ROBERT E. LUCAS, JR.*

Methods and Problems in Business Cycle Theory

1. INTRODUCTION

One of the functions of theoretical economics is to provide fully articulated, artificial economic systems that can serve as laboratories in which policies that would be prohibitively expensive to experiment with in actual economies can be tested out at much lower cost. To serve this function well, it is essential that the artificial "model" economy be distinguished as sharply as possible in discussion from actual economies. Insofar as there is confusion between statements of opinion as to the way we believe actual economies would react to particular policies and statements of verifiable fact as to how the model will react, the theory is not being effectively used to help us to see which opinions about the behavior of actual economies are accurate and which are not. This is the sense in which insistence on the "realism" of an economic model subverts its potential usefulness in thinking about reality. Any model that is well enough articulated to give clear answers to the questions we put to it will necessarily be artificial, abstract, patently "unreal."

At the same time, not all well-articulated models will be equally useful. Though we are interested in models because we believe they may help us to understand matters about which we are currently ignorant, we need to test them as useful

*This paper was prepared for the American Enterprise Institute Seminar on Rational Expectations, held on February 1, 1980, in Washington, D.C. Many of the ideas in it were developed under the stimulus of Don Patinkin’s course on the History of Monetary Thought, taught at the University of Chicago, winter 1979. Allan Drazen, Sherwin Rosen, and Nasser Saidi provided very helpful criticism of an earlier draft. A version of this paper entitled "Economic Policy and the Business Cycle" was presented at The Ohio State University as the Money, Credit, and Banking Lecture, on May 8, 1980.

ROBERT E. LUCAS, JR. is professor of economics, University of Chicago.

© 1980 American Enterprise Institute for Public Policy Research
JOURNAL OF MONEY, CREDIT, AND BANKING, vol. 12, no. 4 (November 1980, Part 2)
imitations of reality by subjecting them to shocks for which we are fairly certain how actual economies, or parts of economies, would react. The more dimensions on which the model mimics the answers actual economies give to simple questions, the more we trust its answers to harder questions. This is the sense in which more "realism" in a model is clearly preferred to less.

On this general view of the nature of economic theory then, a "theory" is not a collection of assertions about the behavior of the actual economy but rather an explicit set of instructions for building a parallel or analogue system—a mechanical, imitation economy. A "good" model, from this point of view, will not be exactly more "real" than a poor one, but will provide better imitations. Of course, what one means by a "better imitation" will depend on the particular questions to which one wishes answers.  

In this paper I wish to review some recent developments in business cycle theory, taking the point of view suggested above. On this view, one would expect developments to arise from two quite different kinds of forces outside the subdisciplines of monetary economics or business cycle theory. Of these forces the most important, I believe, in this area and in economics generally, consists of purely technical developments that enlarge our abilities to construct analogue economies. Here I would include both improvements in mathematical methods and improvements in computational capacity. The neglect in traditional history of doctrine of this force for change in our thinking is a serious omission, and contributes to the common but mistaken sense that everything has been said before or "it's all in Marshall." Marshall's world was enough like ours and Marshall was an astute enough observer of his world that it is difficult to make general observations about our economy which do not have close precedent in some Marshallian observation about his. Our ability to construct analogue economies is, however, much greater, so that we have the capacity to study in detail market interactions about which Marshall could only conjecture.

The second source of theoretical developments is changes in the questions we want models to answer, or in the phenomena we wish to understand or explain. To the journalist, each year brings unprecedented new phenomena, calling for unprecedented new theories (where "theory" amounts to a description of the new phenomena together with the assertion that they are new). Since there is an obvious sense in which this view carries some truth, I will not attempt to refute it. I have argued elsewhere [23] that it is in our interest to take exactly the opposite viewpoint in the study of business cycles, or as close to an opposite view as we can get by with, and will maintain this attitude below. The Great Depression, however, remains a formidable barrier to a completely unbending application of the view that business cycles are all alike.

There is, of course, a third source of developments in our understanding of business cycles: the activities of economists specializing in the field. Making the con-

---

1 I do not know the background of this view of theory as physical analogue, nor do I have a clear idea as to how widely shared it is among economists. An immediate ancestor of my condensed statement is [43].
nections between the technical innovations discovered by our colleagues and the difficult questions thrown at us by the real world is no easy task, and I do not wish to minimize the importance of these efforts. Yet I do think that both amateur and professional historians have tended to go too far in attempting to understand developments in monetary economics in terms entirely internal to the subdiscipline, and that some leaning in the other direction may therefore be useful.

In the next section of the paper, I will review what seem to me the main features of the Keynesian Revolution from the view set out above. There is, I know, a growing feeling that such skeleton rattling is becoming tiresome and old hat. Yet the advancement of theoretical constructs from this era as though they were facts to be explained continues to be a standard mode of debate in macroeconomics, suggesting the existence of still more deeply buried bones. The remainder of the paper will be an attempt to diagnose more recent developments, and perhaps even to extrapolate a little way into the future.

2. BUSINESS CYCLE THEORY THROUGH KEYNES

Business cycle theory (as distinct from monetary economics) is mainly a twentieth century product. For the most part, the major nineteenth century economists set short-term fluctuations to one side in order to focus attention on other issues. The general underlying idea (one might call it a "natural rate hypothesis") must have been that one could understand the main determinants of the average levels and rates of growth in economic activity without understanding the fluctuations.

The task of making precise exactly which aspects of economic life had been so set aside by this fruitful nineteenth century strategy was undertaken at the beginning of the present century by Wesley Mitchell. Mitchell sought an empirical definition of business cycles through the systematic exclusion of those movements in economic time series that appeared likely to be explicable by then existing theory: the general level and pattern of growth in economic activity, and movements in individual series that seemed to arise from supply or demand conditions specific to individual markets. Neither of these exclusion principles is without ambiguity, so Mitchell proceeded cautiously and experimentally, trying a wide variety of data-summarizing techniques on the broadest available collection of series.

For two reasons, it is easy to forget the remarkable character of the regularities that Mitchell succeeded in discovering and documenting. On the one hand, we have lived with them for so long that they seem not so much the product of an imaginative and abstract way of organizing economic time series as simply "facts" that "everyone knows." On the other, they are regularities that, from the point of view most widely adopted since the 1930s, are not especially noteworthy or suggestive. The central finding, of course, was the similarity of all peacetime cycles with one another, once variation in duration was controlled for, in the sense that each cycle exhibits about the same pattern of co-movements among variables as do the others.

Not surprisingly, these empirical findings (guided theoretically in the sense I have indicated), corroborated by other evidence and a variety of less systematically or-
ganized impressions, stimulated a great deal of theoretical work. The idea that one could, with a firm empirical basis, speak of something like a “typical business cycle,” divided into stages invariant in character (if not in duration) suggested that a substantial part of observed fluctuations might be explainable at a fairly abstract, or “simple” level, with a single theoretical explanation of the (i.e., of all) business cycle(s). This is the organizing principle Gottfried Haberler employed, for example, in *Prosperity and Depression* [11]. That so wide a variety of theories as discussed in this book could be fit with such ease into so tight a logical pattern is testimony, I think, to the power of the abstraction in the idea of a single or typical business cycle.

Another example of pre-Keynesian business cycle theory (here I suppose I am inexcusably abusing the language), which is useful for two of my purposes, is John Maynard Keynes’s *Treatise on Money* [16]. Comparison of this work with the *General Theory* [17] is useful both in illustrating the way that limits on our technical ability to construct explicit theory limit our ability to think productively about phenomena, and in illustrating the extent to which thinking in monetary economics is subject to outside or “real world” shocks.

The *Treatise* is sometimes described (and was even by Keynes’s) as a kind of initial, clumsy groping toward the ideas presented in the *General Theory*. I suppose there must be some psychological truth to this view, but emphasizing this aspect of the *Treatise* leads to a very strained reading of what seems to me a fairly straightforward example of pre-Depression thinking on business cycles. The main objective of the book is to try to understand fluctuations in economic activity about a secular trend in which real magnitudes are determined by the real considerations of neoclassical value theory and in which nominal prices are governed by the quantity theory of money. Keynes was convinced, correctly I believe, that attempting to discuss fluctuations in terms of fluctuations in velocity would not be productive and sought instead a point of view that stressed changes in the composition of expenditures over the cycle. He then showed that this point of view could be reconciled with a quantity-theoretic view of longer-term changes.

The accounting system, or set of notational conventions, embedded in his “fundamental equations” was designed to facilitate this reconciliation, and served this function well enough. Beyond this, however, Keynes’s apparatus could not go. He states an identity and discusses it; then he moves a term from the left to the right and discusses the result of that operation; he defines a new variable in terms of previously defined ones and discusses that; then he talks about the identity restated in terms of this new variable; and so on and on. The problem is not that the underlying ideas are trivial, though the algebra certainly is. On the contrary, the book deals in an intelligent way with the fundamental problems business cycles raise. The difficulty is that Keynes has no apparatus for dealing with these problems. Though he discusses them verbally about as well as his contemporaries, neither he nor anyone else was well enough equipped technically to move the discussion to a sharper or more productive level.

---

1In the preface to the *General Theory*, pp. vi–vii.
The onset of the Great Depression did nothing to improve Keynes’s equipment for understanding the business cycle, viewed as a recurrent sequence of booms and depressions. Instead, it permitted him to reformulate the problem itself as one of accounting for the level of output and employment at a point in time, as opposed to one of accounting for a particular pattern repeated in the time series. So reformulated, the problem could productively be studied simply by discarding an equation of static equilibrium theory (the labor supply curve) in contrast to the much more difficult task, undertaken in the Treatise, of supplementing this static theory with suitable short-run dynamics. This simpler problem was one on which progress could be made at the Marshallian level of analysis on which Keynes was a master.3

In reading the General Theory in this way, I am of course simply following the classic exegesis of John Hicks [14] (as well as Hicks’s initial view of the General Theory as “slump economics”) and Franco Modigliani’s [29] pioneering step toward a “neoclassical synthesis.” There is, certainly, much of interest in the General Theory that is not captured either in Hicks’s diagram or Modigliani’s equation system, a fact that led Axel Leijonhuvud (and others, perhaps even Hicks and Modigliani) to view the “Keynesian economics,” which was later based mainly on these early interpretations, as a kind of vulgarization of the General Theory. While there is some truth, forcefully developed in Leijonhuvud’s monograph [18], in this view, it misses what I believe to be the more essential truth, stressed in my introduction, that progress in economic thinking means getting better and better abstract, analogue economic models, not better verbal observations about the world.

The General Theory is, to be sure, a mine of acute and well-phrased remarks about the trials of conducting one’s affairs in an uncertain world. Perhaps there are some, though I would not like to have the task of documenting this, that are both central to business cycle behavior and not prefigured in, say, Mitchell [28], or even a century before in Henry Thornton’s writings. Certainly the likely, central role of profit expectations and investment behavior was stressed by virtually every economist who devoted more than superficial thought to business cycles. Economists who find Keynes’s style congenial will continue to use his writings as Dennis Robertson did Lewis Carroll’s, but surely there is more to the cumulative nature of economics than this!

To extract from the General Theory a simple graphical method for thinking about national income determination is not, I believe, to vulgarize its contribution. Vulgarity in economics would more appropriately be defined as criticizing or caricaturing an abstract (and hence potentially useful) model because it leaves something out.

3. THE NEOCLASSICAL SYNTHESIS

The General Theory was, if the argument of the preceding section is accurate, more successful than the Treatise not because of theoretical advances within mon-
etary economics but because the event of the Great Depression permitted Keynes to restate the problem posed by the business cycle into a form such that the theoretical methods at his disposal admitted genuine progress. It was a fortunate historical accident that at about the same time, and for reasons unrelated to contemporaneous economic events, technical advances in statistical and economic theory occurred, which transformed “Keynesian economics” into something very different from, and much more fruitful than, anything Keynes himself had foreseen.

One of these advances, which one may date from the work of Tinbergen [47] or perhaps from Slutsky [44], was the idea that one might describe an economy as a system of stochastically disturbed difference equations, the parameters of which could be estimated from actual time series. Indeed, in describing a “business cycle theory” as a “fully articulated imitation economy,” I have presupposed the objective of these early econometricians to have by now become common property. The rapid progress of the econometric models toward something that appeared close to imitative perfection was, I have argued elsewhere [22], illusory. Sargent and I [26] have subsequently argued in some detail that useful analogue systems will not be found among variations within the class of these pioneering econometric models. The issues touched on in these earlier papers still seem to me at the center of the question of what we can hope for from a theory of business cycles, but there is no reason to take them up again here.

The second of these advances, like the first, has roots the disentangling of which would challenge an objective scholar’s career. Instead, I will take a subjective course, identifying this development with Paul Samuelson’s Foundations of Economic Analysis [36] and advising the more serious historian to pursue the copious references in that text. In so doing, I follow Don Patinkin’s practice in his Money, Interest, and Prices [32], perhaps the most refined and influential version of what I mean by the term “neoclassical synthesis.” Samuelson advanced, in the first place, the main ingredients for a mathematically explicit theory of general equilibrium: an artificial system in which households and firms jointly solve explicit, “static,” maximum problems, taking prices as parametrically given. He took for granted that such equilibria were nonvacuous in the mathematical sense, a supposition that was later confirmed under fairly broad conditions. Such equilibria can be shown to be equivalent to Pareto-optimal resource allocations.

I refer to this theory as “static” following Samuelson, despite the ambiguity involved in specifying an empirical counterpart to this modifier. The underlying idea seems to be taken from physics, as referring to a system “at rest.” In economics, I suppose such a static general equilibrium corresponds to a prediction as to how an economy would behave should external shocks remain fixed over a long period, so that households and firms would adjust to facing the same set of prices over and over again and attune their behavior accordingly. It is not difficult to think of modifying this idea of “rest” to accommodate slow and fairly predictable secular changes, and this accommodation will be taken for granted in what follows, as it seems to have been in the monetary economics literature.

Now economies experiencing recurrent business cycles are quite evidently not “at rest” so that static general equilibrium theory, though a genuine model in the
sense of being explicit and complete, is not a good imitation of reality for the purpose of understanding these events. To deal with this disparity, the Foundations offered a solution too (though there advanced, it seems to me, as an answer to an entirely different kind of question'). Samuelson proposed a dynamic model of price adjustment in which the rates of change of prices offered in each market were related to the level of "excess demands" in all markets. Whatever the history or underlying objectives of this model of price dynamics (and, implicitly, of quantity dynamics) this theory introduced sufficient additional (to those needed to describe tastes and technology) parameters to the equilibrium system so that, given an initial shock to the system, a wide variety of paths were consistent with its eventual return to equilibrium.

This introduction of additional (to those used to describe preferences and technology) free parameters held out the promise that one could construct a theoretical system the stationary point of which was a general equilibrium in the neoclassical sense but whose movements, out of equilibrium, might replicate the "Keynesian" behavior captured so well by the econometric models. Spelling out these connections became the main program of theoretical research in macroeconomics from the 1940s through the 1960s. The task seemed to involve somehow motivating the introduction of "money" and other financial elements into "real" general equilibrium theory, modifying the purely theoretical system in the direction of the applied econometric theory, and simultaneously reworking the structural equations of the econometric models so as to clarify their theoretical underpinnings. The objective of the enterprise was widely agreed to be "unification" of the two types of theories into which Keynesian ideas were translated in the 1930s and 1940s.

Reviewing the vast amount of useful economics that came out of this attempt at unification would be much too ambitious a task for this paper. Instead, I will make a few general remarks on the attitudes toward stabilization policy that were fostered by the neoclassical synthesis. First, since the synthesis was formed by the addition of free parameters to a static general equilibrium system, the general class of models it suggested admitted a wide variety of possibilities for business cycle behavior. Thus it seemed a framework open enough to contain virtually any point of view toward policy as a special case, an attractive feature to the nondogmatic. I suspect this is one reason why those economists, like Milton Friedman, who made no use of this framework were treated with some impatience by its proponents. Why could he not simply specify which particular parameter values corresponded to the case he believed to fit the facts, let others do likewise, and then the matter could be handed over to the econometricians for a definitive resolution?

Second, since fluctuations about the system's equilibrium represented disequilibrium behavior, standard welfare propositions could be applied only to the average

---

"Samuelson's correspondence principle proposed the use of his stability theory as a criterion to aid in deciding which stationary equilibrium points might actually be observed, and which not: "How many times has the reader seen an egg standing on its end?" Here the idea is clearly to decide which static egg-equilibria are empirically interesting, not to offer an empirically useful dynamic model of rolling or wobbling eggs. Indeed, Gordon and Hynes's [10] criticism of the use of Samuelsonian disequilibrium price dynamics as a description of observed price paths received central support from [37].
behavior of the system, and not to fluctuations about the average. This left one free to apply other criteria in evaluating stabilization policies: “gaps” instead of “triangles,” as James Tobin [50] puts it. The general idea was to use policy tools to keep the actual path of the system “close” in one sense or another to its equilibrium path. Proponents of various stabilization policies were thus free of the burden under which the ordinary welfare economist labors—of “justifying” intervention by some specific “market failure” and tailoring the nature of the intervention to the nature of the failure. Under the neoclassical synthesis, the business cycle was defined to be market failure, and any policy that promised to move the system toward “full employment equilibrium” was viewed as an improvement.

So widely endorsed was the general idea of the neoclassical synthesis described above that its central constructs have become a common shorthand for describing “the facts.” Now when the success, actual and potential, of this synthesis is once again at issue, these constructs no longer facilitate discussion, but rather get in the way. Two examples will illustrate what I mean.

The first is from James Tobin [50], who in summarizing what he calls “the central propositions of the General Theory” begins with: “In modern industrial capitalist societies, prices and wages respond slowly to excess demand or supply, especially slowly to excess supply. Over a long short run, ups and downs of demand register in output; they are far from completely absorbed in prices.” He goes on to cite evidence for this proposition from British economic behavior in the 1920s and the U.S. Great Depression.

What I take Tobin to mean by this complex proposition is something like the following. If one were to try to interpret the British experience of the 1920s, America’s of the 1930s, and business cycles generally, in terms of a static general equilibrium (allowing for secular trend) and dynamic adjustment in prices of the sort described by Samuelson, then we would need to postulate “slow” (say, half-life of many quarters) coefficients describing wage and price responses to excess supply. Qualified in this way, the statement seems to me a true one. Yet qualified in this way, Tobin’s argument cannot shed light on the desirability of attempting to account for business cycles within the framework provided by the neoclassical synthesis as opposed to using some other framework, which is the intended subject of his paper.

In a similar vein, Franco Modigliani [30] characterizes Thomas Sargent’s [38] econometric model of the U.S. as one in which employment declines can be accounted for only by “severe attack[s] of contagious laziness.” He goes on to say that “equally serious objections apply to Friedman’s modeling of the commodity market as a perfectly competitive one . . . and to his treatment of labor as a homogeneous commodity traded in an auction market, so that, at the going wage there is never any excess demand by firms or excess supply by workers.”

As in Tobin in [50], Modigliani is here thinking of a competitive equilibrium in the static sense of the neoclassical synthesis, so that an equilibrium decrease in employment would necessarily have to arise either from a spontaneous technology or taste shift (“an attack of laziness”) and so that a model in which labor markets are continuously cleared is patently in contradiction with observed employment and
unemployment fluctuations. Later on, he refers to the "substantial agreement that in the United States the Hicksian mechanism [his overly modest term for what I am here calling the neoclassical synthesis] is fairly effective in limiting the effect of shocks and that the response of wages and prices to excess demand and supply will also work gradually toward eliminating largely, if not totally, any effect on employment." "These inferences are supported by simulations with econometric models like the MPS."

Now both Modigliani and Tobin are, in the papers from which I have quoted, explicitly defending (via the time-honored tactic of counterattack) their preferred framework for studying business cycles against "monetarist" or "rational expectations" alternatives. Yet both take the correctness of the framework they defend as given and use it as a point of departure in criticizing alternatives. The failure of a simulation of Sargent's econometric model to reproduce the results of simulations of the MPS model is viewed as evidence that Sargent's model does not "fit the facts!" It seems clear that debate at this level cannot advance matters.

4. RECAPITULATION AND ASSESSMENT

It will be useful at this stage to attempt a summary of the theses advanced in the sections above. I began by sketching a view of an economic theory (by which I mean a theory that purports to account for specific observed and as yet unobserved aspects of behavior) as a mechanical, analogue economy. It follows from this view that developments in a particular substantive field, such as the study of business cycles, will be influenced strongly by technical developments in our ability to construct explicit model economies, and by real world developments that alter our view as to the questions we think such models can and ought to be able to help us to answer.

Viewed from this perspective, the main features of the Keynesian Revolution and the neoclassical synthesis into which it evolved in the United States seemed to be the following. They include, in the first instance, the onset of the Great Depression and the consequent shift of attention from explaining a recurrent pattern of ups and downs to explaining an economy apparently stuck in an interminable down. Keynes's *General Theory* is then seen as, first, a recognition of the importance of this change of circumstances, and second, at least as read by Hicks, Modigliani, and others, as the proposal for a simple aggregative account of output and employment determination at a point in time. Developments in British Keynesian macroeconomics since the 1930s give, I think, a reasonably accurate view of what would have become of the revolution had these elements and no others been involved.

In the United States and on the continent, two other elements were involved in essential ways, both of a technical character. One was the development of explicit stochastic descriptions of economic systems. The other was the development of a static general equilibrium theory together with an associated theory of disequilibrium price dynamics. These elements were rapidly combined to provide rigor and clarity to Keynes's account of short-term equilibrium determination, and to add to this theory explicit dynamic elements, which permitted it to fit actual time series in a
fairly literal way. Moreover, they held out a definite promise for additional “unification,” a task that occupied some of the profession’s best talent for three decades.

Toward the close of the 1960s, this orderly progress toward unity was disturbed by two theoretical developments. One of these, Milton Friedman’s presidential address to the American Economic Association [8], was written from the “monetarist” viewpoint, which had continued to pursue the study of business cycles along the line initiated by Mitchell. The other, Edmund Phelps’s [33] and the subsequent “Phelps volume” [34] seemed initially an attempt to complete the unity promised by the neoclassical synthesis through discovery of a microeconomic foundation to the labor market and product pricing side of the standard models. However differently motivated, the papers of Friedman and Phelps both carried the clear implication that “excess demand” was neither necessary nor sufficient for price or wage inflation, and that any average inflation rate was consistent theoretically with any level of unemployment. This conclusion, arrived at via impeccable neoclassical reasoning, conflicted with the prediction of a real output-inflation trade-off, which was at the center of all models based on the neoclassical synthesis.

In attempts to formalize the Friedman-Phelps natural rate hypothesis, it was soon discovered that then-conventional ways of modeling expectations-formation were both central to the issues involved and fundamentally defective. John Muth’s [31] hypothesis of rational expectations, formulated originally to deal with an entirely different set of substantive questions, turned out to be a natural way to formalize the Friedman-Phelps arguments. Subsequent research in macroeconomics has revealed the sweeping implications of this hypothesis, and the extent to which it proves subversive of the main positive and policy presumptions underlying the neoclassical synthesis.

At the present time, these developments retain enough novelty to make them difficult to assess with detachment. Yet I believe it is possible, indeed necessary, to attempt to understand them in terms similar to those I have used in trying to understand earlier developments in monetary economics and business cycle theory. If it was in fact the conjunction of real world shifts in the questions to which people wanted answers together with technical improvements in economic theory which led to the major rethinking of business cycle theory that I have been here calling the neoclassical synthesis, then it is not unlikely that more recent developments can be similarly attributed to forces to these two categories.

The real world event from the recent past which first comes to mind is the combination of inflation with higher than average unemployment that characterized the 1970s. While consistent with the Friedman-Phelps logic, these events were badly misforecast with 1960s vintage econometric models. To what extent this forecast error should be interpreted as a “fatal” error in models based on the neoclassical synthesis or simply as one suggesting some modifications is not so easy to determine. The idea that virtually all of this period was characterized by “excess supply” and hence that virtually all of the inflation must be attributable to “supply shocks” does not seem to be worth taking seriously and I have yet to see a quantitative case

\[5\text{An argument for fatality is in [26].}\]
for this position made. (This is, of course, not to say that there have not been serious supply shocks over the decade.) On the other hand, some recent models stressing the role of contractually fixed nominal prices permit prices to respond to secular or “anticipated” demand without “excess demand” necessarily ever emerging, while retaining a “short-term” role for “excess demands” and supplies. Perhaps these may be interpreted as an attempt to reconcile the experiences of the 1970s with some key features of the neoclassical synthesis. In short, events of the 1970s have been provocative, but perhaps not decisive.

A less spectacular but perhaps ultimately more influential feature of post-World War II time series has been the return to a pattern of recurrent, roughly similar “cycles” in Mitchell’s sense. If the magnitude of the Great Depression dealt a serious blow to the idea of the business cycle as a repeated occurrence of the “same” event, the postwar experience has to some degree restored respectability to this idea. If the Depression continues, in some respects, to defy explanation by existing economic analysis (as I believe it does), perhaps it is gradually succumbing to the Law of Large Numbers.

These new observations have been influential (as new observations should be to empirical researchers), but it seems to me the main outside influences have been, and will continue to be, changes in available theoretical methods. In business cycle theory, it appears not to be the problem that changes but rather the way we look at it. Of changes in methods, certainly the most central have been postwar developments in general equilibrium theory.

5. POSTWAR GENERAL EQUILIBRIUM THEORY

The general equilibrium theory, which Modigliani, Patinkin, and others hoped to integrate with an operational business cycle theory, did not remain frozen in the form it had assumed in the 1930s and 1940s. Indeed, much of what has since developed was sketched out in some detail by John Hicks, in Value and Capital [15]. There Hicks proposed reinterpreting the maximum problems solved by firms and households as involving choices over sequences of dated goods, with choices of specific future goods interpreted as plans and with their prices interpreted as price expectations. Hicks noted that if future goods were viewed as contracted for in advance, the prices of future goods would simply be known numbers, and general equilibrium in a dynamic economy, so modeled, would be equivalent formally to equilibrium in a “static” model. Hicks believed that the presence of uncertainty in real situations rendered a model stressing forward contracts inapplicable in most dynamic situations of real interest, and so put most of his emphasis on a discussion of a sequence of “spot” equilibria.

*See [7, 35, 45]. One way to interpret these models is as attempts to modify the price dynamics of the neoclassical synthesis so as to permit monetary expansion to share at least some of the blame for the 1970s inflation with OPEC, etc. (and for inflation in Argentina, Chile, and innumerable other examples in which high inflation has been associated with subtrend real output and employment). Another way is discussed in note 14.
Kenneth Arrow [2] and Gerard Debreu [5] observed that uncertainty could be incorporated into “static” general equilibrium theory by exactly the same device that Hicks proposed to incorporate the passage of time: namely by indexing goods both by the date on which they are to be exchanged and by the (perhaps stochastically selected) “state of nature” contingent on which the exchange is to occur. Like Hicks’s, the innovation was not initially an extension of general equilibrium theory in a mathematical sense, but rather the observation that the range of applicability of this body of theory could be vastly broadened by some ingenuity in specifying what is meant by a commodity.

One way to interpret a “contingent claim” equilibrium is as a description of an economy in which all state-contingent prices are determined in advance, in the clearing of a single grand futures market. On this interpretation, individual traders may assess the probabilities of the occurrence of future states of nature, but with prices determined in advance, the issue of price expectations does not arise. Alternatively, one may sometimes (though certainly not always) think of a contingent-claim equilibrium as being determined via a sequence of “spot” markets, in which current prices are set given certain expectations about future prices. On this second interpretation, one needs a principle to reconcile the price distributions implied by the market equilibrium with the distributions used by agents to form their own views of the future. John Muth [31] noted that the general principle of the absence of rents in competitive equilibrium carried the particular implication that these distributions could not differ in a systematic way. His term for this latter hypothesis was rational expectations.7

As originally proposed by Arrow and Debreu, this contingent-claim interpretation of a competitive equilibrium model took all information to be simultaneously and freely available to all traders, and many important results (e.g., the extension of the main theorems of welfare economics to uncertain environments) are crucially dependent on this assumption. It was soon recognized by many researchers that the idea of viewing a commodity as a function of stochastically determined shocks is an invaluable one also in situations in which information differs in various ways among traders. Indeed, it is this idea that permits one to use economic theory to make precise what one means by information, and to determine how it is valued economically.

When originally proposed, the contingent-claim formulation tended to be viewed as highly esoteric and remote from practice, no doubt because it arose at the most abstract end of the discipline. However, this formulation rapidly and easily absorbed and clarified a variety of special results in the economics of uncertainty, and facilitated their unification and extension. It is now in standard use in virtually every applied field of economics. It is also in use, though I would not as yet say “standard,” in business cycle theory.

7Muth formulated the hypothesis of rational expectations using the Simon [42]-Theil [46] idea of certainty-equivalence though his introductory discussion makes it clear that the idea is applicable in situations where certainty-equivalence may not be. Its basic logic is not difficult to restate within the Arrow-Debreu contingent claim framework: see [24] for one example.
Would it not be surprising if this were not so? The idea that speculative elements play a key role in business cycles, that these events seem to involve agents reacting to imperfect signals in a way which, after the fact, appears inappropriate, has (as I remarked in section 3) been a commonplace in the verbal tradition of business cycle theory at least since Mitchell [28]. Now for the first time we have at our disposal methods for constructing artificial model economic systems in which these elements play a well-defined role. It is now entirely practical to view price and quantity paths that follow complicated stochastic processes as equilibrium "points" in an appropriately specified space. This is a development that will make a difference in the way we think.

To ask why the monetary theorists of the 1940s did not make use of the contingent-claim view of equilibrium is, it seems to me, like asking why Hannibal did not use tanks against the Romans instead of elephants. There is no reason to view our ability to think as being any less limited by available technology than is our ability to act (if, indeed, this distinction can be defended). The historical reason for modeling price dynamics as responses to static excess demands goes no deeper than the observation that the theorists of that time did not know any other way to do it. It is, of course, conceivable that theorists in full command of newer methods will nonetheless conclude that there are sound reasons for continuing to use the older technology for some purposes (just as there are some purposes for which we continue to prefer elephants to tanks). This seems to me most unlikely if the purpose is to understand business cycles. That a powerful model-building apparatus specifically designed to help us deal with problems involving choice under uncertainty should simply be passed over in favor of an older apparatus, which is (for the most part) incapable of taking these problems into account,\(^8\) would, should it occur, certainly refute rather decisively the point of view toward the development of economic thought that I have advanced in this paper.

6. FUTURE DEVELOPMENTS

Even if one is persuaded, as am I, that the theoretical advances sketched in the last section will have a major impact on our thinking about business cycles, it is obviously impossible to forecast with any accuracy the form this impact will take. What is already clear, however, is that certain characteristics of earlier theories, adopted originally (I have argued) for convenience, are no longer analytically necessary. In particular, it is possible to construct systems in competitive equilibrium, in a contingent-claim sense, which exhibit a vast variety of dynamic behavior. The idea that an economic system in equilibrium is in any sense "at rest" is simply an anachronism.

For a modern theoretical economist, surely the most natural way to read the

\(^8\) I do not mean to suggest that considerations of choice under uncertainty played no role in the neoclassical synthesis. Both portfolio theory and inventory theory were used, for example, to motivate particular hypotheses about demands for money and other assets [4, 48, 49]. Nevertheless, the concept of market equilibrium used was that of static, deterministic general equilibrium theory.
original Friedman and Phelps articles is as attempts to conjecture some of the properties that a successful general-equilibrium model of business cycles will be likely to possess. Indeed, Friedman refers to the natural rate of unemployment as “the level that would be ground out by the Walrasian system of general equilibrium equations, provided there is imbedded in them the actual structural characteristics of the labor and commodity markets” though he is not able to put such a system down on paper. In his introduction to [34], Phelps sketches a specific general equilibrium system in some detail. Much of the research that has been done since may be interpreted as efforts to cast these ideas more explicitly into the contingent-claim framework.

In recent years, a number of economists have worked to develop what I prefer to call equilibrium models of business cycles. These are models that utilize the contingent claim point of view described in the last section in an essential way, and in which prices and quantities are taken to be always in equilibrium. In these models, the concepts of excess demands and supplies play no observational role and are identified with no observed magnitudes. In contrast to the static equilibrium models available in the 1940s, equilibrium models of this new class seem to do about as well in fitting time series as do models based on the neoclassical synthesis.

Now it is clear that these new equilibrium models can, in principle, be “synthesized” with a Samuelson-like model of disequilibrium price adjustment just as could the older equilibrium models. This must be true of any equilibrium model. Moreover, since a synthesis of this kind involves the addition to the model of free parameters, the synthesized version cannot fit facts worse than the original equilibrium version on which it is based. One seems to be led, then, not to equilibrium models as a class, but to a vastly larger class of synthesized disequilibrium models. Now I am attracted to the view that it is useful, in a general way, to be hostile toward theorists bearing free parameters, so that I am sympathetic to the idea of simply capitalizing this opinion and calling it a Principle. In evaluating economic theories that claim to be useful in guiding policy in the way sketched in my introductory section, however, there are important substantive considerations supporting such an attitude of hostility, which seem to me to put it on a sounder basis that can be afforded by a general prejudice in favor of parsimony. I will try to spell these considerations out.

Our task as I see it (to restate my introduction somewhat more bluntly and operationally) is to write a FORTRAN program that will accept specific economic policy rules as “input” and will generate as “output” statistics describing the operating characteristics of time series we care about, which are predicted to result

---

9Some early examples are [20, 21, 41, 38, 3]. Of these, only [20] utilizes the contingent-claim general equilibrium formalism throughout. The others use linear approximations, motivated informally by reference to the contingent-claim models. It seems clear that econometrically operational models will necessarily have to rely on liberal use of approximate, linear methods.

10I have in mind the line of research described by Malinvaud in [27] and recently surveyed by Drazen in [6]. The lesson that two models may be “close” in the sense of fitting the same data about equally well yet have radically different implications for policy was brought home forcefully in [41] and elaborated in a more general way in [39].
from these policies. For example, one would like to know what average rate of unemployment would have prevailed since World War II in the United States had M₁ grown at 4 percent per year during this period, other policies being as they were. (One would like to know the answers to a lot of other questions, too, but this came to mind first, and gives a concrete idea of where this argument is headed.) It must be taken for granted, it seems clear, that simply attempting various policies that may be proposed on actual economies and watching the outcome must not be taken as a serious solution method: Social Experiments on the grand scale may be instructive and admirable, but they are best admired at a distance. The idea, if the marginal social product of economics is positive, must be to gain some confidence that the component parts of the program are in some sense reliable prior to running it at the expense of our neighbors.

How is confidence of this sort earned? This is a question on the answer to which economists are fairly well agreed, yet I cannot recall where I have seen the nature of this agreement articulated. The central idea is that individual responses can be documented relatively cheaply, occasionally by direct experimentation, but more commonly by means of the vast number of well-documented instances of individual reactions to well-specified environmental changes made available “naturally” via censuses, panels, other surveys, and the (inappropriately maligned as “casual empiricism”) method of keeping one’s eyes open. Without such means of documenting patterns of behavior, it seems clear that the FORTRAN program proposed above cannot be written. Suppose, on the contrary, that such means are available, or that we have some ability to predict how individual behavior will respond to specified changes. How, if at all, can such knowledge be translated into knowledge of the way an entire society is likely to react to changes in its environment?

To be more concrete, consider the question: How will a monkey that has not been fed for a day react to a banana tossed into its cage? I take it we have sufficient previously established knowledge about the behavior of monkeys to make this prediction with some confidence. Now alter the question to: How will five monkeys that have not been fed for a day react to one banana thrown into their cage? This is an entirely different question, on which the knowledge of preferences (each monkey wants as much of the banana as he can get) and technology (banana consumption in total cannot exceed unity) gives us scarcely a beginning. We clearly need to know something about the way a group of monkeys interacts, in addition to their individual preferences, in order to have any hope of progress on this complicated question.

People interested in the way groups of monkeys solve problems of allocating scarce resources satisfy their curiosity by assembling groups of monkeys and tossing them scarce resources. I have taken it as given that we economists cannot proceed in this way, yet the allocation of scarce resources is something we are admired for being experts at. Economics is sometimes characterized as the working out of the implications of the idea that individuals pursue their self-interest, yet we have just observed how empty an idea this is, applied in isolation to even the most trivial animal experiment. Can we imagine that it gains power in some mysterious way when applied to human societies with millions of participants?
The ingredient omitted so far is, of course, competition. Let us take our banana, cut it into five pieces, give each of the five monkeys one piece, and impose on them the rule they may interact only by exchanging banana pieces for minutes of backscratching, at some fixed rate. (I confess to having no idea how this imposition might be effected in practice.) Now in this situation, and given sufficient information as to how individual monkeys are willing to trade-off backscratching and banana eating, we can predict the outcome of this interaction (equilibrium price and quantities exchanged), at least given sufficient computational ability. Notice that, having specified the rules by which interaction occurs in detail, and in a way that introduces no free parameters, the ability to predict individual behavior is nonexperimentally transformed into the ability to predict group behavior.\textsuperscript{11}

I have emphasized that it is the hypothesis of competitive equilibrium which permits group behavior to be predicted from knowledge of individual preferences and technology without the addition of any free parameters. This needs further illustration, in contexts closer to our interests than animal experiments. Employment and nominal wages are, in an immediate sense, determined by some very complicated labor market interactions involving employees and employers. It is possible, we know, to mimic the aggregate outcome of this interaction fairly well in a competitive equilibrium way, in which wages and manhours are generated by the interaction of "representative" households and firms.\textsuperscript{12} The parameters in this model describe either households' willingness to substitute goods and leisure contemporaneously and intertemporally or the technology available to firms.

It is also possible, in the manner of the neoclassical synthesis, to fit a quite different model—a Phillips curve—to these same aggregate outcomes. Here one also uses a parametric description (different, of course, in this case) of preferences and technology and, in addition, a parameter describing the speed with which an "auctioneer" adjusts the nominal wage to excess demands and supplies. Now the introduction of a fictional auctioneer is not a defect of this second way of looking at things, relative to the first. All models are fictitious according to the viewpoint I am taking, and in any case, an auctioneer is presupposed in the first model too, but one operating so rapidly that he is not noticed. Nor can it be a disadvantage of this second way of modeling wage and employment determination that it cannot fit data as well as the first: the addition of a free parameter cannot hurt in this sense.

\textsuperscript{11}This is a case for the use of explicit game theory in general, not for the use of competitive theory in particular. The case for the use of competitive theory in modeling business cycles would, if I were to develop it here, be based entirely on convenience, or on the limits imposed on us by available technology for working out the implications of other equilibrium definitions.

This qualification to the text must also qualify the claim that equilibrium theory requires the introduction of no free parameters other than those used to describe individual tastes and technology. That is to say, insofar as there is considerable latitude as to which equilibrium concept is being used, one can think of selecting one of them as the fixing of a "free parameter." This does not seem to me a serious issue in practice at the present time, but one can imagine it becoming so through further developments in theories of noncompetitive games.

This observation that noncompetitive games may someday prove to be of use in business cycle theory, which seems difficult to quarrel with, is sometimes used to rationalize wholly arbitrary models unrelated to any well defined game. I hope it is clear that this note is not intended as a defense of this practice.

\textsuperscript{12}See [25] though parts of this study need updating. A modern replication of this study would utilize ideas in [40, 12].
The disadvantage of the second model is this: there is no way to obtain information on the rate at which the wage is adjusted, this added parameter, except by observing the entire system in operation. If this parameter changes in reaction to changes elsewhere in the system (as we know that in fact it does), there is no way to predict the nature of these responses short of experimenting with the system as a whole. Yet it is precisely the attempt to avoid having to do this that leads us to use economic theory in the first place.

In the case of the equilibrium account of wage and employment determination, parameters describing the degree of intertemporal substitutability do the job (in an empirical sense) of the parameter describing auctioneer behavior in the Phillips curve model. On these parameters, we have a wealth of inexpensively available data from census cohort information, from panel data describing the reactions of individual households to a variety of changing market conditions, and so forth. In principle (and perhaps before too long, in practice, for there is a good deal of very promising research going on on just this topic\(^{13}\)) these crucial parameters can be estimated independently from individual as well as aggregate data. If so, we will know what the aggregate parameters mean, we will understand them in a sense that disequilibrium adjustment parameters will never be understood. This is exactly why we care about the "microeconomic foundations" of aggregate theories.\(^{14}\)

Researchers familiar with current work alluded to in the preceding paragraph will appreciate the extent to which it describes hopes for the future, not past accomplishments. These hopes might, without strain, be described as hopes for a kind of unification, not dissimilar in spirit from the hope for unification which informed the neoclassical synthesis. What I have tried to do above is to stress the empirical (as opposed to the aesthetic) character of these hopes, to try to understand how such quantitative evidence about behavior as we may reasonably expect to obtain in society as it now exists might conceivably be transformed into quantitative information about the behavior of imagined societies, different in important ways from any which have ever existed. This may seem an intimidatingly ambitious way to state the goal of an applied subfield of a marginally respectable science, but is there a less ambitious way of describing the goal of business cycle theory?

\(^{13}\)For example, [9, 13, 19, 1]. I hasten to add that these and related studies were not motivated or intended as support or confirmation of Rapping's and my results, nor is at all clear at this stage that they are leading to estimates consistent with ours. My point is simply that equilibrium aggregate models carry implications for a variety of other kinds of data, raising the possibility of independent confirmation (and contradiction) of estimates from aggregative time series.

\(^{14}\)Do the models cited in note 6 constitute a third class, intermediate to those just discussed? I do not believe so. If contract length in these papers is viewed simply as a free parameter (as note 6 interprets them) then this parameter is as unintelligible (in the sense of this paragraph) as that describing the rapidity of an auctioneer's adjustment. If, on the other hand, contract length is viewed as emerging from a decision problem solved by agents then these models, so elaborated, would be equilibrium models (with a different commodity space than those discussed above) and would not necessarily serve to reinforce the point of view toward policy taken by the neoclassical synthesis.

Whether this observation should be taken as severe criticism of contract-based models depends, of course, on one's views as to the likelihood of our being able to account for business cycles using no free adjustment parameters. Certainly this question must be regarded as open at present, and however it may be resolved, it must surely be the case that models with one or two free adjustment parameters represent analytical progress over models with dozens, or hundreds, of them.
7. CONCLUDING REMARKS

This paper has been an attempt to understand and clarify the nature and origins of some recent developments in business cycle theory. In taking an historical point of view, it may be that I have inadvertently adopted the Marxist tactic of describing what I would like to have happen as something that History has already ordained. I have, certainly, emphasized what seems to me the decisive importance of improvements in the analytical equipment we have at our disposal, but with the intent of stressing the expansion of our opportunities these improvements offer, not particular directions they dictate. For this reason, I have tried to avoid claiming too much for the particular examples of equilibrium models that now exist. There is no point in letting tentative and, I hope, promising first steps harden into positions that must be defended at all cost.

If an historical approach cannot guarantee an ability to foresee the future, it does seem to me to aid in distinguishing those elements in past thinking that remain useful from those that do not. The neoclassical synthesis arose, as does all useful economics, from a compromise between what we would like to have known and what the methods at our disposal seemed to make it possible to know. Nothing could be more detrimental to the productive use of methods more recently developed than to view the categories and constructs that were produced by this compromise as constraints on the way we think about business cycles today.

LITERATURE CITED


50. ———. “How Dead is Keynes?” *Economic Inquiry*, 16 (1977), 459–68.