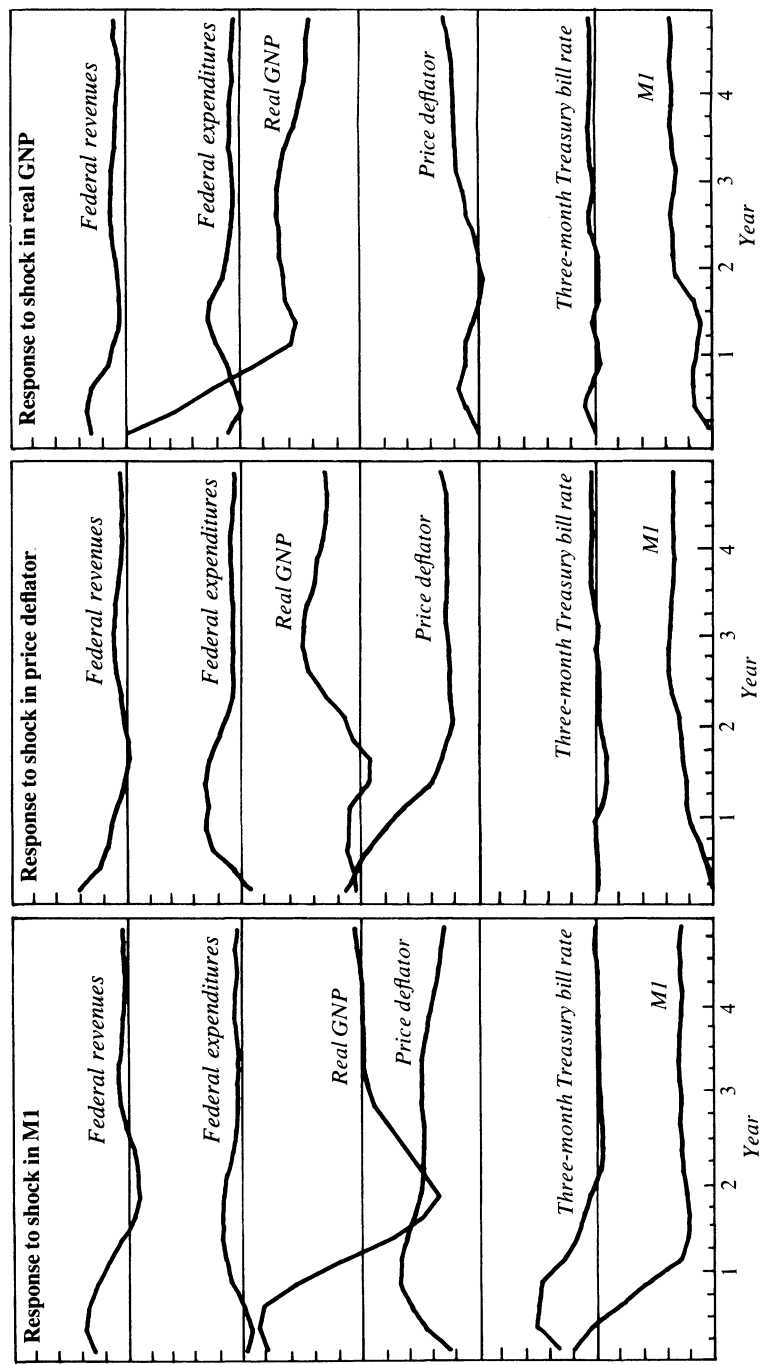
Sources: Estimations by the author based on data from the national income and product accounts and the Board of Governors of the Federal Reserve System.

a. Entries show the percentage of forecast error variance, six and fourteen quarters ahead, that is accounted for by innovations in the variables in the first column. Entries in parentheses are for the three-variable system; all others are for the six-variable system.

quarter of 1981, allowing for parameter drift.<sup>18</sup> M1 emerges as Granger causally prior to a close approximation, while accounting for substantial parts of the variance in GNP and the deflator. Figure 1 shows that the estimated response of the system to a disturbance in money stock is just

18. Because the system allows for time varying parameters, it is not linear. Stochastic parameters multiply the random variables on the right-hand side of each equation. Thus the summary statistics in table 1 are not the model's implications about actual forecast errors. They are instead its implications about forecast errors assuming the 1981:4 estimates of the coefficients were to remain fixed. If parameters actually changed rapidly in the estimated model, such statistics would be of limited interest, but the model implies the parameters change slowly, so these tables are comparable to similar tables summarizing the structure of fixed-parameter linear models estimated by myself and others. In general, actual historical forecast errors many periods ahead behave quite differently from what is implied by taking current parameter estimates as known exactly, whether one uses fixed or varying parameter models.

Figure 2. Dynamics of the Six-Variable System, Responses to Shocks in M1, Prices, and Real GNP<sup>a</sup>



a. See the discussion of table 1 in the text and appendix for details. The variables are defined in the notes to table 2. The vertical scale for all except the fiscal variables makes the distance from one horizontal axis to the one above it correspond to one-half percent of the original level of the variables before the shock (one-half percentage point for the Treasury bill rate). For federal expenditures and federal revenues this distance corresponds to 5 percent of their original level to allow for their greater variability.

what a monetarist would predict: a substantial temporary response in output and a more slowly developing, more persistent response in prices.

As has recently been documented in my own work and that of Litterman and Weiss, among others, this pattern disappears in systems that include an interest rate.<sup>19</sup> Table 1 also shows the decompositions for a six-variable system that adds the three-month Treasury bill rate and federal government expenditures and revenues. Here it is shown that much of the variance in M1 can be accounted for by disturbances to the interest rate and much of the variance of GNP is now attributed to the interest rate and federal revenues, with substantial reduction in the share attributed to M1. Figure 2 shows that the responses of M1 itself and of GNP to disturbances in M1 are weaker in the system with an interest rate, though price responses are strengthened. Figure 3 shows that simultaneous downward movements in M1 and GNP emerge as responses to interest rate disturbances in the larger system.<sup>20</sup>

The arguments given above against making too much of the explanatory power of variables determined in financial markets apply to the interest rate with full force. The fact that one finds that it is precisely the part of M1 that is predictable in advance from knowledge of the interest rate that is most strongly connected with real activity does, however, seem to be inconsistent with simple rational expectations monetarist models that imply only the unpredictable part of the money stock affects real activity. Also, it was shown above that M1 can behave as if it were directly determined in financial markets if the monetary authority is passive. The fact that the interest rate absorbs explanatory power from M1 is consistent with the view that M1 behaves as Granger causally prior in systems including no other financial variables because it acts as a kind of proxy financial variable.

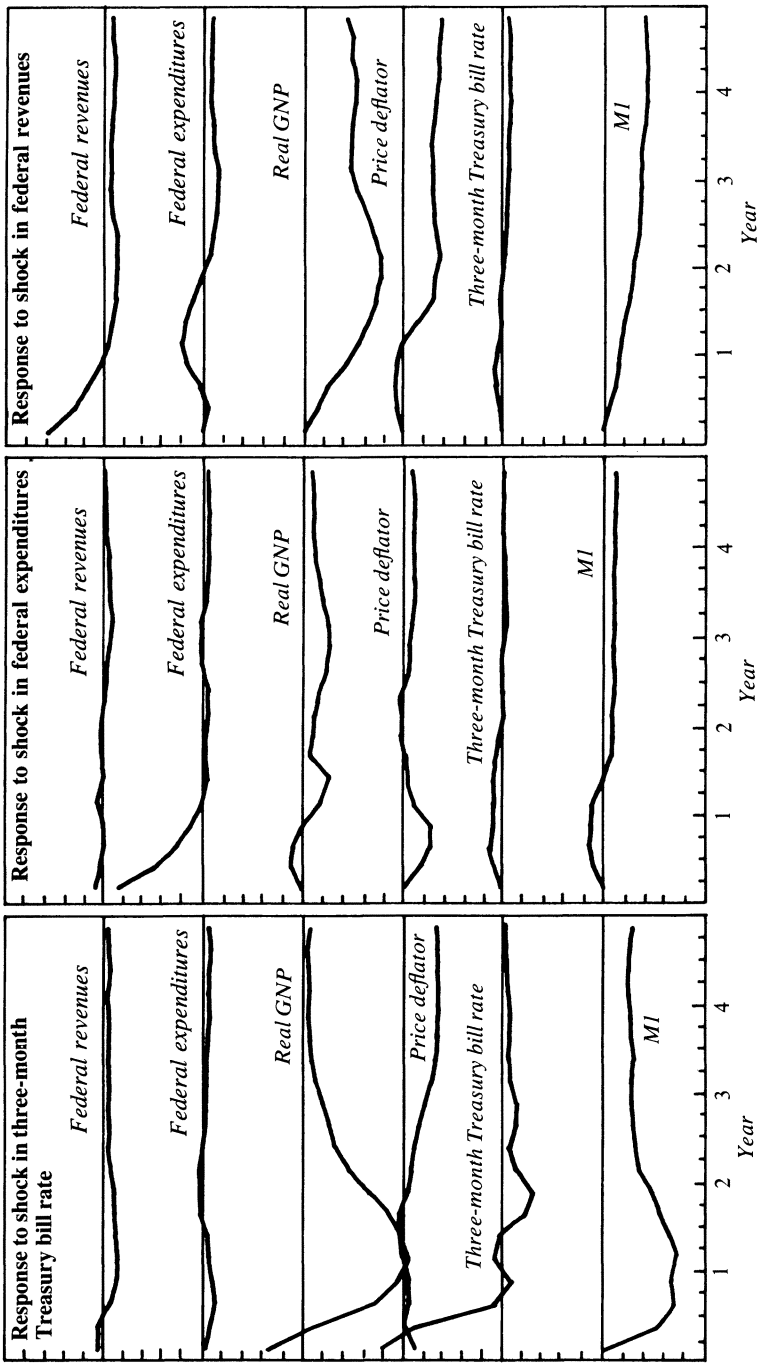
One interpretation of these results might be as follows. Suppose the monetary authority exerts tight control in the short run over interest rates but cannot influence M1 except indirectly through those rates.<sup>21</sup>

19. Christopher A. Sims, "Comparison of Interwar and Postwar Business Cycles: Monetarism Reconsidered," *American Economic Review*, vol. 70 (May 1980, *Papers and Proceedings*, 1979), pp. 250–57 and "International Evidence on Monetary Factors in Macroeconomic Fluctuations," Discussion Paper (University of Minnesota, 1980); and Robert B. Litterman and Laurence Weiss, "Money, Real Interest Rates, and Output," Working Paper 179 (Federal Reserve Bank of Minneapolis, 1981).

20. The appendix provides a formal explanation of the figures.

21. This possibility was suggested to me in a private conversation with Robert Litterman.

Figure 3. Dynamics of the Six-Variable System, Responses to Shocks in Three-Month Treasury Bill Rate, Federal Expenditures, and Federal Revenues<sup>a</sup>



a. See note a to figure 2.

Suppose further that the response to interest rate disturbances is as displayed in figure 2. Then monetary policy has large but slow effects on real GNP and M1. Monetary policy is effective, but the authority has little control over the disturbances to the M1 equation. Instead the authority produces predictable smooth swings in M1 over several months by influencing the interest rate. Because of the shape of the responses to interest rate disturbances, it can be shown that it will be impossible to offset short-run disturbances in M1 or real GNP via manipulation of the interest rate. Attempts to do so will lead to what is known as instrument instability—explosive oscillations in the interest rate that can only be damped at the cost of large oscillations in M1 and GNP. This view could explain a historical pattern in which the monetary authority stabilized interest rates, at least in the short run, and thus why money stock might move mainly in response to demand. As in the example above, this could generate statistical exogeneity of the money stock. This view would also explain why attempts to stabilize the path of the money stock have coincided with large movements in the interest rate and large short-run oscillations in GNP and M1. But though this view is appealing as a working hypothesis to suggest further research, it is not the only way to look at the data. One does not know how much of the historically observed disturbances in interest rates can be attributed to policy choice. The recent change in the volatility in interest rates might stem from changes in the structure of nonpolicy disturbances.

Thus the theoretical model and empirical results under discussion do not, unfortunately, lead to any simple resolution of the main disputed points in macroeconomics. As has been shown, a passive monetary authority would create behavior in the money stock that would make money seem exogenous and make it appear to explain prices and real activity. But the same result would emerge if the money stock did in fact explain real activity and the monetary authority moved it around erratically.

### **Policy Evaluation When the Structure is Unknown**

Even to project the effects of policy choices, not to speak of arriving at the best choice, one must resolve the uncertainty about which variables policy choices can actually affect and how their historical variation

reflects those choices. The time-series data will not resolve the uncertainty on their own. One will have to make judgments on the basis of common sense and evidence other than time series. This process is vulnerable to prejudice, and its results will inevitably be controversial. More or less objective statistical analysis of the time-series data can nonetheless play an important role.

The U.S. postwar data contain enough information to give a useful characterization of the conditional distribution of the future of major macroeconomic aggregates given the past. Although there is evidence that this structure changes over time, there is also evidence that it does not change suddenly, so that a model fit to the whole postwar period as if parameters were fixed over that whole period is not badly biased because of parameter changes. This means that, while there is much room for judgment and prejudice in the choice of macroeconomic policy, careful attention to the historical data exerts an important discipline on what can be plausibly asserted about the way the economy works. The discipline should be applied in two ways. Those with a particular view about the way the economy works should construct their policy recommendations using an econometric framework that limits the exercise of prejudice and informal reasoning to areas in which the data are in fact uninformative. In sorting out the claims of different policymakers about the likely effects of their proposals, projections of policies and their effects should be checked for plausibility against predictive distributions derived from the data.

This may sound like a modest and uncontroversial set of standards for policy debate, but it does run counter to some views. It is sometimes claimed that the probabilistic structure of the economy is so unstable that there is no point in econometric evaluation of policy. This is sometimes seen, I think, as an implication of Lucas's critique of econometric policy evaluation and of the fact that econometric models have in recent years made forecast errors that have been large in absolute terms.<sup>22</sup> But Lucas's critique implies only that if policy rules change often the structure will also shift often. As I have argued above, policy rules in the relevant sense of that term have not changed frequently or by large amounts. The large forecast errors of recent years do not seem to be attributable mainly to shifts in the structure of predictive equations.

22. Lucas, "Econometric Policy Evaluation."



Statistical models allowing for drift in predictive structure estimate best when the change in that structure is assumed to be slow, so that recent large predictive errors are interpreted as large random shocks to the equations, not mainly as the effect of parameter changes.

Another view that conflicts with what is proposed here is the idea that economists should focus their attention on choice of rule and that since the rule change will shift the structure, projections using the existing probabilistic structure are beside the point. Again, as I have argued above, disputes about the optimal rule are no more important in principle than disputes about how to implement the existing “rule” as it emerges from existing institutions and interests. Nonetheless, even when confronting economists who insist that the rule ought to be changed and that their proposed actions are part of a change in rule, one should use the existing probabilistic structure to evaluate such proposals. Precisely because those vying for control of policy will propose to make permanent changes in the rule much more often than they will succeed in doing so, the public is likely to discount their rhetoric and react to the actual course they set for policy as if it were a disturbance to the existing probabilistic structure. If those proposing the change in rule succeed in implementing it for some years, and if they have not been similarly successful in the past, the statistical model will give little reliable information about the effects of their persistent success. But the immediate consequences of the proposed course of action for the next two or three years, as projected from the existing probabilistic structure, are likely to be much more reliably determined and deserve to be weighed carefully against the uncertain claims of long-term gains.

When a policy projection emerges as implausible in projections using the historical probability structure, there are at least three possible interpretations. The obvious one is that the proposed policy actions are unlikely to have the claimed effects. Another is that the proposed actions are very different from what has been done under similar circumstances in the past. A third is that the proposed paths of policy variables are probably not attainable. Determining which of these interpretations is correct requires partial identification of the model—an interpretation of it that explains how policy actions affect its disturbances. The data alone cannot produce behavioral interpretations, but they can, by locating the source of improbability in one variable or another, help to decide which interpretations are plausible. Furthermore, even if it is claimed that the

only reason a proposed course of policy action appears implausible is that it represents a deliberate break with past patterns of policy choice, it is reasonable to take the implausibility as at least one strike against the proposal. A policy choice very different from what the historical probability structure predicts will have consequences that are difficult to predict. This will occur not only because of adjustment of expectational mechanisms but also because any use of a statistical model to make projections from conditions unlike those observed in the historical sample is subject to large errors in sampling and specification.

The practical implications of the view put forward here are best illustrated by an example of the proposed type of analysis. Consider the results from estimating the probability structure of the U.S. quarterly postwar data for six variables: M1, the three-month Treasury bill rate, real GNP, the GNP deflator, federal government expenditures, and federal government revenues.

A linear vector autoregressive model was fitted to the data, allowing for parameter drift and beginning from a fairly loose Bayesian prior on the parameters.<sup>23</sup>

The model can be written as

$$X_{t+1} = \sum_{s=0}^6 B_t(s)X_{t-s} + e_{t+1}.$$

The next period's vector of data,  $X_{t+1}$ , is determined by current and past data; by the current coefficient matrices,  $B_t(s)$ ; and by an unpredictable innovation,  $e_{t+1}$ . The model assumes that data more than six quarters old are not relevant to the determination of  $X_{t+1}$ . The coefficient matrices themselves drift over time, following a random walk:

$$B_t(s) = B_{t-1}(s) + V_t(s) \quad s = 0 \dots 6.$$

Even if one knew  $B_t(s)$  precisely at time  $t$ ,  $X_{t+1}$  could not be forecast perfectly because  $e_{t+1}$  is not predictable from knowledge of current and past  $X$  and current  $B$ . But  $B_t(s)$  is not known precisely, at time  $t$  or even after the event. Since the  $B$  parameters change over time, it is not the

23. The time variation in the parameters is the methodological novelty here. In other respects the approach is similar to that used in my "Macroeconomics and Reality," and in Stanley Fischer's "Relative Shocks, Relative Price Variability, and Inflation," *BPEA*, 2:1981, pp. 381-431. A more detailed description of the technique used here is available from the author in a technical appendix.

case, as it would be with fixed parameters, that a sufficiently long sample period would enable one to estimate  $B$  with an arbitrary degree of precision. Even if certain knowledge of  $B_{t-1}(s)$  were available, one would only be able to estimate  $B_t(s)$  approximately because of the random innovation,  $V_t(s)$ .

It is assumed that  $V_t(s)$  is uncorrelated with  $e_t$  for all  $s$ . Both innovations are by construction serially uncorrelated. This assumption allows application of the technique known as the Kalman filter to form

$$G_t(s) = E[B_t(s) \mid X(u), u = 1 \dots t],$$

the best estimate of the current value of the parameter vector available at time  $t$  by a convenient recursive formula.<sup>24</sup> The formula generates  $G_t(s)$  from  $G_{t-1}(s)$  and  $X(t)$ ; it can be applied period by period through the sample at reasonable computational cost. The Kalman filter requires outside information on the covariance matrices of  $V$  and  $e$  to decompose its error in forecasting  $X(t)$  with  $t - 1$  information into  $X$  error and  $B$  drift. The greater is the variance of unpredictable  $X$  innovations relative to parameter changes, the less the filter will change its estimate of  $B$  in response to a forecast error.

To apply the Kalman filter it is necessary to have a previous estimate of the  $B_t(s)$  parameters for the beginning of the sample period, which can then be modified by subsequent data. The prior was chosen to make each variable on the  $X$  vector an independent random walk. It is also necessary to specify the covariance matrices for the prior, for  $V$ , and for  $e$ . This was done by means of a somewhat unsystematic grid search to determine the most likely degree of variation in the parameters. The criterion used in the search was the sum of squares of the recursive residuals—one-step forecast errors from the model estimated with data only up through the time of the forecast. While the grid for this search was not fine enough to produce precise answers, the best fit was obtained with a model in which the standard error of the change in an individual parameter over the whole sample period was about two-tenths of a percent. Because the model has many coefficients (six lags on each of six variables in each equation), this is enough parameter variation to affect the probability structure of the model, with one-half to two-thirds

24. See Ralph Deutsch, *Estimation Theory* (Prentice-Hall, 1965), chap. 12, or A. C. Harvey, *Time Series Models* (Halsted Press, 1981), chap. 4.

of one-step-ahead forecast error estimated to be accounted for by parameter variation. Estimates allowing for this degree of parameter variation differ noticeably, but not by a very large amount, from those estimated on the assumption of no parameter variation at all. The degree of parameter variation is small enough that studying parameter variation by estimation using separate subsamples would not be profitable—the sampling error would dominate the estimated parameter change. Doubling or halving the rate of parameter drift increases prediction error variance by only about three or four percent.

When the model estimated from data through 1981:4 is used, one obtains the results displayed in table 2. The table also presents the administration's projections for the same variables.<sup>25</sup>

The administration proposed a much slower expansion of the money supply than the model deemed likely as of 1981:4.<sup>26</sup> It also proposed less rapid growth of both expenditures and revenues than the model. The administration is more optimistic on both inflation and real growth, especially on inflation.

It is interesting to examine the discrepancies between the forecasts of the model and those of the administration, based on historical experience with forecast errors from the model. Because this model was estimated recursively, it was not too difficult to generate a complete set of forecast errors 1 through 12-steps ahead for it for each date in the sample. The covariance matrix of these historical forecast errors then provides a standard against which to measure the gap between forecasts by the administration and the model.<sup>27</sup> A natural way to display the

25. Projections were obtained from the *Economic Report of the President, February 1982*, and from the Congressional Budget Office, *An Analysis of the President's Budgetary Proposals for Fiscal Year 1983* (Government Printing Office, 1982).

26. The fourth quarter of 1981 showed a sharp drop in the interest rate, which the model had not predicted. This large forecast error shifted the parameters of the model quite substantially in the direction of producing a lower level of interest rates and a higher trend growth rate of the money stock in the forecast. An early version of this paper used parameters estimated through 1981:3 in the time-series model. That model projected M1 growth and inflation very similar to that projected by the administration, but was much less optimistic about real growth than the administration. However, for exercises like those in table 4, in which the administration's projected paths for policy variables or target variables are imposed on the time-series model, results are similar in both versions of the model.

27. This method does not use the assumed probability structure of the model to generate forecast error covariances for reasons of computational convenience. The procedure used does have the advantage that it gives a reliable picture of the model's forecasting properties even if the model is misspecified, so long as those properties show some stability over time.

**Table 2. Alternative Forecasts for Selected Economic Indicators, 1982–84**  
Percent unless otherwise specified

Forecast and year	Projected variable						
	M1 <sup>a</sup>	Three-month Treasury bill rate <sup>b</sup>	GNP deflator <sup>a</sup>	Real GNP <sup>a</sup>	Federal expenditures <sup>c,d</sup>	Federal revenues <sup>c,d</sup>	Deficit (billions of dollars) <sup>d</sup>
Autoregressive time-series							
1982	9.3	11.7	8.7	2.8	16.2	5.9	136
1983	8.8	11.6	8.2	4.1	16.3	14.9	170
1984	8.6	12.3	8.9	3.0	12.6	14.3	177
Administration							
1982	4.6	11.7	7.2	3.0	9.9	4.5	95
1983	4.3	10.5	5.5	5.2	4.4	6.1	88

Source: Autoregressive time-series forecast—same as table 1; administration's forecast—*The Budget of the United States Government, Fiscal Year 1983*.

a. Change from fourth quarter to fourth quarter. The M1 variable refers to M1B spliced to M1 for periods before M1B is available. Real GNP is GNP in 1972 dollars.

b. Calendar year average.

c. Change from fourth quarter of the preceding year to third quarter of the current year.

d. National income accounts budget figures. When the administration's projections are used, they are applied to fiscal year averages of national income accounts budget figures as percent changes, although the projections were originally prepared for the unified budget. This also accounts for discrepancies in the deficit projections shown that are national income accounts deficits derived by applying administration projections of percentage changes in expenditures and revenues to the national income accounts versions of those figures.

measure is to order the variables and then express the forecast error in variable  $j$  at the  $t$ th step ahead as the sum of a "shock,"  $u_{jt}$ , and the regression of the forecast error on  $u_{jt}$  for earlier dates or the same date for the  $j$  earlier in the ordering. These  $u_{jt}$  can be thought of as the components of the  $t$ -step forecast error, which would be "news"  $t$  steps ahead, compared to the parts that would already have been revised based on information arriving in step 1 through step  $t - 1$  ahead. Each  $u_{jt}$  is normalized to have a unit equal to its own standard deviation. Table 3 shows an analysis of the discrepancy between the model and the administration's forecasts using this decomposition. Many different sets of  $u_{jt}$ 's could generate the administration's forecast; the table displays that set which does so with minimum sum of squares.

The disturbances shown in table 3 are quite small; none of them are as large as two standard deviations, and only three exceed one standard deviation. The decreasing size of the disturbances necessary to match the administration's forecast, as the forecast horizon increases, is consistent with the presence of parameter drift during the sample period. Disturbances like those displayed in the top half of table 3 are well within the expected range of deviations from the model forecast as measured simply by the sum of squared disturbances.

**Table 3. Standardized Disturbances Required to Generate the Administration's Forecast from the Autoregressive Time-Series Model, 1982:1 through 1983:4<sup>a</sup>**

<i>Forecast and period</i>	<i>Disturbance</i>					
	<i>M1</i>	<i>Three- month Treasury bill rate</i>	<i>Federal expendi- tures</i>	<i>Federal revenues</i>	<i>GNP deflator</i>	<i>Real GNP</i>
Administration's forecast						
1982:1	-1.4	0.1	-0.8	0.6	-0.8	1.0
1982:2	-1.2	0.2	-0.5	0.3	-0.2	0.6
1982:3	-0.7	0.1	0.1	0.3	0.0	0.5
1982:4	-0.3	0.2	0.0	0.0	-0.2	0.7
1983:1	-0.1	0.2	0.0	0.4	0.1	0.7
1983:2	-0.2	0.2	-0.1	0.2	0.0	0.5
1983:3	-0.2	0.2	0.0	0.2	0.1	0.4
1983:4	-0.2	0.1	0.0	0.2	0.0	0.3
Administration's forecast for output and inflation only <sup>b</sup>						
1982:1	0.1	-1.1	0.5	-0.6	...	...
1982:2	0.0	-0.9	0.2	-0.2	...	...
1982:3	-0.2	-0.4	0.5	0.1	...	...
1982:4	-0.5	-0.2	0.5	-0.3	...	...
1983:1	-0.5	-0.2	0.6	-0.5	...	...
1983:2	-0.2	0.0	0.3	-0.2	...	...
1983:3	-0.1	0.0	0.0	-0.2	...	...
1983:4	0.0	0.1	0.0	-0.1	...	...

Source: Simulations by the author.

a. Entries show the set of disturbances (with size normalized to have a unit equal to their own standard deviation) that generate the administration's forecast with the minimum sum of squares of the recursive residuals.

b. With no disturbances to GNP and inflation equations.

Taking into account 1984, there are seventy-two disturbances, and their sum of squares is only 9.9. However, the persistent signs of the disturbances suggest a systematic bias in the administration's forecasts. Under a normality assumption, the probability of a sequence of four positive errors in real GNP forecasts, each of which is at least as large as the corresponding one of the first four positive shocks on the real GNP column in the top half of table 3, is about three-tenths of a percent.

Furthermore, this bias appears to be an optimistic one. The probability of an outcome showing at least as much real growth as the administration's forecast while maintaining at least as low a money-growth path and expenditures path, for example, is much lower than the probability

of an outcome with sum of squared forecast disturbances 9.9 or more. A heuristic “plausibility index” (PI) for a projection, designed to show the probability of all outcomes that are further away from the central forecast than the projection in the same direction, is the probability of a normal random variable exceeding the square root of the sum of squares of the disturbances required to fit the projection. The administration’s forecast has a PI of 0.0008.

The implausibility of the administration’s forecast, as far as the model is concerned, concentrates on M1 and real GNP. Money growth as low and real GNP growth as high as projected are quite unlikely, at least as a combination. The first two rows of table 4 show what the model produces as the most likely outcome for real GNP and inflation conditional on the administration forecasts for M1, interest rates, and the fiscal variables being realized. The model’s inflation forecasts are then as low as those of the administration but the model’s real GNP forecasts are far more pessimistic. A table analogous to table 3, showing the forecast disturbances needed to generate the table 4 projections (not shown) would be very similar to table 3 except the last two columns. One might have expected the money forecast to have appeared much more plausible with the GNP forecast not imposed, but in fact it does not turn out that way. Leaving the inflation and real GNP forecasts out of the projection raises the PI to 0.011—still a low figure, but the original PI was smaller by a factor of 0.07.

The results of using the model to generate the administration’s real GNP and inflation forecasts through disturbances in only M1, interest rates, and fiscal variables are reported in the third and fourth rows of table 4. The GNP and inflation forecasts imply a moderate monetary and tremendous fiscal stimulus. The bottom half of table 3 shows that the model does not view these policy settings as extremely unlikely, though again they require patterns of sustained shocks that are less likely when their signs are taken into account. A succession of four negative disturbances in the interest rate forecast, each as large as the corresponding entry in the bottom half of table 3, has a probability under normality of only about four-tenths of a percent. The PI for this projection is 0.013, roughly the same as for the administration’s projected setting of policy variables.

In an experiment not shown in the tables the model was checked for the most likely way for the administration’s forecasts for inflation and

**Table 4. Projections from the Autoregressive Time-Series Model When Constrained to Match Alternative Features of the Administration's Forecast, 1982-83**

Percent unless otherwise specified

Characteristic of projection and year	Projected variable <sup>a</sup>						
	GNP deflator	Real GNP	M1	Three-month Treasury bill rate	Federal expenditures	Federal revenues	Deficit (billions of dollars)
Constrained to match forecast for M1, three-month Treasury bill rate, and fiscal variables <sup>b</sup>							
1982	6.7	-1.0	(4.6)	(11.7)	(9.9)	(4.5)	(95)
1983	5.7	0.4	(4.3)	(10.5)	(4.4)	(6.1)	(88)
Constrained to match administration's forecast for output and inflation <sup>c</sup>							
1982	(7.2)	(3.0)	10.2	10.7	20.3	0.8	177
1983	(5.5)	(5.2)	9.9	10.1	25.2	5.6	389
Constrained to match administration's forecast for M1 <sup>d</sup>							
1982	9.1	-0.4	(4.6)	15.4	13.0	8.7	90~
1983	8.6	-2.5	(4.3)	13.6	9.2	7.0	116

Source: Same as table 2.

a. The administration's forecasts are in parentheses. See also notes to tables 2 and 3.

b. All equations in the system reflect variables undergoing disturbances.

c. With disturbances to equations for M1, three-month Treasury bill rate, and fiscal variables.

d. With disturbances to the three-month Treasury bill rate equation only.

real GNP to be realized. These target forecasts in themselves emerge as not terribly unlikely, as would be expected from the relatively small discrepancies between the unconstrained forecast and the administration's projections shown in table 2 for these two target variables. The largest disturbance required is only 0.39 standard errors, though there are a succession of disturbances of this size of the same (negative) sign required in the price forecasts. The PI for this projection is 0.10. The outcome for real growth and inflation projected by the administration is not itself extremely implausible; it is this outcome conjoined with the projected path of policy variables that conflicts with historical experience.

The experiments up to this point treat M1, interest rate, and fiscal variables as "policy variables." It is unlikely that policy authorities can set all four of these variables independently. There are many possibilities



one could explore assuming that some of the projections for these variables will have to be abandoned. As one example, one can ask what the model predicts if only the administration's money forecast must be fulfilled, and interest rates are used by the administration to achieve the money target. The results are shown in the last two rows of table 4. Persistently high interest rates are required, producing declining output and little improvement in inflation. The initial forecast disturbance (for 1982:1) to interest rates is 4.6 standard deviations, so this outcome appears highly unlikely. The fact that such persistent inflation is shown in the face of such a gloomy output forecast suggests that treating interest rate disturbances as determined by policy might be a mistake. Interest rate increases may at times have anticipated adverse supply shocks that raise the price level. The model may therefore associate them with more subsequent inflation than would actually occur if they were generated by policy action in the absence of a supply shock.

The projections for the time-series model in which either money growth or price and GNP growth were unconstrained, with the exception of table 4, show a slow upward drift in velocity—the sum of growth rates in output and prices is close to or slightly higher than the growth rate of the money stock. The administration's forecast requires that velocity grow much faster, by roughly 5.6 percent in 1982 and 6.4 percent in 1983. The forecast in which the slow money growth is generated by interest rate disturbances generates a rapid increase in velocity, mainly in 1982, but does so by an understandable mechanism—very high interest rates. That history gives no good reason to suppose one will find a rapid increase in velocity in a period of stable or declining interest rates has been pointed out before, for example, by the Congressional Budget Office.<sup>28</sup>

The analysis carried out here makes the case that the administration's forecast is implausible. It also gives some insight into what is implausible about it. But in the discussion the two central difficulties with such analysis have not been touched upon—the identification problem and the problem of extrapolating a model under unprecedented conditions. The four potential "policy variables" in the system, M1, interest rate, expenditures, and revenues, behave as if they are predetermined. There is little evidence of strong feedback from target variables to these four

28. Congressional Budget Office, *An Analysis of the President's Budgetary Proposals*.

variables.<sup>29</sup> But it has already been shown that policy variables, especially monetary policy variables, could behave this way even if they have moved mainly passively in the historical data.

If policy actions have in fact had little influence on M1 and interest rates historically, and if from 1982 on policy will exert strong influence on them, the projections of the model could be misleading. If historically these variables have been influenced both by policy and by other factors, the model will make conditional projections that average the “response” to policy actions and the response following other sources of movement in M1 and interest rate. However, in this case there would be great uncertainty surrounding the system’s response in the sample period, which should show up in large forecast error variances. If one concedes that policy-induced periods of tight and easy money are a substantial part of the historical record, the implausibility of the administration’s projection for real GNP and inflation given the policy variable projections should be convincing evidence that the proposed policy is not likely to have the intended effects.

The projections for policy variables are in themselves implausible, however. As discussed above, this raises questions about the reliability of the model in forecasting the consequences of policy actions putting these variables on the projected paths. The implausibility does not arise from the size of any one shock being outside the historically recorded range, but rather from sequences of shocks of the same sign being required. This is the kind of disturbance to policy about which rational expectations warns us to be suspicious. If periods of sustained below-prediction money growth have occurred before, one could in principle use a model that systematically adapts its forecasting mechanism to such periods. The model under consideration does not do this. It attempts to avoid large errors by adapting its linear forecasting structure to recent history. The administration’s long-term policy of steady reduction in money growth may not yet be reflected well in the model’s coefficients.

This being said, however, the question is whether there is reason to suspect anything systematically wrong with model projections—that is,

29. Such feedback would have showed up, for example, in a substantial difference between the disturbances in the first four columns of the top half of table 3 and those in the same four columns of a similar table (not shown) computed for the first two rows of the table 4 forecast. Such a table is not provided because it differs little from table 3 in the first four columns.

is there a priori knowledge that would allow one to correct the model? I doubt that there is any professional consensus on such knowledge, but there are some hypotheses among economists that suggest systematic corrections, most notably rational expectations combined with a monetarist view of business cycles. This view would suggest that sustained dampening of money growth will eventually be understood as a firm, long-run policy; that when it is understood, expectations of inflation, and thereby actual inflation and nominal interest rates, will fall without the need for mediating movements in real variables. As far as I can see, such a view does little to make the administration's forecast appear more plausible. A neoclassical, rational expectations world view can explain why nominal interest rates, inflation, and money growth could all decrease at the same time. It does not suggest that, with inflation and interest rates declining, velocity should increase; most models along that line would suggest the opposite, that to the extent that interest rates affect velocity, declines in interest rates should reduce velocity. If the administration believed that announcement of its monetary policy intentions would bring inflation down only at the moderate rate given in its forecast, it should not have projected such high real growth. A projection of much more rapid deceleration in inflation would have been even more implausible from the point of view of the model, but would have been defensible on the assumption of strong and rapid announcement effects.

The results give no support to the idea that the administration's projection represents an unprecedented fiscal policy. In fact, the top half of table 3 shows that the forecast disturbances in expenditures are all less than one standard error, and they do not form a sequence of values with the same sign. The projected revenues are higher than the model expects, even given the administration's real GNP forecast. The tax cuts thus represent no unusual fiscal stimulus.

The same issues of identification and reliability of the model under unusual policy paths cast doubt on the wisdom of attempting to aim at the policy variable paths shown in the third and fourth rows of table 4. It is doubtful that policy can hold interest rates down below 11 percent while the deficit explodes as in that table. The 1984 forecast (not shown) for that part of the table has inflation jumping back up to 8.5 percent and real growth dropping back to 4 percent. To keep inflation down and real growth up beyond 1983 would probably require a still more implausible and persistent set of forecast disturbances in policy, and there is good

reason to suppose that continued large deficits would refuel and accelerate inflation.

The forecast of table 2 seems to me both plausible and desirable: a moderately declining rate of growth in the money stock, flat interest rate path, revenues starting to catch up with expenditures in 1984, modest but fairly steady GNP growth, and steady and moderate inflation. Reasonable debate might surround the question of whether some more overt deflationary pressure would be desirable—by modest moves toward fiscal and monetary restriction that are different from what is shown in table 2. There is also the question of whether the substance of the expenditure increases justifies their cost. Such a reasonable debate, conditioned by historical experience, should be possible and in my view is not far from what was actually achieved with the sensible use of econometric models in much of the postwar period.

## **Conclusion**

The procedures used above are in a way only marginal modifications of the conventional use of econometric models in macroeconomic policymaking. They do involve making projections of the likely effects of various paths for policy variables and using the plausibility and desirability of those projections to evaluate the policy proposals, which is a common procedure. They differ from the usual procedures in two respects: they take account of policy endogeneity and they avoid constructing behavioral stories about each individual equation in the model.

They take account of policy endogeneity by generating true conditional projections, given specified paths for policy variables. The usual procedure assumes that the specified path for a policy variable is generated by disturbances to the policy equation, with all other disturbances held at zero. With endogenous policy, this is a potentially misleading way to generate projections.

The procedures also follow the lead of the rational expectations approach to macroeconomics because they use a statistical model that brings previous information to bear only on the system as a whole, not

by claiming to have useful knowledge that allows restrictions to be placed on the model one equation at a time.<sup>30</sup>

These technical differences from standard procedures are not the main point of the policy evaluation examples or of this paper. The main point is that careful statistical modeling of the historical probability structure can properly be used to make conditional projections that will be useful in policy analysis. The ambitious, probably unattainable goal of the rational expectations school—to identify parameters of behavior that would be invariant to unprecedented permanent changes in rule—should not condition an entire research plan. We should be improving our methods for estimating and using statistical models that do not require identifying such parameters. Most policy analysis does not require that kind of identification.

## APPENDIX

### *Notes to the Figures*

THE MODELS underlying the figures have the form

$$B_t(L)Y(t) = u(t),$$

where  $B_t$  is a matrix polynomial in the lag operator with lead coefficient  $B_{t0} = I$ , the identity matrix. The residual vector,  $u$ , is assumed to be uncorrelated with  $Y(s)$  for  $s < t$ . The model for evolution of  $B_t$  assumes that each of its coefficients is a random walk and that the best estimate of future  $B_t$  is the same as the best estimate of the current  $B_t$ . The dynamics implied by  $B_t$  are easier to understand by examining  $B_t^{-1}(L)$  than by looking at  $B_t(L)$  itself. To see why, suppose that at time  $t$  there is a forecast from the model of  $Y(t + s)$  for  $s > 0$  based on data through time  $t$ . One then asks what change in the forecast will result if some nonzero equation disturbance is assumed at  $t + 1$ ,  $u(t + 1) = v$ . It is not

30. This aspect of the rational expectations movement is given particular emphasis by Robert E. Lucas in the introduction to his book, *Studies in Business-Cycle Theory* (MIT Press, 1981), pp. 1–18.

hard to check that the modification in the forecast at  $t + s$  is  $G_{ts}v$ , where  $G_t(L) = B_t^{-1}(L)$ . More generally, given an arbitrary set of assumed values for the disturbances,  $u(t + s)$ ,  $s = 1, \dots$ , the implied modification of the forecast at  $t + s$  is given by  $G_t(L)u(t + s)$ , where in applying this formula one takes  $u(t + s)$  to be zero for nonpositive  $s$ . Thus the inverse of  $B_t(L)$  can be thought of as summarizing how the system responds to equation disturbances. The original  $B_t$  contains the same information but in less comprehensible form. Typically the elements of  $B_t$  are lag distributions with oscillatory or erratically varying weights whose implications for the system's dynamics are hard to unravel.

When  $B_t$  is fixed so the  $t$  subscript is unnecessary, all the system's dynamics consist of responses to equation shocks. When  $B_t$  is itself a stochastic process, the system also receives a shock by changes in  $B_t$ . With stochastically varying  $B_t$ , therefore, examination of  $B_t^{-1}$  can reveal only part of the system's dynamics. In the model estimated for this paper, the  $B_t$  are implied to vary slowly, so examination of  $B_t$  itself captures a large part of the dynamics.

Each panel of the figures shows the response to a single disturbance,  $u(t) = v$ , where  $v$  is one of the columns of the lower triangular factorization,  $W$ , of the variance-covariance matrix of  $u$ ,  $WW' = \text{Var}(u(t))$ , with the variables ordered in the opposite sequence from the way they appear in the figure—that is, with M1 at the top. Thus the charts can be thought of as displaying responses to shocks of typical size, with correlations among the  $u(t)$  being treated as generated by responses of elements of  $u(t)$  to disturbances higher in the ordering.

Somewhat inconsistently, the figures take  $W$  from the covariance matrix of the model's sample recursive residuals—that is, from the sample covariance matrix of actual errors one step ahead. Because this covariance matrix includes the effects of shocks to  $B_t$ , it is larger than any good estimate of the variance of  $u(t)$ , the equation disturbances. Use of the recursive residual covariance matrix may help bring the figures closer to representing full-system dynamics.

## *Comments and Discussion*

**Stephen M. Goldfeld:** Christopher Sims has presented an interesting paper that addresses the important issue of how one should conduct formal policy analysis. Sims observes that the rational expectations critique advanced by Robert Lucas and others has cast a pall over the use of econometric models in policy analysis. As one whose intellectual capital was in danger of technological obsolescence, I certainly find congenial Sims's punch line that the rational expectations critique should be viewed as a "cautionary footnote" to policy analysis. I also like Sims's paper for its thoughtful discussion of policy endogeneity, a topic that worried Alan Blinder and me some years ago.

Sims starts out by confronting the paradox of viewing policy as "random" at the same time that one thinks of it as purposeful, even, perhaps in some fairy-tale world, chosen in an optimal way. While acknowledging the valid elements of the Lucas approach, Sims points out that the rational expectations critique typically treats the choice of a policy rule as if it occurs *de novo* each time—that is, by failing to consider the nature of choices embodied in historical data, this line of argument ignores the issue of policy endogeneity.

As a footnote to Sims's argument, it has always seemed to me that there was a certain asymmetry to the Lucas critique. In particular, changes in policy rules hardly tend to occur at random times. One could well imagine that changes in rules are often brought about by changes in the "model," that is, the real world, rather than the other way around. The history of the gold standard provides a case in point. This sort of policy endogeneity suggests a need for rethinking the rational expectations critique.

This is precisely what Sims does. He concludes that policy rule

changes of the sort envisaged in the Lucas critique are rare events and that conventional econometric models can be used sensibly to analyze policy changes in a wide variety of cases. He does, however, have some unkind words to say about conventional models and some nice things to say about the use of rational expectations to improve existing models.

This is all so preposterously reasonable that I have little quarrel with it. Nevertheless, while I am delighted to have all this said, I wonder whether it will convince even a quasi-believer in the strong form of the rational expectations critique. It is, after all, an empirical question as to what is a sensible modeling strategy, and Sims provides only limited evidence to address this issue. Furthermore, much of the discussion seems to gloss over the distinction between estimation problems and policy analysis and I, for one, would have liked a bit of clarification on this score.

The second part of Sims's paper discusses the identification problem in models with rational expectations. As a general proposition, imposing rational expectations on the model seems to make things worse and, in one sense, this should not be very surprising. About twenty years ago the econometrician, T. C. Liu, argued that virtually all variables appearing in econometric models are endogenous so that models are generally underidentified. Although this view has been shown to be a bit misleading, it does have relevance for the present case. In a crude sense, replacing some ad hoc expectations mechanism, such as some distributed lag, by rational expectations might be thought of as endogenizing more things; therefore, this creates difficulties in identification.

Sims actually talks about the identification problem in a slightly different way. More specifically, he provides a number of clever examples of models in which there are variables that appear predetermined and have explanatory power for variables we care about such as GNP, and yet these predetermined variables are passive and have no policy significance. As he readily acknowledges, these examples suggest that caution is needed in using and interpreting tests for exogeneity—tests that Sims himself has pioneered. He illustrates these difficulties by examining two small multivariate vector autoregressive models. In one, the money stock appears to be causally prior while in the second model, which includes an interest rate, this result no longer holds.

Sims also hypothesizes that one view consistent with the data is that the Federal Reserve controls the interest rate directly and the money



stock indirectly. But because of the shape of the dynamic responses of the economy to policy, if the Federal Reserve attempts to track the money stock too closely, very large fluctuations in interest rates will result. There is, in fact, some empirical evidence suggesting that this is a quite plausible scenario. More particularly, Peter Tinsley has recently examined the relation between the variability of interest rates and the variability of the money stock around some long-term target path. As the Federal Reserve attempts to return more rapidly to the target path, Tinsley finds that interest rate variability steadily increases. Moreover, if the attempted gap closing is more than about 30 percent a month, actual monetary control is not improved so there is no compensation for greater interest rate variability.

The final part of the Sims paper uses the technique of vector autoregressions (VAR) to do policy analysis and to analyze the plausibility (or rather lack thereof) of the forecasts of the current administration. As many have observed, it requires a bit of new math to reconcile the various components of the administration's forecast. Among others, the forecast has a "velocity problem" because healthy growth in nominal GNP is coupled with an assumed stingy growth in the monetary aggregates. The implied increase in velocity is off the charts by historical standards, even ignoring the fact that the forecast also contains a marked decline in interest rates. By use of the VAR technique, Sims elaborately demonstrates these same difficulties. Because, as I have suggested, a few great ratios of economics should be adequate to cast serious doubts on the administration's forecast, for me the main virtue of this part of Sims's paper is that it demonstrates how one can use VAR for policy analysis. Nevertheless, even here a few questions remain.

As compared with conventional VAR applications, Sims advances the state of the art by allowing for drift in his parameter estimates. I would have guessed that this would be of some quantitative importance. One reason is that there seem to be some dimensional difficulties with the specifications in Sims's VAR analysis since he mixes both nominal and real variables. Because the coefficient of a nominal variable regressed on a real variable might be expected to reflect an upward movement in the price level, one would think this might show up in parameter drift. A second reason is that, although Sims cautions us against looking at a single equation, I could not help but anticipate that the well-documented instability in conventional money-demand equations would somehow

show up in Sims's estimates. Yet despite these expectations, Sims reports that parameter drift does not appear to be very consequential. I am not sure I fully understand or believe this result. Perhaps it is the case that the particular way Sims chooses to model changing parameters—namely, as a random walk—is not sufficiently rich to capture the kinds of changes which might have occurred. Just as Sims has cleverly illustrated in the second part of his paper, I suspect one can construct examples of changing parameters that are important and yet trick the Sims approach into thinking there is little change in the parameters. This strikes me as worthy of further investigation.

A second set of issues concerns the interpretation of the standardized disturbances from the VAR model that form the basis for assessing the plausibility of the administration forecast. One difficulty is that these estimates are not invariant to the ordering of the equations in the VAR model. This is a troublesome feature of the VAR approach, and it would be nice to know the sensitivity of the estimates to the causal ordering.

It can be argued that a related problem of interpretation of the standardized disturbances results from treating government spending and taxes as ordinary variables in the VAR approach. The tax equation strikes me as particularly problematical since it makes taxes a function of such variables as lagged real GNP and the lagged GNP deflator, but nowhere includes a variable for legislated tax rates. What this means is that we end up judging the implausibility of the administration's forecast partially on the basis of the implausibility of estimated tax revenues, but only given real GNP and the like. However, it could be argued that what is most implausible is the passage of the cuts in tax rates, and it seems strange to ignore the fact that this implausibility, has already taken place. Rather, it would seem that one should judge the plausibility of revenue estimates given tax rates. This strikes me as a good example of how the nonstructural VAR approach can run into logical problems in policy analysis. Be this as it may, however, Sims certainly provides a variety of evidence reinforcing the implausibility of the administration's forecasts.

One last point I would like to take up is first mentioned in the introduction to Sims's paper. He begins by suggesting that it is widely believed that econometric models badly misled policymakers by mischaracterizing the nature of the relation between inflation and unemployment—the Phillips curve. Later in the paper, Sims seems to argue

against this view, put forth for example by Lucas and Thomas Sargent, who assert that there has been a “spectacular failure” of econometric models. Having seen the follies of many econometric models at close hand, I would hardly go to the wall to defend model performance. Nevertheless, I believe this particular charge to be a bit unjust. Evidence on this point, at least for the MPS model, is contained in a recent paper by Albert Ando, who shows that the model is perfectly capable of tracing the historical pattern of wages and prices for the 1970s as well as it did in the 1960s.<sup>1</sup> This result is examined with various vintages of the model, and it suggests that those who wish to assert spectacular failure had best be a bit more specific.

In the spirit of sorting out paradoxes that motivate much of the Sims paper, I would like to conclude with a thought from an earlier pioneer in the field of rational expectations, Aldous Huxley, who observed: “that men do not learn much from the lessons of history is the most important of lessons history has to teach.”

**Jeffrey D. Sachs:** Christopher Sims has written an important and challenging paper that grapples with one of the central issues in empirical economics: the use of econometric models for policy simulation. In many ways it deals with the same issues taken up by Blinder and Goldfeld ten years ago.<sup>2</sup> With the very rapid development in recent years of time-series analysis, to which Sims himself has made major contributions, it is fitting that the Brookings panel reconsider this issue. Much of the past decade’s work in macroeconometrics has been devoted to showing how badly we can mislead ourselves when performing standard simulation analyses on macroeconomic models. It is noteworthy, therefore, that Sims ends with almost the same confidence that Blinder and Goldfeld did concerning the feasibility of policy simulation. Unfortunately he himself gives us enough reasons to doubt this confidence.

The central issue in policy simulation is the relation between correlation and causation. A macroeconometric model summarizes in a highly sophisticated way the historical correlations of economic time series. Correlation per se may be enough for forecasting, assuming that the

1. Albert Ando, “The ‘Failure’ of Keynesian Economics and the Phillips Curve in the 1970’s: A Fact or a Fiction?” (University of Pennsylvania, September 1981).

2. Alan S. Blinder and Stephen M. Goldfeld, “Some Implications of Endogenous Stabilization Policy,” *BPEA*, 3:1972, pp. 585–644.

future is like the past, and for this reason, the vector autoregression model studied by Sims is useful for forecasting even though it embodies little economic theory. Policy simulation, however, explicitly assumes that the future may be unlike the past, at least with regard to the dynamics of the policy variable itself. In that case we need to know whether the historical correlations will hold up with policy changes. Sims calls the model structural if the correlations are invariant to the policy changes under consideration. In this terminology policy simulations will be meaningful only if the estimated model is structural.

Consider, for example, the link between income,  $Y$ , and a monetary aggregate,  $M$ . It has been known for a long time that whether a correlation between  $Y$  and  $M$  will hold up to policy changes depends on how  $Y$  and  $M$  were generated in the sample period. If  $M$  was changed exogenously, say according to a roll of the dice, then the correlation would tend to be structural, because one would be observing randomized experiments; if  $M$  was changed *because* of  $Y$  or because of a common third factor that also affected  $Y$ , the relation is not likely to withstand policy changes. This much is well known. In 1921, Governor Strong of the Federal Reserve System of New York complained about the staff of his research department for using correlations to argue “that an increase in bank loans and currency is the inevitable cause of higher prices, [while they] . . . are unwilling to accept the view that sometimes bank loans and currency expand in response to prices, which arise from other causes than the ‘quantity’ of money . . . .” In 1970 James Tobin strengthened this case by showing that even if  $M$  *leads*  $Y$ —that is,  $M$  is correlated with future  $Y$ —no causality is implied. And Sims himself shows in this paper that even if  $M$  “Granger-causes”  $Y$ —that is, helps to predict  $Y$  given past  $Y$ —it need not cause  $Y$  in the policy-invariant sense. Rather,  $M$  may be correlated with unobservables (real productivity shocks in Sims’s case), which themselves cause  $Y$ .

To understand Sims’s defense of policy simulation it is useful to consider a specific example. Suppose that the econometrician, perhaps the research department of the Federal Reserve, estimates

$$(1) \quad Y = \hat{a}_1 Y_{t-1} + \hat{b}_0 M_t + \hat{b}_1 M_{t-1} + \hat{c}_1 Z_{t-1}.$$

Suppose further that no other lags of  $M$  or  $Y$  are statistically significant and that the equation appears to be well estimated by the usual standards of  $t$ -statistics, Durbin-Watson, and so forth. Is equation 1 structural in

the sense that a policy simulation would yield consistent estimates of  $Y$  conditional on  $M$ ? That depends on the true structure of the economy and on the process that generated  $M$  in the sample period. Suppose that 1 is of the correct functional form,

$$(2) \quad Y_t = a_1 Y_{t-1} + b_0 M_t + b_1 M_{t-1} + c_1 Z_{t-1} + u_t,$$

and that policy has been set by

$$(3) \quad M_t = \alpha_1 M_{t-1} + \beta_1 Y_{t-1} + \gamma_1 Z_{t-1} + v_t.$$

Estimates of 2 exist that will be structural in the sense used here under two well-known conditions. Either  $v_t$  is uncorrelated with  $u_t$ , or  $c_1 = 0$  and  $Z_{t-1}$  is uncorrelated with  $u_t$ . (In the latter case,  $Z_{t-1}$  is used as an instrumental variable.) The first condition says that history has given us pure money shocks or randomized experiments. The second condition states that even if the experiments are not random (because they are conditional on  $Z_t$ ), they are informative about the role of money because  $Z$  does not directly affect income.

Conventional structural models have traditionally made such assumptions on variables like  $Z$  (that is, exclusion restrictions). In Sims's vector autoregression approach such restrictions are frowned upon as arbitrary, and indeed they might be. But an even more astounding assumption is then necessarily made in order to use VAR to forecast the effect of policy actions: that  $v_t$  and  $u_t$  are uncorrelated. The approach is summarized as follows: all variables *in* the model are allowed to affect both  $M$  and  $Y$ ; all variables *outside* the model are captured by the error term and are allowed to affect  $M$  or  $Y$  only in an arbitrary specified way. A terrific weight is carried by the residual if one tries to argue for structural identification in a VAR.

This assumption leads to the following conundrum, which Sims discusses at length in the first part of the paper: if  $M$  is set purposefully by monetary authorities, presumably with some goals in mind, how can we claim that  $v_t$  is in fact a purely random outcome? The answer that ultimately emerges is that we cannot. In the second part of the paper Sims presents a model that can be interpreted as showing that  $v_t$  and  $u_t$  will be correlated if the monetary authority has had interest rate targets over the sample period. He offers strong evidence that they have in fact had such targets when he shows that  $i$  helps to predict, or Granger-cause,  $M$ , but not vice versa. The Blinder-Goldfeld study focused on the case

in which  $v_t$  and  $u_t$  are in fact correlated because of conscious stabilization policy. They showed the important and paradoxical result that the better the stabilization, the smaller would be the estimated effect of  $M$  on  $Y$  in 1.

Sims argues that changes in the policymaker's goals, say because of alternating political administrations, can give the independent variance in  $M$  needed for statistical identification. This argument is only partially correct, because it is fair to assume that the alternating policymakers are still responding to interest rates or to other variables, even though they have different feedback rules. Also, the change in political administration is at least partially correlated with economic events.

On the point of structural invariance then, my reading is that Sims is unconvinced in his solution to this vexing problem.

He is more convincing on a second point, however. Suppose that instead of 2, the true income process is given by a Lucas-Barro supply function:

$$(4) \quad Y_t = A_1 Y_{t-1} + B_0(M_t - M_t^e) + u_t,$$

where  $M^e$  is expected money, so that income fluctuates according to unanticipated monetary shocks. Assuming that agents know the economic structure in 4 and the money-supply process of 3, one may combine 3 and 4 to obtain 4':

$$(4') \quad Y_t = (A_1 - B_0\beta_1)Y_{t-1} - B_0\alpha_1 M_{t-1} + B_0 M_t - B_0\gamma_1 Z_{t-1} + u_t.$$

It can be immediately seen that the links between  $Y$  and  $M$  depend on the monetary rule itself. Although  $Y_t$  is in fact independent of the money-supply rule, a naive estimation of 4' will suggest otherwise. The reduced-form estimate, 1, will say nothing useful about the effects of a change in money-supply rule (that is, a change in  $\alpha_1$ ,  $\beta_1$ ,  $\gamma_1$ ) on income. This problem is the famous Lucas critique.

Sims offers convincing reasons for doubting the overwhelming relevance of this concern: policy shifts are likely to be short-lived given the political system; many policy announcements are never carried out and are likely never believed; large policy changes are rare, tinkering is the norm; true policies are never known, so that agents are always inevitably smoothing their estimates using recent history. And practically speaking, because discrete changes in regime are not clearly perceived, the econometric task in accounting for regime changes is formidable, and in fact has never been carried out.

But Sims then oversteps his case by arguing that actual policymaking is therefore merely selecting the errors (such as  $v_t$ ) in a stable policy rule such as 3. Policy changes *do* occur, as the current monetary policy illustrates. And even if policy changes are short-lived, they are probably not representable by a *random* sequence of drawings on  $v_t$  at quarterly intervals. Operationally, I would wager Sims that out-of-sample errors on the policy equations in his VAR display strong serial correlation in many subperiods, reflecting short-run changes in regime. Sims is certainly correct that certain aspects of economic behavior will be almost invariant to policy change if the change is perceived as short run and reversible. (Wage-setting behavior and wage outcomes seem to be one example.) But other aspects of behavior, particularly in financial markets such as the foreign exchange market, will react quickly to even short-run changes. Generally speaking, the Lucas critique will apply when there is a large role for expectations, and when technological or institutional factors do not weigh heavily against a change in operating procedures.

While many may share Sims's judgment call here, I doubt that one may usefully confront the challenge of the Lucas critique by *a priori* judgments. Sims's empirical evidence on structural stability that I discuss below is useful in this regard. But more important, the proposition should be tested straight on. Have widely perceived policy changes, such as the switches from fixed to flexible exchange rates or from interest rate to money targets, affected *other* structural relations in the predicted way? This debate simply cannot be settled by assuming that there has been a single policy regime.

In the last section of the paper, Sims's preferred methodological approach is put on display. He modifies a standard vector autoregression to allow for drift in the coefficients and tests for the plausibility of the Reagan administration's macro program. It is ingenious econometric work. Sims seems to rely on two principles: first, it is usually plausible to assume that policy rules will not change, even despite policy pronouncements to the contrary; and second, since there is little reason to expect a shift in structure, one may as well use a VAR for forecasting, since it parsimoniously describes the relevant historical experience.

To jump to Sims's conclusion, the administration forecasts look different from the past—indeed, precisely what the administration has been arguing all along. Sims cleverly formalizes this notion by asking what shocks across equations are needed to achieve the forecasts.

Presumably by studying the pattern of shocks one can determine whether they might plausibly arise from a rule change of the sort that President Reagan has proposed.

A few questions can be raised about the specific example in this section.

First, one does not get a good feel, unfortunately, for the differences in forecast variance introduced by the drifting parameter estimates. How robust is Sims's finding of little drift? Are the shifts in the "policy" equation any bigger than in the rest of the model, for example? How confident can one be that there has not been a change in the rate of drift?

Second, Sims's method of detecting drift may indicate little change in coefficients over time even when major policy changes have occurred, if those changes are not well characterized by a random drift of the policy coefficients. Suppose that the policy rule alternated regularly between Democratic and Republican administrations. The random walk model might best be summarized by low drift even if policy is regularly bouncing between two poles. The "errors" from the policy equation estimated in this case would display strong serial correlation. It is conceivable that agents' behavior might change according to the administration in power.

Finally, the shifts in coefficients are by design uncorrelated with the relevant economic time series, and with each other. There is no feedback from economic outcomes to policy, or from policy to structure. There are extremely strong assumptions that may reduce the evidence of structural drift or its importance. The Lucas critique holds that drift in the policy equations should be correlated with drift elsewhere in the model. Sims might offer some evidence on this point.

### **General Discussion**

Many discussants agreed with Sims's criticism of the Lucas critique. James Duesenberry suggested that the domain in which the critique applies is quite limited. First, only a fraction of policy developments could be interpreted as changes that, even theoretically, should actually cause parameters in the model to vary. Furthermore, there are many economic relations in which expectations about macroeconomic policy variables are simply not that important. To someone selling hot dogs at a ball game, it might be worth a couple of bucks to buy a weather forecast, but almost nothing to buy a money-supply forecast. This



situation characterizes most individual equations in large models. Finally, even when expectations about prices or output may be important to consumption or investment decisions, those expectations are far removed from forecasts of the money supply. Charles Holt made several related points in the spirit of Sims's paper. He noted that policy actions contain a great deal of randomness, like treatments assigned in an agricultural experiment, and that this randomness should help in determining the rest of the structure. He also argued that policy rules are mainly economists' abstractions of a complicated reality and are quite foreign to economic agents. If policy rules are obscure, so are changes in them; the impact of changes on expectational relations must consequently be very slight.

Lawrence Klein disagreed with the rational expectations view that the parameters in structural equations would be altered as a result of policy changes. Even if people's method of forming expectations changed, the primary implication is that expectations variables would have to be adjusted, not that the other parameters of the system would vary. Stanley Fischer disagreed with Klein. In Fischer's view, some parameters in many macro models vary with policy changes. He argued that this requires more complicated structural equations in which the effects of policy rules on agents' behavior are allowed for explicitly.

Some of the discussants commented on Sims's analysis of macroeconomic modeling and identification. One point emphasized in his paper is that there are rarely abrupt or complete changes in policy rules. John Taylor thought this emphasis was misplaced with respect to the usefulness of policy analysis. Economists still need to provide guidance to policymakers about the effects of various policy rules and the mechanisms for changing these rules. It is possible to use statistical methods that focus on policy systems and their effects even if dramatic changes have not occurred in the past. Sims cautioned that it would be difficult to apply such methods when observed policy movements could simply represent random drift.

Klein disagreed with Sims's emphasis on optimal control exercises as a test of the adequacy of conventional structural macro models. He argued that when the models are actually used in policy applications, alternative discrete policies are explored simply to provide guidance for moving in an advantageous direction. He also disputed Sims's claim that the conventional models are predicting more poorly now than they did

in the early 1970s. After some deterioration in 1973–74, the normalized errors are no worse now than they were before 1973. Klein also disputed the view that vector autoregressions are to be preferred to conventional models for simple forecasting. Referring to work by Stephen McNees, Klein maintained that vector autoregressive (VAR) models are all right for predictions one quarter ahead, but VAR predictions quickly deteriorate so that conventional models offer superior predictions further in the future. Sims responded that if one reestimates the VAR models quarter by quarter and projects ahead two to six quarters, the VAR predictions are approximately as reliable as conventional forecasts.

William Poole questioned whether policy regimes are as endogenously determined as Sims's paper suggests. Pointing to the shift from fixed to flexible exchange rates in 1971, Poole argued that if such regime changes are endogenous, there is no chance in principle of using data in nonexperimental science to draw conclusions about how the world works under alternative regimes. But Sims said that it is one point of this paper to show that even endogenous policy shifts can have effects which, in theory, can be identified.

There was spirited discussion of Sims's estimates of the plausibility of the Reagan administration's economic forecast. George von Furstenberg thought those forecasts should not be taken seriously in the paper. According to von Furstenberg, they were dictated by political considerations without regard for the likelihood or internal consistency of the predictions. Charles Schultze raised a more fundamental objection. He doubted that Sims's VAR methodology could be applied to assessing the plausibility of the administration's forecast. Sims estimated a system in which policy is endogenous—that is, determined by previous values of variables in the system and by random error. But policy may not be generated by the previous decision rule; the present administration would certainly contend as much. Consequently, the outcomes predictable solely on the basis of the historical record might not apply. Given that the decision rule is now significantly changed, what is the probability that the forecast will turn out to be accurate? This question is inherently unanswerable using the VAR assumptions. In the present case, a forecast that incorporates the information that tax rates have in fact been radically changed cannot be evaluated by a VAR approach in which tax rates are themselves implicitly a forecast based on past tax rates.