Volume Title: Econometric Models of Cyclical Behavior, Volumes 1 and 2

Volume Author/Editor: Bert G. Hickman, ed.

Volume Publisher: NBER

Volume ISBN: 0-870-14232-1

Volume URL: http://www.nber.org/books/hick72-1

Publication Date: 1972

Chapter Title: Introduction and Summary to "Econometric Models of Cyclical Behavior, Volumes 1 and 2"

Chapter Author: Bert G. Hickman

Chapter URL: http://www.nber.org/chapters/c2780

Chapter pages in book: (p. 1 - 22)

INTRODUCTION AND SUMMARY

BERT G. HICKMAN · Stanford University

IN ONE sense the title of this Conference is a misnomer. With a single exception, the authors included herein did not set out deliberately to construct econometric models which are mathematical versions, or translations, of particular business-cycle theories. Rather, their purpose was to specify and quantify empirically valid behavioral hypotheses about the decisions and actions of various economic agents, and to integrate the estimated relationships into a complete system capable of determining the values of all the current endogenous variables for known, or assumed, values of the predetermined variables. Once built, the dynamic properties of the resulting models are a proper subject for investigation, but the models were not intended to test any specific theory of the cycle and, indeed, were constructed primarily for short-term forecasting and policy analysis.

In another sense, however, the tests of dynamic properties reported at the Conference are natural extensions of the mathematical literature on business fluctuations. Thus, it is well known that self-contained systems of linear difference equations, such as those specified in multiplier-accelerator and similar analytical models, may or may not have cyclical (complex) roots, depending on the numerical values assumed for the parameters. Moreover, such systems may, or may not, exhibit maintained cycles in the absence of stochastic shocks. Since modern econometric models are both large and nonlinear, it is usually necessary to resort to simulation techniques rather than direct mathematical analysis of the equation systems, but the aim of the simulation exercises is to answer the same questions about dynamic properties as may be asked of the simplest multiplier-accelerator model.

The four papers on business-cycle simulations in Part I were planned as a unit. The prototype for this set of complementary papers was the seminal 1959 *Econometrica* article: "The Dynamic Properties of the Klein-Goldberger Model," by Irma and Frank Adelman.

Just as in the Adelman study, the dynamic properties of the contemporary generation of macroeconometric models are studied in both a deterministic and stochastic context by the use of simulation techniques. The scale of the present effort is much greater, however, because more and bigger models are studied, the stochastic assumptions are more highly elaborated, and the simulation results are subjected to a larger battery of tests. The entire enterprise was feasible only because of the cooperative endeavors of the several model-building groups and the NBER team of cycle analysts.

In the first three papers of Part I, the OBE, Wharton, and Brookings model-builders describe and analyze stochastic and nonstochastic simulations prepared according to a common plan. The original program had called, also, for simulations to be done by the FMP (FRB-MIT-Penn) and the NBER (Chow-Moore) Models, and for all five sets of simulations to be turned over to researchers at the NBER for an independent analysis of their cyclical characteristics. It was recognized from the outset, however, that the last two models might still be in the developmental stage when the deadline for submission of the simulations for NBER cycle analysis was reached. As it turned out, the FMP group completed some of the simulations in time for the NBER analysis but were unable to complete the entire set, or to prepare a paper for the Conference, whereas Gregory Chow and Geoffrey Moore finished their model in time to describe it in a paper included in Part II, but not early enough to do the simulations. Finally, the Brookings group completed all simulations and submitted a paper discussing them, but again not in time for inclusion in the NBER analysis by Victor Zarnowitz, Charlotte Boschan, and Geoffrey Moore. The final outcome is that the NBER analysis deals only with the OBE, Wharton, and FMP Models and is incomplete in its coverage of the latter. Even so, it is a comprehensive and stimulating analysis, and the NBER plans to continue applying the methods developed in the paper in its future work on model-testing and evaluation.

Apart from the charge to supply simulations on a common scheme to the NBER, the model-builders were left free to structure their own analyses as they wished. Thus, the papers on the OBE, Wharton, and Brookings Models include materials that augment and complement the independent analysis of the NBER group. These papers can, and should, be read as self-contained and highly informative pieces on the structures and cyclical properties of the several models, as they existed circa 1967–69. For the purposes of this Introduction, however, it will be more enlightening to discuss certain common features of the simulation studies than to attempt an independent commentary on each paper.

The first point to emphasize is that the models are dynamically stable when treated as deterministic systems. This feature emerges clearly from the control solutions used for the long-run postsample simulations of the OBE, Wharton, and Brookings Models. For the most part, the exogenous variables were projected to follow smooth trends over the twenty-five-year interval utilized in these simulations. (The purpose was to provide a control solution for comparison with the stochastic simulations, of course—not to make an unconditional forecast of U.S. economic performance in the next quarter century.) Given the initial conditions at the beginning of the simulation period and the time pattern of the exogenous variables, the models were solved for the endogenous variables, which were found, in general, to follow trendlike paths rather than fluctuations. It may be objected that this result would follow even for an unstable model structure if the initial conditions happened to be consistent with the equilibrium solution. This would occur only by accident, however. In the case of the Wharton and Brookings Models, moreover, temporary perturbations are introduced in some of the exogenous and policy variables early in the simulation period to reflect an assumed Vietnam settlement during 1970–71. The economic response to this disturbance is perceptible but highly damped in both models. This is partly because effective compensatory policies are assumed, but it also reflects the stable dynamic structures of the models.

Further evidence on the point is found in the deterministic long-run simulations for the sample period. These are complete model-solutions, using the actual values of lagged endogenous variables at the beginning of the sample period as initial conditions, and setting the exogenous variables at their actual values throughout the period. It should be noted that these ex post simulations are rather stringent tests of the model structures, despite the use of observed values for the exogenous variables, and of parameters estimated from the same data that is being predicted, since errors can cumulate both across equations and, through the influence of the lagged endogenous variables, over time. Thus, "even if the model were perfectly specified, the neglect of stochastic elements would in itself give rise to errors which, due to the presence of lagged terms, would be carried forward," as George Green, Maurice Liebenberg, and Albert Hirsch note in the OBE paper.

In the course of summarizing their examination of the continuous sample-period simulations, Zarnowitz, Boschan, and Moore observe:

> Each of the models shows the economy . . . as declining during the first recession period covered (1948–49 for Wharton, 1953–54 for OBE, and 1957–58 for FMP). . . . The three models also have GNP^{58} contracting, or at least flattening out during the contractions in 1953–54, 1957–58, and 1960–61, respectively. The Wharton Model does not produce a fall in GNP^{58} during the recession of 1957–58, and neither the Wharton nor the OBE Model produces one in the 1960–61 recession. . . . The important conclusion is that there appears to be a progressive dampening of the fluctuations the further away a model's simulation proceeds from its initial-conditions period.

This same generalization might be made about the Brookings Model sample-period simulation for 1957–65, as reported by Gary Fromm, Lawrence Klein, and George Schink, in which "real *GNP* does follow the 1957–58 period, but fails to decline in 1960–61."

On one interpretation, the tendency toward progressive dampening could be attributed to damped cyclical roots. This may indeed be partly valid, but it cannot be the sole explanation. The cyclical roots are quite weak, to judge from the nonstochastic postsample simulations already discussed. The better cyclical performance of the models during the sample period must in large measure reflect the different treatment of exogenous variables. These variables are entered at their actual values in the sample-period simulations, where their movements represent an external source of disturbance to the model systems. The Wharton Model actually does better in reproducing the contraction of 1953–54 than that of 1948–49, despite the fact that the initial conditions had been left far behind by 1953. This superior performance during 1953–54 doubtless occurs because a sharp reduction in military spending was a prominent feature of that contraction, and such spending is exogenous in the Wharton Model, as in the others. Similarly, the OBE Model benefits from the exogenous fluctuation of government spending in 1953–54, so that its better performance in that recession than in those of 1957–58 or 1960–61 may be due as much, or more, to exogenous factors as to the proximity of the initial conditions.

Whether or not the models should be interpreted as endogenously generating progressively damped cycles in the deterministic sample-period simulations, there appears to be little doubt that they are highly stable as specified. If it is provisionally assumed that the models are correctly specified, cycles could nonetheless result from a damped cyclical response-mechanism which was kept going by erratic shocks, as in the Wicksell–Slutsky–Frisch theory. This hypothesis was extensively tested in the stochastic simulations prepared for the Conference by the OBE, Wharton, and Brookings groups. The general procedure was to generate random shocks to be applied to the endogenous behavioral variables in the long-run postsample simulations. The shocks were generated by a method developed by Michael McCarthy, and are explained in the Appendix to the paper on the Wharton Model by Michael Evans, Lawrence Klein, and Mitsuo Saito. The method allows for intercorrelation of the errors in different equations and, in some applications, for serial correlation in the errors of individual equations. In these respects, it is more realistic than the one employed earlier by the Adelmans, in which it was assumed that the disturbances in individual equations of the Klein-Goldberger Model were independent across equations and over time. On the other hand, their procedure achieved greater realism in other respects by scaling shocks to maintain the same ratio of the standard deviation of residuals to the value of the dependent variable as observed in the sample period, and by experimenting, also, with random shocks to the exogenous variables.

Fifty stochastic simulations were made for the OBE and Brookings Models and one hundred for the Wharton Model. In all runs for a given model, the exogenous variables and initial conditions were the same as in the nonstochastic control solution. Serially correlated disturbances were assumed for all fifty Brookings simulations, and for half of the runs for the OBE and Wharton Models. In all experiments, the shocks were chosen so as to reproduce the statistical properties of the sample-period residuals of the various models.

The stochastic simulations were analyzed in two different ways at the Conference. First, each model-building group studied its set of simulations by spectral methods to ascertain whether crucial endogenous variables displayed cyclical periodicities. Second, the simulations from the OBE and Wharton Models were subjected to NBER cycle analysis by Zarnowitz, Boschan, and Moore.

The principal conclusions reached by Green, Liebenberg, and Hirsch from their analysis of the stochastic simulations with the OBE Model are as follows: the real *GNP* series from the fifty different runs rarely showed absolute downturns, presumably because of the strong growth trends assumed in the exogenous variables. When expressed as deviations from control-solution values, however—i.e., approximately, as deviations from trend—the shocked *GNP* series show definite cycles. The "cycles" from the runs using serially uncorrelated shocks are unrealistically ragged, as compared with observed business cycles, however; and the average power-spectrum for these runs fails to reveal significant peaks at business-cycle periodicities. The runs with serially correlated shocks are considerably smoother, and a spectral analysis based on a preferred method of trend-removal reveals peaks at periodicities consistently falling in the range of two- to five-year cycles.

The findings are much the same for the Wharton Model as reported by Evans, Klein, and Saito. That is to say, serially independent shocks do not lead to average spectral-density functions with distinct peaks, whereas the application of serially correlated disturbances does produce a high concentration of spectral peaks for real *GNP* in the range 26.7 to 10 quarters, with a mode at 16 quarters—or the average duration of business cycles in the NBER chronology.

The NBER analysis of the stochastic simulations confirms the general findings of the OBE and Wharton groups, and adds a great deal of information on the cyclical attributes of the simulated series. Thus, in an analysis of the average durations of rises and declines in all of the simulation runs for current- and constant-dollar *GNP*, Zarnowitz, Boschan, and Moore conclude that many fluctuations do occur (especially for the runs with serially uncorrelated shocks), but they are in large part too short to qualify as cyclical movements. When the stochastic simulations are expressed as deviations from the control solutions, however, they reveal characteristics which are closer to the cycles of experience. Again, as in the earlier papers, it is found that the simulations based on autocorrelated shocks are substantially smoother than those with serially uncorrelated shocks, and are generally of more plausible appearance.

The NBER group did not attempt to identify business-cycle peaks and troughs in the detrended *GNP* simulations, although they did compute measures of the average duration, and amplitude of rises and declines, where declines as short as one quarter were included. Of considerably greater interest in the present context is the more exhaustive analysis they undertook, employing a random sample of three runs each from the OBE and Wharton Models. The plans for the Conference called for simulation results for a specified list of endogenous variables, so that each stochastic simulation for the Wharton Model included output for 17 variables, and 22 variables were available in the OBE simulations. For each of the six simulation sets examined, the NBER group dated the specific cycle-peaks and troughs in all of the detrended series, and computed a diffusion index of the percentage of the series undergoing expansion in each successive quarter. The diffusion indexes were expressed in cumulative form to derive a relatively smooth index, whose peaks and troughs would be centered on periods of greatest concentration of peaks and troughs in the component series. The cumulated diffusion indexes for both models show distinct cyclical fluctuations, whose average durations are similar to those of the cycles in trend-adjusted *GNP* during 1948–68. This finding provides interesting independent confirmation of the existence of cyclical periodicities in the detrended stochastic simulations, as established by the spectral analyses of the OBE and Wharton groups.

Does all this evidence amount to confirmation of the hypothesis that cyclical fluctuations are caused by stochastic disturbances impinging on a dynamically stable response-mechanism? Certainly, the evidence is consistent with at least one version of that hypothesis, but some major qualifications are in order. First, on the evidence of this Conference, it is necessary to reject the classical version of the hypothesis. Second, inadequate attention was paid at the Conference to the effects of shocks or fluctuations in exogenous variables. Third, it can be argued that the apparent success of the stochastic simulations is due to mis-specification of the model structures. Finally, if the models are indeed seriously mis-specified, the hypothesis of a deterministic cyclical-mechanism cannot be rejected on their account.

It is convenient to discuss the first two qualifications in the context of the elegant analysis of the dynamic properties of the Wharton Model by E. Philip Howrey. This paper, included in Part II, is a completely independent attack on the same problems that were studied in the stochastic simulations of Part I. Howrey uses an analytical technique based on the spectral representation of a stochastic process to determine whether a condensed, linearized version of the Wharton Model exhibits cyclical properties. The method has some drawbacks – especially the need to linearize the model, with unknown effects on its properties for large departures from the neighborhood of linearization – but it permits a rigorous statement and analysis of the alternative hypotheses on the nature of cyclical fluctuations.

In Howrey's notation, the linear econometric model can be written in the form

$$B(L)y(t) = C(L)x(t) + u(t) \tag{1}$$

where $y(t)$ and $x(t)$ are vectors, respectively, of endogenous and exogenous variables at time t; $B(L)$ and $C(L)$ are matrices of polynomials in the lag operator L, with coefficients for the various lagged values of y and x, as estimated in the structural model; and $u(t)$ is a vector of random disturbances. The solution of the model is

$$y(t) = P(t) + B(L)^{-1}C(L)x(t) + B(L)^{-1}u(t) \tag{2}$$

where $B(L)^{-1}$ is the inverse of the matrix $B(L)$, and where $P(t)$ is a vector of functions giving the solutions to the transient part of the system in terms of initial conditions and characteristic roots.

According to equation (2), the time path of the endogenous variables depends on the structure of the model and on the nature of the forces impinging on the system. The structure of the model determines the values of the characteristic roots of $P(t)$, and it also determines the response of the system to external forces, as summarized in the matrices $B(L)^{-1}C(L)$ and $B(L)^{-1}$, which represent systems of weights to be applied to the various current and lagged values of the exogenous variables and random disturbances.

Although Howrey is primarily concerned with testing the stochastic part of the system for cyclical effects, he has also performed the arduous task of extracting more than 40 characteristic roots of the model; this work was later checked independently by Kei Mori, using a different computational algorithm. The linearized condensed version of the Wharton Model apparently yields one pair of cyclical roots with only moderate damping. Further analysis by Howrey leads to the conclusion that the complex roots do not impart discernible cyclical properties to the solution, however, since their effects are swamped by the contribution of larger positive real roots. One real root is slightly greater than unity, incidentally, implying that the system may be unstable in a growth sense, although this is uncertain because of the sampling variability to which the estimates are subject.

Thus, Howrey's results confirm the implication of the simulation studies that the Wharton Model is stable in its deterministic part, at least insofar as an endogenous cyclical mechanism is concerned. What about its response to random disturbances? To answer this question, Howrey analyzes the spectral representation of the stochastic process $B(L)^{-1}u(t)$ implied by the model. He concludes that the lag structure of the model does not impart the sort of smoothing that is required to convert a sequence of random shocks into cyclical fluctuations in the endogenous variables. "Any cyclical behavior that this model might exhibit is, therefore, due to serial correlation in the disturbance process or to business-cycle variations in the exogenous variables."

Once again, these results confirm the earlier findings. Random disturbances were not enough to generate fluctuations with cyclical properties in the stochastic simulations of the Wharton and OBE Models—rather, it was necessary to introduce serial correlation in the shocks to accomplish that result. It appears that the classical Wicksell–Slutsky–Frisch hypothesis—that business cycles are the result of a stream of erratic shocks operating on a dynamic system which otherwise would exhibit only damped oscillations—must be rejected on the evidence of this Conference. On the other hand, broader versions of the hypothesis, in which the admissible class of shocks or impulses is enlarged, cannot be rejected.

Indeed, the first possibility to be considered—that the exogenous variables may be subject to random shocks—can be regarded as falling within the classical hypothesis on erratic shocks. If the shocks were random and serially uncorrelated, the moving average process represented by $C(L)$ could, nevertheless, impart a cyclical path to the exogenous variables, and hence to the endogenous ones, even if the lag structure of the model itself were noncyclical.

Second, as Frank de Leeuw emphasizes in his discussion of the Wharton Model simulations, the impulses impinging on the economy include identifiable and measurable forces, such as wars, monetary disturbances, and exogenous changes in policy variables or parameters. Perhaps, the business cycles of experience reflect sporadic occurrences or variations in such exogenous forces, rather than truly stochastic disturbances that cannot be measured directly or isolated in time. This hypothesis would argue for systematic historical investigations of the role of identifiable exogenous impulses, and of the responses to such impulses during particular cyclical episodes or epochs. Such studies can and have been done as model simulations, but they did not fall within the purview of this Conference. Incidentally, to the extent that some external events, such as wars, have a generalized impact on the economy, allowance should be made for covariation between disturbances to exogenous variables and stochastic equations in simulation experiments.

Third, shocks to either the exogenous variables, or the equations, may be serially correlated. As we have seen, the latter possibility was extensively tested in this Conference, and the general finding is that the model simulations did exhibit "business cycles" in response to serially correlated disturbances. Moreover, the OBE group also ran five stochastic simulations in which serially correlated shocks were applied to the exogenous variables, as well as to the endogenous equations, with the result that the cycles were increased in amplitude; in addition, they frequently showed absolute declines in real *GNP*, lasting three to five quarters. As the authors observe, "These brief results suggest that movements commonly considered exogenous in large-scale models may play a crucial role in the determination of business cycles."

Thus, it appears that some classes of shocks may generate cycles when acting upon the models studied at this Conference. It should be emphasized, however, that broadening the class of shocks to include perturbations in exogenous variables—and to allow for serial correlation in the disturbances to equations and exogenous variables—diminishes the role of model structure as a cycle-maker. If the real roots dominate the cyclical ones and the lag structure does not convert serially independent random shocks into cycles in the endogenous variables, the model structure becomes simply a multiplier mechanism for amplifying shocks of any kind. There is still an impulse-response mechanism, but the cycles are inherent in the impulses rather than in the responses.

The question of whether serial correlation should be expected in real-world shocks to the economy naturally arises. It is easy to think of reasons why this could be so. Important political events, such as wars, may impinge on the economy for several quarters, or even years, before being reversed. Similarly, policy actions may be maintained over several quarters or more. It is well to recall, also, that Frisch suggested the Schumpeterian theory of innovations as a possible source of impulse energy, which would impinge on the economy in a sustained fashion during certain phases of the cycle, while being absent in others.

Unfortunately, it is possible, too, that the observed presence of serial correlation in the sample disturbances may merely reflect mis-specification of the econometric models. At the Conference, this position was stated most forcefully in de Leeuw's perceptive comment, where he concludes that "the principal simulation results reported in the Evans-Klein-Saito paper and some of the other Conference papers could, it seems to me, just as easily result from mis-specification as from the historical validity of the Slutsky-Frisch theory." He supports this conclusion partly by arguing that the actual errors during the sample period from a poorly specified model will tend to be larger than the true magnitude of stochastic forces and, hence, may exaggerate their power to generate fluctuations. This could happen, of course, but note should also be taken of the observation of Green, Liebenberg, and Hirsch that the sample-period errors were not rescaled to reflect the much larger size of the economy that is implied in the control-solutions for the postsample simulations, so that the shocks could be understated on that account.

One final caveat concerning the stochastic simulations is in order before we turn to other topics. The simulations prepared for the Brookings Model failed to reveal systematic cyclical periodicities in the average spectral density function for fifty stochastic runs, despite the inclusion of serial correlation in the disturbances. Hence, the simulation results are consistent with the stochastic disturbance theory only for the OBE and Wharton Models. On the present evidence, the theory would be rejected for the Brookings Model, insofar as it applies to shocks to equations rather than to exogenous variables. Finally, the theory has not yet been tested for the FMP Model.

Let us consider now another major topic that cuts across several papers in this volume—the subject of forecasting. It is convenient to organize the discussion under three headings: ex post predictions, ex ante forecasts, and forecasts of cyclical turning points. Ex post predictions provide a test of model specification and structural stability, whereas ex ante forecasts involve, as well, the skills of the forecaster in supplementing the model with extraneous predictions of such items as policy actions, changes in noncontrolled exogenous variables, strikes and political events.

The paper by Ronald L. Cooper, "The Predictive Performance of Quarterly Econometric Models of the United States," is a major effort to evaluate and compare the structural specifications of alternative econometric models. This is done in two ways: by testing the models for structural change over time, and by comparing their ex post forecasts against predictions by naive auto-regressive models. His general conclusions are that all the models tested are structurally unstable, and that none could forecast better than purely mechanical models with no economic content. Apparently skeptical of further structural model-building, he suggests that forecasting performance might be improved by combining instrumental variables from econometric models with mechanical auto-regressive schemes. These are sobering conclusions, indeed; because of their sweeping nature, they will require careful evaluation by econometricians and noneconometricians alike.

Cooper's procedure was to make a series of ex post single-period forecasts from the reduced forms of the models, given the actual values of the predetermined variables; and then, to compare them with similar predictions from auto-regressive equations fitted to the individual endogenous variables. An auto-regressive scheme is simply a mechanical regression of a variable on its own lagged values, and in the present paper, Cooper chose for each variable the best-fitting equation from among candidates with up to eight lagged quarterly values. One or another of the econometric models could predict some of the endogenous variables better than the corresponding auto-regressive equations, but the latter had the highest over-all score for both the sample and postsample forecasts. Seven econometric models were included in the tests, including earlier versions of the OBE and Wharton-EFU Models than those used in the other papers in this volume.

One important issue concerns the length of the forecast period. As already noted, Cooper's tests were restricted to single-period forecasts. All the discussants of his paper stress that this choice not only puts the auto-regressive schemes in the most favorable light—since they tend to deteriorate rapidly when forecasting more than one period ahead—but is also irrelevant to realistic forecasting situations, which usually involve forecasts of four to eight quarters ahead.

Thus, there is a presumption that structural models are capable of better multiperiod forecasts than are auto-regressive schemes. In his comment on Cooper's paper, Michael McCarthy cites results for the current Wharton Model to demonstrate that it outperforms the best auto-regression for real *GNP* over the sample period, when using only the initial information available at the start of the period—even though the standard deviation of one-quarter forecasts was smaller for the naive model over the same period. In similar vein, the OBE group quotes average errors for real *GNP* obtained over the sample period for one to six quarter forecasts, made with a later version of their model than the one tested by Cooper. In this case, the model outperforms the naive auto-regressive form even for first-period forecasts, and the improvement gets progressively larger as the forecast horizon is extended. As Cooper notes in his rejoinder, these results are for the sample period only, and it is possible that the relative performance of the structural and naive models would be reversed for multiperiod predictions beyond the sample period. Further testing will be required to settle the issue definitively, but it seems likely that structural models will prove superior to auto-regressive schemes when it comes to multiperiod ex post forecasting.

Some additional evidence on ex post forecasts is included in the paper by Michael Evans, Yoel Haitovsky, and George Treyz, "An Analysis of the Forecasting Properties of U.S. Econometric Models." Their analysis deals with newer versions of the Wharton and OBE Models than were studied by Cooper. In the case of the Wharton Model, the errors for one-quarter predictions during the sample period were generally larger than those of a naive model extrapolating the last observed quarterly change; and those for postsample forecasts were larger than a naive model of no-change. These results are consistent with Cooper's findings. On the other hand, the newer OBE Model outperforms the no-change and same-change naive models for the sample period (which – in contrast to the Wharton Model – includes the trend-dominated years, 1965–66), as distinct from Cooper's finding for the earlier version, which did, however, involve comparison with a more stringent naive model.

Evans, Haitovsky, and Treyz also made multiperiod (six-quarter) ex post predictions with both the OBE and Wharton Models, including some postsample forecasts for the latter. In general, both models outperform the naive forecasts after one or two quarters.

It appears from the preponderance of evidence, then, that the models perform better than mechanical schemes in multiperiod forecasts with known exogenous variables. This is not to say that the generally poor record on single-period ex post forecasts should be lightly dismissed. Until an econometric model is brought to a point where it can make better ex post forecasts over a short horizon than can a mechanical scheme, the structural specification or stability of the model must always be in question.

Critics and model-builders alike are agreed on the pervasiveness of structural change and the difficult problems it poses for specification and forecasting with econometric models. Cooper's study was designed to test the specifications of alternative models by refitting them through a statistically equivalent estimation technique to a common sample-period, in order to run a fair race between them. His discussants stress, however, that such a mechanical reestimation of the models meant that some of them were fitted to historically revised national-income data which differed considerably from those originally used; that others were fitted to Korean War data which had been omitted or treated with dummy variables in the original models; that tax-rate changes over the Cooper sample period were ignored by him in refitting the revenue and depreciation functions; and that other of the modifications made by Cooper in the reestimation of several of the models implicitly or explicitly changed their structures. In short, structural change is a persistent phenomenon, and econometricians attempt to deal with it as carefully and explicitly as possible when specifying and estimating a model. Cooper's negative findings, it is asserted, may stem in large part from an inadvertent introduction of structural mis-specifications due to his wholesale approach and mechanical reestimation procedure. As Stephen Goldfeld remarked in his comment: "More recent evidence such as the FRB-MIT results cited earlier, and the results with the OBE Model presented at this Conference, suggest that even for one-period forecasts, carefully estimated large econometric models outperform the auto-regressive standards."

Be that as it may, Cooper's work presents a serious challenge to model-builders to redouble their efforts to improve methods of specification, estimation, and evaluation in order to assure more accurate and stable structures. At the Conference, a number of constructive suggestions were made in this connection. A new estimation technique, suggested earlier by Dale Jorgenson and employed by Cooper, was further tested on a partial basis in the paper by Evans, Haitovsky, and Treyz. It was found to lead to estimates of the structural parameters which reduced the ex post forecast error for two quarters, and the authors suggest that an as yet untried modification of the method suggested by Klein may improve the results for longer forecast periods. The new techniques are intended to reduce the propagation of single-equation errors across equations, and to reduce error-buildup over time through the lagged endogenous variables.

In his comment on the paper by Evans, Haitovsky, and Treyz, Alvin Karchere conjectures that the principal problem is single-equation error itself, particularly as it leads to systematic bias in the postsample forecasts. He urges model-builders to choose among alternative specifications of structural equations according to their ex post error properties over the standard forecast period, rather than by their sample-period characteristics. Another suggestion for dealing with individual equations was made from the floor: namely, to use bloc simulation to isolate those equations, or sectors, contributing the most to forecast errors, and then to concentrate on improving the weak equations. Finally, as shown in several of the papers, improved ex post forecasts are obtained from the models when automatic adjustments are made to the constant terms of the normalized structural equations to take account of serial correlation in the calculated residuals.

A few comments are now in order on ex ante forecasting with econometric models. It is here that science shades into art. Anyone who has the mistaken impression that an econometric model is simply a black box used to convert ex ante predictions of exogenous variables into ex ante forecasts of endogenous variables, will find the paper by Evans, Haitovsky, and Treyz enlightening indeed. They distinguish three classes of judgmental inputs to ex ante model forecasts: (1) the selection of values for exogenous variables, including informed guesses about future monetary and fiscal policies, export developments, and so forth; (2) adjustments to the constant terms of individual structural equations to allow for known structural changes since the estimation period, to correct for substantive data revisions, to offset observed autocorrelation in the estimated residuals, and to incorporate extraneous information about future exogenous developments, such as strikes; and (3) changes in the initial decisions made about constant adjustments or exogenous variables if preliminary forecasts with these assumptions lead to a forecast for some variables that is out of the range of the forecaster's a priori concept of a reasonable forecast.

In their section on evaluation of ex ante forecasts, Evans, Haitovsky, and Treyz compare actual ex ante forecasts from the Wharton, Evans, and OBE Models with ex post forecasts, and with hypothetical ex ante forecasts using alternative mechanical schemes to adjust the constant terms of structural equations for autocorrelation. They demonstrate that the true ex ante forecasts are much better than the ex post forecasts. For example, the average forecast error for the true ex ante first-quarter forecasts of the Wharton-EFU Model for real *GNP* during 1966–68 was 3.0 billion dollars, as compared with

6.0 billion for a mechanically adjusted ex post forecast and 16.3 billion for unadjusted ex post forecasts. They also show that the true ex ante forecasts are better than hypothetical ex ante forecasts generated by mechanical methods for the Wharton Model, and for most of the OBE forecasts. Evidently, the use of judgmental adjustments contributes substantially to the reduction of ex ante forecast errors. On the basis of their empirical analysis of the sources of ex ante forecast error, the authors conclude that it is the third class of judgmental adjustment—changing the preliminary assumptions on exogenous variables and constants until the resulting forecast falls within the range thought to be reasonable—which is principally responsible for the improvement in ex ante forecasts. From this point of view, the model serves primarily to assess judgmentally the general implications of the forecaster's assumptions on future exogenous developments, including his ad hoc adjustments for anticipated changes in structure since the sample period, and for the correction of apparent specification errors.

In summary, Evans, Haitovsky, and Treyz recognize the current limitations of econometric forecasting techniques but are sanguine about future improvements:

> This study has shown that econometricians have had a better forecasting record to date than an analysis of the econometric models that they used would have led us to predict. Our results offer no substantive evidence that the same econometricians, forecasting without the "benefit" of an econometric model, would have done any better or any worse in their predictions. This recognition of the limitations of current models need not lead to pessimism about the future development of accurate econometric forecasting models. With a finer understanding of how changes in monetary and fiscal policy actually influence economic activity, closer attention to the short-run specifications and lag adjustments of the system, possible improvements in the National Income Accounts, and refinement of existing estimation and forecasting techniques, the next few years could offer substantial advances in the art and the science of econometric forecasting.

The discussion thus far has abstracted from the prediction of business-cycle turning points. There was no analysis of ex ante forecasting of turning points at the Conference. However, the business-cycle simulations of Part I include relevant material concerning ex post forecasts of upturns and downturns. In addition to the long-run simulations discussed earlier, these papers contain a series of short-run ex post simulations around the observed postwar cyclical peaks and troughs, as dated by the National Bureau of Economic Research. For each turning point, three sets of six-quarter simulations were prepared, beginning respectively three quarters, two quarters, and one quarter before the business-cycle peak or trough. Results are available for varying sample periods for the Wharton, OBE, FMP and Brookings Models. However, the Brookings simulations were not completed in time for analysis by the NBER team.

In the judgment of Zarnowitz, Boschan, and Moore, the Wharton, OBE, and FMP Models were fairly successful in duplicating cyclical turns, with two-thirds or more of the actual turns being matched by turns in the model simulations. The success rates were about the same at peaks and troughs. Also, there is not much to choose between the simulations started one, two, or three quarters before the business-cycle turns, despite the implied differences in the amount of foreshadowing information included in the initial conditions. Finally, although two-thirds or more of the business-cycle turns were matched by the simulations, the simulated turns did not always coincide with the actual peaks and troughs, although a substantial majority were either coincident or within one quarter of the actuals.

It should be noted that the short-period simulations analyzed by Zarnowitz, Boschan, and Moore were made without correction for serial correlation of the residuals. However, the OBE and Wharton groups discovered that the short-term simulations were marginally improved by making automatic adjustments to the constant turns to correct for autocorrelation. The discussion of sample-period turning point behavior in the OBE paper is based on the adjusted simulations, but no comparison is attempted with the unadjusted simulations analyzed by the NBER group. It would be interesting to make such comparisons for all models.

The limited objectives of these experimental simulations should be borne in mind when appraising the results. The principal focus of the Conference was on the dynamic properties and structural stability of econometric models, rather than on forecasting technique. Hence, there was no attempt to provide a comparison of the errors of the turning point simulations with those from ex post cyclical forecasts with naive auto-regressive models. Similarly, no comparison was attempted with ex post forecasts by other methods, let alone between ex ante forecasts by persons using econometric models and those using other techniques. The paper by D. J. Daly, "Forecasting with Statistical Indicators," contains a judicious appraisal of the strengths and weaknesses of one of the principal forecasting alternatives to econometric models, but it is not intended to provide the basis for a systematic comparison of either ex ante or ex post forecasts of turning points by the two methods.

The last two papers to be introduced stand somewhat apart from the others. The first one is an attempt to model a particular literary theory of business cycles, whereas the second deals with the methodological problem of the effects of aggregation over time on the estimated lag structure of dynamic econometric models.

In "An Econometric Model of Business Cycles," Gregory Chow and Geoffrey Moore present a progress report on their efforts to specify and estimate a mathematically explicit model containing the major elements of the Mitchell-Burns theory of the business cycle. They point out that they have included many, though not all, of the important elements in Arthur F. Burns' recent article on business cycles in the *International Encyclopedia of the Social Sciences* (1968). "Hence, this is by no means a perfect translation. In general, the material we present is a simplified, aggregative version of the earlier text." The model contains 20 stochastic equations and 5 identities.

As mentioned at the outset of this Introduction, it was originally hoped that the new "NBER Model" would be completed in time for simulations to be made and analyzed on the common plan for Part I. The existing large quarterly models include many of the variables emphasized in the analytical descriptions of business-cycle processes by Burns, Mitchell, Moore, and others at the NBER. The structural hypotheses incorporating the variables in the models are generally different from those suggested in the NBER writings, however, and in some cases markedly so. Thus, even if these models were capable of simulating well the characteristics emphasized in NBER empirical studies—timing, amplitude, conformity, diffusion, and so forth—this would not be verification of the Mitchell-Burns theory of cyclical processes, since many structures are consistent with the same reduced form. It seemed desirable, therefore, to construct a model with a structural specification reflecting the Mitchell-Burns hypotheses, in order to see if it could match, or exceed, the successes of other specifications in capturing the empirical regularities isolated in the long history of NBER cycle studies.

Owing to limitations of time, this program still remains to be carried out. Meanwhile, the progress report of Chow and Moore must be evaluated in terms of its structural specification and statistical methodology. The lively debate between the authors and discussants can scarcely be summarized here, beyond noting that R. A. Gordon, especially, has questioned the extent to which this initial specification embodies the essential features of the "NBER approach" to an explanation of business cycles.

The effects of aggregation over time on the parameter estimates and dynamic properties of econometric models is a subject of great interest to model-builders. The paper by Robert Engle and Ta-Chung Liu is an important attack on this highly technical and intractable problem. Their basic approach is to assume that a given model with a particular time unit—in the empirical application, Liu's monthly model—is the true one, and then to investigate the biases that may result from aggregating the observations into longer time-periods. Unfortunately, there are several different effects working in different directions; the net result depends on the time structure of the exogenous variables and cannot readily be generalized. The analytic results presented by Engle and Liu are based on a rational distributed-lag model, which is aggregated to a Koyck-Nerlove form. The predicted results are then tested by estimating quarterly and annual versions of Liu's model for comparison with the monthly estimates. The authors conclude that the empirical results are consistent with the theory and that biases are, indeed, introduced by aggregation over time.

Apart from the restrictive assumptions necessary to produce predictable outcomes, the major issue concerning the Engle-Liu results is the validity of assuming that the monthly equations are the correct ones. This assumption is questioned, especially, in the Comment by Zvi Griliches, who argues that the Koyck-Nerlove lag distribution does not make much sense for monthly time units, and that the Liu Model is a pioneering effort to test the feasibility of building a monthly model, rather than being a well-shaken-down final version. An important related point is the prevalent belief that in actual applications with calendar-year data, autocorrelation is less prevalent than when calendar time is sliced more finely. Perhaps, the annual model is the more nearly correct, and the additional "bias" resulting from the introduction of serial correlation in the disturbances by aggregation of, say, the monthly time unit, is less serious than that stemming from the serial correlation initially existing in the model estimated from monthly, instead of quarterly or calendar-year, data.

In conclusion, the papers and comments included in this volume reflect a high order of analytical insight and technical competence. Because many of the issues are both technical and controversial, the reader is urged to study the comments as carefully as the papers. As in all good conferences, the proceedings have opened as many issues as they have settled. Scientific testing of econometric models is still in its infancy, but the effort is a continuing one, and the papers in this volume have substantially advanced the subject by clarifying the issues and sharpening the methodology for future work in this important field.

Insofar as prediction is concerned, research on alternative forecasting schemes will doubtless continue and broaden. These schemes can, and should, include auto-regressive models, direct reduced-form systems and, perhaps, other methods which also eschew economic analysis and provide a forecasting standard for econometric models. Nonetheless, these methods must, it seems to me, be regarded as spurs for structural model-building, rather than as substitutes. Unless economic theory is truly a set of empty boxes, it should be possible to improve on mechanical methods by incorporating structural hypotheses into our quantitative models. In this respect, all economists have a stake in the econometricians' credo or, at least, his hopes. Moreover, structural models have much more to offer than alternative methods. In forecasting applications, they provide a framework into which extraneous information on structural changes—especially those involving fiscal and monetary policy—and predictable shocks can be incorporated. They afford, also, a vehicle for analysis of the historic causes of economic instability. It is quite possible, for example, that the postwar models are strongly damped because of the high degree of built-in flexibility which now characterizes the economy; this is a hypothesis which can be tested on a structural model by simulation experiments. Finally, structural models which have been constructed with an eye to policy analysis, and which, therefore, incorporate explicit quantitative policy instruments and parameters, can be used not only to forecast, but also to study alternative policies to change the course of the economy if the forecasted outcome is unsatisfactory.[1]

[1] For a discussion of policy uses of econometric models, the reader is referred to an earlier conference, which was also sponsored by the Social Science Research Council Committee on Economic Stability. The papers have been published in *Quantitative Planning of Economic Policy*, Bert G. Hickman, ed. (the Brookings Institution, 1965).

PART ONE

BUSINESS-CYCLE SIMULATIONS