Volume Title: Business Cycles: Theory, History, Indicators, and Forecasting

Volume Author/Editor: Victor Zarnowitz

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-97890-7

Volume URL: http://www.nber.org/books/zarn92-1

Conference Date: n/a

Publication Date: January 1992

Chapter Title: Econometric Model Simulations and the Cyclical Characteristics of the Economy

Chapter Author: Victor Zarnowitz

Chapter URL: http://www.nber.org/chapters/c10380

Chapter pages in book: (p. 265 - 280)

9 Econometric Model Simulations and the Cyclical Characteristics of the Economy

9.1 Questions, Methods, and Data

This paper grows out of a comprehensive study which addresses some old but unsolved questions concerning several econometric models (Zarnowitz et al. 1972). The major substantive issue involved here was raised early in the literature: Do business cycles consist mainly of endogenous or exogenous movements? This is presumably an empirical problem, in the usual sense of being amenable to scientific treatment through formulation of suitable hypotheses that can be tested against the data. There is no dearth of either endogenous or exogenous or "mixed" theories, some of which can be and have been tested, though not always adequately or persuasively. Over the years, the subject of business cycles has attracted much systematic research and observation, which has added much to our factual knowledge. Nonetheless, the issue still resists a solution, and perhaps not surprisingly so, as it requires understanding the modern economy in motion to a degree not yet achieved.

It seems quite natural that a close connection should exist between the problem of how business cycles are generated and the method of studying the economy through building and analyzing econometric macromodels. Indeed, interest in testing various cyclical hypotheses first motivated Tinbergen to construct such models (1938–39). However, even now, more than five decades and several generations of aggregative econometric models later, disagreement abounds on how best to scale, specify, and estimate such models. The great gains in theory, information, and computational techniques and capacities have yet to be fully reflected in comprehensive systems of proved superi-

Reprinted from *The Business Cycle Today,* edited by Victor Zarnowitz, pp. 241–59 (New York: NBER, 1972).

The author is greatly indebted to Charlotte Boschan for helpful comments, to Josephine Su for valuable statistical assistance, to H. Irving Forman for the preparation of the figure, and to Gnomi Schrift Gouldin for editorial improvements.

ority. This presents a grave problem since what an econometric model suggests about the nature of business cycles may not be dependable if the model itself is not.

The study on which I report is, therefore, more properly described as a search for answers to these questions: Do the models under review generate cyclical behavior as defined and observed in the empirical business cycle studies, notably those of the National Bureau of Economic Research (NBER), which provide the main documentation on the subject? If so, to what extent are such fluctuations in the estimated series produced endogenously by the models, and to what extent are they attributable to external impulses? The aims of the study, then, will be recognized as very similar to those of the 1959 analysis of the annual Klein-Goldberger model by Irma and Frank Adelman (1959). Its scope, however, is substantially larger as the materials now available are much richer. Four different quarterly models are examined, to be labeled Wharton, OBE, FMP, and Brookings.[1] It is generally recognized that quarterly data are far more adequate in business cycle analysis than are annual data.

The methods employed also largely parallel the techniques used in the pioneering study by Adelman and Adelman (1959). Three types of complete-model simulations are analyzed, namely:

(*a*) Nonstochastic simulations over six-quarter periods beginning, alternatively, one, two, and three quarters before each of the business cycle turns that occurred during the model's sample period.[2] Each of these runs starts from new, correct initial conditions and uses ex post values for the exogenous variables.

(*b*) Nonstochastic simulations over the entire sample period covered by each model; also based on the initial conditions (actual value) at the beginning of that period and on the historical values of the exogenous variables.

(*c*) Stochastic simulations projecting the models for a period of 25 years starting at the end of the sample period. In these experiments, the exogenous variables are generally continued along smooth growth trends based on their compound interest rates of growth during the sample period.

One set of short and one of long nonstochastic simulations (*a* and *b*) was required for each model, but for the stochastic simulations (*c*) as many as 50 computer runs per model were made, so as to gain information on the variability of responses to different configurations of shocks and to avoid excessive reliance on any particular, and possibly idiosyncratic, shock distribution.

---

[1] The abbreviations refer to the Wharton–Econometric Forecasting Unit model; the Office of Business Economics of the U.S. Department of Commerce model; the Federal Reserve Board-MIT-Penn. model; and the Brookings-SSRC model. The model variants on which this analysis is based are those developed by the summer of 1969 and explained in several papers prepared for the Harvard Conference on Econometric Models of Cyclical Behavior, November 1969.

[2] The business cycle peaks and troughs are dated according to the NBER reference chronology (in quarterly terms) and are also referred to as reference turns.

Completed work covers nonstochastic simulations of type (*a*) for three models (Wharton, OBE, FMP), those of type (*b*) for all four models, and the stochastic runs for three (Wharton, OBE, and Brookings).

Regrettably, the results for the different models are not strictly comparable, for at least two reasons. First, the sample periods differ: the Wharton model covers 79 quarters, from 1948:3 through 1969:1; the OBE model covers 55 quarters from 1953:2 through 1966:4; the FMP model covers 44 quarters from 1956:1 through 1966:4; and the Brookings model covers 36 quarters from 1957:1 through 1965:4. Thus, the Wharton period includes four of the completed contractions or recessions in the postwar economic history of the United States (as well as such milder retardations as those of 1951–52, 1962–63, and 1966–67), the OBE period includes three, and the FMP period and the Brookings period each include two of these contractions. Such differences can strongly affect the relative performance of the models, and as a task for the future, it would be very desirable to recalculate the simulations with one common sample period for all included models. Second, models differ in coverage: in particular, what is endogenous in one of them may be exogenous in another. This must be accepted and only some partial remedies are available here depending on the cooperation of the model builders; but this study reduces the problem by concentrating upon a subset of selected variables that are basically common to, and endogenous in, all of the models covered.

The endogenous variables used in the simulations are listed in table 9.1, which classifies the series according to their typical timing at business cycle turns, as historically determined. The list includes eight series from the national income accounts, of which five are in constant and three in current dollars; five series relating to employment and unemployment, hours of work, and unit labor costs; four relating to commitments to produce durable goods and invest in equipment and housing; and three relating to interest rates and money. The main sources are the U.S. Department of Commerce, the Bureau of Labor Statistics (BLS), and the Federal Reserve Board (FRB). Most series are used after seasonal adjustment. In addition, three variables unclassified by cyclical timing and not included in table 9.1 were also represented in the simulations for all four models: the implicit price deflator for the GNP (P); private wage and salary compensation per labor-hour in dollars (W); and net exports in billions of 1958 dollars (NE). Some variables were selected because of their importance for macroeconomic theory in general and business cycle analysis in particular, some in view of their cyclical sensitivity and timing, and some for both reasons. With relatively few exceptions but frequent modifications, they appear in most of the recent econometric models of intermediate or large size.

Although simulation is a powerful tool of economic analysis, its inherent limitations are substantial. Inferences drawn from simulation results about the properties of the economic system are only as good as the model that is used as the analogue of that system. However, evidence from studies based on

Table 9.1 **List of Variables and Data Definitions for Simulations of Four Models**

Abbreviation	Variable by Timing Group	Available for Models:
	Leading Series	
1. *IH*	Investment in nonfarm residential structures[a]	Wharton, OBE, FMP, Brookings
2. *II*	Change in nonfarm business inventories[a]	Wharton, OBE, FMP, Brookings
3. *CPR*	Corporate profits before taxes and inventory valuation adjustment[b]	Wharton, OBE, FMP, Brookings
4. *AWW*	Average workweek, private employment, hours per week, BLS	Wharton, OBE, FMP, Brookings
5. *LH*	Total hours per person per annum in nonfarm private domestic sector, BLS	FMP
6. *OMD*	New orders, durable manufacturers' goods[a]	OBE
7. *UMD*	Unfilled orders, durable manufacturers' goods, end of quarter[a]	Wharton, OBE, Brookings
8. *OUME*	Unfilled orders, machinery and equipment industries, end of quarter[b]	FMP
9. *HS*	Private nonfarm housing starts, annual rate, thousands, Census	OBE
10. *M*	Demand deposits adjusted and currency outside banks[c]	OBE, FMP
	Roughly Coincident Series	
11. *GNP*	Gross national product[b]	Wharton, OBE, FMP, Brookings
12. *GNP58*	Gross national product in constant dollars[a]	Wharton, OBE, FMP, Brookings
13. *C*	Personal consumption expenditures[a]	Wharton, OBE, FMP, Brookings
14. *YP*	Personal income[b]	Wharton, OBE, FMP, Brookings
15. *LE*	Total civilian employment, millions of persons, BLS	Wharton, OBE, FMP, Brookings
16. *UN*	Unemployment rate, percentage (of labor force), BLS	Wharton, OBE, FMP, Brookings
	Lagging Series	
17. *ISE*	Investment in nonresidential structures and producers' durable equipment[a]	Wharton, OBE, FMP, Brookings
18. *RS*	Average yield, 4–6 months prime commercial paper, percentage per annum, FRB	Wharton, OBE, FMP, Brookings
	Government bill rate, percentage	Brookings
19. *RL*	Average yield, corporate bonds, Moody's, percentage per annum	Wharton, OBE, FMP, Brookings
	Government bond rate, percentage	Brookings
20. *L/CO*	Private employee compensation per unit of private GNP in constant dollars, OBE	OBE

[a] Annual rate, billions of 1958 dollars. National income accounts; OBE, Bureau of the Census, seasonally adjusted.

[b] Annual rate (except line 8), billions of current dollars. National income accounts; Bureau of the Census, seasonally adjusted.

[c] Daily average of quarter, billions of current dollars. Currency is exogenous; deposits are endogenous.

different models and applications to different periods may to some extent cumulate and reduce this weakness. This argues in favor of comprehensive and diversified coverage of econometric model simulations in business cycle analysis.

9.2 Six-Quarter Simulations around Business Cycle Turns

The determination of cyclical turning points in these nonstochastic simulations (described as type [a] above) presents considerable difficulties because the data refer to short, unconnected periods and it is sometimes uncertain whether the observed changes in direction are cyclically significant or merely reflect short random movements. This is particularly true when the suspected turns fall close to the beginning or end of the six-quarter period. Consideration of events outside this period—turns in the actual series that occurred shortly before or after—may be helpful, but it too is not always clearly legitimate. Alternative measures were, therefore, computed, one set including and the other excluding comparisons between outside actual and inferred simulated turns.[3] In some cases, doubts remained but were met by deciding in favor of recognizing turns in the simulations if this seemed at all reasonable.

Two models succeeded fairly well, and one (FMP) rather better, in reproducing the turns in the actual series at business cycle peaks and troughs. When the inferred turning points are included, the percentages of the turns matched are 60–67 for Wharton, 66–73 for OBE, and 76–90 for FMP (see table 9.2, lines 1–3, for the underlying numbers).[4] When they are excluded, the corresponding percentages are lower, averaging 58, 66, and 75 for the respective models.

The evidence does not indicate that the simulations beginning closer to the reference turn are systematically more successful than those beginning earlier. (The former, it may be noted, cover fewer specific-cycle turns in the actuals than do the latter.) Neither is the expectation that troughs are better reproduced than peaks definitely met, although troughs are often more sharply defined and more closely clustered. This may be due to the constancy of the lag structure used by the models. However, in the simulations that start one or two quarters prior to the reference dates, the percentages of troughs matched do tend to be somewhat higher than the corresponding figures for peaks.

Coincidences with the actual turns account for 21%, 39%, and 45% of the simulated turns in the six-quarter periods for the Wharton, OBE, and FMP

---

[3] To illustrate such comparisons, if the actual series showed a peak shortly before the beginning of the simulation period and the simulated series continued downward locally, the latter was presumed to have produced a peak.

[4] The higher attainment rate of the FMP model cannot be discounted simply on the presumption that fluctuations are more easily simulated for the period 1957–61; the better performance of the FMP model is retained also if the comparisons for all three models are limited to the turning points of this shorter period. Still, it is possible that the fits are better for 1957–61 and that this at least partly explains the superiority of the FMP model.

Table 9.2 **Nonstochastic Six-Quarter Simulations around Reference Turns, Selected Statistics Relating to Cyclical Conformity, Timing, and Amplitudes of Simulated and Actual Series, Three Models**

	Wharton Model (1949–61)		OBE Model (1954–61)		FMP Model (1957–61)	
	Actual (1)	Simulated (2)	Actual (3)	Simulated (4)	Actual (5)	Simulated (6)
	Frequency of Turning Points (number)					
Simulations starting						
1. 3 quarters before reference turns	95	64	88	58	50	38
2. 2 quarters before reference turns	95	61	88	60	50	44
3. 1 quarter before reference turns	95	57	88	64	50	45
	Relative Frequency of Leads and Lags (percentage of all turns)[a]					
Leading series						
4. Leads	46	80	55	70	45	33
5. Coincidences	44	2	36	10	31	30
6. Lags	10	17	9	20	24	37
Roughly coincident series						
7. Leads	39	88	38	55	32	47
8. Coincidences	48	2	50	21	53	19
9. Lags	13	10	12	24	15	34
Lagging series						
10. Leads	11	36	14	12	12	32
11. Coincidences	30	40	21	46	53	20
12. Lags	59	23	64	41	35	48
	Average Percentage Amplitudes, by Cycle Phase[b]					
13. Expansions	7.8	6.4	7.4	3.5	7.6	3.7
14. Contractions	−7.0	−3.6	−6.6	−3.1	−6.5	−4.8

Note: Observations in lines 1–3 include, and those in lines 4–12 exclude, the inferred turning points in simulations corresponding to the known actual turns that occurred outside the simulation period. See text. For the classification of series by cyclical timing (lines 4–12), see table 9.1.

[a] The entries in cols. 2, 4, and 6 refer to all simulations regardless of starting date.

[b] The figures in cols. 2, 4, and 6 refer to simulations starting two quarters before reference turns (the results for the other simulations are similar). Expansions and contractions are phase movements in the actual and simulated series within the six-quarter periods.

models, respectively. The corresponding figures for leads of the simulated, relative to the actual, turns are 54%, 36%, and 19%, whereas for lags the percentages are 25%, 25%, and 36%. When the series are classified by historical timing groups and the comparisons are made relative to the business cycle peaks and troughs, leads are found to be much more frequent in the simulations than in the actuals for the Wharton simulations in all groups (table 9.2, lines 4–12, cols. 1–2). The same statement applies to the OBE model, though less strongly and not for the lagging series, but there is no comparable bias toward early turns in the FMP simulations (lines 4–12, cols. 3–6). On the whole, these simulations discriminate but weakly (and much less effectively than the actuals) between the historically leading, coincident, and lagging variables.

The simulated series show rises and falls that tend to be substantially smaller than their counterparts in the actual series within the selected turning-point segments (table 9.2, lines 13–14). Of course, nonstochastic simulations must be expected to vary less than the actuals on the average, because they do not include the component of random disturbances that is present in the actuals. However, this factor often seems to explain only a part of the observed underestimation of amplitudes. Good estimation and simulation of systematic, cyclical amplitude components is in any event desirable. The Wharton series approximate relatively well the average size of the actual rises; the FMP series give better results for the declines.

9.3 Sample-Period Simulations

In figure 9.1, each of the models shows the real GNP (taken to represent the aggregate economic activity) as declining during at least some portion of the first recession period covered. (For Wharton, this means the 1948–49 recession; for OBE, the one in 1953–54; and for FMP and Brookings, the one in 1957–58.) The Wharton and the FMP model also have *GNP58* contracting during the second recession, in 1953–54 and 1960–61, respectively. Neither Wharton nor OBE produces a fall in *GNP58* during either the 1957–58 or the 1960–61 recession. Although the FMP model does produce such declines in these two periods, it would be wrong to conclude that it is therefore better, because the initial conditions for this model, being as of 1956:1, are much closer to these episodes than the initial conditions for Wharton and OBE. Where the simulated series fail to match the declines in *GNP58*, they at least flatten off, however (e.g., Wharton and OBE in 1957–58, Brookings in 1960–61).

This leads to the important inference 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 type of movement would be characteristic of a hypothetical economy representing a stable macrodynamic system insulated from external disturbances.[5] It is the response of such a system to the irregular but persistent outside shocks that is supposed to convert the damped fluctuations into a maintained movement of the type historically observed as the recurrent "business cycles."

---

[5] The diminishing oscillations in this model originate in the divergencies from equilibrium that are likely to exist in any initial state of the system; they tend to disappear as the system approaches its equilibrium rate of growth. This hypothesis, completed by the notion that external disturbances or "erratic shocks" do in fact impinge upon the economy continually, gained influence following the important contribution by Frisch (1933).

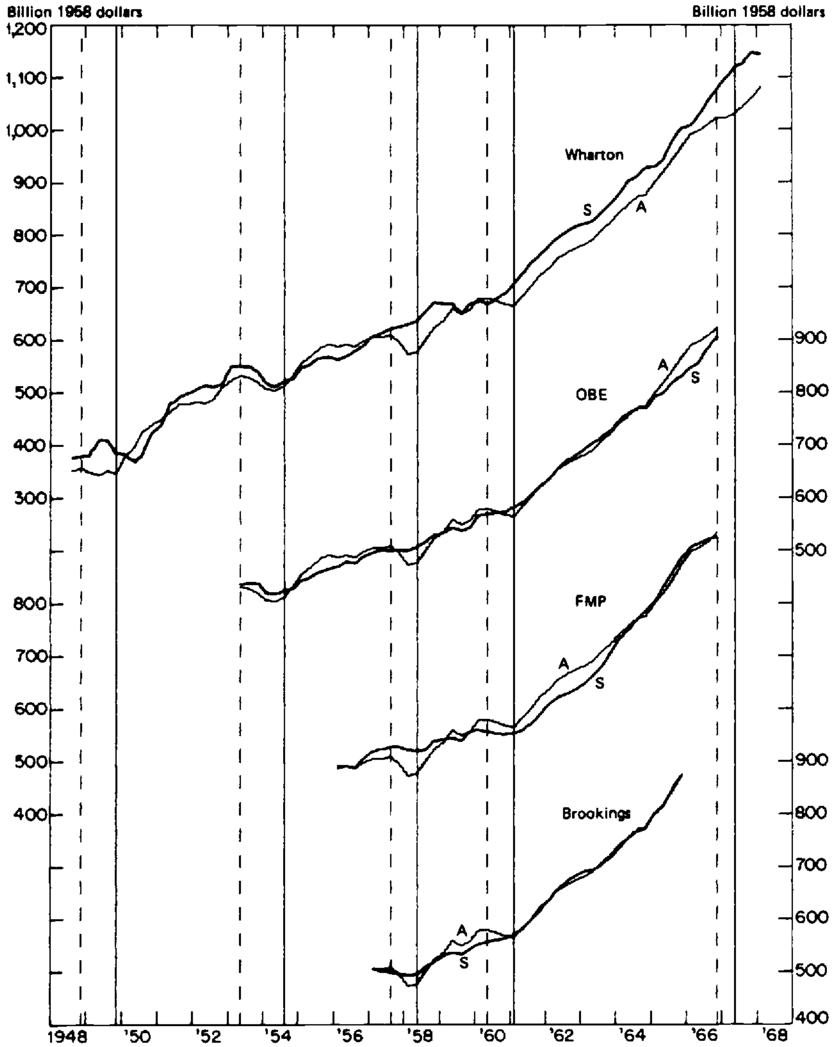

Fig. 9.1 Nonstochastic sample-period simulations of GNP in constant dollars, four models

Note: A = actual; S = simulated. Dashed vertical lines indicate business cycle peaks; solid lines, troughs. The last pair of such lines, however, refers to a business retardation in 1966–67, which did not develop into another recession.

Under this hypothesis, therefore, the failure of nonstochastic sample-period simulations to re-create the continuous cyclical developments that did actually occur need not constitute any adverse evidence about the structure of the underlying model. Instead, such results could be due to the suppression of the disturbance terms. It must be noted, however, that the simulations here reviewed use ex post values of exogenous variables. Changes in the latter include a large subset of "autonomous" shocks—variations in government expenditures, tax rates, monetary base, reserve requirements, population, exports, etc. Important effects of monetary and fiscal policy changes are thus incorporated. What these simulations suppress, then, is essentially the stochastic components of the endogenous variables. We cannot be certain that it is the disregard of this source of variability that is predominantly responsible for the errors of the nonstochastic sample-period simulations. There are undoubtedly misspecifications in the models, which could be just as important. The autocorrelations of the disturbance terms in some of the original structural equations are high enough to be disturbing. The failures of the simulations to track major cyclical movements can often be traced to the weakness of certain specific relations, for example, those for inventory investment or the price levels.

Nonstochastic simulations, which refer to the periods to which the models were fitted and use the correct ex post values of the exogenous variables, do not provide tests of the predictive powers of the models. They do, however, subject the models to rather demanding tests of a different kind, since in simultaneous estimation, errors are liable to cumulate across a model, and through the effects of lagged dependent variables, errors are also liable to cumulate over time. There is evidence that the calculated values do tend to drift away, though not necessarily continuously, in simulations that cover more than one or two business cycles. For trend-dominated variables such as *GNP, GNP58* or *C*, the drift appears sometimes as an increasing overestimation but more often as an increasing underestimation of the levels of the series. In fig. 9.1, the former is illustrated by the Wharton simulation for the 1960s and the latter by the OBE and FMP simulations in part of that decade.) Generally, the discrepancies between the levels of the simulated and actual (*S* and *A*) series are much greater than those between the corresponding quarterly changes. The reason lies in autocorrelated errors, which cumulate, thus throwing off base the long multiperiod predictions that are here involved.

Common to both short and long nonstochastic simulations is a strong tendency to underestimate the amplitudes of the observed cyclical movements. Contractions in the series, however, are often missed altogether by the simulations rather than merely underestimated. About one third of the recorded turning points are not matched by the sample-period simulations.

In table 9.3 are some measures of the kind that would be helpful to answer the question, how do the models compare with one another in terms of the relative accuracy of their simulations? (See lines 1–5.) However, because of

Table 9.3 Nonstochastic Sample-Period Simulations for Four Models, Average Error Statistics and Relative Frequency Distributions of Leads and Lags at Business Cycle Turns

	Wharton Model		OBE Model		FMP Model		Brookings Model	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	MAERC % points	MAERC/ MAARC	MAERC % points	MAERC/ MAARC	MAERC % points	MAERC/ MAARC	MAERC % points	MAERC/ MAARC
Selected Variables[a]								
1. *GNP*	1.17	0.681	0.70	0.459	0.61	0.377	0.57	.363
2. *GNP58*	1.12	0.852	0.64	0.518	0.65	0.524	0.57	.428
3. *P*	0.27	0.453	0.24	0.488	0.22	0.429	0.22	.544
4. *ISE*	3.12	1.036	1.90	0.812	1.79	0.746	1.40	.557
5. *UN*	17.80	2.502	6.00	0.890	6.26	1.155	5.63	.895
	Relative Frequency of Leads and Lags (percentage of all turns)							
	Actual	Simulated	Actual	Simulated	Actual	Simulated	Actual	Simulated
Leading series[b]								
6. Leads	62	56	73	74	68	67	56	60
7. Coincidences	32	17	20	15	14	17	38	30
8. Lags	5	26	7	12	18	17	6	10
Roughly coincident series[b]								
9. Leads	35	44	38	42	31	30	27	33
10. Coincidences	51	9	50	17	50	30	60	33
11. Lags	14	48	12	42	19	40	13	33
Lagging series[b]								
12. Leads	8	32	8	11	17	30	18	60
13. Coincidences	42	23	21	28	42	20	45	20
14. Lags	50	46	71	61	42	50	37	20

[a] For meaning of the abbreviations, see table 9.1.

[b] For the classification of series by cyclical timing, see table 9.1.

the (already noted) differences in coverage among the models, this question cannot be answered conclusively. The errors of the Wharton simulations are on the average considerably larger than those of either the OBE or the FMP or the Brookings simulations, except for the price level, *P*, where the differences are small (compare cols. 1, 3, 5, and 7). But the Wharton simulations cover a much longer period than the others, including the unsettled and difficult-to-fit developments of the late 1940s and the Korean War.

Dividing the mean absolute errors of relative change (MAERC) by the mean absolute values of actual relative change (MAARC) is a standardizing procedure which probably tends to correct for the differences in the sample periods but does not guarantee an unbiased comparison.[6] The resulting ratios

---

6. As elsewhere in the analysis of predictive accuracy. the comparisons with changes are on the whole much more meaningful than those with levels. The smaller the ratio (MAERC)/(MAARC). the better it speaks of the model; and a ratio that exceeds unity signifies that the errors are on the

(cols. 2, 4, 6, and 8) show smaller differences between the models than do the MAERC figures, but the models would be ranked rather similarly according to the two measures. (Brookings comes out somewhat better than FMP and OBE, and Wharton ranks fourth for most variables; for the price level, however, FMP and Wharton show the lowest ratios and Brookings the highest.)

The second part of table 9.3 (lines 6–14) shows that the simulations do discriminate broadly between the groups of leading and lagging indicators, but they do not carry this differentiation nearly as far as the actual timing distributions do. The OBE model yields good approximation for both leaders and laggers; the FMP and Brookings models for the leading series only. Brookings is particularly weak on the timing distribution for the laggers. The worst results are obtained for the six roughly coincident indicators, where exact coincidences make up 50%–60% of the timing observations for the actual series but only 9%, 17%, 30%, and 33% of the observations for the Wharton, OBE, FMP, and Brookings models, respectively.[7] It is for this category, too, that the simulations have the poorest record on cyclical conformity: the *S* series for *GNP* and other comprehensive aggregates of income, employment, and consumption show few turning points and frequently "skip" the peaks and troughs of business cycles.

9.4 Hundred-Quarter Ex Ante Simulations

These simulations (see type [*c*] above) have been computed only for the Wharton, OBE, and Brookings models, and their analysis is incomplete. Each of them covers a period of a hundred quarters, beginning past the space of sample experience (in 1968:3 for Wharton, in 1966:1 for OBE and Brookings). The "control solutions" (nonstochastic simulations) produce, over these long future periods, smooth series with uninterrupted growth trends for the comprehensive indicators of overall economic activity such as GNP, personal income, and employment. The trendlike control series contrast with the nonstochastic sample-period simulations that do show some recurrent, if damped, fluctuations. A probable reason for the contrast lies in the fact that in these control solutions, the exogenous variables are projected along smooth monotonic upward trends without any fluctuations. The historical series for the same variables, which were used in the nonstochastic sample-period simulations, often show considerable short-term fluctuations. However, this need not be the only or the main reason: another one may be provided by the specification errors on the models.[8]

---

average larger than the recorded changes; that is, the model does worse than a type of "naive" extrapolation.

[7] Note that the large shares of leads and lags tend to approximately balance each other in this group (table 9.3, lines 9–11).

[8] The control solutions suggest that at least in this context of long-term projections, all three models are confronted with difficult problems of internal consistency. They include some series

The lack of fluctuations in the control series for the comprehensive aggregates (GNP, etc.) indicates that none of these models generates cyclical movements endogenously. Evidently the models contain no mechanisms that would cause the simulated aggregates to fluctuate in the absence of shocks in either the exogenous quantities or in the relationships with endogenous variables.

The random shocks used in the stochastic simulations for the OBE and Wharton models were generated so that the expected value of the variance-covariance matrix of the shocks over the simulation period is equal to the variance-covariance matrix of the observed residuals over the sample period. In another set of experiments, serially correlated shocks were used, their lag correlations for a sufficiently large number of observations also being equal to the corresponding sample values obtained from the residual matrix.[9] For the OBE model, 25 simulations use serially uncorrelated random shocks and 25 use serially correlated shocks; for the Wharton model, the number is twice as large in each set. Only autocorrelated shocks were used in the 50 simulations for the Brookings model.

The stochastic simulations are strongly trend dominated for *GNP* in current and constant dollars and several other comprehensive aggregates (*YP, C, LE, P, W,*and *M*). There are systematic differences between the series with nonautocorrelated shocks (S_u) and those with autocorrelated shocks (S_c). The latter are far smoother than the former and hence tend to have larger average durations and smaller average percentage amplitudes of rises and declines.[10] The Wharton S_u series for *GNP* and *GNP58* show somewhat shorter and smaller declines than the sample-period actuals (*A*), while the S_c series show many fewer declines, all of them short and separated by overly long rises. In the corresponding OBE simulations of either type, declines are altogether rare, short, and small. The same can be said about the Brookings S_c series for *GNP* (in those for *GNP58*, declines are also small and short but more frequent). The simulated series that have weaker trends and stronger fluctuations (relating to investment processes, orders, unemployment, average workweek, and interest rates) tend to have shorter movements than the corresponding *A* series, in either direction. The S_c series often underestimate the length of the recorded movements of *A* less than the S_u series do.

---

that either are made to behave in a more or less arbitrarily predetermined fashion or are permitted to behave in ways that would seem difficult to rationalize. Such questionable simulations (as illustrated particularly by the control series for unemployment and interest rates) are perhaps best viewed as concomitants of the search for a broadly satisfactory control solution for the overall aggregates. In short, to get a plausible projection for GNP, the simulation of, say, the unemployment rate may have had to be compromised.

[9] The method of generating the shocks is that of McCarthy (1972a).

[10] A rise (decline) is used to denote any upward (downward) movement in a series, however short or small. In this analysis such changes are distinguished from cyclical movements that must be sufficiently long and pronounced to qualify as "specific-cycle" expansions and contractions (as defined by NBER).

These experiments suggest that the use of autocorrelated shocks is helpful in many but by no means all cases and that it works better for the more volatile series than for the comprehensive aggregates with dominant growth trends and subdued fluctuations. The declines in S_c tend to be longer but also smaller than those of S_u. The criterion of duration is presumably more important than that of amplitude.[11] When this is taken into account, the balance of the comparisons favors the S_c over the S_u simulations for most variables, but not without some important counterexamples (notably for *GNP* and *GNP58* in the Wharton model). In general, the cyclical aspects of the simulated series are much weaker than those observed in the historical series, in contrast to the long trends and short erratic variations that are often considerably stronger in the S than in the A series.

That the cyclical movements get blurred in the stochastic simulations could be due in large measure to the inadequate handling or scaling of the shocks, in particular to the neglect of disturbances in the exogenous variables. Hence we have also analyzed the relative deviations of shocked from control series, in the expectation that they would be more indicative of the cyclical effects of relatively weak impulses. This expectation was confirmed, but the ratios of the stochastic to the control series are also much more erratic than the shocked series proper, reflecting not only greater sensitivity to the effects of the shocks but presumably a telescoping of "measurement" errors as well. It is particularly the ratios of S_u to the control series that tend to be highly erratic; the ratios of S_c are much smoother and generally appear more plausible.

Ratios of the historical series to their exponential trends were computed to provide measures for the sample-period actuals that correspond to the measure for the simulated ratio series. As shown in table 9.4, lines 1 and 2, the trend-adjusted *GNP* series are better approximated by the S_c than by the S_u ratios, in terms of the durations (and therefore also the frequencies) of rises and declines. Comparisons of amplitudes alone would point to the reverse (lines 3 and 4), but, again giving more weight to the duration than to the amplitude criterion, the results for the ratio series generally favor the S_c over the S_u simulations, and do so rather more strongly than the findings based on the level comparisons. This conclusion also applies to the simulations for *GNP58* and other variables.

Using the ratio series, cumulated diffusion indexes (CDI) were constructed

---

[11] The random shock hypothesis here considered asserts, in the formulation by Frisch (1933), that "the majority of the economic oscillations . . . seem to be explained most plausibly as free oscillations. . . . The most important feature of the free oscillations is that the length of the cycles and the tendency towards dampening are determined by the intrinsic structure of the swinging system, while the intensity (the amplitude) of the fluctuations is determined primarily by the exterior impulse." This suggests that the amplitudes of movements in the stochastic S series would depend mainly on the simulator's decision as to the magnitude of the shocks applied. They may be quite different from the amplitudes of the actuals, not because of any failure of the model to reproduce the basic structure of the economy, but because the impulses (shocks) have not been properly scaled.

Table 9.4 **Stochastic 100-Quarter Simulations (Ratio Series) for Three Models, Selected Summary Measures of Duration and Amplitude and Relative Frequency Distributions of Leads and Lags at Turns in CDIs**

	Wharton Model			OBE Model			Brookings Model	
		Simulations with			Simulations with			Simulations with
	Sample-Period Actuals (*A*) (1)	Nonautocorrelated Shocks (S_u) (2)	Autocorrelated Shocks (S_c) (3)	Sample-Period Actuals (*A*) (4)	Nonautocorrelated Shocks (S_u) (5)	Autocorrelated Shocks (S_c) (6)	Sample-Period Actuals (*A*) (7)	Autocorrelated Shocks (S_c) (8)
				Durations (quarters)[b]				
Quarterly movements in GNP (actual and simulated)[a]								
1. Rises	3.4(2.1)	1.7(0.8)	2.2(1.4)	3.1(3.4)	1.9(1.1)	2.7(1.9)	3.3(1.7)	2.4(1.5)
2. Declines	2.6(1.6)	1.7(0.8)	2.3(1.9)	2.3(1.9)	1.9(1.1)	2.6(1.8)	2.5(1.3)	2.3(1.4)
				Amplitude (percentage)[c]				
3. Rises	0.9	0.9(0.1)	0.4(.05)	0.6	0.4(.05)	0.3(.04)	0.8	0.34(0.06)
4. Declines	1.0	0.9(0.0)	0.4(.05)	0.6	0.4(.05)	0.3(.05)	0.9	0.33(0.06)

Relative Frequency of Leads and Lags (percentage of all turns)[d]

Group of variables[e]	*Leads*	*Coincidences*	*Lags*	*Leads*	*Coincidences*	*Lags*	*Leads*	*Coincidences*	*Lags*
5. Leading	48	26	26	43	28	29	36	21	43
6. Roughly coincident	23	43	34	22	47	30	33	44	23
7. Lagging	28	21	51	24	25	51	31	24	45

Note: See text for the explanation of simulations with nonautocorrelated and autocorrelated shocks (S_u and S_c), of the form in which they are used here (as ratios to the corresponding control series), and of the timing comparisons underlying the entries in line 5–7 (observations at reference—CDI—peaks and troughs).

[a] Actual (*A*): relative deviation of *GNP* from its exponential trend. Simulated (S_u and S_c): relative deviation of shocked from control series for *GNP*. On the meaning of "rises" and "declines," see n. 10. The entries in lines 1–4, cols. 2, 3, 5, and 6 are based on all available simulation runs.

[b] For *A*: mean duration, with standard deviation in parentheses. For *S*: mean duration per run, with mean standard deviation of the durations "within the run" in parentheses.

[c] For *A*: mean amplitude. For *S*: mean amplitude per run, with standard deviation of means per run in parentheses (in percentages, at quarterly rate).

[d] Based on comparisons of the turning points in the simulated ratio series with the corresponding reference dates—peaks and troughs in the CDIs.

[e] Classified according to the timing of the historical series. See table 9.1.

for three randomly chosen runs of the Wharton model and three of the OBE model. For either model, the selection includes one set of series based on S_u and two based on S_c simulations. For the Brookings model, the CDI indexes were computed for two sets of the S_c series. The indexes are of the "historical" type: after the cyclical turning points have been identified in each of the simulated ratio series in a given set, the percentage of the series undergoing specific-cycle expansion is calculated for each quarter and then the deviations of these percentage figures from 50 are cumulated. Each of the CDIs shows reasonably well defined cyclical movements, whose turning-point dates can be used as a reference chronology with which to compare the timing of the simulated series in the given set. The average durations of the specific cycles in the CDIs (about 13, 15, and 18 quarters for Brookings, OBE, and Wharton, respectively) are smaller than those of the postwar (1948–68) cycles in trend-adjusted *GNP* and *GNP58* (18–20 quarters). This reflects mainly short expansions in the indexes, but the overall differences for some of the runs are not large.

In general, the series resulting from simulations with autocorrelated shocks conform better to the reference indexes (CDIs) than the series resulting from simulations with nonautocorrelated shocks, because the former have fewer "extra" turns than the latter. The comprehensive indicators of national product, income, and expenditures, which historically rank high on conformity, also score relatively well according to these comparisons.

There is considerable correspondence between the relative timing of the ex ante stochastic simulations and of the historical data for the same variables, as indicated by the average leads and lags of the ratio series at the major turns in the CDIs. Indeed, the distributions of the timing observations for the ratio series (table 9.4, lines 5–7) appear to be appreciably better than those for the sample-period simulations in identifying the coinciders. However, they are not so sharp in differentiating between the groups of typical leaders and laggers (see table 9.3, lines 6–14), particularly because of discrepancies relating to several of the leading series. Also, the total picture is less favorable than the distributions alone would imply, for many turns in the more volatile ratio series (particularly from the S_u runs) cannot be matched with the reference turns, and some that can be are difficult to date, so that the timing comparisons are rather uncertain.

9.5 Concluding Remarks

To produce any cyclical movements, the models included in this study seem to require perturbations in either the exogenous variables or the relationships with endogenous variables or both. Even the best stochastic simulations here obtained—those with serially correlated shocks to the equations—show only residual cyclical elements, much weaker than those observed in the historical series used in the estimation of the models. This is a disappointing result, assuming that it is reasonable to expect the stochastic simulations to reproduce the recent pattern of the economy's movement at least over several years beyond the sample period. Errors in either the estimates of the disturbances or in the structure of the models could account for this finding.

The absence of shocks or fluctuations in the projected exogenous variables is an unrealistic feature that is likely to be partly responsible for the weakness of the cyclical elements in the stochastic ex ante simulations. Further experiments should test whether this weakness can be remedied by imposing more or less sporadic disturbances on the exogenous factors—or, better, to what extent it can be reduced. There are some indications that the role of such exogenous movements may be large, but the evidence is still very fragmentary (Green, Liebenberg, and Hirsch 1972). Moreover, it is possible that the general picture conveyed by the simulations is seriously distorted by specification errors in the models; certainly, important errors of this sort would tend to obscure the meaning of the evidence that the simulations can provide.[12] Future simulation studies, therefore, should be combined with a comparative analysis of misspecifications in the models covered.

A more limited task that could be readily accomplished with the materials already collected is to examine larger samples of the stochastic simulations. Also, to compare the models with regard to their ability to approximate the main characteristics of major short-term fluctuations of the economy, there is need for more standardized simulations—at least for a suitable common sample period for the different systems. Finally, the simulation studies should be extended to other recent models and to revised versions of the included models. The more varied the assortment of the represented systems, the more we are likely to learn from this research.

---

[12] This point was repeatedly made in discussions at the 1969 Harvard conference. See Bert G. Hickman's (1972) introduction to *Econometric Models of Cyclical Behavior.*