Volume Title: Conference on Business Cycles

Volume Author/Editor: Universities-National Bureau Committee for

Economic Research

Volume Publisher: NBER

Volume ISBN: 0-87014-193-7

Volume URL: http://www.nber.org/books/univ51-1

Publication Date: 1951

Chapter Title: Business Cycle Research and Business Policy Chapter

Author: C. Ashley Wright

Chapter URL: http://www.nber.org/chapters/c4767

Chapter pages in book: (p. 337 - 374)

Part Three

BUSINESS CYCLE RESEARCH AND POLICY

BUSINESS CYCLE RESEARCH AND BUSINESS POLICY

C. ASHLEY WRIGHT, *General Economics Division, Standard Oil Company (N.J.)*

This paper is divided into three parts, at some cost in terms of expositional unity. Its form was dictated by two considerations. First, the beneficent powers controlling the destinies of the Conference on Business Cycle Research gave me to understand that I was expected to discuss the topics covered in the first and third sections. Second, a review of the literature on the topics assigned and an examination of the subjects to be considered at the Conference suggested that the discussion of the second section — a survey of forecasting from the viewpoint of one who must consider it in terms of the needs of business — would serve a useful purpose.

The first section treats the limitations of our knowledge of the role of information concerning the business cycle in business decisions, and outlines the principal fields of business policy in which business cycle research might be of value in reducing both social and business costs. The second section, as indicated above, classifies and discusses the techniques of business cycle forecasting that have been or are being used. The discussion in this section deals only with forecasting the turning points of the so-called '40 months' or Kitchin cycles since attempts to forecast longer cycles and amplitudes of any type appear to have been based upon methods that are obviously inadequate quantitatively or at best without an established presumption in their favor.

The third section deals with an attack on the problem of forecasting turning points that is primarily empirical in spirit. Its principal virtue, if any, lies in the fact that initial experiments suggest it may be capable of reducing the risks of error in the field of its application to within just tolerable limits for policy purposes.

1 BUSINESS CYCLE RESEARCH AND BUSINESS POLICY

The literature is of little help in an examination of the subject of this section — for several reasons.

In the first place, business cycle research is related to business policy mainly in so far as the results are used in policy making, that is in the formulation of decisions that will be executed in the future and depend upon expectations. It appears to follow from this that a direct understanding of the role of research in policy making and of the relation between business policy making and the business cycle can be obtained only by examining the behavior of the individual firm, that is, by 'microscopic' analysis.

Much has been done on the 'microscopic' level, of course, in analyzing the behavior of the individual firm but this work, in general, has been based on assumptions of conditions of equilibrium. Business cycle theory, on the other hand, has been developed almost entirely in terms of aggregates and throws little direct light upon the detailed behavior of the individual firm. In so far as the literary theories of the business cycle discuss business behavior at all, they do so in terms of 'anticipations', 'expectations' and other general factors that are assumed to behave in certain ways as aggregates. In so far as attempts have been made to explain the behavior of these aggregates in terms of the individual firm and the behavior of the entrepreneur, there has been only a simple extension of the assumptions of individual behavior and motivation from the field of equilibrium analysis to that of the business cycle. It is perhaps not an exaggeration to say that business cycle theory, strictly defined, has told us nothing concerning the behavior of the individual firm and the entrepreneur; indeed, it has done no more than offer another opportunity for using assumptions formulated for the purposes of equilibrium theory.

In the second place, relatively little empirical work has been done on the conditions considered relevant by business policy makers to a choice between the alternatives that are open to them or on the limitations imposed upon them by their economic environment.[1] The obstacles to such empirical work are at least as great as in the field of empirical verification of the theory of the individual firm, and probably much greater. Such an investigation not only faces the difficulties of finding data and interpreting their relevance to the categories of theory but must deal immediately with questions of behavior and motivation concerning which direct data are inherently unobtainable.[2] Business policy makers themselves probably

[1] See, however, Ruth P. Mack, *The Flow of Business Funds and Consumer Purchasing Power* (Columbia University Press, 1941), especially pp. 242-305.

[2] It is possible to argue that motivation, being subjective, may not be susceptible to direct examination. Perhaps the most that can be done is to examine empirical data, formulate a theory of behavior that is consistent with the observations, from the theory draw conclusions about other empirical data, and verify the conclusions by an examination of the latter data. If so, we may accept the theory as 'true' when the conclusions agree with the data, but we shall have said nothing about the truth of the behavior axioms on which the theory is based. There may always be an alternate set of axioms leading to the same conclusions. Such considerations lead to skepticism concerning behavior conclusions drawn from empirical correlations – involving lagged profits, for example – however useful they may be empirically.

could not solve the theoretical problems of motivations — and the role of research in policy making — not only because of the difficulties of arriving at truth through introspection but also because the processes of formulating policy depend upon the subjective evaluation of meticulously prepared empirical data, expectations concerning the future, and the amount (in some ill defined sense) of uncertainty that must be associated with those expectations. Moreover, the formulation of business policy depends upon the subjective weighing of almost as many variables and almost as many interrelations among those variables as are contemplated by economics itself — and this in full knowledge of the fact that the interrelations are at the very best but dimly understood. Not only must economic variables be considered over, say, the next 5 or 10 years but so also must be social and political trends over longer periods, say, 20 to 30 years. Indeed, the process of formulating business policies for the 'maximization of future expected profits' must contemplate trends in the ethical demands of society upon business as well if profits are to be obtained at all.

Finally, the literature on expectations and their role in economic phenomena, to which one would logically turn in the light of the above considerations, has been so far mainly devoted to the definition of concepts and the formulation and clarification of deductive theory conceived for the most part in terms of conditions of equilibrium. Important and valuable as this work is to the development of a broad and consistent body of thought, it does not yet appear to have reached the stage of inductive verification or a point at which it can tell us much about empirical phenomena not yet empirically investigated.

With these limitations in mind, it will be apparent that the following paragraphs must be in the nature of an exploratory survey of the manner in which business cycle research may be useful in business policy making and that the opinions expressed must be derived largely from personal observation.

Since business policy formation is in part a process of formulating expectations concerning the future and since business cycle fluctuations are the greatest single source of changes in income flows, in the relationships at any given time between costs and prices, and in trends in these relations, it is at once apparent that business cycle research can contribute most to business policy formation by developing techniques of forecasting the turns and amplitudes of business fluctuations. While this is obvious and has been the tacit assumption behind most business cycle research, it is worth mentioning because it implies that the study of how interrelations among economic variables change in different phases of the cycle is of distinctly secondary importance for policy making, however important it may be for a satisfactory theory of the cycle. If in the expanding phase of the cycle costs or certain classes of costs are rising more rapidly than the prices of different classes of output, it is usually possible to ascertain the facts with considerable accuracy from the current data. Such information would probably be sufficient to make appropriate adjustments in input, output, and proportions of the productive factors if it were not for the fact that any existing trend in relationships among these variables may be changed or even reversed at some uncertain time and to some uncertain extent in the immediate future. While there may be some time lag in appropriate policy adjustments, where current data must be relied upon, the social and business costs of failures to make adjustments promptly for this reason are no doubt negligible compared with the costs of failure to adjust promptly to turns in the business cycle.

If it were possible and necessary to make a choice between having reasonably accurate forecasts of turning points and reasonably accurate forecasts of amplitudes, the former would no doubt be chosen. Once having concluded that a fall in demand, cyclical or other, for his products is to be expected, the policy maker is rather severely limited in the adjustments he can accomplish. Hiring can be suspended, labor forces can be reduced, purchasing can be curtailed, inventories may be allowed to decline, administrative and technical efficiency can be improved, and long-term investment programs can be halted, but the critical step is probably reversing an existing and well elaborated policy and initiating a new one. Moreover, when the appropriate policy decisions have been made and put into execution, say during a business decline, little more can be done beyond detailed supervision, continuous review, and further refinement of policy and execution regardless of how far the decline goes. On the other hand, if turning points could be foreseen far enough in advance, new policies could be formulated and set into operation with a much shorter time lag for the elimination of uncertainty and for the administrative processes necessary for both planning and execution.

If, as has been supposed and is probably the case, long-term investments are based less on consideration of cyclical fluctuations in profits over the period of investment than on the fear of entering prosperity with obsolete and high cost equipment or the fear of having to maintain one's cash position, in a period when costs exceed income, by selling assets at a severe capital loss, accurate forecasts of turning points would probably tend to stabilize investment activity through the reduction of uncertainty associated with these fears. The stabilizing influence obtained through the elimination of this sort of uncertainty might not, and probably would not, greatly reduce the amplitude of prolonged declines such as 1929-33, but it might go a long way toward moderating mild fluctuations in which inventory changes may predominate.

In the field of investment policy there is perhaps another way in which business cycle research might contribute to the formulation of business policies and, more important, possibly to the reduction of cyclical fluctuations and their impact upon society: through the careful investigation of the possibility that in some types of industry and in at least some situations a contra-cyclical investment policy might be more profitable than the usual cyclical one. In an economy in which profits, at least in a broad sense, provide the motivation for action the most promising way to obtain the cooperation of business in limiting fluctuations in investment would be to show under what conditions it might be profitable. I have seen at least one inconclusive study made by a business man for his own concern in which the profitability of a contra-cyclical investment policy was suggested. A preliminary and aggregative study prepared by the staff of which I am a member, based upon a simplified model, suggested that a contra-cyclical policy might differ negligibly from a cyclical one in mild depressions and might be slightly more profitable than a cyclical one in severe or prolonged depressions. The full investigation of a problem of this kind is usually not appropriate to the staff of a private concern but its potential importance to society, if it can be shown that enough industries could profitably adopt contra-cyclical investment policies, makes it suitable for the private investigator.

While there is some incentive to minimize capital outlays when costs are high and to maximize them when costs are low, many considerations other than those mentioned in the above paragraphs must be recognized in formulating an investment program. Some may be influenced by the business cycle while some may influence the business cycle itself in the role of possible 'random' autonomous impacts. For example, an industry that supplies necessities or semi-necessities may find itself compelled to embark upon a large investment program during a business boom, under conditions of high construction costs and a cyclical high in the demand for its products, in a period following one in which investment was curtailed by, say, a war or a severe depression.[3] It may find itself in a similar position if it has failed to foresee accurately the demand for its products at full employment levels as demand increases cyclically or otherwise.

In the case of some products with low income and price elasticity of demand, full employment consumption may perhaps be estimated with a fair degree of accuracy over a moderately wide range of price assumptions. In these cases secular factors are usually more important than those of the business cycle, and business cycle research can make only a minor contri-

[3] See statement by E. Holman, President of the Standard Oil Company (N.J.), before the Investment Subcommittee of the Joint Committee on the Economic Report, Washington, D.C., September 27, 1949.

bution to their demand aspects. In others, demand fluctuates sharply with price or income and business cycle research can contribute materially not only to the formulation of appropriate long-term investment policies but also to those applicable to current operations. In some cases the demand for a particular product may not vary cyclically in a simple fashion but may be determined by the cyclical fluctuations of some other good with which it is used. For example, new installations of household oil heating equipment may increase sharply in the upswing of the cycle but decline sharply on the downswing. The consumption of household heating oil may move on an upward trend in a series of connected more or less S-shaped curves.

Finally, improved knowledge of the influence of cyclical factors would aid materially in the isolation of secular and other movements and contribute at least in some measure to the reduction of uncertainty in long-term investment decisions.

In the field of business pricing policies, questions of cyclical timing are of major and perhaps for practical purposes of sole importance in relation to business cycle research. Prices are, of course, the subject of the most intense continuous scrutiny and consideration. When a small or moderate price decline follows a long continued rise, it is of immediate importance to determine whether the decline is due to a 'random' movement, a seasonal factor, a significant change in the conditions of supply, or a significant cyclical or noncyclical change on the demand side. Different answers to the question of what is causing the price change will lead to very different policy decisions. Random and seasonal influences are usually not too difficult to detect, at least if sufficient data are available and are subjected to careful statistical analysis, and changes in the conditions of supply are likely to be fairly well understood by acute minds continuously immersed in the affairs of their business. It is the factors influencing demand and particularly changes in demand that are most likely to be surrounded by uncertainty and it is in this connection that business cycle research could contribute to the elimination of uncertainty and the reduction in the lag of policy decisions behind current events.

Let it not be assumed that this is easy. It is well known that the demand for many commodities is closely correlated with disposable income. For such commodities random deviations from regression lines are likely to be small, but before the economist has plotted and observed a deviation and tested its significance statistically, if he can, he may find that the policy maker, with his more detailed knowledge of and exclusive concentration on the day to day behavior of his market, has already arrived at a decision which, more often than not, is correct.

Price decisions, if erroneous, are usually not too difficult to change;[4]

hence errors are not likely to be as costly as in other fields. Major errors in inventory policy are an obvious example but there are others less obvious.

Cyclical shifts in demand and price may vary greatly in timing and amplitude for different products in joint supply. For example, it is likely that the demand for gasoline turns with a lag, on the average, and that the amplitude of its cyclical fluctuations is small. On the other hand, the demand and price of residual fuel oil may turn early, on the average and with a large variance, and may have a large average cyclical amplitude.[5] Often it is possible to make appropriate shifts in output but such shifts usually involve substantial time lags and often necessitate considerable expenditure for the reconditioning of old equipment, or long-term investments in new. These lags are likely to depend on the time necessary, first, to diagnose an observed change; second, to formulate and convey to others an appropriate policy;[6] and finally, to put such a policy into action. One may well find that before an appropriate policy has been executed it may have to be reversed.

In the field of personnel relations it is difficult for one who has only the slightest acquaintance with its problems to do more than single out that still unsettled question of the relation of wage changes to the attainment of prosperity from a position short of full employment. Professor Viner's point that wage changes will increase, decrease, or leave income unchanged according to the elasticities of demand for labor in each industry, has been strangely neglected by empiricists and theorists alike. Perhaps this neglect is attributable in part to the trend toward aggregate analysis. In any case it is not impossible that aggregate analysis might benefit by increased knowledge in this connection, and business cycle research, theoretical and empirical, might well turn to it. The social gains to be obtained are obvious, and there can be little doubt of its importance to industries in which the demand for labor fluctuates widely in the cycle, always provided that the results of research are sufficiently conclusive to reduce uncertainty to

[4] Except, of course, where long- and medium-term price contracts or large inventory changes are involved. The role of business cycle research in connection with these could be similar to its role in assisting in the formulation of investment policy.

[5] In addition to the point under discussion, these may involve embarrassing and costly storage problems where they are not foreseen sufficiently in advance and where production is a continuous process, as is the case in an oil refinery.

[6] Investment in advance planning is probably discouraged, to a considerable extent, in a manner similar to that in which investment is discouraged by uncertainty concerning the time at which it will be needed. There is always the risk that a plan made too far in advance will become obsolete or not be needed at all. Perhaps this is one reason why business men have an instinctive lack of interest in certain types of 'planning' even though they use long-term 'plans' (constantly revised) as an aid in policy making.

manageable proportions. In the presence of uncertainty short-term interests are likely to prevail.

In industries in which the demand for labor is relatively stable throughout the business cycle, the opportunities are fewer, but reduced costs, business and social, are possible. Here, as in so many fields, gains may be obtained through improved timing. In certain industries at least, it is often sufficient to cease hiring, and not necessary to commence firing, when a cyclical decline in the need for labor is recognized. Attrition may be counted on to reduce the labor force to appropriate size. In so far as the timing of policy changes can be improved, costs can be reduced, morale preserved, and security extended.

For industries that have extensive interests in the foreign field, the long view is particularly important. While constant attention is devoted to the immediate future, careful consideration is given to expected market demand and supply trends over periods up to five years at least. Problems of long-term investment in plant and equipment and its location are studied in relation to world wide shifts in economic sources of supply and markets, to long-term trends in productivity and the standards of living in countries of possible location, to the implication of these long-term trends for the stability of money and exchange rates, and to political trends that may lead to the loss or destruction of investment or inability to withdraw profits and amortization if liquidation should later seem desirable. Neither the business man, who by experience and study has broadened his horizon to deal with worldwide economic and political problems, nor the specialists who assist him have any illusions concerning the difficulties of foreseeing these great movements accurately over long periods. But they must be dealt with, and there is always the presumption that careful study and painstaking analysis will result in a higher proportion of correct decisions in the long run and on the average than would methods less systematic.

The general classes of problems that present themselves in foreign operations are the same as those which are met in the domestic field. There are, however, additional complications. In general, data are scarce and essential information is often entirely lacking. Except for a very few countries, there is almost literally no information on business fluctuations. Very little information is to be had about the influence of business cycles in the industrialized countries of the western world upon less industrialized countries. Very little detailed information has been assembled concerning the manner in which business cycles are transmitted or the conditions under which they may be transmitted from country to country, beyond a few obvious but important generalizations.

The shorter-run problems of foreign exchange are obviously important. Business is continuously concerned over the risks of loss associated with inconvertibility and such risks are an important factor in investment decisions. The existence of blocked accounts plus the possibility that inflation may eliminate their value is an inducement to investment that might otherwise be uneconomic. In this connection there remains a great unsolved problem: how can the effects of business cycles be curbed and full employment be maintained, and at the same time the gains to be obtained from relatively stable foreign exchange rates be maximized?

2 BUSINESS CYCLE FORECASTING

It is proposed in this section to classify and survey briefly the various techniques developed from time to time for the purpose of forecasting turns in business. A satisfactory solution to this problem would be the greatest single contribution business cycle research could make to the formulation of business policy or government policy. It might even be argued that the fate of our institutions and perhaps the form of our society itself hinges upon it, for a failure to foresee and take appropriate measures to forestall a depression such as we experienced in the 1930's would place these in peril. It is in this field, moreover, that business cycle research has had its most systematic application.

Forecasting methods can be placed in three classes: intertemporal, intratemporal, and heterotemporal.[7] The first class includes methods that involve comparisons of economic variables considered to be relevant and 'structural' relationships assumed to exist over or between different periods.

The simplest method in this class is that of simple historical analogy. Its limitations and the obstacles to its success are almost obvious, but the stubborn fact remains that when one is dealing with questions of forecasting cyclical amplitudes one is forced to rely heavily upon it, with qualitative allowances for the effects of 'structural' changes known to be or believed to be important.

The second group, in order of sophistication, includes methods that assume some sort of simple empirical historical relationship. In some of these, such as that formerly used by the Babson Service, variation from an assumed 'normal' relationship has been used. In others, an average cyclical pattern of some kind is assumed to have empirical existence and used as a basis of forecasting. In still others, such as the old Harvard Barometer and Brookmire methods, a supposed empirical relationship, usually a lead or lag between time series, is used. In so far as these methods are rationalized in terms of business cycle theory, as was the case with the Harvard Barometer, they move toward the third and most sophisticated group in this class, that of the theoretical dynamic hypothesis.

[7] A heterogeneous combination of roots, not without precedent, which we hope will be forgiven. The classification given here owes much to that of C. O. Hardy and G. V. Cox, *Forecasting Business Conditions* (Macmillan, 1927).

The simplest of this third group is illustrated by such attempts to deal with the problem as that of Dewey and Dakin. Their hypothesis includes a mechanical relationship between Kondratieff, Juglar, and Kitchin cycles and is presumably derived more or less directly from empirical data. Jevon's sunspot theory, if it has ever been used for forecasting purposes, would fall in this group but the hypothesis on which it is 'based' is physical rather than mechanical and from this viewpoint one step forward in sophistication; it is hardly necessary to point out that in Jevon's theory a whole chain of hypotheses is omitted and reliance is placed on empirical or supposed empirical evidence to provide the missing links. Hawtrey's monetary theory of the cycle, if it also has ever been used for forecasting, is still more sophisticated as a theoretical elaboration or construct, as are the models of Kaldor, Kalecki, Metzler, and Samuelson. As one glances back over this list, one has the feeling that there is not only a steady progression toward theoretical subtlety but a striking regression from the relevant data.[8]

Intratemporal methods are essentially static in spirit. They fall into two principal classes. The first is the simple enumeration of favorable and unfavorable factors together with implicit or more or less explicit assignment of weights to each factor. As a forecasting method simple enumeration has obvious and probably insuperable faults but it is almost essential as a first step in a more elaborate analysis.

The second class contains the two main types of static models. The first is the model that assumes a set of simple empirical or theoretical relationships between single dependent variables on the one hand and separate groups of independent variables on the other. The second group, that of more complex models, assumes a set of interdependent relationships between a set of dependent (endogenous) and a set of independent (exogenous) variables, the models of the Cowles Commission being the outstanding examples. While these, like all similar methods, rely on data extending over periods of time, they are static in the sense that the relations (and parameters of equations defining them) are assumed as given and unchanging over the periods covered by the data.[9] Time series are used

[8] No criticism is implied. I am aware of the advantages as well as the limitations of specialization and of the difficulties of the subject.

[9] In so far as lagged variables are introduced, as they must be for statistical reasons, these models verge into the heterotemporal class. They appear to be static in spirit, however, and their heterotemporal relationships link only very short periods. The same may be said of models, such as Clark's, which make adjustments for discontinuous 'structural' changes, such as changes in tax rates, or include partly discontinuous and partly continuous changes in parameters which depend on the variables of the system. Though static in spirit, models that include lagged variables can in theory lead to values of the variables over time that show damped, undamped, or explosive oscillations.

to estimate the values, assumed to be fixed, of the parameters of the equations that define these relationships. It is sometimes said that no variable *in such a system is 'causal' and that all are mutually determining.* This statement contains an important element of truth, but we may, if we like, regard the independent variables, if any, as 'causal'. When the system of relationships contains as many linearly independent equations as there are unknowns, the autonomous 'random' variations may be regarded as 'causal'.[10] The point appears to be largely one of definition and terminology.

It is hard to say very much about heterotemporal methods because the problems that arise in their construction have scarcely been discussed or considered and the difficulties appear to be great. They might include theories of the time changes in exogenous variables, the separation of endogenous variables into endogenous and exogenous parts with theories concerning the behavior of the latter, theories concerning the behavior over time of variables now regarded as autonomous and random, and theories concerning changes in the parameters of equations describing 'static' relationships. In so far as changes in parameters are expressed in terms of measurable quantities, one can either regard them as changes in relationships or as fixed relationships with the addition of new variables. One would expect, however, that such new variables would play a role different from that of the old ones.

'Modigliani's constant' and the similar suggestions of Duesenberry provide simple examples of this method. It seems plausible to suppose that discontinuous shifts in parametric values might prove useful in refining and utilizing the suggestions of literary theories of the cycle that have to do with cumulative processes in the vicinity of turning points.

Finally, if a plausible case can be made for tentatively assuming that there are autonomous cyclical fluctuations in certain sectors of the economy, say inventories or building, or damped or explosive tendencies associated with 'random' impacts, it would be interesting to see what could be done to include equations describing these variables in a simple model *which for constant time was essentially static.*

Simple historical analogy and historical methods in general face extremely serious obstacles as forecasting techniques. In the first place, they must deal with a multiplicity of variables and relationships and the practical impossibility of finding situations sufficiently alike to justify conclu-

[10] It may be argued that variations are 'random' when their behavior may be described by a frequency distribution and their variation cannot be explained in terms of some nonrandom variable or variables. When they can be explained, they cease to be 'random'. The fact that most 'random' variables in economic data show serio- and inter-correlations suggest that they could be explained, at least in part, if we only knew how. In theory, of course, this need not be so.

sions based on historical comparisons. This difficulty would remain even if it were known what variables and what relationships were relevant. In the second place, the variations economic series show from cycle to cycle are so great as to make conclusions based on simple methods of handling them most uncertain. For example, the leads and lags of series that are regarded as having the most stable ones relative to the turning points of the cycle vary so much from cycle to cycle that it is difficult or almost impossible to use them in a simple way to derive information concerning a current situation. Methods that rely on patterns averaged over several cycles are hardly better for the same reason. Cyclical variations about averages are too wide to be of much assistance in a particular cycle.

It does not follow that the search for systematic patterns should be abandoned. Perhaps a more sophisticated attack which dealt with these questions as problems in frequency distributions might be productive of useful results.[11] Mathematical model building has more grandiose aims and offers more promise for the advancement of the broad frontiers of economic theory, but in a humbler sphere much of value could probably be done along these lines and discoveries might be made that would clarify our knowledge of empirical behavior and ultimately contribute to the improvement and refinement of models.

The physical sciences have developed largely through the discovery of empirical relationships that could not be explained or theoretically integrated with current thought. The latter has then been modified to explain the new facts and later used to make new and subsequently verified hypotheses concerning empirical data.

A principal and unsolved difficulty in the use of theoretical dynamic hypotheses for forecasting purposes is the discovery of a set of assumptions that can be depended upon to hold reasonably well in the future. The objective here is *not* to discover which of several hypotheses is consistent with the data to the exclusion of alternatives but rather to find one (or a set) that will work fairly well and be accurate enough to be useful.[12] There are probably several alternative and mutually contradictory sets of hypoth-

[11] See for example R. C. Vining, Location of Industry and Regional Patterns of Business Cycle Behavior, The Region as a Concept of Business Cycle Analysis, *Econometrica,* January and July 1946; also *Proceedings of the American Economic Association,* December 1948.

[12] T. C. Koopmans has very properly stressed the role of mathematical models in the construction and statistical screening of business cycle theories. One gets the impression, however, that their usefulness for screening is still in about the uncertain state that their usefulness for forecasting was two or three years ago. This should not be taken to imply any skepticism concerning the essential soundness of Professor Koopmans' position. See The Econometric Approach to Business Fluctuations, *Proceedings of the American Economic Association,* December 1948.

eses that would satisfy these simple criteria. We would be very much the gainers if we could find *one even if it were inconsistent with itself in some respects*. The physicists are getting along quite nicely with two inconsistent theories of light and some ten to fifteen elementary 'particles', some of which are of doubtful reality, some postulated on theoretical grounds, and some on empirical grounds unexplained by theory.

The theoretical dynamic hypothesis method faces essentially the same obstacles as those faced by historical analogy or empirical historical relationship methods in the diversity of variables and relationships to be considered and in the variation of measures used to represent them. In addition, it shares with static models the problem of formulating a theoretical structure with precision and testing the significance of the structure itself.[13] The principal difficulties of testing are large variations associated with few observations and the possibility that good agreement with the data is due to the flexibility of the theoretical explanation rather than any causal or empirical significance it may have for the past or for the future. There is no solution for the first difficulty except to await the future or dig patiently in the past, and none for the second except to divide our already limited data into two or more parts, using one as a basis for devising our theories and the others for testing them, and to apply rigid aesthetic standards of simplicity.

Turning to intratemporal methods and omitting simple enumeration, the writer has no coals of fire to add to the heap of cinders now cooling on the brows of the transition forecasters. While such critics as Woytinsky, Hart, Garfield, and others can hardly be said to have made short work of them, since the debate was extended, time appears to have driven home the lesson and the model builders, Klein and Hagen in particular, have accepted its verdict with remarkable grace. They have turned to the task of improvement and refinement. Perhaps the most substantial contribution to this task is Michael Sapir's paper in *Studies in Income and Wealth, Volume Eleven* (NBER, 1949): Review of Economic Forecasts for the Transition Period.

On the whole, the critics were right but when all is said and done, model building remains the ultimate goal of economics if it is to be treated scientifically, for any logical system, if it is self-consistent and accurately descriptive of empirical data, is a model, and these two characteristics are the objectives of rational thought. We are all model builders and in the long run, if economics can become even partly 'exact', we shall probably all end up mathematical model builders. Meanwhile there remains the possibility that mathematical models will have to be broadened in view-

[13] These problems are not peculiar to the methods under discussion but are common to all scientific analysis. Where they do not arise, they have been ignored.

point as well as improved in technical detail. Possibly the mathematical model builders themselves are the best judges of these matters but one gets the impression that techniques are being studied more intensively than the broader and equally difficult problems involved in their useful application.

In general, model building faces many of the obstacles faced by other methods and as these are surmounted, it discloses obstacles shared with and often unrecognized by other branches of theory, together with technical difficulties peculiar to itself. It would be presumptuous to offer an opinion as to what problems should be tackled first, especially when one considers the present state of that obviously important relation the consumption function. Instead, this section will limit itself to a few remarks on models and a few suggestions.

An examination of Colin Clark's model, the only one for which quarterly data and calculations are easily available, gives an impression of great simplicity and considerable accuracy in the period of fit.[14] Applied to the postwar period it fails but the errors on the whole seem to be systematic and, if systematic, should be susceptible to correction if only on an empirical basis. Perhaps they can be, but their elimination does not seem easy. They may be due in part to biases in estimation and if so the methods of Anderson and Rubin, and the results and suggestions of Cochrane and Orcutt,[15] may contribute to their reduction but it is likely that their fit, in the sense of least squares, in the period covered by the data used for the calculation of parameters will be worsened.[16]

If one examines Clark's data on income in the period of fit, 1921-41, out of 9 turning points roughly agreeing with National Bureau reference dates one finds 3 cases in which the calculated values show no turns at all; 1 case in which a turn may be detected but would probably have been ignored by any reasonable test of significance; 1 case that might have been used as a basis of a forecast but with a lead so long as to make it doubtful that it could be regarded as a success; and 4 cases that might reasonably be

[14] A System of Equations Explaining the United States Trade Cycle, 1921 to 1941, *Econometrica,* April 1949.

[15] T. W. Anderson and H. Rubin, Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations, *Annals of Mathematical Statistics,* March 1949; D. Cochrane and G. H. Orcutt, Applications of Least Squares Regression to Relationships Containing Auto-Correlated Error Terms, *Journal of the American Statistical Association,* March 1949, and A Sampling Study on the Merits of Autoregressive and Reduced Form Transformations in Regression Analysis, *ibid.,* September 1949.

[16] It is likely that the residuals will be serio-correlated and the use of this fact might greatly improve average accuracy. It is likely also that accuracy will be worsened at turning points, just where accuracy is most wanted. We shall have to await a trial, however.

regarded as successful on the assumption that all calculated values had been estimated in advance. In short, successes might have been achieved in about 4 or 5 cases out of 9. This is not the whole story, however, because in 2 cases false turns would probably have been indicated.

It seems likely that, had these equations been available in 1920 and had they been relied upon, we would have been generous if we had given ourselves 5 successes and 6 failures for 9 turning points. The analogy with coin flipping is obvious. In fairness it should be stressed that Clark made no claims for his model as a forecasting device and that his article has overtones of modesty.

These remarks are based on a consideration of the period of fit and on the tacit assumption that his autonomous variables could have been perfectly forecasted. Even if unbiased estimation techniques had been used, random variations in the slopes of his hyperplanes and errors of forecasting his autonomous variables would probably have worsened these results in a projection to the postwar period.

Nor is this all. The general magnitude of his errors of fit is about the same as the change in income from one year to the next. That is, if annual data were used, it would be difficult to test the significance of any calculated value that showed a turn from the preceding year, a vital matter if a forecast is to be used in policy making. If data were available and calculations were made on a quarterly basis, this might not be insurmountable but the fact that the Department of Commerce quarterly data on gross national product go back only to 1939 is a handicap. It is reported that the Cowles Commission is extending its work to the study of models using quarterly data and we shall await their results with great interest.

Finally, it has been frequently noted, we are estimating by these methods not gross national product or its components but rather the Department of Commerce estimates of gross national product. Presumably the variance of our final estimates is made up of the variances associated with the methods of the model builder and the variances of those used by the Department. The size of the latter and their behavior through time is by no means clear but the indications are that they are substantial. This vastly complicates the problem, for they may be large enough to be comparable with quarterly, annual, or longer period changes in what we are really supposed to be estimating.

A number of questions arise. Model errors appear to be serially correlated. Are those of the Department of Commerce estimates? What effect would serio-correlation have on our ability to forecast? One suspects that there might be a large adverse effect at turning points but, if we knew enough about these errors, the existence and use of serio-correlation information might be possible.

Relevant to the question of what we are estimating appears to be the behavior of Clark's 'observed' income data (in terms of wage units) relative to the National Bureau reference dates. It is possible, without worrying too much about questions of significance, to identify turns in the general vicinity of the latter dates.[17] In 5 cases out of 9 Clark's figures show turns not more than 1 quarter from the Bureau's dates. In 3 cases they lead by 2 or 3 quarters. One case is ambiguous in the sense that there is about an even choice between a point that coincides and one that lags by 2 quarters.

On the face of it, this is fair agreement but from the viewpoint of policy making we would be generous in assuming that agreement was good enough in 5 cases out of 9. The usefulness of a forecast for policy making depends in part upon the confidence of business men in the forecast. If we could be substantially correct in spotting turns (say within a year) nearly every time they occurred and if we could say they were not coming when they are not, an error of 6 to 9 months in timing might be tolerable now and then. However, if we associate errors of about 6 months with failing to foresee turns that occur and foreseeing turns that do not occur, 4 errors of 6 or more months is too many for policy purposes. This suggests that, in our present state of knowledge, our techniques have got to be very perfect indeed to be very useful.

It is not assumed that the National Bureau dates are necessarily of final significance for business cycle purposes. Rather, the purpose is to indicate that different methods yield different results and that the variances about whatever we are trying to measure are likely to be substantial. The turning points of gross national product and its components may be more important for some purposes and the Bureau dates for others. When Clark's 'observed' turning points differ markedly from those of the Bureau, they seem to lead in every case and this may be a useful fact, if it could be conclusively established. On the other hand, the turning points of the Bureau of Labor Statistics index of factory employment appear to cluster closely about the National Bureau dates, as does the Federal Reserve index, with certain notable exceptions which are perhaps explainable.

Despite these criticisms model building is obviously moving toward the right goals and it has already contributed to the development of theory that is useful in policy making, as well as to the elimination of inconsistencies in theory, to a more precise understanding of the limitations of theory, and to a more precise understanding of the proper weights to be assigned to various economic variables. As Klein and others have pointed out, models can be used to obtain a more precise understanding of the effects

[17] Incidentally, this gives a rough confirmation of the general soundness of National Bureau methods.

of 'structural' (policy) changes; this contribution is not all net gain, since we had techniques for dealing with some of these problems, and its effectiveness has not yet been fully examined. Without ironical intent, it may be added that models are very helpful in evaluating the effects of autonomous changes not only in autonomous but also in endogenous variables. The effects of changes in both autonomous and endogenous variables can often be estimated in terms of inequalities that are sufficient for the purpose at hand. These are substantial gains and small only in relation to what might have been and what yet may be.

The following suggestions are offered in all modesty. I am not sufficiently immersed in the study of models and their construction to have firm convictions as to where the greatest effort should be placed or to be sure that I have even recognized all the principal problems to be solved.

First, it would seem desirable, at least at this stage, to define more precisely what we are trying to do. If we are trying to say something about unemployment in order to contribute to the formulation of sound public policy on this most important matter, could not models be constructed to forecast unemployment with greater accuracy at perhaps the cost of theoretical elegance and accuracy in estimates of other variables?[18] As is well known, a moderate error in estimating gross national product leads to large proportional errors in unemployment estimates, too large, in my opinion, to be of much use for policy making. It may be of critical importance in the future for us to know whether unemployment is going to be less than 5 or more than 7 million. If such an attempt is feasible, the model builders might derive much help from the materials and methods of the National Bureau.

Second, work is already being done on the problem of eliminating estimation bias due to serio-correlation in error terms. An approximate solution at least does not seem to be too difficult and the indications are that it might materially improve estimates and especially projections.[19] A more difficult problem, which may also be a more important one, is the elimination of bias due to intercorrelations between error terms of sets of equations where the parameters are simultaneously estimated.

A perusal of the literature and a fairly superficial examination of the most recent models suggests that work along these two lines, together with the development and use of quarterly data, might contribute materially to the accuracy of estimates useful for forecasting. In addition, there remains the possibility of utilizing empirical serio-correlations or inter-correlations with lags to improve short-term projections for forecasting purposes.

[18] I am well aware that the transition forecasters were trying to forecast unemployment. One gets the impression, however, that the emphasis has shifted.

[19] The Cowles Commission is reported to be studying this problem also.

I have not thought it desirable to offer suggestions on the details of such important work as improving the consumption function, developing investment functions, and introducing expectations. They are mentioned only lest they appear to have been overlooked. In some ways these problems are more important than those mentioned above but perhaps they are also more difficult. In any case they would probably be much easier to deal with if the technical problems of statistical method were solved and if more efficient techniques were generally available and commonly used.

3 USE OF STATISICAL INDICATORS IN FORECASTING

The purpose of this section is to describe a method of forecasting turning points in business, where this term is used to refer to turning points as defined by the National Bureau reference dates. The method may have some interest since it appears to offer the possibility of forecasting in advance turning points of the business cycle with a fair probability of success. In terms of the classifications used in Section 2, it falls in the intratemporal group and relies upon a simple empirical hypothesis concerning the behavior of turning points in individual business series, treated as a frequency distribution.

The method as so far developed utilizes only information specific to the current situation for the purpose of forecasting the next turning point in business, whether a peak or a trough, wasting certain information concerning the intercycle behavior of business series. It is hoped to correct this deficiency and utilize both types of information but attempts along this line are still in an experimental stage.

It appears to be well established that some business series have significant leads or lags on the average from cycle to cycle in the vicinity of turning points and the question arose as to whether or not it was possible to utilize this fact for forecasting purposes. Investigation made it at once apparent that the turning points of individual series show significant average leads or lags but that the variance of these leads or lags is so great from cycle to cycle as to make any single series or simple combination of series useless for forecasting purposes. This made it necessary to cast about for some method that was not dependent upon the behavior of individual series in the vicinity of any turning point.

The National Bureau methods of determining reference dates suggested an alternative. Ignoring a number of important technicalities and qualifications we may describe the Bureau method of identifying peaks and troughs of business cycles in three principal steps:

First, it defines business cycles in such a manner as to restrict the general field of investigation to fluctuations occurring "at about the same time in many economic activities" with a total duration of more than a year and less than 10 or 12 years.[20] With the possible exception of a few doubtful cases such as the peaks in 1918 and 1926, this definition appears to restrict 'business cycles' to fluctuations that are of major importance to business and interest to the forecaster.[21]

Second, business annals are examined in order to discover rough intervals in which peaks and troughs fall.

Third, a mass of information consisting of peaks and troughs in individual series [22] is examined in order to narrow these intervals to (in some sense) an optimum date representing, roughly at least, a clustering tendency in the distribution of specific turning points.[23] While some of the dates so chosen are open to further refinement and some 'cycles' may be of doubtful practical significance,[24] they appear to provide as satisfactory a standard of comparison as is available and one that is accurate enough for most purposes.

The Bureau's work indicated that turning points are not characterized by sudden changes in the majority of series in a short period but that specific turning points, that is, turning points in individual series, are distributed along a time axis in a frequency distribution that tends to show a pronounced clustering tendency about a reference cycle date. If one could find some means of identifying significant turning points currently in the individual series of a representative sample, if one could fit the tail of a unimodal distribution to the early observations, and if one then could calculate the mean or mode of the distribution, one would have a means of estimating the date of the next cyclical turn. This method, however, does not utilize any information about the behavior of specific turning points in business series from cycle to cycle but only the behavior of the frequency distribution of specific turning points in the current cycle.

The first step in the development of an approach of this kind was to investigate the current identification of turning points. Enough has been done to indicate that turning points in individual series can be identified

[20] Arthur F. Burns and Wesley C. Mitchell, *Measuring Business Cycles* (NBER, 1946), pp. 3-8.

[21] *Ibid.*, p. 96.

[22] Called 'specific' turning points (peaks, troughs, cycles), *ibid.*, p. 11.

[23] I can find no explicit statement in *Measuring Business Cycles* that such a clustering tendency is sought, often found, and selected as a reference date as a general procedure, but this seems to be a clear inference from several passages (pp. 6, 7, 77, 80, 83, etc.). Explicit attention is given to cases in which a single clustering tendency is not clearly evident or does not exist, however, and special methods are used in these cases (pp. 81 ff).

[24] *Ibid.*, pp. 80, 90-4.

soon after they occur with a reasonably small probability of error. Generally speaking, the methods used in this connection involve the fitting of a trend to recent data and the establishment of critical limits on each side of this trend for the acceptance of the hypothesis that the trend has changed. Certain improvements in these techniques appear to be possible but after concluding that errors in identifying turns could be reduced to fairly small proportions, the problem has been laid aside for the time being. Thus, the identification of turning points does not appear to be a serious obstacle and errors in selecting dates of turning points will do little more than increase the variance of the final estimate of the date of the next turn.

Setting aside this question and assuming that specific turning points can be identified, with a reasonably high probability of success, shortly after they occur, one selects a sample of various business series as generally representative of economic activity as the data will permit. The next step is to determine the cyclical turning points of the individual series in the sample and to examine the distribution of the turning points through time. More explicitly, one counts the number of series in the sample that show turns in each month and plots these numbers against months. One finds certain months, usually but not always, close to the Bureau reference peaks and troughs about which these numbers show a clustering tendency. Exceptions to this generality will be discussed briefly below.

Limiting himself for the moment to the vicinity of a peak, the observer who identifies peaks in his sample shortly after they occur will see first one or two in a certain month, one or two a little later, then three or four, and so on until a clustering point has been reached and passed, after which the number of peaks successively observed in each month will tend to fall off. The upper part of Figure 1 illustrates this behavior for the individual peaks of an experimental sample of 28 series at the 1923 turning point. This example was selected because it appeared to be reasonably representative of the behavior of the sample and convenient for illustrative purposes.

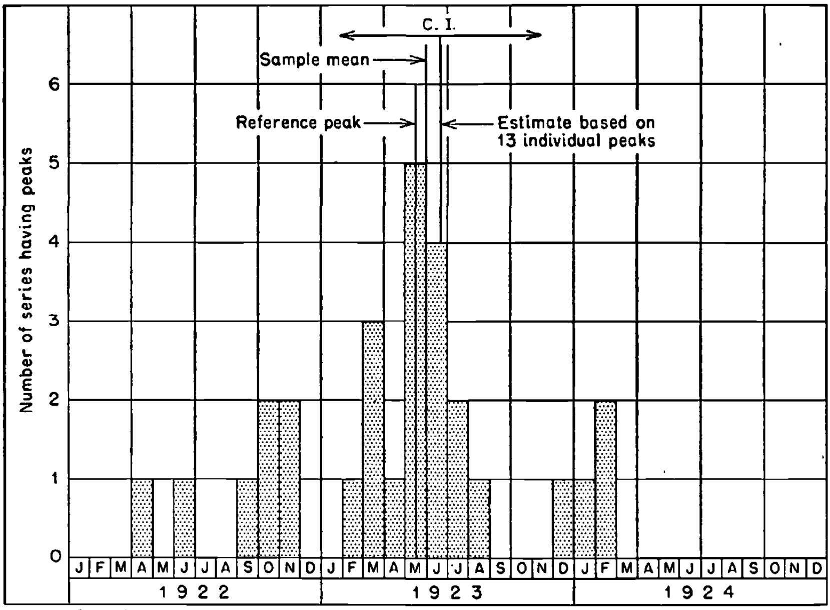
If one considers only the left side of the distribution when, say, only six or seven peaks have been identified, the question naturally arises of fitting some unimodal frequency distribution to these observations and estimating the date of the clustering tendency from the distribution by estimating its mean.

This raises the question of choosing the distribution. Since a distribution used in this way is essentially a special smoothing device and since it is fitted to only one side of the sample distribution, it seems reasonable to suppose that a fairly large class of more or less symmetrical unimodular distributions could be used without greatly affecting the location of the estimated mean. Because tables, numerical values, and asymptotic formulas useful in estimating the parameters of the distribution were available

Figure 1

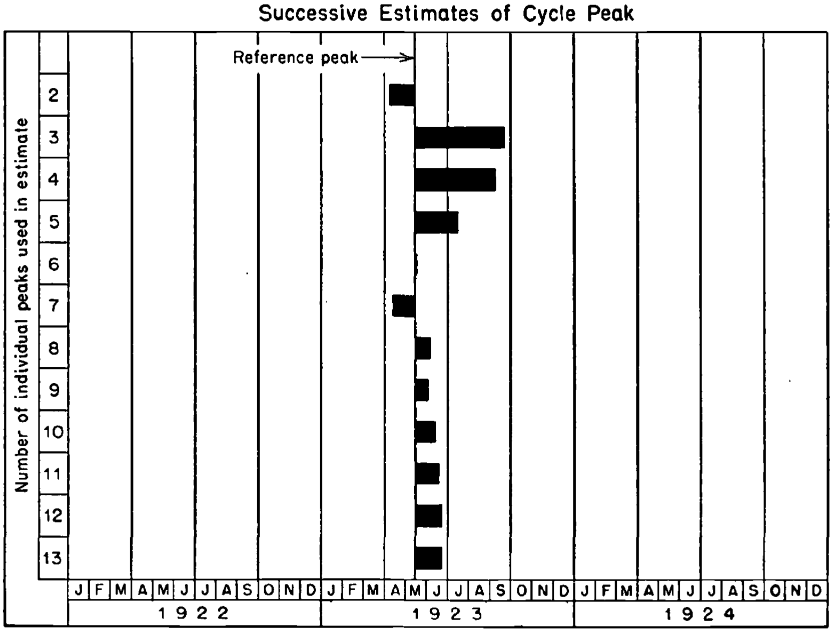
Distribution of Individual Peaks in 28 Series and Successive Estimates (1923 Cycle Peak)

Distribution of Individual Peaks



C. I = 95% confidence interval

Successive Estimates of Cycle Peak



for the normal curve, this distribution was used.[25] The binomial is another obvious choice and still others could probably be used. Possibly some minor improvement in the final results might be obtained by further investigation of this point but substantial changes do not appear likely.

The next step is to devise a method for estimating the parameters of the normal distribution from the truncated sample that is available at any time before the turning point is reached. It is not difficult mathematically to construct various estimates of these parameters but the mathematical problems involved in finding efficient estimates appear to be great. The estimates used to obtain the results described below are easily shown to be asymptotically unbiased but their efficiency remains unknown (see the Appendix). There are intuitional reasons for thinking that the estimates are highly efficient, however, and empirical results tend to confirm this view.

The procedure involved in using this approach in forecasting turns in business activity is then: (1) to select a sample that tends to cluster around business cycle turning points; (2) to watch the individual series of the sample and identify their turning points as soon after they occur as possible; (3) to plot the number of turning points which occur in each month, serially against time; (4) after a few turns have been observed, to fit the normal distribution to the tail of the sample distribution by estimating the parameters (mean and variance) of the normal distribution. The mean so estimated is an estimate of the reference cycle turning point.

In practice after observing a few turning points one revises one's calculations as new turning points are identified. Obviously, when there are only a few observations the variance of the estimated mean is likely to be large. A large variance, say 6 to 12 months, in the estimated mean, implies great uncertainty concerning the estimated date of a business turn and considerably impairs the usefulness of the forecast at the time it is made whether or not it turns out to be correct. To get around this difficulty one can estimate an approximate confidence interval, say at the 95 per cent level, for each revised estimated date and adopt a rule of thumb for deciding when to base a *forecast* upon an *estimate*. The selection of such a rule depends upon the purpose of the forecast, the accuracy needed, and the cost of waiting until enough observations are obtained to reduce the variance of the mean and of the confidence interval to any required size. If one assumes that an interval of 6 to 8 months is narrow enough for practical purposes, one revises one's estimates successively until one obtains a con-

[25] C. Hastings, Jr., F. Mosteller, J. W. Tukey, and C. P. Winsor, Low Moments for Small Samples: A Comparative Study of Order Statistics, *Annals of Mathematical Statistics,* September 1947, pp. 413 ff. See also F. Mosteller, On Some Useful 'Inefficient' Statistics, *ibid.,* December 1946, pp. 377 ff.

fidence interval (using the normal curve) of this size, then bases a forecast on it. The larger the interval one is willing to accept, the earlier one can make an initial forecast but the more inaccurate it is likely to be. The behavior of successive estimates in the 1923 case is illustrated in the lower half of Figure 1.

No great faith or reliance should be placed on such a confidence interval, of course, since the actual distribution of individual peaks may be skew and since certain approximations have to be made in order to calculate it. It is well known that both the size of a confidence interval and the location of its boundary points are sensitive to changes in the shape of the distribution function used. Nevertheless, a rough rule for deciding whether or not a forecast is sufficiently accurate to be useful for policy purposes seems obviously valuable and this rule appears to be sufficient in practice. Confidence intervals, calculated at the 95 per cent level in this fashion for 16 reference peaks, cover the Bureau's dates in 14 cases and fail to cover them in 2 cases, one of which was the atypical 1918 peak where the sample distribution appears to be rectangular instead of normal and begins to develop as early as 1916. Incidentally, a confidence interval of 6 to 8 months assumes that plus or minus errors of 3 or 4 months for a successful forecast of a turning point is about the tolerable limit of error for policy purposes.

Experiments have been conducted with samples, of various sizes and average leads, of business series. Most have been directed to the verification of theoretical methods for improving accuracy and the lead of initial forecasts (as distinct from estimates) and will not be discussed in detail. What these possibilities are will be indicated later. The following paragraphs will discuss briefly the results of an initial experiment with a sample of 28 business series. The discussion will be in terms of reference peaks and peaks in individual items of the sample.

The experimental sample of 28 series was originally selected for the purpose of using the fact that certain series show significant leads in the vicinity of the Bureau reference dates. The only systematic study of leads known to me at the time these experiments were initiated was that of Mitchell and Burns.[26] The authors compiled a list of series having certain characteristics, including that of reliable average leads at lower turning points. The experimental sample was based on this list with some substitutions. Since the list was biased in favor of series that lead at the lower turning point, it was expected that on the average specific peaks of all 28 series would cluster about a point that would lead at least at troughs. This turned out to be the case. It was similarly expected that a lead would

[26] Wesley C. Mitchell and Arthur F. Burns, Statistical Indicators of Cyclical Revivals, *Bulletin 69* (NBER, 1938).

show up at cyclical peaks but that it would be shorter since the series were selected to lead at troughs, not at peaks. This also turned out to be the case and the lead was so small as to warrant treating the sample as unbiased at peaks, and this was done.

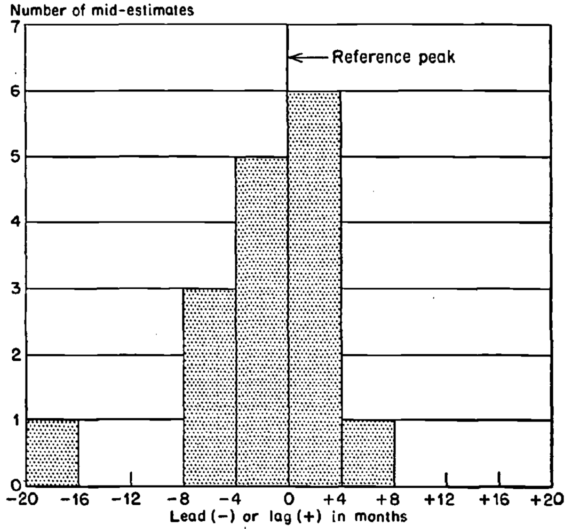
To check the accuracy of the method, estimates were made for each cyclical peak (and trough) for the 16 cycles beginning with the peak of 1882 and ending with that of 1937. Since the 28 series vary in length, it was necessary to base the calculations on enough observations and to start with a late enough cycle so that enough observations would be available to get a reasonably stable estimate of the turning point. For the purposes of this test it was decided to calculate what were called 'mid-estimates', that is, estimates based on the first half (or one less than half when the number was odd) of the series available at each turn that show specific turning points. The number of series used in each calculation declines as one gets back to the earlier turns until the estimates are based on only four or five observations.

Figure 2 shows the frequency distribution of peak 'mid-estimates', as defined above, about their respective reference peaks. Of 16 peaks, 11 are within 4 months or less of the reference peak and 4 lead the reference peak by 4 or more months. The latter were estimates for 1893, 1907, 1918, and 1926 and are regarded as 'errors'; 1893 and 1907 are so classified on the pragmatic ground that errors of more than 4 months were too great for policy purposes. It seemed likely that one or more 'errors' so defined could be attributed to sampling fluctuations and this was confirmed by the use of a larger (and somewhat modified) sample. 1918 is an obvious failure from any point of view. Peaks began to develop in individual series as early as 1916 but the entry of the United States into World War I appears to have been accompanied by a 'stretching out' of the distribution of peaks until it was approximately rectangular. The 1926 'error' may be attributable to sampling fluctuations but this is not likely. There is evidence that the turning period in this case partook of the nature of a plateau rather than a clearly defined peak. This implies a breakdown in the assumptions used in estimation and a failure to get a good estimate through the use of the normal curve.

If 1918 is excluded, the distribution is not very different from what one would expect by chance, and 95 per cent confidence intervals cover all but one of the Bureau reference peaks. Even if 1918 is included, failures of the confidence intervals to cover their corresponding peaks do not differ significantly from their expected number and the average lead or lag of the estimates does not differ significantly from zero.

If we accept 4 failures, the distribution of mid-estimates suggests that the estimation technique can provide sufficiently accurate estimates of

Figure 2
Distribution of Mid-estimates



business cycle peaks about three times out of four, based on the 28 series used. It seems likely that a larger sample would provide greater accuracy.

Estimates based on about a quarter of the observations for 5 cycle peaks from 1920 to 1937 give substantially the same results but the cycles are too few to be of much significance.

Mid-estimates for troughs, based on the same sample, showed a significant tendency to lead reference troughs by about 3½ months, as was to be expected from the nature of the sample. If a correction is made by adding this amount to each mid-estimate, the estimates at troughs show about the same proportion of 'successes' and 'failures' as they do at peaks.

The above considerations are pertinent to an examination of successes and failures, in the vicinity of turning points, of the techniques of estimation, as distinct from the making of forecasts. To evaluate the latter, it is necessary to determine also whether the techniques would have indicated turning points when in fact none occurred. In addition, one must consider the time necessary for the publication of data, the identification of specific turning points in individual series, and the accumulation of enough observations on the truncated curve to assure reliable estimates. If such estimates can be made only after the event, the method might have diagnostic but not prognostic value.

An examination of the chronological behavior of specific turning points indicates 3 cases in which peaks not classified as reference peaks might have been indicated: in 1877, 1933, and 1947. In these cases the behavior of the sample might have led to estimates that would have been regarded as cyclical peaks if only the data provided by the estimates had been considered and all other information excluded.

The first two of these 'false' peaks followed short-lived upward movements from the bottoms of the two long and severe declines 1873-79 and 1929-33, the 1933 false peak being accompanied by a subsidiary peak in the Federal Reserve Board index of industrial production. It will be recalled that a spasmodic upsurge in many series followed the bank holiday in 1933 and that the movement appeared to be largely speculative at the time. It is thought that resort to very little evidence outside the scope of the method under discussion would have identified these peaks for what they were. If so, the forecaster would have identified these turns accurately but would have added that they ought to be regarded as minor, though identifiable, turns, not cyclical peaks.

1947 offers more difficulty since a turn was indicated (and occurred) during prosperity. The estimate was accurate but a forecast would have been wrong, not because a turn failed to develop but because the ensuing recession did not go far enough to be classified as a cyclical decline or to be of practical importance. While this sort of thing may happen, there appears to be no evidence that it happens often.

To obtain some idea of the forecasting lead to be expected from the technique here described, it was assumed that 5 months would be necessary for the publication of data and the identification of peaks in the individual series of the sample, and that about 7 peaks on the average would give useful estimates. These assumptions, while rough, were not entirely arbitrary but were based on a detailed study of the sample.

It was found on these assumptions that: (a) the distribution of dates on which initial forecasts could have been made was skewed sharply in the direction of leads, as was to be expected; (b) that three initial forecast dates including that for 1937, the distribution of which was closely concentrated (small standard deviation), followed their reference cycle peaks by less than 2 months; and (c) that the average lead of the initial forecasting date was about 2 months. Not much faith was placed in these results but they were considered sufficient to indicate the desirability of improving the lead of the initial forecasts.

There seemed two principal ways of doing this. The first was to select a sample that was biased in the direction of leads and to find a way of correcting for the bias. This would use part of the information concerning the inter-cycle behavior of lead series, information neglected by the method applied to the 28 series. Samples of about 50 and about 80 series were constructed and some gains were made along these lines, but it was found that limitations in the number of series that were available and regularly and promptly published necessitated the acceptance of series with shorter and shorter leads as the sample was enlarged. In other words, the mean of the sample turning point distribution moved toward the reference turning point, and the advantage of using a lead sample was lost as its size increased.

The second principal way in which it was hoped to increase the lead of initial forecasts was to enlarge the sample. As a sample increases the probability of obtaining observations far out on the tail of the distribution of specific turning points increases. It was hoped, therefore, to obtain enough observations from a larger sample to improve the leads of initial forecasts materially. Theoretical calculations applied to data from the 28 series suggested that initial forecasting leads, estimated as indicated above, could be lengthened to about 8 or 10 months with a distribution highly skewed toward longer leads. Experiments with the sample of 80 series almost exactly confirmed these theoretical expectations, and this tended to strengthen confidence in both theory and method.

Work has not progressed sufficiently far along these lines to enable one to draw any final conclusions concerning the leads of initial forecasts obtainable. At present two improved, and partly overlapping samples, one for peaks and one for troughs and each of about 100 series, are being constructed from a set of series screened for conformity to cycle patterns and classified with respect to lead.[27] Experiments are also being conducted with a view to dividing these two samples into a series of subsamples, each with an average lead with respect to reference turning points. If this can be done, as appears likely, it will be possible to use information concerning the inter-cycle average leads of individual series in combination with information specific to the current situation. The method is to estimate the mean for each subsample, correct this estimate to the reference turning point, and construct a weighted average of the subsample estimates so corrected. Should this method fail to improve the lead of initial forecasts with given accuracy, it will still be possible to treat each of the two samples as a single distribution. As stated in the preceding paragraph, theoretical considerations applied to data obtained from the earlier samples indicate the possibility of a distribution of forecasting leads averaging about 8 to 10 months and skewed in the direction of longer leads.

In short, after allowing for the fact that a larger sample will reduce

[27] This would have been impossible but for the courtesy of Geoffrey H. Moore in making available unpublished results of extended studies of his own. The most important of these have since been published in Statistical Indicators of Cyclical Revivals and Recessions, *Occasional Paper 31* (NBER, 1950). I welcome this opportunity to acknowledge my indebtedness to Mr. Moore and the National Bureau's staff for their invaluable assistance.

somewhat the number of 'errors' as defined above and for the fact that a few peaks, not significant for practical purposes, will be forecast, it seems possible to develop a method along these lines that will forecast turns 'correctly', i.e., within plus or minus 4 months, about 3 times out of 4 with a lead of initial forecast long enough to be useful.

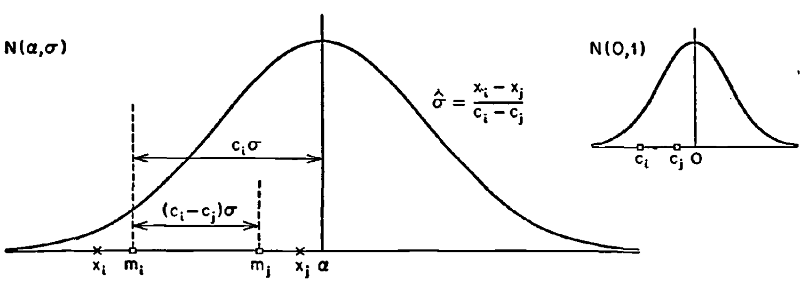
Appendix

Estimation Methods

If α is a random variable and X_1, X_2 - - - X_n are the observations of a sample of n observations *arranged in order of size*, X_i is said to be the i-th order statistic of the sample.[28] The date of the first specific turn observed in the sample frequency distribution of the text may be regarded as the first order statistic, the date of the second turn as the second order statistic, and so on. A problem of grouping arises when two or more turns fall in the same month but this can be handled with a small loss of accuracy by distributing the peaks throughout the month in which they occur.

It is assumed that the sample is from $N(\alpha, \sigma)$ a normal population with mean α and standard deviation σ. Let $m_{i}/_{n}$ (m_i) be the expected value of the i-th order statistic in a sample of size n from this population (Fig. 3).

Figure 3
Estimation of Standard Deviation



Let $c_{i}/_{n}$ (or c_i) be the *expected* value of the i-th order statistic in a sample of size n drawn from $N(O, 1)$, a normal population with zero mean and unit standard deviation[29] (Fig. 3). Then:

$$m_i = \alpha + c_i\sigma$$

$$m_j = \alpha + c_j\sigma; \text{ and}$$

$$m_i - m_j = (c_i - c_j)\sigma$$

[28] Cf. note 25. See also S. S. Wilks, *Mathematical Statistics* (Princeton University Press, 1943), p. 89.

[29] Exact values of $c_i/_n$ for $n \leq 10$ and an asymptotic formula for $c_i/_n$, $n > 10$ are given by Hastings, Mosteller, Tukey, and Winson, *op. cit.*

$$\sigma = \frac{m_i - m_j}{c_i - c_j}$$

Since X_i and X_j are distributed about m_i and m_j, their expected values respectively, we may substitute the first for the second to get an *estimate* $\hat{\sigma}$ of σ:

$$\hat{\sigma} = \frac{X_i - X_j}{c_i - c_j}$$

where the X's and c's are known.

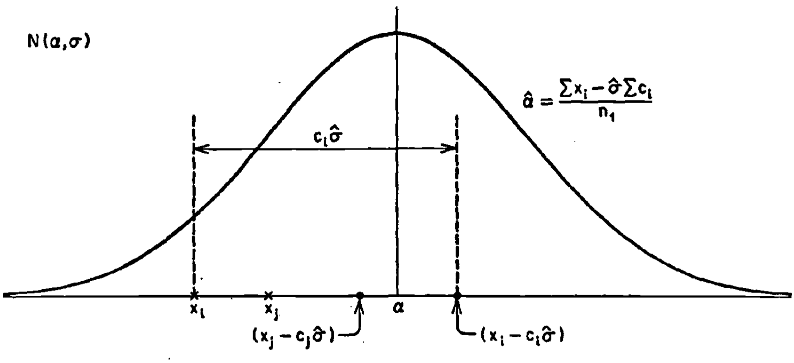
By forming differences between the observations in this way many estimates of σ may be obtained, then combined in many ways to improve the estimate of σ. The particular estimate used in this study was:

$$\hat{\sigma} = \frac{\Sigma_1 X_i - \Sigma_2 X_j}{\Sigma_1 c_i - \Sigma_2 c_j} \tag{1}$$

where Σ_1 and Σ_2 indicate summations over the first and second halves of the (ordered) observations. (When the number of observations is odd, one may be omitted and the observations re-ordered.) This estimate is easily shown to be asymptotically unbiased. It is not 'efficient' but its 'efficiency' is probably high and it is simple to calculate.

Figure 4

Estimation of Mean



Since $m_i = \alpha + c_i\sigma$,

$\Sigma m_i = n_1\alpha + \sigma\Sigma c_i$, where n_1 ($\leq n$) is the number of available observations. Then:

$$\alpha = \frac{\Sigma m_i - \sigma\Sigma c_i}{n_1}$$

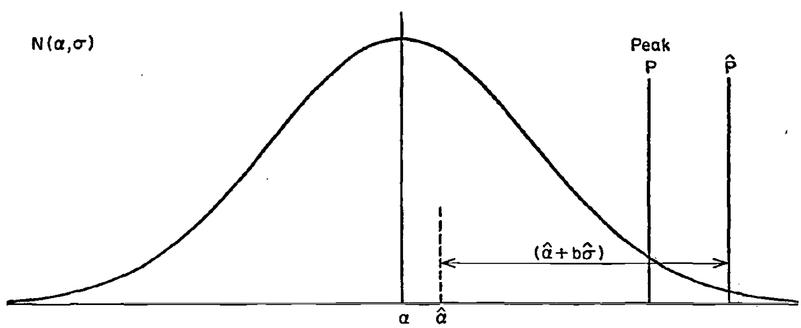
Substituting X's for m's as above and substituting $\hat{\sigma}$ as calculated from formula (1), we get

$$\hat{\alpha} = \frac{\Sigma X_i - \hat{\sigma}\Sigma c_i}{n_1} \tag{2}$$

as an estimate of the mean. In effect, this formula displaces each observation toward the mean by the amount of its estimated expected value, then averages the results (Fig. 4). This estimate also is asymptotically unbiased and its efficiency is believed to be high.

When samples with a significant average lead are used, the estimated sample mean is corrected to bring it to the vicinity of the reference turn by adding a correction calculated from the regression of the inter-cycle lead of the sample mean, relative to reference dates, on the estimated standard deviation $\hat{\sigma}$ (Fig. 5).

Figure 5
Estimation, 'Biased' Sample



COMMENT

MILLARD HASTAY, *National Bureau of Economic Research*

The aspects of Mr. Wright's researches with which I am personally familiar are set out in Part III of his paper, and it is to this Part that I shall confine my comments.

I think that every statistician will find intriguing the application of order statistics which Mr. Wright proposes here. The very conception of an ordering performed in advance and invariant to subsequent events is novel. In ordinary applications of ranking, one must assemble his sample and determine the order-identification of separate items after all are known; in Mr. Wright's case, the order of an observation is determined simultaneously with its magnitude and cannot be affected by subsequent developments. This is the essential fact of his method which makes a forecasting procedure possible. It is not, of course, a mysterious fact because it follows naturally from his interest in time itself as a variable. But neither is it an obvious fact, and Mr. Wright is to be credited with one of those original and fruitful insights which serve to extend to seemingly unrelated problems the basic techniques of statistical estimation.

In putting this insight to work Mr. Wright has had to make compromises of both a theoretical and practical kind; and one of these, the use of the normal or Gaussian distribution as a model of the empirical clustering of turning points, may need fuller justification than he has given it. For let us recall the usual conditions for expecting a normal distribution to emerge. These may be formulated as a single dominant factor which tends to establish the central tendency of a distribution, operating together with a multiplicity of small, independent factors — no one of them dominant — which determine the clustering of items about this center. In such a milieu, the order or rank of any particular event (specified somehow without regard to its magnitude) is a chance variable and will differ unpredictably from one experiment to another. Is this a fair approximation to what we expect in the ordering of turning points in different time series? Most business cycle theorists would, I think, say no. They believe — and their researches could hardly be justified otherwise — that there is a genuine sequence in the timing of various processes; that orders for industrial equipment, for example, tend to turn down early while consumption of nondurable goods usually turns down later, and that this pattern is repeated in most, if not all, cycles. The sequence of particular turns is determined, therefore, not by a multiplicity of small, independent causes but by a complex system of interdependent causes tending to produce a recurrent pattern of timing relations in which the early downturns are taken by roughly predictable processes and the later ones by others not less predictable.

I think that Mr. Wright would answer that in using normal distribution theory, he is not endorsing any particular model of business cycle causation. For him the normal distribution serves as a purely formal construction by which he seeks to approximate certain empirical distributions which he has observed in the past and expects to emerge in the future. If the approximation suffices, his prediction mechanism will give valid results; otherwise it will not. His problem is therefore to demonstrate by empirical tests that the formal model is a sufficient approximation for his purposes; and it is this attempt that gives chief significance to the detailed experiments which he presents in his paper. How conclusive these experiments are is a matter on which disagreement is likely. Unfortunately, the burden of proof that such experiments must bear is necessarily greater when a model is proposed merely on formal grounds than when general economic considerations can be adduced to support its applicability.

RUTH P. MACK, *National Bureau of Economic Research*

Mr. Wright's paper dangles the threat of good forecasting before our eyes. This invites consideration of what would happen were it actually possible to anticipate turns in business.

We know a little about how business men arrive at prognostication of the future, but virtually nothing, and I am sure Mr. Wright would be the first to agree, about how the guesses are actually used. To my mind, the second question — how forecasts are put to work — is the more important of the two. For market behavior could easily be more significantly influenced by differences from time to time in the degree of confidence with which a given expectation is held than by variation in expectations proper, other things the same. Moreover, not only for the understanding of how much, but also of *when,* expectation-based activity will be undertaken, very earthy knowledge of how forecasts are used is necessary. We lack such knowledge.

The little work that I have done touching this matter makes me wonder whether, if business men felt confident that they could actually forecast cyclical troughs and peaks, they might not flock to order plant and equipment only sufficiently prior to when the plant was expected to get really busy (and/or equipment cost expected to rise materially) to have the installations operating efficiently by the time things really started to hum. Most typically this moment would follow, not precede, the turn in sales and thus accentuate a rise rather than terminate a fall in output. The best time to cease equipment buying, on the other hand, might often be quite a long while before a confidently expected peak. In connection with speculative buying of stock in trade, the prospect of good forecasts of the price of materials as well as of the volume of demand raises questions too. For one thing, the amount of speculation that is undertaken is certainly a function of the confidence with which forecasts can be made, and the forecast of price is probably the more critical. Speculation of this sort seems to move in shorter waves than the forty-month cycle. It affects, as well as depends on, prices.

I urge, in other words, that efforts to improve our technology of forecasting be paralleled by study of how forecasts are used by business men. Otherwise we may find ourselves threatened by a new type of technological unemployment.

GEOFFREY H. MOORE, *National Bureau of Economic Research*

I have a few remarks to make on Section III of Mr. Wright's paper, which presents his ingenious method of forecasting business cycle turning points.

First, the assumption that the distribution of turning points in a group of series is akin to the normal distribution of a random variable seems dubious, because of the interrelationships among the economic activities represented by the series. The process by which a turning point in the aggregate economic activity of the country is reached in a particular month, when one views that process as a sequence of turning points reached by different components or aspects of the aggregate, is not, I think, analogous to the way random shots distribute themselves on either side of a target. Indeed, the interrelationships seem to me to be the essence of the matter; I cannot conceive of a business cycle without them. Possibly it makes little practical difference how one regards a frequency distribution to be generated, but I should like to be convinced of that by a demonstration.

I think Wright's forecasting method can be expressed as follows. During an expansion of business activity, when one is interested in forecasting the position of the next peak, the sample of series is examined for peaks. When a certain number of series have turned down, one is in a position to make the forecast. Let us say this happens in month x; that is, it is now month x and peaks can be recognized in, say, 25 of the 100 series. Wright's proposition is, as I understand it, that if these 25 peaks are spread out over a large number of months in the past, the general business peak will follow month x by a considerable interval; whereas, if the 25 peaks are clustered close before month x, the business peak is near at hand. In the one case the estimated variance of the distribution will be large, in the other, small.

It should be possible to test this proposition directly, and perhaps Mr. Wright has. It is not obviously true, though there may be reasons to expect it. It seems to me that the downward pull on the economy exerted by the 25 series that have turned down depends largely on what series they are and how fast they are going down, and very little on how far back in the past they began to fall. In fact, one might believe that the longer their average period of fall before month x the greater their depressive influence; downturns in the next 25 series might be expected to follow sooner rather than later. Thus the correlation between the pre-forecast variance and the forecast interval may be small, and it might even be negative.

The point is, is the forecast improved significantly by estimating the variance anew for each turning point, using the evidence provided by one tail of the distribution? If not, why not use a single estimate based on a set of distributions? The procedure would certainly be simpler.

My second remark is that I am a bit troubled by Wright's confidence that by enlarging his sample, his forecasting lead can be lengthened. Does that not depend on how the sample is enlarged? If it is enlarged by adding series that have a less regular relation to business cycles, which is almost inevitable since one is bound to try to choose the best series first, the 'quality' of the sample will deteriorate. In other words, although enlarging the sample may tend to push the tail of the distribution of turning points further out, it may also increase the variance of the estimate; while the former effect is an advantage, the latter is a disadvantage, and one must be weighed against the other.

Third, I have some doubts about our ability to identify cyclical turns in individual series, with only a brief lag. Not having seen how Mr. Wright's method works in practice, I cannot criticize it; but if he has solved the problem it will be an extremely useful contribution. One matter that he will have to deal with is the possibility of making multiple errors. That is, because of inter-correlation among series, he may identify downturns in several series, all of which prove abortive. I imagine this was what happened in 1947.

In short, I see some limitations in the approach. But while these may reduce, they do not destroy its usefulness, and I think Mr. Wright's endeavor to utilize a significant and hitherto neglected feature of business cycles is most promising. I hope he will continue and expand his experiments.

REPLY BY MR. WRIGHT

The comments of Mr. Hastay, Mrs. Mack, and Mr. Moore appear to raise technical and methodological questions on the one hand and questions for further research in other but allied fields on the other. The commentators have been so courteous that if any criticisms are made they refer to omissions rather than mistakes. If such criticisms are implied, I find myself agreeing in principle with the critics. However, it may be possible for me to throw some light on points that appear to have been obscure in my original discussion.

If I understand them correctly, Messrs. Hastay and Moore are troubled by the fact that my method of forecasting uses the normal probability curve when the data to which it is applied are not subject to a multitude of small random forces. It is natural to have doubts on this point because the normal curve can be and has been derived by postulating a large number of small random forces determining the detailed results of any sampling experiment. Such a postulate is intuitively plausible in many fields, the classical example being the analysis of the results of many refined and accurate physical measurements. It is now known, of course, that even here skew distributions are not uncommon.

I do not think that this consideration should raise any considerable doubts concerning the use of the normal curve as I have used it. My method merely postulates that the form of the distribution through time of turning points of the individual items in a sample of business series is unimodular and can be approximated by the normal distribution. The curve is used essentially as a smoothing and extrapolation formula. It makes no difference for this purpose, always provided the normal curve gives a good fit to the data, whether the turning point in any particular series is randomly determined relative to the reference turn or whether it always tends to fall in the same general section of the distribution. Order relations between turning points of particular business series are irrelevant to the problem of estimating the center of the distribution. This, of course, has nothing to do with the use of order statistics in the estimation problem.

Perhaps my point can be intuitively clarified by an example from another field. Conic sections may be derived as the path of a particle, in a gravitational field, from Newton's law of gravitation, but this does not require the statistician to postulate or to demonstrate the existence of a gravitational field in order to justify the use of, say, a parabola for smoothing or extrapolation purposes.

While my method is not affected by the existence of timing patterns specific to individual series, the utilization of such timing patterns might, in theory at least, improve it as a forecasting device by utilizing relevant information now ignored. I have experimented along such lines and the possibility remains open. However, the random element in timing patterns appears to be so great relative to the stable element in timing that I do not anticipate much improvement in this direction. The postulate on which my method of estimation is based occasionally breaks down in particular cycles. For example, the distribution of peaks in the vicinity of the National Bureau 1918 reference peak turns out to be practically rectangular. In these cases the method fails and I have counted such failures in estimating the long run probability of succeeding and failing by the use of this forecasting method.

It seems practically certain that other unimodular distributions could be substituted for the normal distribution, possibly with some improvement in forecasting turning dates accurately. The binomial distribution is an obvious candidate. I do not think, however, that improvements along these lines are likely to be great. Any forecast by these methods turns out to be essentially a mean and this is not likely to deviate from the mode of the sample distribution, nor will either of these depart materially from reference cycle turning points except in cycles where my specific postulate breaks down.

All this explains why I am not very worried about rationalizing the use of the normal curve in terms of random economic variables. My postulate is essentially an empirical one which can be examined against the data from cycle to cycle. It does not appear to me to be necessary to postulate random forces. However, if it were necessary to make such a postulate, my feeling is that it would be near enough to the truth to give fairly accurate results. I agree with Mr. Moore that the order in which the particular series turned is not random. There are significant leads and lags in the turning points of individual business series. I also agree with Mr. Moore that there are inter-correlations among the series of my sample. A detailed study of timing relations makes me feel that their random aspects 'swamp', so to speak, their nonrandom aspects. To some extent the same can be said about correlations among the turning points of the individual series in my sample. If these had substantial inter-correlations and if this fact were ignored, I should, in effect, have fewer degrees of freedom than my estimation methods assume. In so far as this is the case, the accuracy of my estimates is impaired. I have evaded this problem in some degree by discarding from my sample series that are obviously related to other series that have been retained. No doubt there remains an inter-correlation effect which is theoretically a flaw in the method but I can judge its importance along with other sources of error only by looking at the accuracy of my final results. This accuracy appears fairly (and only fairly) high, so I infer that this and other sources of error have been reduced to moderate proportions.

Significant leads and lags in the timing of turning points in individual series and the existence of inter-correlations among series provide information that is presumably relevant to the problem of forecasting turning points. This information is ignored by my method. Perhaps someone can devise a superior technique which combines the information I have used and that which I have ignored.

Mrs. Mack's point that variations from time to time in the degree of confidence with which a given expectation is held might be more influential than variations in the expectation itself is well taken. I am in entire agreement with her concerning the desirability of tackling the great obstacles that surround the development of a realistic theory of expectations and the degrees of confidence associated with them. These are questions of great importance in a wide sphere of economic analysis. My feeling is that from this viewpoint the field of forecasting is a narrower and much humbler one, and that it will be a long time before forecasting techniques can be made sufficiently accurate to gain a widespread influence in determining either expectations or the degrees of confidence with which they are held. Even if the theorist could develop a methodology that warranted his confidence, the business man would require empirical verification of his theories, and empirical post-verification of theories pertaining to the '40-month cycle' is a time-consuming process.