UNDERSTANDING BUSINESS CYCLES*

Robert E. Lucas, Jr.
University of Chicago

I.

Why is it that, in capitalist economies, aggregate variables undergo repeated fluctuations about trend, all of essentially the same character? Prior to Keynes' General Theory, the resolution of this question was regarded as one of the main outstanding challenges to economic research, and attempts to meet this challenge were called business cycle theory. Moreover, among the interwar business cycle theorists, there was wide agreement as to what it would mean to solve this problem. To cite Hayek, as a leading example:

> [T]he incorporation of cyclical phenomena into
> the system of economic equilibrium theory, with
> which they are in apparent contradiction, remains
> the crucial problem of Trade Cycle Theory;[1]

> By 'equilibrium theory' we here primarily under-
> stand the modern theory of the general inter-
> dependence of all economic quantities, which has
> been most perfectly expressed by the Lausanne
> School of theoretical economics.[2]

A primary consequence of the Keynesian Revolution was the redirection of research effort away from this question onto the apparently simpler question of the determination of output at a point in time, taking history as given.[3] A secondary consequence of this Revolution, due more to Tinbergen than to Keynes, was a rapid increase in the level of precision and explicitness with which aggregate economic theories were formulated. As a result, Keynesian macro-

---

[1]Hayek (1933), p. 33n.

[2]Hayek (1933), p. 42n.

[3]This redirection was conscious and explicit on Keynes' part. See, for example, the first sentence of his chapter on the trade cycle. "Since we claim to have shown in the preceding chapters what determines the volume of employment at any time, it follows, if we are right, that our theory must be capable of explaining the phenomena of the trade cycle " (1936), p. 313.

economics has benefited from several decades of methodological improvement whereas, from this technical point of view, the efforts of the business cycle theorists appear hopelessly outdated.

Yet from another point of view, they seem quite modern. The observation that macroeconomics is in need of a microeconomic foundation has become commonplace, and though there is much confusion about the nature of this need and about what it would mean to satisfy it, it is likely that many modern economists would have no difficulty accepting Hayek's statement of the problem as roughly equivalent to their own. Whether or not this is so, I wish in this essay to argue that it should be so, or that the most rapid progress toward a coherent and useful aggregate economic theory will result from the acceptance of the problem statement as advanced by the business cycle theorists, and not from further attempts to refine the jerry-built structures to which Keynesian macroeconomics has led us.

Honoring one's intellectual ancestors is a worthwhile aim in itself, but there is a more immediate reason for interpreting the contemporary search for a theoretically sound aggregative economics as a resumption of the work of pre-Keynesian theorists. Accompanying the redirection of scientific interest occasioned by the Keynesian Revolution was a sharp change in the nature of the contribution to policy which economists hoped to offer and which the public has come largely to accept. The effort to "explain business cycles" had been directed at identifying institutional sources of instability, with the hope that, once understood, these sources could be removed or their influence mitigated by appropriate institutional changes. The process envisaged was the painfully slow one of public discussion and legislative reform; on the other side, there was the hope of long-term or "permanent" institutional improvement. The abandonment of the effort to explain business cycles accompanied a belief that policy could effect immediate, or very short-term, movement of the economy from an undesirable current state, however arrived at, to a better state.

The belief that this latter objective is attainable, and that the attempt to come closer to achieving it is the only legitimate task of research in aggregate economics is so widespread that argument to the contrary is viewed as "destructive," a willful attempt to make life more difficult for one's colleagues who are only trying to improve the lot of mankind. Yet the situation is symmetric. If the business cycle theorists were correct, the short-term manipulation on which much of aggregative economics is now focused only diverts attention from discussion of stabilization policies which might actually be effective; such postponement is, moreover, accompanied by the steady and entirely understandable erosion in the belief on the part of noneconomists that

aggregative economics has anything useful to say.

In the next section, I will review some of the main qualitative features of the events we call business cycles, and then turn to the Keynesian response to these facts, to the progress made along the line Keynes and Tinbergen initiated, and finally to the severe limits to this progress which have now become apparent. The remainder of the essay will consider the prospects of accounting for cyclical phenomena by an economic theory, in the narrow sense in which Hayek and other business cycle theorists have used that term.

II.

Let me begin to sharpen the discussion by reviewing the main qualitative features of economic time series which we call "the business cycle." Technically, movements about trend in gross national product in any country can be well described by a stochastically disturbed difference equation of very low order. These movements do not exhibit uniformity of either period or amplitude, which is to say, they do not resemble the deterministic wave motions which sometimes arise in the natural sciences. Those regularities which are observed are in the co-movements among different aggregative time series.

The principal among these are the following.[4] (i) Output movements across broadly defined sectors move together. (In Mitchell's terminology, they exhibit high conformity; in modern time series language, they have high coherence.) (ii) Production of producer and consumer durables exhibits much greater amplitude than does the production of nondurables. (iii) Production and prices of agricultural goods and natural resources have lower than average conformity. (iv) Business profits show high conformity and much greater amplitude than other series. (v) Prices generally are procyclical. (vi) Short-term interest rates are procyclical; long-term rates slightly so. (vii) Monetary aggregates and velocity measures are procyclical.

---

[4]The features of economic time series listed here are, curiously, both "well known" and expensive to document in any careful and comprehensive way. A useful, substantively oriented introduction is given by Mitchell (1951), who summarizes mainly interwar, U.S. experience. The basic technical reference for these methods is Burns and Mitchell (1946). U.S. monetary experience is best displayed in Friedman and Schwartz (1963). An invaluable source for earlier British series is Gayer, Rostow, and Schwartz (1953), esp. Vol. II. The phenomena documented in these sources are, of course, much more widely observed. Most can be inferred, though with some difficulty, from the estimated structure of modern econometric models.
An important recent contribution is Sargent and Sims (1976), which summarizes postwar U.S. quarterly series in several suggestive ways, leading to a qualitative picture very close to that provided by Mitchell, but within an explicit stochastic framework, so that their results are replicatable and criticizable at a level at which Mitchell's are not.

There is, as far as I know, no need to qualify these observations by restricting them to particular countries or time periods: they appear to be regularities common to all decentralized market economies. Though there is absolutely no theoretical reason to anticipate it, one is led by the facts to conclude that, with respect to the qualitative behavior of co-movements among series, business cycles are all alike. To theoretically inclined economists, this conclusion should be attractive and challenging, for it suggests the possibility of a unified explanation of business cycles, grounded in the general laws governing market economies, rather than in political or institutional character-istics specific to particular countries or periods.

I have omitted the behavior of foreign trade statistics from the above catalogue of phenomena-to-be-explained, in part because, for a large economy like the U.S., trade statistics do not exhibit high enough conformity to be cyclically interesting. For a smaller country, to be sure, export movements would do much to "explain" cycles, but to focus on open-economy explanations would, I think, beg the more difficult and crucial question of the ultimate origins of cyclical movements.

Also omitted, but too striking a phenomenon to pass over without com-ment, is the general reduction in amplitude of all series in the twenty-five years following World War II. At this purely descriptive level, it is impossible to distinguish good luck from good policy. Nevertheless, so long a period of relative stability strongly suggests that there is nothing inherent in the workings of market economies which requires living with the level of instability we are now experiencing, or to which we were subject in the pre-World War II years. That is, attempts to document and account for regular cyclical movements need not be connected in any way to a presumption that such movements are an inevitable feature of capitalist economies.

III.

The implications of Keynesian macroeconomic models conform well to the time series features reviewed above. Early versions (for example, by Hicks, 1937, and Modigliani, 1944) fit well qualitatively; the econometric models which developed from this theory and from Tinbergen's largely independent early work[5] conform well quantitatively. These models located the primary

---

[5]For example, see Tinbergen (1939). This work was not explicitly Keynesian; indeed, it was conceived as an empirical complement to Haberler's review and synthesis of theoretical work on business cycles (1936). Keynes, on his part, was actively hostile toward Tinbergen's work. See Moggridge (1973), pp. 285-320. In referring to those who built in part on Tinbergen's work as "Keynesian" I am, then, contributing to the continuation of an historical injustice.

disturbances in investment behavior, linked via lags (in Tinbergen's U. S. model) to the highly volatile profit series. Movements in these high-amplitude series then induce general movements in output and employment. Since these disturbances were, in Hicks' terms, "IS shifts," they were consistent with procyclically moving interest rates and velocity. The assumption of rigid wages and prices was a good empirical first approximation. Later on, a wage-price sector (still later called a Phillips curve) was added to fit observed procyclical wage and price movements.[6]

In this description, movements in money play no important role in accounting for cycles. This feature certainly did not result directly from the theoretical models; Keynes, Hicks, and Modigliani all gave great emphasis to monetary forces. The de-emphasis on money was on empirical grounds: econometricians from Tinbergen on discovered that monetary factors did not seem very important empirically.[7]

The empirical success of these developments was measured in an original and historically apt way by Adelman and Adelman (1959) in their simulation of the Klein-Goldberger model of the U. S. economy. The Adelmans posed, in a precise way, the question of whether an observer armed with the methods of Burns and Mitchell (1946) could distinguish between a collection of economic series generated artificially by a computer programmed to follow the Klein-Goldberger equations and the analogous series generated by an actual economy. The answer, to the evident surprise of the Adelmans (and, one suspects, of Klein and Goldberger, who had in no way directed their efforts to meeting this criterion) was no.[8]

This achievement signaled a new standard for what it means to understand business cycles. One exhibits understanding of business cycles by constructing a model in the most literal sense: a fully articulated artificial economy which behaves through time so as to imitate closely the time series behavior of actual economics. The Keynesian macroeconomic models were the first to attain this level of explicitness and empirical accuracy; by doing so, they altered the meaning of the term "theory" to such an extent that the older business cycle theories could not really be viewed as "theories" at all.

These models are not, however, "equilibrium theories" in Hayek's sense. Indeed, Keynes chose to begin the General Theory with the declaration (for

---

[6]Klein and Goldberger (1955).

[7]Tinbergen (1939), pp. 183-185. Tinbergen, as did most subsequent macroeconometricians, used the significance of interest rates to test the importance of money.

[8]It is not correct that a search for "good fits" would have led to a model satisfying the Adelmans' criteria; think of fitting polynomials in time to "explain" each series over the sample period.

Chapter II is no more than this) that an equilibrium theory was unattainable: that unemployment was not explainable as a consequence of individual choices and that the failure of wages to move as predicted by the classical theory was to be treated as due to forces beyond the power of economic theory to illuminate.

Keynes wrote as though the "involuntary" nature of unemployment were verifiable by direct observation, as though one could somehow look at a market and verify directly whether it is in equilibrium or not. Nevertheless, there were serious empirical reasons behind this choice, for nowhere is the "apparent contradiction" between "cyclical phenomena" and "economic equilibrium" theory sharper than in labor market behavior. Why, in the face of moderately fluctuating nominal wages and prices, should households choose to supply labor at sharply irregular rates through time? Most business cycle theorists had avoided this crucial problem, and those who addressed it had not resolved it. Keynes saw that by simply sidestepping this problem with the unexplained postulate of rigid nominal prices, an otherwise classical model could be transformed into a model which did a fair job of accounting for observed time series.

This decision on the part of the most prestigious theorist of his day freed a generation of economists from the discipline imposed by equilibrium theory, and, as I have described, this freedom was rapidly and fruitfully exploited by macroeconometricians. Now in possession of detailed, quantitatively accurate replicas of the actual economy, economists appeared to have an inexpensive means to evaluate various proposed economic policy measures. It seemed legitimate to treat policy recommendations which emerged from this procedure as though they had been experimentally tested, even if such policies had never been attempted in any actual economy.

Yet the ability of a model to imitate actual behavior in the way tested by the Adelmans (1959) has almost nothing to do with its ability to make accurate conditional forecasts, to answer questions of the form: how would behavior have differed had certain policies been different in specified ways? This ability requires invariance of the structure of the model under policy variations of the type being studied. Invariance of parameters in an economic model is not, of course, a property which can be assured in advance, but it seems reasonable to hope that neither tastes nor technology vary systematically with variations in countercyclical policies. In contrast, agents' decision rules will in general change with changes in the environment. An equilibrium model is, by definition, constructed so as to predict how agents with stable tastes and technology will choose to respond to a new situation. Any disequilibrium model, constructed by simply codifying the decision rules which agents have found it useful to use over some previous sample period, without explaining why these rules were used, will be of no use in predicting the consequences of nontrivial policy changes.

12

The quantitative importance of this problem is, of course, a matter to be settled by examination of specific relationships in specific models. I have argued elsewhere[9] that it is of fatal importance in virtually all sectors of modern macroeconomic models, primarily because of the faulty treatment of expectations in these models. Rather than review these arguments in detail, let me cite the most graphic illustration: our experience during the recent "stagflation."

As recently as 1970, the major U. S. econometric models implied that expansionary monetary and fiscal policies leading to a sustained inflation of about 4 percent per annum would lead also to sustained unemployment rates of less than 4 percent, or about a full percentage point lower than unemployment has averaged during any long period of U. S. history.[10] These forecasts were widely endorsed by many economists not themselves closely involved in econometric forecasting. Earlier, Friedman (1968) and Phelps (1968) had argued, purely on the basis of the observation that equilibrium behavior is invariant under the units change represented by sustained inflation, that no sustained decrease in unemployment would result from sustained inflation. In this instance, the policy experiment in question was, most unfortunately, carried out, and its outcome is now too clear to require detailed review.

It is important that the lesson of this episode not be lost. The issue is much deeper than the addition of a few new variables to econometric Phillips curves (though this is the only revision in macroeconomic models which has followed from it), as Friedman made clear in his Presidential Address. Friedman's argument did not proceed on the basis of a specific aggregative model, with a better "wage-price sector" than the standard models. On the contrary, it was based on a general characteristic of economic equilibrium: the zero-degree homogeneity of demand and supply functions. Thus, without using any very specific model, and without claiming the ability to forecast in any detail the initial response of the economy to an inflation, one can, in the case of sustained inflation, reason that, if the unemployment rate prior to the inflation were an equilibrium (or "natural") rate, then the same rate will be an equilibrium once the inflation is underway.

The case of sustained inflation is a relatively simple one (though apparently not too simple, as it is still highly controversial). For other kinds of policy questions, one would need a more explicit model. How would the variance, and other moments, of real output change if a policy of 4 percent monetary growth were adopted? Under a balanced budget fiscal rule? Under flexible rather than fixed exchange rates? One can generate numerical answers

---

[9] Lucas (1976).

[10] Hirsch (1972), de Menil and Enzler (1972).

to questions of this sort from current macroeconomic models, but there is no reason for anyone to take these numbers seriously. On the other hand, neither can quantitative answers be obtained by purely theoretical reasoning. To obtain them, one needs an explicit, equilibrium account of the business cycle.

IV.

I have summarized, in section II, the main features of the cyclical behavior in quantities and prices. In section III, I have argued the practical necessity of accounting for these facts in equilibrium (that is, non-Keynesian) terms. That is, one would like a theory which accounts for the observed movements in quantities (employment, consumption, investment) as an optimizing response to observed movements in prices.

In the next section, I will describe the general point of view toward individual decision making to be taken in the remainder of the paper, and will explain, in particular, why the recurrent character of business cycles is of central importance. Given this general view, I shall consider in sections VI and VII the way in which relative price movements induce fluctuations in employment and investment. Sections VIII, IX, and X examine the conditions under which these same quantity responses may be triggered by movements in general, or nominal, prices. Not surprisingly, the source of general price movements is located, in section XI, in monetary changes.

V.

The view of the prototypical individual decision problem taken by modern capital theory is a useful point of departure for considering behavior over the cycle, though it is in some respects highly misleading. An agent begins a period with stocks of various kinds of capital accumulated in the past. He faces time paths of prices at which he can trade in the present and future. Based on his preferences over time paths of labor supplied and goods consumed, he formulates a plan. Under certainty, he is viewed as simply executing a single plan without revision; with uncertainty, he must draw up a contingency plan, saying how he will react to unforeseeable events.

Even to begin to think about decision problems of this general form, one needs to imagine a fairly precise view of the future in the mind of this agent. Where does he get this view, and how can an observer infer what it is? This aspect of the problem has received rather offhand treatment in traditional capital theory, and no treatment at all in traditional macroeconomics. Since it is absolutely crucial for understanding business cycles, we must pursue it here in some detail.

14

At a purely formal level, we know that a rational agent must formulate a subjective joint probability distribution over all unknown random variables which impinge on his present and future market opportunities. The link between this subjective view of the future and "reality" is a most complex philosophical question, but the way it is solved has little effect on the structure of the decision problem as seen by an individual agent. In particular, any distinction between types of randomness (such as Knight's (1921) distinction between "risk" and "uncertainty") is, at this level, meaningless.

Unfortunately, the general hypothesis that economic agents are Bayesian decision makers has, in many applications, little empirical content: without some way of inferring what an agent's subjective view of the future is, this hypothesis is of no help in understanding his behavior. Even psychotic behavior can be (and today, is) understood as "rational," given a sufficiently abnormal view of relevant probabilities. To practice economics, we need some way (short of psychoanalysis, one hopes) of understanding which decision problem agents are solving.

John Muth (1961) proposed to resolve this problem by identifying agents subjective probabilities with observed frequencies of the events to be forecast, or with "true" probabilities, calling the assumed coincidence of subjective and "true" probabilities rational expectations. Evidently, this hypothesis will not be of value in understanding psychotic behavior. Neither will it be applicable in situations in which one cannot guess which, if any, observable frequencies are relevant: situations which Knight.[11] called "uncertainty." It will most likely be useful in situations in which the probabilities of interest concern a fairly well defined recurrent event, situations of "risk" in Knight's terminology. In situations of risk, the hypothesis of rational behavior on the part of agents will have usable content, so that behavior may be explainable in terms of economic theory. In such situations, expectations are rational in Muth's sense. In cases of uncertainty, economic reasoning will be of no value.

These considerations explain why business cycle theorists emphasized the recurrent character of the cycle, and why we must hope they were right in doing so. Insofar as business cycles can be viewed as repeated instances of essentially similar events, it will be reasonable to treat agents as reacting to cyclical changes as "risk," or to assume their expectations are rational, that they have fairly stable arrangements for collecting and processing information, and that they utilize this information in forecasting the future in a stable way, free of systematic and easily correctable biases.

---

[11]Knight (1921). I am interpreting the risk-uncertainty distinction as referring not to a classification of different types of individual decision problems but to the relationship between decision maker and observer.

In moving from these general considerations to more specific theory, it will be helpful to consider as an example a "representative" agent.[12] Imagine a single worker-producer, confronted each period with a given market price for a good which he then makes to order, at a fixed rate of output per hour. That is, he comes to his place of work, observes his current selling price, determines how many hours to work that day, sells his produce, then goes home to relax.

The good he receives in exchange for the effort is "money"; I shall not be concerned with the historical reasons for this arrangement, but simply take it for granted. This money, in turn, is spent on a wide variety of goods, different from day to day. Some purchases he makes on his way home, in an hour's break from work, or several days later. I assume for now that he holds no other securities. I assume also that this agent lives in a cycle-free world, in which the general or average level of prices does not change, though individual prices fluctuate from day to day.

Now let us postulate an increase of 10 percent in today's selling price, as compared to the average of past prices. How will this hypothetical producer respond? The answer given by economic theory must be: who knows? At this point, I have said nothing which would enable one to imagine what the producer thinks this price movement <u>means</u>. If he believes the price change signals a permanent change in his selling price, we know from much evidence that he will work no harder, and probably a little less hard. That is, we know that "long run" (very unfortunate terminology, since the "long-run" response to a permanent price change will be <u>immediate</u>) labor supply elasticities are zero or negative.

What if, at the opposite extreme, the price change is transitory (as would be the case if each period's price were an independent drawing from a fixed distribution)? The answer in this case amounts to knowing the rate at which the producer is willing to substitute labor today for labor tomorrow. If "leisure" is highly substitutable over time, he will work longer on high price days and close early on low price days. Less is known about actual labor supply responses to transitory price movements than about the "long-run" response, but what we do know indicates that leisure in one period is an excellent substitute for leisure in other, nearby periods. Systematic evidence at the aggregate level was obtained by Rapping and myself (1970); Ghez and Becker (1975) reached

---

[12] Many of the arguments in this and subsequent sections have been developed more explicitly elsewhere. The closest single parallel treatment is in Lucas (1975). See also Phelps, et al. (1970), Barro (1976), Sargent and Wallace (1975), Sargent (1976). In what follows, I will not document particular arguments, nor will I attempt to apportion credit (or blame) for ideas discussed.

the same conclusion at a disaggregative level. The small premiums required to induce workers to shift holidays and vacations (take Monday off instead of Saturday, two weeks in March rather than in August) point to the same conclusion, and this "casual" evidence is somewhat more impressive because of its probabilistic simplicity: holidays are known to be transitory. On the basis of this evidence, one would predict a highly elastic response to transitory price changes .

Before dealing with complications to this example, let us note its promise for business cycle theory. I have described a producer who responds to small price fluctuations with large fluctuations in output and employment: exactly what we observe over the cycle. The description rests on economically intelligible substitution effects, not on unintelligible "disequilibria." Yet let us go slowly: our aggregative observations refer to co-movements of output and prices generally; the example refers to relative price movements in a stationary environment.

Before facing this difficult issue, let us consider some variations on the example just considered. First, from a descriptive point of view, it often seems more realistic to think of demand information being conveyed to producers by quantity changes: new orders, inventory rundowns, and the like. There seems to be no compelling substantive reason to focus exclusively on prices as signals of current and future demand. At this verbal level, it seems to me harmless and accurate to use the terms price increase and sales increase interchangeably. Somewhat surprisingly, however, rigorous analysis of equilibrium determination when producers set prices is extremely difficult, and no examples relevant to business cycle behavior exist.

A second variation is easy to carry out. Rather than consider a worker-entrepreneur, one could separate these functions, introduce firms, and consider labor and product markets separately. In the present context, this would introduce a distinction between wages and prices, and raise the issue of risk-allocating arrangements between employers and workers.[13] It would also permit the study of possibly different information sets for firms and workers. None of these questions is without interest, but all are, in my opinion, peripheral for business cycle theory. Observed real wages are not constant over the cycle, but neither do they exhibit consistent pro- or countercyclical tendencies. This suggests that any attempt to assign systematic real wage movements a central role in an explanation of business cycles is doomed to failure. Accordingly, I will proceed as though the real wage were fixed, using the terms "wages" and "prices" interchangeably.

---

[13]One such arrangement is the practice of "laying off" workers. See Azariadis (1975).

17

Additional variations can be obtained by distinguishing among various uses of the worker-producer's time when he is not working. Many writers have attempted, for example, to interpret measured unemployment as time engaged in job search. Certainly, if one substitutes away from work one substitues into some other activity, and experience shows that one's belief in the importance of substitution is bolstered by some plausible illustrations. Nevertheless, there is little evidence that much time is spent in job search, that search is less costly when unemployed than when employed, or, for that matter, that measured unemployment measures any activity at all. Economically, the important issue is the magnitude of the elasticity of employment with respect to transitory wage and price movements, not the reasons why that elasticity is what it is.

Indeed, I suspect that the unwillingness to speak of workers in recession as enjoying "leisure" is more a testimony to the force of Keynes' insistence that unemployment is "involuntary" than a response to observed phenomena. One doesn't want to suggest that people like depressions! Of course, the hypothesis of a cleared labor market carries with it no such suggestion, any more than the observation that people go hungry in cleared food markets suggests that people enjoy hunger.

VII.

More complex variations on this example arise when capital of various kinds is introduced. Let us do this, retaining still the assumption of stability over time in the general level of prices.

Three possibilities of interest arise. First, suppose that current production can be stored as finished goods inventory. This possibility seems to work against the account of price-output co-movements sketched above. The producer will surely produce in low price periods for sale later when price is high, smoothing labor supply relative to the case where storage is precluded. On the industry level, however, this behavior also dampens price movements. The net result is likely to be a reduction in the elasticity of employment-production with respect to price, and an increase in the real sales-price elasticity.

As a second possibility, suppose the producer can use a part of his current production to acquire a machine which will raise his output-per-hour in all future periods. As a third, suppose he can take a course in school which will have the same effect. Since these two possibilities do not differ economically, they may be considered as one. In the example of purely transitory price movements, discussed earlier, it is clear that neither of these options will ever be exercised -- provided the producer was satisfied with his original stock of capital. By the time the new capital can be applied to production, the price movement which made it appear profitable will have vanished.

Current relative price movements will have their maximal effect on capital accumulation when, at the opposite extreme, they are regarded as permanent. In this case, however, as I have noted, employment will be insensitive to price movements. Thus, to observe investment and employment moving systematically in the direction of relative price movements, it must be the case that such movements are a mix of transitory and permanent elements. In such a situation, the producer will find himself obliged to engage in what engineers call "signal processing": he observes a single variable (price) changing through time; these movements arise from movements in more fundamental variables (the transitory and permanent components of price) which cannot be observed directly; from these observed price movements, together with his knowledge of the relative importance of the two unobserved sources of price change, he imperfectly infers the movements in the two components. Based on his solution to this implied conditional probability calculation, he takes a decision. Not surprisingly, the decision turns out to be an average of the decisions appropriate to the two extremes.

To recapitulate, our hypothetical producer is taken to face stochastic price variability, which is describable as a mix of transitory and permanent components, both unobserved. His optimal response to price movements depends on two factors: the way he interprets the information contained in these changes, and his preferences concerning intertemporal substitution of leisure and consumption. Under assumptions consistent with rational behavior and available evidence, his response to an unforeseen price increase is a sizable increase in labor supplied, a decline in finished goods inventory, and an expansion in productive capital accumulation of all kinds. This behavior is symmetric; the responses to price decreases are the opposite.[14]

VIII.

It is time to think of situating this representative producer in an economy comprised of similar agents, though of course producing different goods and subject to different individual price movements. To do this, one must go behind price movements to the changes in technology and taste which underlie them. These changes are occurring all the time and, indeed, their

---

[14] What is happening to consumption expenditures as these employment and investment responses take place? In his critique of equilibrium business cycle models, Grossman (1973) argues that consumption must necessarily move in the opposite direction from labor supplied. Since this is not what is in fact observed over the cycle, it would indeed by a serious paradox if a negative correlation were a consequence of utility theory. One can derive it for special cases (see Lucas, 1972, Fig. 1) but this implication is certainly not a general fact for optimizing households; it does not, for example, follow from Rapping's and my (1970) theory or from that of Ghez and Becker (1975, ch. 4).

importance to individual agents dominates by far the relatively minor movements which constitute the business cycle. Yet these movements should, in general, lead to relative, not general price movements. A new technology, reducing costs of producing an old good or making possible the production of a new one, will draw resources into the good which benefits, and away from the production of other goods. Taste shifts in favor of the purchase of one good involve reduced expenditures on others. Moreover, in a complex modern economy, there will be a large number of such shifts in any given period, each small in importance relative to total output. There will be much "averaging out" of such effects across markets.

Cancellation of this sort is, I think, the most important reason why one cannot seek an explanation of the general movements we call business cycles in the mere presence, per se, of unpredictability of conditions in individual markets. Yet this argument is not entirely tight. It is surely possible for a large number of agents spontaneously to feel an urge to increase their work weeks and expand investments. More seriously, there have been many instances of shocks to supply which affect all, or many, sectors of the economy simultaneously. Such shocks will not cancel in the way I have described, and they will induce output fluctuations in the aggregate. They will not, however, lead to movements which fit the description sketched in section II: all supply shifts will lead to countercyclical price movements (other things being equal) in contrast to the procyclical movements we observe.

It is, then, possible to situate our hypothetical producer in a general equilibrium setting, in which his price and output fluctuate, yet aggregate levels do not. His responses to these relative prices movements will mimic the aggregate responses to general price movements which constitute the business cycle. We have then a coherent model, but not one which as yet accounts for the general phenomena to be explained. This model can, without difficulty, be modified to permit general, supply-induced output fluctuations, but these bear no resemblance to the modern business cycle.

Before leaving this world of stable aggregates, it is worth stressing that most of the risk which troubles and challenges economic agents would be present in such a setting. Will consumers take to a novel automobile design, or will it become a national joke? Will a dozen years of training in piano lead to the concert stage, or just a pleasurable hobby? Will this week's overtime wages help finance a child's education, or tide the family over next month's strike? By the time one has acquired the information necessary to resolve questions like these, it is too late; one way or the other, one is committed.

Compared to risks of this nature and magnitude, the question of whether the hours actually worked in the year ahead will be 1.03 times what one plans

for now, or .97, seems a minor one, and seems so because it is. In aggregative economic theory, we are accustomed to think of business cycles as a kind of risk imposed on an otherwise stable environment. Such habits of thought reflect the transfer of abstractions useful for some purposes into contexts where they involve fatal distortions of reality.

IX.

Let us now drop the assumption of stability in average prices. From the point of view of the individual producer, this involves only a slight change in the nature of the signal processing problem which must be solved. Before, a given movement in his "own price" could mean a permanent relative price change or a transitory one. Now, it can also mean that all prices are changing, a situation which, if correctly diagnosed, would lead to no real response on the producer's part. Yet, for the same reason that permanent and transitory relative price movements cannot be sorted out with certainty at the time, neither can relative and general movements be distinguished. General price increases, exactly as will relative price increases, will induce movements in the same direction in employment and investment.

Unlike the responses to taste and technology changes described earlier, these responses to general price increases will not tend to cancel over markets. To be sure, some producers will observe declines in demand even during price expansions, but more will observe increases (this is what a general price increase means), and therefore more will be expanding in real terms than will be contracting. The net effect will be co-movements in prices, output, and investment at the aggregate level, just as is observed over the actual cycle.

It is essential to this argument that general price movements not be perceived as such as they are occurring. Within the context of the aggregative models ordinarily used, this assumption may seem implausible: how could traders not know the price of goods? In the reality of a multi-commodity world, however, no one would want to observe all prices every day, nor would many traders find published price indices particularly useful. An optimizing trader will process those prices of most importance to his decision problem most frequently and carefully, those of less importance less so, and most prices not at all. Of the many sources of risk of importance to him, the business cycle and aggregate behavior generally is, for most agents, of no special importance, and there is no reason for traders to specialize their own information systems for diagnosing general movements correctly.

By the same reasoning, one can see that sustained inflation will not affect agents' real decisions in the way that transitory price movements do. Nothing

is easier than to spot and correct systematic bias in forecasts. Such corrections involve no changes in agents' information systems or in the costs of processing information. There may, of course, be some lag in diagnosing sustained inflation for what it is; about as often, agents will incorrectly perceive a transitory inflation as though it were sustained.

Changes in the degree of price variability will have more fundamental effects on agents' information processing behavior, because they affect the "weights" placed on price information in forecasting future prices. The general idea is that one trusts "noisy" price signals less.

X.

The aggregate or average response to general price movements becomes more complex as one considers investment as well as employment responses. Investment decisions will be distorted by general price movements, for the same reasons as will employment, and in the same direction as the responses induced by relative price movements.

Further complications follow, however, from the observation that current investment affects future capacity, and hence future prices. This effect can be seen to extend in time, perhaps even to amplify, the initial effects of general price movements.

To spell this out in more detail, imagine that some event occurs which would, if correctly perceived by all, induce an increase in prices generally. Sooner or later, then, this adjustment will occur. Initially, however, more traders than not perceive a relative price movement, possibly permanent, in their favor. As a result, employment and investment both increase. Through time, as price information diffuses through the economy, these traders will see they have been mistaken. In the meantime, however, the added capacity retards price increases generally, postponing the recognition of the initial shock. In this way, unsystematic or short-term shocks to prices can lead to much longer swings in prices.

In addition, there is a downturn automatically built in to this expansion of capacity. When recognition of general inflation does occur, investment will have to become less than normal for a time while capacity readjusts downward. There is no reason to expect this readjustment to come rapidly, or to be describable as a "crash," or "bust."

This scenario, like the earlier description of the employment response, depends crucially on the confusion on the part of agents between relative and general price movements. This is especially clear in the case of investment, since optimal investment policy has a great deal of "smoothing" built into it:

22

since investment is a long-term commitment, it will respond only to what seem to be relatively permanent relative price shifts.

This observation has led, on serious grounds, to skepticism as to the importance of accelerator effects in the business cycle. How can moderate cyclical movements in prices lead to the high-amplitude movements in durable goods purchases which are observed? Here again, one must insist on the minor contribution of economy-wide risk to the general risk situation faced by agents. For individual investment projects, rates of return are highly variable, often negative, and often measured in hundreds of percent. A quick, current response to what seems to others a weak "signal" is often the key to a successful investment. The agent who waits until the situation is clear to everyone is too late; someone else has already added the capacity to meet the high demand. What appears, at the aggregate level, to be a high-amplitude response pattern to low-amplitude shocks is, at the level at which decisions are made, a high-amplitude response to still higher amplitude movements in returns to individual investments.[15]

<p style="text-align:center">XI.</p>

I began section II with a definition of business cycles as repeated fluctuations in employment, output, and the composition of output, associated with a certain typical pattern of co-movements in prices and other variables. Since in a competitive economy, employment and output of various kinds are chosen by agents in response to price movements, it seemed appropriate to begin by rationalizing the observed quantity movements as rational or optimal responses to observed price movements. This has been accomplished in the preceding five sections. I turn next to the <u>sources</u> of price movements.

For explaining <u>secular</u> movements in prices generally, secular movements in the quantity of money do extremely well. This fact is as well established as any we know in aggregative economics, and is not sensitive to how one measures either prices or the quantity of money.[16] There is no serious doubt as to the direction of effect in this relationship; no one argues that the anticipation of

---

[15]"Austrian" or "monetary-over-investment" business cycle theory (see Haberler, 1936, or Hayek, 1933) was based on this same idea of mistaken investment decisions triggered by spurious price signals. However, the price which this theory emphasized was the rate of interest, rather than product prices as stressed here. Given the cyclical amplitude of interest rates, the investment-interest elasticity needed to account for the observed amplitude in investment is <u>much</u> too high to be consistent with other evidence.

[16]Friedman and Schwartz (1963).

sixteenth-century inflation sent Columbus to the New World to locate the gold to finance it. This evidence has no direct connection to business cycles, since it refers to averages over much longer periods, but the indirect connections are too strong to be ignored: we have accounted for the pattern of co-movements among real variables over the cycle as responses to general price movements; we know that, in the "long run," general price movements arise primarily from changes in the quantity of money. Moreover, cyclical movements in money are large enough to be quantitatively interesting. All these arguments point to a monetary shock as the force triggering the real business cycle.

The direct evidence on short-term correlations between money, output, and prices is much more difficult to read. Certain extreme episodes appear to indicate that depressions and recoveries are money-induced.[17] In general, however, the link between money and these and other variables is agreed to be subject, in Friedman's terms, to "long and variable lags."

Paradoxically, this weakness in the short-term evidence linking money to economic activity, and in particular to prices, is encouraging from the point of view of monetary business cycle theory. To see why, recall the theoretical link between general price movements and economic activity as sketched above. This connection rested on the hypothesis that the signal processing problem of identifying general price movements from observations of a few individual prices was too difficult to be solved perfectly by agents. Now suppose it were true that one could describe short-term general price movements by a simple, fixed function of lagged movements in some published monetary aggregate. Then, far from being difficult, the signal processing problem to be solved by agents would be trivial; they could simply observe current monetary aggregates, calculate the predicted current and future price movements they imply, and correct their behavior for these units changes perfectly. The result would be a very tight relationship between money and prices, over even very short periods, and no relationship at all between these movements and changes in real variables.

These remarks do not, of course, explain why monetary effects work with long and variable lags. On this question little is known. It seems likely that the answer lies in the observation that a monetary expansion can occur in a variety of ways, depending on the way the money is "injected" into the system, with different price response implications depending on which way is selected. This would suggest that one should describe the monetary "state" of the economy as being determined by some unobservable monetary aggregate, loosely related to observed aggregates over short periods but closely related secularly.

---

[17] Again, see Friedman and Schwartz (1963).

XII

Let me recapitulate the main features of the business cycle theory sketched in the preceding sections. We began by imagining an economy with fluctuating tastes and technology, implying continually changing relative prices, and studied the co-movements in quantities and prices which would emerge if agents behaved in their own interest and utilized their incomplete information effectively. We then superimposed on this economy sizable, unsystematic movements in a monetary aggregate, adding an additional source of "noise" to individual price movements. The result is to generate a pattern of co-movements among aggregate series which appears to match the observations summarized in section II.

In retrospect, this account seems rather embarrassingly simple: one wonders why it seems to be necessary to undo a Revolution to arrive at it. Yet one must be careful not to overstate what has, in fact, been arrived at. I think it is fairly clear that there is nothing in the behavior of observed economic time series which precludes ordering them in equilibrium terms, and enough theoretical examples exist to lend confidence to the hope that this can be done in an explicit and rigorous way. To date, however, no equilibrium model has been developed which meets these standards and which, at the same time, could pass the test posed by the Adelmans (1959). My own guess would be that success in this sense is five, but not twenty-five years off.[18]

The implications for economic policy of a successful business cycle theory of the sort outlined here are, I think, easy to guess at even when the theory itself is in a preliminary state. Indeed, much of the above is simply an attempt to understand and make more explicit the implicit model underlying the policy proposals of Henry Simons, Milton Friedman, and other critics of activist aggregative policy. By seeking an equilibrium account of business cycles, one accepts in advance rather severe limitations on the scope of governmental countercyclical policy which might be rationalized by the theory. Insofar as fluctuations are induced by gratuitous monetary instability, serving no social purpose, then increased monetary stability promises to reduce aggregate, real variability and increase welfare. There is no doubt, however, that some real variability would remain even under the smoothest monetary and fiscal policies. There is no prima facie case that this residual variability would be better dealt

---

[18]Proceeding further out on this limb, it is likely that such a "successful" model will be a close descendant of Sargent's (1976).

with by centralized, governmental policies than by individual, decentralized responses.[19]

In view of this lack of novelty in the realm of policy, it seems a fair question to ask: why do we need the theory? The general answer, I think, is that in a democratic society it is not enough to believe oneself to be right; one must be able to explain <u>why</u> one is right. We live in a society in which the unemployment rate fluctuates between, say, 3 and 10 percent. It follows that both situations are attainable, and it is clear that most people are happier at three than at ten. It is also clear that government policies have much to do with which of these situations prevails at any particular time. What could be more natural, then, than to view the task of aggregative economics as that of discovering which policies will lead to the more desirable situation, and then advocating their adoption? This was the promise of Keynesian economics, and even now, when the scientific emptiness of this promise is most evident, its appeal is understandable to all who share the hope that social science offers more than elegant rationalization of the existing state of affairs.

The economically literate public has had some forty years to become comfortable with two related ideas: that market economies are inherently subject to violent fluctuations which can only be eliminated by flexible and forceful governmental responses; and that economists are in possession of a body of scientifically tested knowledge enabling them to determine, at any time, what these responses should be. It is doubtful if many who are not professionally committed hold, today, to the latter of these beliefs. This in itself settles little in the dispute as to whether the role of government in stabilization policy should be to reduce its own disruptive part or actively to offset private sector instability. As long as the business cycle remains "in apparent contradiction" to economic theory, both positions appear tenable. There seems to be no way to determine how business cycles are to be dealt with short of understanding what they are and how they occur.

---

[19]That is to say, active countercyclical policy would require the same kind of cost-benefit defense used in evaluating other types of government policies. See Phelps (1972), and also Prescott's review (1975).

REFERENCES

1. Adelman, I., and Adelman, F. L., "The Dynamic Properties of the Klein-Goldberger Model," <u>Econometrica</u>, 27, No. 4, (October 1959), 596-625.

2. Azariadis, C., "Implicit Contracts and Underemployment Equilibria," <u>Journal of Political Economy</u>, 83, No. 6, (December 1975), 1183-1202.

3. Barro, R. J., "Rational Expectations and the Role of Monetary Policy," <u>Journal of Monetary Economics</u>, 2, No. 1, (January 1976), 1-32.

4. Burns, A. F., and Mitchell, W. C. <u>Measuring Business Cycles</u>. New York: National Bureau of Economic Research, 1946.

5. Friedman, M., "The Role of Monetary Policy," Presidential Address to the American Economic Association, <u>American Economic Review</u>, 58, No. 1, (March 1968), 1-17.

6. Friedman, M., and Schwartz, A. J. <u>A Monetary History of the United States, 1867-1960</u>. Princeton: Princeton University Press for the National Bureau of Economic Research, 1963.

7. Gayer, A. D., Rostow, W. W., and Schwartz, A. J. <u>The Growth and Fluctuation of the British Economy, 1790-1850</u>. Oxford: The Clarendon Press, 1953.

8. Ghez, G. R., and Becker, G. S. <u>The Allocation of Time and Goods Over the Life Cycle</u>. New York: National Bureau of Economic Research, 1975.

9. Grossman, H. I., "Aggregate Demand, Job Search, and Employment," <u>Journal of Political Economy</u>, 81, No. 6, (November/December 1973), 1353-1369.

10. Haberler, G. <u>Prosperity and Depression</u>. Geneva: League of Nations, 1936.

11. Hayek, F. A. von. <u>Monetary Theory and the Trade Cycle</u>. London: Jonathan Cape, 1933.

27

12. Hicks, J. R., "Mr. Keynes and the 'Classics': A Suggested Interpretation," Econometrica, 5, (1937), 147-159.

13. Hirsch, A. A., "Price Simulations with the OBE Econometric Model," in The Econometrics of Price Determination Conference. (ed. O. Eckstein), Washington, D. C.: Board of Governors of the Federal Reserve System and Social Science Council, 1972.

14. Keynes, J. M. The General Theory of Employment, Interest and Money. London: Macmillan, 1936.

15. Klein, L. A., and Goldberger, A. S. An Econometric Model of the United States, 1929-52. Amsterdam: North Holland, 1955.

16. Knight, F. H. Risk, Uncertainty and Profit. Boston: Houghton Mifflin, 1921.

17. Lucas, R. E., Jr., "Expectations and the Neutrality of Money," Journal of Economic Theory, 4, No. 2, (April 1972), 103-123.

18. _____, "An Equilibrium Model of the Business Cycle," Journal of Political Economy, 83, No. 6, (December 1975), 1113-1144.

19. _____, "Econometric Policy Evaluations: A Critique," in The Phillips Curve and Labor Markets, (eds. K. Brunner and A. H. Meltzer), Carnegie-Rochester Conference Series on Public Policy, 1, Amsterdam: North Holland, 1976, 19-46.

20. Lucas, R. E., Jr., and Rapping, L. A., "Real Wages, Employment, and Inflation," in Microeconomic Foundations of Employment and Inflation Theory, (eds. E. S. Phelps, et al.), New York: Norton, 1970.

21. de Menil, G., and Enzler, J. J., "Price and Wages in the FR-MIT Econometric Model," in The Econometrics of Price Determination Conference, (ed. O. Eckstein), Washington, D. C.: Board of Governors of the Federal Reserve System and Social Science Council, 1972.

22. Mitchell, W. C. What Happens During Business Cycles. New York: National Bureau of Economic Research, 1951.

23. Modigliani, F., "Liquidity Preference and the Theory of Interest and Money," Econometrica, 12, No. 1, (January 1944), 45-88.

24. Moggridge, D. (ed.) The Collected Writings of John Maynard Keynes, Vol. XIV. London: Macmillan, 1973.

25. Muth, J., "Rational Expectations and the Theory of Price Movements," Econometrica, 29, No. 3, (July 1961), 315-335.

26. Phelps, E.S., "Money Wage Dynamics and Labor Market Equilibrium," Journal of Political Economy, 76, No. 4, II, (July/August 1968), 687-711.

27. _____. Inflation Policy and Unemployment Theory: The Cost-Benefit Approach to Monetary Planning. London: Macmillan, 1972.

28. Phelps, E. S., et al. Microeconomic Foundations of Employment and Inflation Theory. New York: Norton, 1970.

29. Prescott, E. C., "Efficiency of the Natural Rate," Journal of Political Economy, 83, No. 6, (December 1975), 1229-1236.

30. Sargent, T. J., "A Classical Macroeconometric Model for the United States," Journal of Political Economy, 84, No. 2, (April 1976), 207-237.

31. Sargent, T. J., and Sims, C. A., "Business Cycle Modeling Without Pretending to Have Too Much A Priori Economic Theory," University of Michigan working paper, March 1976. ·

32. Sargent, T. J., and Wallace, N., " 'Rational' Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply Rule," Journal of Political Economy, 83, No. 2, (April 1975), 241-254.

33. Tinbergen, J. Business Cycles in the United States of America, 1919-32. Geneva: League of Nations, 1939.