*J. W. Elliott**

A Direct Comparison of Short-Run GNP Forecasting Models

Over the past several years, empirical evidence has accumulated suggesting that GNP forecasting may be performed as well by small-scale Keynesian econometric models such as the Friend-Taubman system as by larger models of the Wharton variety.[1] More recently, the development of distinctly non-Keynesian aggregate econometric models by the Saint Louis Federal Reserve and by the Office of Management and Budget has produced statements both in and out of print as to their probable capacities for short-term economic forecasting. At the same time, other work has appeared suggesting that theory-inspired models of Keynesian or monetarist inclination may offer no improvement over essentially ad hoc structures defensible only as the reduced forms of unknown structural systems. All these developments invite objective comparison of the relative forecasting capacities of associated econometric structures, leading to conclusions as to the strengths of claims made in their behalf.

Such a comparison is the purpose of the research reported in this paper. The models involved in the comparison (all involving quarterly data) are (1) the Friend-Taubman small-scale Keynesian model, (2) an ad hoc model built from key variables in the Wharton model by M. K. Evans, (3) a quarterly version of a model proposed by G. Moore based on leading NBER indicators, (4) the monetarist model of the Saint Louis Federal Reserve, and (5) the Office of Management and Budget model developed by A. Laffer and D. Ranson.

GROUND RULES FOR THE COMPARISON

To compare these models, several constraints and ground rules were established in order to increase the objectivity and fairness of the results.

First, the same data period and sample size were used to establish the coefficients in all models. For example, the first data period was

* School of Business Administration, University of Wisconsin—Milwaukee. The author is grateful to Keith Carlson for data and advice supplied regarding the St. Louis model, to M. E. Echols and members of the faculty research seminar at the School of Business Administration for their critique of this work, and to S. Salnas for his assistance with the computations.

1. For example, see M. K. Evans, *Macroeconomic Activity: Theory, Forecasting, and Control* (New York: Harper & Row, 1969); and Herman O. Stekler, "Forecasting with Econometric Models: An Evaluation," *Econometrica* 36, nos. 3–4 (July-October 1968): 437–63.

the 30 quarters beginning with the third quarter of 1955 (1955.3) and ending with 1962.4 (fourth quarter of 1962), leading to a forecast for 1963.1, 1963.2, and 1963.3. All models were estimated with the 30 data points of this period.[2] Then, the prediction time frame was repositioned one quarter ahead, with the data period now comprised of the 30 quarters from 1955.4 to 1963.1, and forecasts were made for 1963.2, 1963.3, and 1963.4. The purpose of this overlapping definition was to reflect in the forecasts the results of the short-term forecaster's actual ability to update his coefficients every quarter by adding current data. Such appears to be the actual practice involved in most short-term econometric forecasting systems. Thus, we did not forecast more than three quarters ahead with a given set of coefficients. Each of the forecasts for one quarter ahead thus results from a different set of updated coefficients. In all, 32 successive overlapping data fields were employed, with the final field running from 1963.2 to 1970.3, producing forecasts for 1970.4, 1971.1, and 1971.2. This gave 32 successive one-quarter-ahead forecasts for each of the five models, along with an equal number of two-quarter-ahead and three-quarter-ahead forecasts with each model.[3] The forecast period was limited to the 1963–70 period by data limitations for some variables for periods prior to 1955.3. This forecast period is relatively well balanced in terms of rapid and slow growth periods, with 16 quarters showing real GNP increases of 5 billion or greater and 16 quarters with growth of less than 5 billion. However, only three quarters had absolute declines in real GNP, thus limiting the ability to draw well-supported conclusions about quarterly turning-point performance of the models compared.

A second constraint in the comparison deals with the admissible information assumed known during the forecasting period. In order to achieve highly comparable predictive comparisons for the model structures, it was necessary to achieve a high degree of alignment among models regarding the information assumed given during the forecast period. A completely ex ante approach was neither possible nor desirable as a basis for testing the models, since it would be next to impossible to separate model predictability from predictability of the unlagged exogenous variables, even if a history of ex ante predictions could be compiled.[4] Thus, the constraint established was homogeneity in the unlagged exogenous variables assumed given during the forecasting period. Initial consideration of the five models revealed that the

2. However, some variables reach further back into the data due to their definition as lagged or as weighted or moving averages. These additional data needs are not included in the stipulation of 30 observations. The latter refers to the number of data points used in the analysis.

3. In real GNP predictions with the five models, various other variables were predicted as steps in deriving real GNP estimates. These data are available upon request from the author.

4. See Herman O. Stekler, "Forecasting with an Econometric Model: Comment," *American Economic Review* 56, no. 5 (December 1966): 1241–48.

level of government expenditures was the only such exogenous variable common among the models.

As a result, government expenditures were determined to be the only unlagged predetermined variable considered "given." Thus, the forecasting results reported here must be considered conditional or ex post in nature due to the government expenditure term. All other exogenous variables in the five models, if not already lagged, were lagged sufficiently to enable their value to be known at the beginning of the forecasting period. This condition resulted in some modifications in the original time lags specified in the models considered. While this was an undesirable modification to introduce into the comparison, it achieved an exact alignment of the assumed known variables among all models. Comparisons of residual variation among original and modified models consistently indicated that the goodness-of-fit effect of the time lag changes made was minimal, due undoubtedly to the normally high degree of serial correlation present among involved variables. In what follows, we assume that goodness-of-fit comparisons between original and modified models will approximate forecasting comparisons between original and modified models, that is, that such goodness-of-fit comparisons are a useful basis for judgments as to whether the modified models are faithful representations of the original versions.

The model comparisons reported on here deal exclusively with real GNP. Since real GNP blends prices and nominal expenditure behavior, it reflects the combined uncertainty in these measures and is thus an interesting forecasting candidate. The use of a single output responded (*a*) to the single-dependent variable constraint of the Evans and NBER models and (*b*) to the need to limit the already huge computational requirements of the project.

MODELS COMPARED

The first short-term model in our comparison is that of I. Friend and P. Taubman.[5] The Friend-Taubman (FT) model is of interest for several reasons. First, an earlier comparison by Stekler of forecasting results obtained from the FT model with those of much larger Keynesian-type models, including the Klein, Office of Business Economics (OBE), Fromm, and Liu models, concludes that the FT model proves very successful compared with those larger models for comparable forecasting periods.[6] Second, a comparison by Evans of the forecasting properties of the FT model relative to those of the much larger Wharton model in predicting real GNP (actually, nominal GNP with prices assumed known) led him to the conclusion that important differences could not be isolated in the accuracy of the two models in predicting

5. Irwin Friend and Paul Taubman, "A Short-Term Forecasting Model," *Review of Economics and Statistics* 46, no. 3 (August 1964): 229–36.

6. Stekler (n. 4, above), pp. 1242–43.

real GNP over the nine-quarter period from 1963.2 to 1965.2.[7] However, as Evans points out, this period was characterized by continuous economic expansion, which may not provide an adequately rigorous test of the predictive differences of the two models.

Stekler and Evans's work suggests that the small-scale, crude Keynesian (i.e., no liquidity-preference monetary linkages) FT model might be expected to perform comparably with larger-scale Keynesian models in real GNP predictions. Of course, the greater range of variables predicted by larger models enhances the general usefulness of these models compared with small models. But for the real GNP predictions of interest in this comparison, the available evidence suggests we would gain little by including a much less computationally tractable larger Keynesian model in our analysis over the results obtainable by use of the FT model.

The FT model is given by the authors as follows:

$$\Delta C = C_{11} + C_{12}\,\Delta\widetilde{Y} + C_{13}\Delta C_{-1} + \mu_1, \tag{1}$$

$$\Delta H = C_{21} + C_{22}(\Delta\widetilde{Y} - \Delta Y_{-1}) + C_{23}\Delta HS_{-\frac{1}{2}} + C_{24}\Delta PE^e + \mu_2, \tag{2}$$

$$\Delta PE = C_{31} + C_{32}(\Delta\widetilde{Y} - \Delta Y_{-1}) + C_{33}\Delta PE^e + \mu_3, \tag{3}$$

$$\Delta I = C_{41} + C_{42}\Delta S^e + C_{43}I_{-1} + C_{44}\Delta PE^e + \mu_4, \tag{4}$$

$$\Delta Y = \Delta C + \Delta H + \Delta PE + \Delta I + \Delta G', \tag{5}$$

where: ΔC = change in personal consumption expenditures, constant dollars; ΔY = change in GNP, constant dollars; $\Delta\widetilde{Y}$ = computed ΔY from first stage; ΔH = change in residential construction expenditures, constant dollars; ΔHS = change in housing starts; ΔPE^e = plant and equipment anticipations, Commerce Department; ΔPE = plant and equipment expenditures, constant dollars; ΔI = change in inventory investment, constant dollars; ΔS^e = business sales anticipations, Commerce Department; $\Delta G'$ = total government purchases of goods and services plus net exports, constant dollars; and where $\Delta G'$, ΔHS, ΔPE^e, and ΔS^e are considered autonomous. Given the constraints of this comparison on the admissible nonlagged autonomous variables, the original FT model specification had to be somewhat modified. First, the assumption contained in the $\Delta G'$ term that net exports are given was modified. This was accomplished by adding an equation for net exports, in which net exports are explained by the volume of world trade, by current real GNP, and by past patterns of net export expenditures.[8] Since ΔPE^e and ΔS^e are anticipatory variables available at the start of the forecast period, they need not be lagged for one-period-ahead forecasts. Likewise, hous-

7. Evans, pp. 490–91.

8. The choice of variables for the explanation is similar to a simplified reduced form of the export-import sector of the Wharton model. See Evans, pp. 221–43.

ing starts are lagged sufficiently to be known at the beginning of the forecast period. For two- and three-quarters-ahead forecasts, lagged values of ΔPE^e and ΔS^e as well as further lags in ΔHS are required to maintain the constraint on variables assumed known, resulting in three versions of the FT model, one each for one-, two-, and three-quarters-ahead forecasts. Accordingly, the modified structure is as follows:

$$\Delta C = C_{11} + C_{12}\Delta\widetilde{Y} + C_{13}\Delta C_{-1} + \mu_1, \quad \text{(1) M1,M2,M3}$$

$$\begin{aligned}\Delta H &= C_{21} + C_{22}(\Delta\widetilde{Y} - \Delta Y_{-1}) \\ &\qquad + C_{23}\Delta HS_{-1/2} + C_{24}\Delta PE^e + \mu_2 \quad \text{(2) M1} \\ &= C_{21}^* + C_{22}^*(\Delta\widetilde{Y} - \Delta Y_{-1}) \\ &\qquad + C_{23}^*\Delta HS_{-3/2} + C_{24}^*\Delta PE_{-1}^e + \mu_2^* \quad \text{(2) M2} \\ &= C_{21}^{**} + C_{22}^{**}(\Delta\widetilde{Y} - \Delta Y_{-1}) \\ &\qquad + C_{23}^{**}\Delta HS_{-5/2} + C_{24}^{**}\Delta PE_{-2}^e + \mu_2^*, \quad \text{(2) M3}\end{aligned}$$

$$\begin{aligned}\Delta PE &= C_{31} + C_{32}(\Delta\widetilde{Y} - \Delta Y_{-1}) \\ &\qquad + C_{33}\Delta PE^e + \mu_3 \quad \text{(3) M1} \\ &= C_{31}^* + C_{32}^*(\Delta\widetilde{Y} - \Delta Y_{-1}) \\ &\qquad + C_{33}^*\Delta PE_{-1}^e + \mu_3^* \quad \text{(3) M2} \\ &= C_{31}^{**} + C_{32}^{**}(\Delta\widetilde{Y} - \Delta Y_{-1}) \\ &\qquad + C_{33}^{**}\Delta PE_{-2}^e + \mu_3^{**}, \quad \text{(3) M3}\end{aligned}$$

$$\begin{aligned}\Delta I &= C_{41} + C_{42}\Delta S^e + C_{43}I_{-1} + C_{44}\Delta PE^e + \mu_4 \quad \text{(4) M1} \\ &= C_{41}^* + C_{42}^*\Delta S^e{}_{-1} + C_{43}^*I_{-1} \\ &\qquad + C_{44}^*\Delta PE_{-1}^e + \mu_4^* \quad \text{(4) M2} \\ &= C_{41}^{**} + C_{42}^{**}\Delta S^e{}_{-2} + C_{43}^{**}I_{-1} \\ &\qquad + C_{44}^{**}\Delta PE_{-2}^e + \mu_4^{**}, \quad \text{(4) M3}\end{aligned}$$

$$\begin{aligned}\Delta NEX &= C_{51} + C_{52}\Delta\widetilde{Y} + C_{53}WT_{-1} \\ &\qquad + C_{54}\sum_{i=1}^{3} W_i\Delta NEX_{t-i} + \mu_5 \quad \text{(5) M1} \\ &= C_{51}^* + C_{52}^*\Delta\widetilde{Y} + C_{53}^*WT_{-2} \\ &\qquad + C_{54}^*\sum_{i=1}^{3} W_i\Delta NEX_{t-i} + \mu_5^* \quad \text{(5) M2} \\ &= C_{51}^{**} + C_{52}^{**}\Delta\widetilde{Y} + C_{53}^{**}WT_{-3} \\ &\qquad + C_{54}^{**}\sum_{i=1}^{3} W_i\Delta NEX_{t-i} + \mu_5^{**}, \quad \text{(5) M3}\end{aligned}$$

$$\Delta Y = \Delta C + \Delta H + \Delta PE + \Delta I + \Delta NEX + \Delta G$$

where new variables are WT = quantum index of world trade, 1964 = 100; ΔNEX = change in net exports, constant dollars; W_i = exponentially declining normalized weights,

$$\sum_{i=1}^{3} W_i = 1$$

ΔG = change in government purchases of goods and services, and where M1, M2, and M3 refer to the one-quarter-ahead, two-quarters-ahead, and three-quarters-ahead versions of the model, respectively. In evaluating the relationship of forecasting results previously obtained by the original FT model with those obtained from the M1, M2, and M3 modifications, it is useful to compare goodness-of-fit statistics. The results of such an analysis are given in table 1. In addition to the standard

Table 1
Goodness of Fit, Original and Modified Friend-Taubman Model

		S.E.			
	Data Period	ΔC	ΔH	ΔPE	ΔI
Original FT	1953.1–1960.4	2.06	0.58	0.95	1.98
Reestimated FT	1958.3–1965.4	1.70	0.48	0.92	1.58
Modified FT (M1)	1958.3–1965.4	0.93	0.51	0.61	1.79
Modified FT (M2)	1958.3–1965.4	0.85	0.43	0.63	1.82
Modified FT (M3)	1958.3–1965.4	0.97	0.60	0.60	1.84

errors associated with each of the original four structural equations as reported by FT, the table includes this measure for the identical model fit to more recent data representing approximately the middle of the present sample period. Finally, the table shows comparable standard errors for the MI, M2, and M3 modifications. When the original and reestimated models are compared, a reasonable correspondence is observed for every equation. When the modified models are compared with the reestimated original over the same time period, no sizable loss in goodness of fit is apparent, although something of an expected deterioration in goodness of fit appears to occur as the more distant time lags of M3 are approached. However, the overall conclusion suggested by table 1 is that M1, M2, and M3 remain quite reasonable representations of the FT system in terms of goodness of fit. These three versions of the FT model were each estimated for the 32 successive overlapping time bases.

A one-quarter-ahead forecast was made with M1, one- and two-quarters-ahead forecasts with M2, and one-, two-, and three-quarters-ahead forecasts with M3 for the period just beyond each time base. In the original FT predictions, the equations were estimated for overlapping semiannual periods, resulting in overlapping semiannaul predictions.

This procedure was followed precisely in the present analysis, with quarterly predictions extracted from these semiannual projections by algebraic solution. Finally, two-stage least-squares estimating procedures were employed following FT.

The second model in our comparison was proposed by M. K. Evans.[9]

In commenting on the comparative forecasting properties of large-scale versus small-scale econometric models, Evans states, "If one is primarily interested in predicting the level of GNP, there is no particular reason a five-equation model should represent the optimal size. For simplicity and ease of computation, it could be argued that the reduced form of such a model, presented as a compact one-equation annual model, might be even better." [P. 492]

Evans has formulated such a reduced form built on selected key variables from the Wharton model. In an estimate of this model based on annual data from 1948–64, quite reasonable goodness-of-fit statistics were obtained by Evans. In addition, favorable forecasting results were obtained for 1965 and 1966.[10] The Evans model is given as follows:

$$\Delta X = C_1 + C_2\Delta G + C_3\Delta I_{-1} + C_4\Delta iL_{-1} + C_5\Delta U_{-1} + C_6\Delta CNS_{-1} + C_7 STXR + \mu_6, \quad (6)$$

where: ΔX = change in GNP, constant dollars; ΔG = change in government purchases, constant dollars; ΔI = inventory investment excluding farm inventories, constant dollars; ΔiL = change in long-term bond yield (Moody's); ΔU = change in unfilled orders, durables, end of year, constant dollars; ΔCNS = change in consumption, nondurables and services, constant dollars; $STXR$ = starting (first bracket) tax rate for personal income taxes and where all lags are in years.

The inclusion of Evans's model in the present comparison will provide an interesting quarterly test of his proposition now evaluated only with annual data. To effect this comparison, it is necessary to develop a quarterly version of his model which maintains our limitations on known exogenous variables. All the variables included in (6) are available quarterly. Therefore, it is possible to estimate the model with quarterly rather than annual data.[11] This procedure was followed with one exception. Initial ordinary least-squares estimates produced noticeably poor results for the starting tax rate ($STXR$) variable, confirming suspicions that this variable may not remain the most appro-

9. Evans, pp. 491–93.

10. Evans, p. 493.

11. In Evans's analysis, the annual lags in nondurables and services, inventory change, unfilled orders, and interest rates seem more directed at the serial relationships in the annual series than a reflection of specific distributed lags. Therefore, it appeared more appropriate to represent the same serial relationships in the quarterly model than to assume that a lag of −1 in the annual model translates to a −4 in the quarterly model.

priate measure of the effects of tax structure for a quarterly framework. As an alternative, the average tax rate was employed, representing a blend of legislative and economic fluctuations. This variable was retained on the basis of its more favorable empirical results. This gave the one-quarter-ahead Evans model as

$$\Delta X = C_1 + C_2 \Delta G + C_3 \Delta I_{-1} + C_4 \Delta iL_{-1} + C_5 \Delta U_{-1} + C_6 \Delta CNS_{-1} + C_7 ATX_{-1} + \mu_6, \quad (6)\ \mathrm{M1}$$

where ATX = the ratio of personal taxes to disposable personal income and where lags are in quarters.

Accordingly, the two-quarters-ahead version (M2) is identical except for an additional quarter's lag in ΔI, ΔiL, ΔU, ΔCNS, and ATX. The M3 version is similar, employing an additional lag in these five variables, thus retaining our constraint on the variables assumed known during the forecasting period. Some indication as to the goodness-of-fit consequences of these successive lags can be seen by examining the standard error for each of the three versions. Again, picking a 30-quarter period about in the middle of our time frame, 1958.3–1965.4, we obtain the following: for M1, S.E. = 4.95; for M2, S.E. = 4.76; for M3, S.E. = 5.00, which indicates no clear pattern of decay in the goodness of fit as more distant lags are substituted into the model.[12] As in the FT case, this stability suggests the effects of serial correlation in the lagged variables is to make the model's fit somewhat insensitive to the choice of lag period within our limits.

A third approach to quantitative aggregate economic forecasting, proposed by Geoffrey Moore, depends entirely on the NBER leading indicators.[13] Using annual data, Moore fits regressions involving an index of leading indicators as the single independent variable and both nominal and real GNP as alternative dependent variables. Whereas the Evans model may be described as only indirectly related to the Keynesian theoretical apparatus via the variables having arisen essentially from the Wharton model, Moore's approach explicitly rejects reliance on any theoretical framework, as the leading indicators which comprise the index have emerged in a manner independent of any underlying theory.

Moore has developed some evidence suggesting that the annual real GNP forecasts of a simple model built on the basis of an index of leading indicators performs nearly as well or better than (1) a sample of eight noneconometric business forecasts, (2) the forecasts of the President's Council of Economic Advisors as published in the *Economic Report of the President,* and (3) the econometric model developed by

12. For the model fit to the entire time period 1955.3–1970.3, the results are similar: M1, 5.10; M2, 4.98; M3, 5.35.

13. See Geoffrey H. Moore, "Forecasting Short-term Economic Change," *Journal of the American Statistical Association* 64, no. 325 (March 1969): 1–22.

D. Suits.[14] In commenting on these results, Moore states, "For the most part, the forecasts of change in GNP based upon the leading index compare favorably in terms of magnitude of error and degree of correlation. The experiment suggests, therefore, that this mechanical use of the leading index can produce a standard of forecasting accuracy which is not easy to surpass" (p. 17). Moore's work raises an interesting question as to how well we may expect a quarterly model based entirely on leading indicators to perform. To develop evidence on this question, a quarterly model was developed based on his ideas and constructed in accordance with the criteria established for the present comparison. In this effort, the initial approach was to form the simple regression of real GNP on the index of leading indicators in the month preceding the forecast period. Statistical results were discouraging. This approach involves accepting the weights assigned by the index to each of the 12 component indicators, a potentially risky and unnecessary limitation. To explore a more intuitively attractive alternative, all 12 leading indicators on the NBER's short list were incorporated into a model explaining real GNP, with the weights established for each of the 32 time bases by regression. The choice of time lags in the operation of these variables responds to an earlier study by Moore and Shiskin in which the authors defined the median period by which each of the 12 short-list leading indicators leads NBER business cycle peaks and valleys.[15] Their results are shown in table 2.[16]

Table 2
Moore-Shiskin Lag

Indicator	Median Lead (Months)
1. Average work week, production workers (*HR*)	5
2. Nonagricultural placements (*PL*)	3
3. Index of net business formations (*FM*)	7
4. New orders, durable goods (*OD*)	4
5. Contracts and orders, plant and equipment (*PE*)	6
6. New buildings permits, private housing units (*BP*)	6
7. Change in book value, manufacturing, and trade inventories (*IV*)	8
8. Industrial materials prices (*MP*)	2
9. Stock prices, 500 common stocks (*SK*)	4
10. Corporate profits after taxes (*CP*)	2
11. Ratio, price to unit labor cost, manufacturing (*PC*)	3
12. Change in consumer installment debt (*ID*)	10

In constructing a model employing these 12 variables, the translation between monthly independent variables and a quarterly dependent variable is accomplished by assuming the second month of the

14. Moore, p. 18, table 7.

15. See Geoffrey H. Moore and Julius Shiskin, *Indicators of Business Expansions and Contractions*, NBER Occasional Paper 103 (New York: Columbia University Press, 1967).

16. See Moore and Shiskin, pp. 36–39, table 6, for details.

quarter represents the average value of real GNP over the quarter. Thus, a 5-month lead period would translate into a lead of 4 months prior to the start of the quarter. In addition, in the absence of government expenditures among the leading indicators, a method had to be found to represent the admissible assumption of government expenditures as known. This was handled by subtracting the known government expenditure change from the change in real GNP and predicting the remainder. Accordingly, the M1 version of the NBER model is given as

$$\begin{aligned} X' = C_1 + C_2HR_{-4} + C_3PL_{2-} + C_4FM_{-6} + C_5OD_{-3} \\ + C_6PE_{-5} + C_7BP_{-5} + C_8IV_{-7} + C_9MP_{-1} \\ + C_{10}SK_{-3} + C_{11}CP_{-3} + C_{12}PC_{-2} + C_{13}ID_{-9} + \mu_7, \end{aligned} \quad (7)\ \text{M1}$$

where $\Delta X'$ = change in real GNP less change in real government purchases, where all other variables are as defined in table 2, and where all lags are in months. The only exception to Moore and Shiskin's median lags in equation (7) occurs for corporate profits (CP), which is a quarterly series and does not allow the 2-month lead specified by Moore and Shiskin.

Following the procedure used in the first two models, I introduced sufficient lags in the two- and three-quarters-ahead versions of the model to enable knowledge of all variables. Specifically the changes were for M2, PL_{-4}, OD_{-4}, MP_{-4}, SK_{-4}, CP_{-6}, PC_{-4}; for M3, HR_{-7}, PL_{-7}, FM_{-7}, OD_{-7}, PE_{-7}, BP_{-7}, MP_{-7}, SK_{-7}, CP_{-9}, PC_{-7}. The goodness of fit (1958.3–1965.4) of the M2 and M3 versions of the model do not particularly deteriorate as a result of introducing longer lags than those of Moore and Shiskin in M1;[17] S.E. for M1 = 4.67; for M2, 5.05; for M3, 4.29.

The fourth model in our comparison has been developed by the Saint Louis Federal Reserve to represent a monetarist approach to economic forecasting.[18] In the author's words, "The general monetarist view is that the rate of monetary expansion is the main determinant of total spending, commonly measured by Gross National Product (GNP). Changes in total spending, in turn, influence movements in output, employment, and the general price level" (p. 8). In accordance with this view, the Saint Louis model first predicts nominal GNP, then employs this value to predict prices. Real GNP is then derived through an identity. The specific equations are

$$\Delta Y_t = a_0 + \sum_{i=0}^{4} m_i \Delta M_{t-i} + \sum_{i=0}^{4} e_i \Delta E_{t-i} + \mu_8, \quad (8)$$

17. Other time periods show similar results.

18. See Leonall C. Anderson and Keith M. Carlson, "A Monetarist Model for Economic Stabilization," *Federal Reserve Bank of St. Louis Review* 52, no. 4 (April 1970): 7–21.

$$\Delta P_t = b_0 + \sum_{i=0}^{5} d_i D_{t-i} + b_1 \Delta P_t{}^A + \mu_9, \tag{9}$$

$$D_t = \Delta Y_t - (X_t{}^F - X_{t-1}), \tag{10}$$

$$\Delta P_t{}^A = Y_{t-2} \left\{ \left[\left(\sum_{i=1}^{17} p_i \frac{\dot{P}_{t-i}}{U_{t-i/4}} \right) .01 + 1 \right]^{1/4} - 1 \right\}, \tag{11}$$

where

$$\Delta X_t = \Delta Y_t - \Delta P_t, \tag{12}$$

$$U_t = c_0 + c_1 G_t + c_2 G_{t-1} + \mu_{10}, \tag{13}$$

and where p_i in (11) are obtained through regression performed on (14),

$$R_t{}^L = d_0 + d_1 \dot{M}_t + d_2 Z_t + \sum_{i=0}^{16} x_i \dot{X}_{t-i} + \sum_{i=0}^{16} p_i \left(\frac{\dot{P}_{t-i}}{U_{t-i/4}} \right) + \mu_{11}. \tag{14}$$

Variables are: Y = nominal GNP; ΔY = change in nominal GNP; ΔM = change in money supply, narrowly defined; $\dot{M}$ = compounded rate of change in money supply; ΔE = change in high-employment government expenditures,[19] D = demand pressure; ΔP^A = anticipated price change; ΔP = dollar change in total spending due to price change; X^F = potential output; X = GNP in constant dollars; $\dot{X}$ = compounded rate of change in constant dollar GNP; ΔX = dollar change in Y due to real output change; $\dot{P}$ = annual compounded rate of change in GNP deflator; U = unemployment rate, R^L = Moody's corporate Aaa bond rate in quarter *t; Z* = credit dummy variable (0 for 1955.3–1960.4, 1 for 1961.1–1970A); G_t = full-employment output gap;[20] x, m, e, p, d = Almon weights, computed with various polynominal constraints.[21]

In contrast with earlier models discussed in this comparison, the Saint Louis prediction of real GNP movements is largely conditioned by money supply changes, particularly since the m coefficients in (8) which attach to the money supply are considerably larger than the e coefficients which attach to government expenditure changes.

19. For details of this data series, see Keith M. Carlson, "Estimates of the High-Employment Budget: 1947–1967," *Federal Reserve Bank of St. Louis Review* 49, no. 6 (June 1967): 6–14.

20. $G_t = (X_t{}^F - X_t)/X_t{}^F$.

21. See Anderson and Carlson for details of the constraints and on the application of Almon weights.

$$\sum_{i=0}^{4} m_i = 5.57$$

versus

$$\sum_{i=0}^{4} e_i = 0.17$$

for their data. Thus, including this model in our comparison should provide direct evidence as to the value of monetary data in short-term economic forecasting relative to that of alternative economic variables assigned key explanatory roles in other models. In addition, the present comparative test of the Saint Louis model will enable comments upon certain of the author's conclusions, for example, "the suggestion is offered that a small model constructed within a monetarist framework may yield as much information about the key aggregates as a large structural model" (p. 18).

In addition, a recent study by Nelson has produced the conclusion that the large-scale FRB-MIT Penn monetarist-type model predicts no better than conventional naïve models under conditions where current values of the exogenous variables are assumed known.[22] Measuring the Saint Louis model against naïve base lines can thus effect an indirect comparison of the Saint Louis results with those obtained by Nelson.

The versions of the model tested differ from the original Saint Louis model only in the time lags involved, where were lengthened where necessary to keep with our known-information ground rule. Thus, the M1 version of the model tested is

$$\Delta Y_t = a_0 + \sum_{i=1}^{5} m_i \Delta M_{t-i} + \sum_{i=0}^{4} e_i \Delta E_{t-i} + \mu_8, \qquad \text{(8) M1}$$

$$\Delta P_t = b_0 + \sum_{i=0}^{5} d_i D_{t-i} + b_1 \Delta P_t{}^A + \mu_9, \qquad \text{(9) M1}$$

$$R_t{}^L = d_0 + d_1 \dot{M}_{t-1} + d_2 Z + \sum_{i=1}^{17} x_i \dot{X}_{t-i} + \sum_{i=1}^{17} p_i \left(\frac{\dot{P}_{t-i}}{U_{t-i/4}} \right) + \mu_{11}. \qquad \text{(14) M1}$$

Expressions (10), (12), and (13) remain as in the original model. In (11), the p_i Almon weights depend on the results of the regression

22. See Charles Nelson, "The Prediction Performance of the FRB-MIT-Penn Model of the U.S. Economy," Report 7117, Center for Mathematical Studies in Business and Economics, University of Chicago, April, 1971.

in (14). Since the original (14) relies on current values of $\dot{M}$, $\dot{X}$ $\dot{P}$, and U, three versions (M1, M2, M3) are necessary for the computation of the appropriate p_i. Accordingly, three differing estimates of ΔP^A are built from the three sets of p_i and used in the regressions of (9) M1, (9) M2, and (9) M3.

In the M2 version of (8), the summation of the ΔM term runs from two to six and in the M3 version, from three to seven. In the M2 version of (9), the summation on the D term goes from one to six, while in the M3 version the summation goes from two to seven. In addition, the differing ΔP^A values are incorporated in each version as described earlier. Finally, in (14) M2, $\dot{M}_{t-1}$ is replaced by $\dot{M}_{t-2}$ and by $\dot{M}_{t-3}$ in (14) M3.[23]

As in the previous cases, the effect of incorporating these longer time lags on the model's goodness of fit is of interest in judging whether we may expect the M1, M2, and M3 configurations to approach the explanatory properties of the original model. The standard errors shown in table 3 are taken from approximately the middle of the present sample period and are typical of those found throughout the 32 time bases.

Table 3
Saint Louis Goodness-of-Fit Comparisons

Saint Louis Model	Time Base	S.E. ΔY	S.E. ΔP	S.E. U	S.E. R^L
Original	1953.1–1969.4	3.84	1.07	0.30	0.28
M1	1958.3–1965.4	4.21	0.70	0.25	0.13
M2	1958.3–1965.4	4.26	0.70	*	0.14
M3	1958.3–1965.4	4.64	0.71	*	0.13

* Same equation for M2 and M3 versions.

As seen, little if any deterioration in goodness of fit occurs in the tested versions of the model over the original model on this criterion, suggesting that fit of their equations is rather insensitive to the choice of time lags among our three versions. On this basis, I conclude that the comparative forecasting results for M1, M2, and M3 are likely to be representative of those of the original Saint Louis model.

The final model of this comparison was developed by Arthur Laffer and David Ranson for use by the President's Office of Management and Budget. We shall refer to it as the OMB model.[24] Unlike all preceding models, the OMB model employs seasonally unadjusted and unannualized data wherever possible. This is in response to the authors' contention that conventional seasonal adjustment (1) constitutes an

23. This substitution is what gives rise to a separate estimate of the p Almon weights for each version of the model.

24. See Arthur B. Laffer and David R. Ranson, "A Formal Model of the Economy," *Journal of Business* 44, no. 3 (July 1971): 247–70.

artificial smoothing of the data, (2) may introduce autocorrelation, and (3) fails to adjust correctly for degrees of freedom. As an alternative, the OMB approach coadjusts for seasonality through the use of dummy variables. The OMB model is expressly not tied to a particular theoretical construct. In the words of the authors:

With respect to the development of the . . . (OMB) model, the data are left as free as possible to tell their own story. Many of the models in the literature are constructed subject to constraints which represent the *a priori* theoretical beliefs of the model builders. They possess some of the characteristics of simulation as well as the characteristics of hypothesis tests. In other words, inferences drawn from the resutls are conditional on builtin assumptions as to the "true" structural relationships. The OMB model does not fall into this category. Constraints on the form and nature of the equations have been held to the minimum necessary to allow construction of explicit relationships. [P. 249]

In this passage, the OMB model is characterized as a kind of reduced form of an unknown underlying structure. Variables included supposedly represent Keynesian, quantity theory, and efficient markets theories, but in no structured way. The authors contend this agnostic view may lead to better forecasts, along with their unconventional approach to handling seasonality. The present analysis will provide an extended test of these propositions, on both an absolute level as measured against naïve baselines and on a relative level with respect to those models in our comparison which involve explicit theoretically inspired structural restrictions.

In the OMB model, two approaches are employed in predicting real GNP. The first is to predict nominal GNP and the implicit deflator separately and then derive real GNP. The second approach is to predict real GNP directly by an equation which contains all the variables found in either of the two separate equations. In commenting on the real GNP computation resulting from these two approaches, the authors state that the results are virtually independent of the choice between the two. Accepting this finding, I have chosen the direct approach for this analysis, since it is slightly simpler. The OMB equation is

$$\begin{aligned}\Delta Ly = {} & a_0 + a_1 D_1 + a_2 D_2 + a_3 D_3 + a_4 \Delta LM1 \\ & + a_5 \Delta LG + a_6 \Delta LG_{-1} + a_7 \Delta LG_{-2} + a_8 \Delta LG_{-3} + a_9 \Delta SH \\ & + a_{10} \Delta L\text{S\&P}_{-1} + a_{11} \Delta LP_{-1} + a_{12} \Delta LP_{-2} + a_{13} i_{-1} + \mu_{12}, \quad (15)\end{aligned}$$

where ΔLy = quarterly change in the log of real GNP; D_1, D_2, D_3 = seasonal dummies for the first, second, and third quarters; $\Delta LM1$ = quarterly change in the log of the conventional money supply (currency plus demand deposits, revised); ΔLG = quarterly change in the log of federal government purchases of goods and services; ΔSH = quarterly change in the ratio of the proportion of man-hours lost due to strikes to manufacturing employment; $\Delta LS\&P$ = quarterly change in the log Standard and Poor's composite index of common stock prices; ΔLP = quarterly change in the log of the GNP price deflator; i = market yield on 13-week Treasury bills (percentage per annum) converted to continuous compounding at a quarterly rate. In this model, all variables except ΔLP are assumed autonomous by the authors. For one-quarter-ahead forecasts ΔLP is indeed available as it is lagged. For forecasts beyond one quarter, computed ΔLP values are available from the OMB model's price equation, given as follows:

$$\Delta LP = b_0 + b_1 \Delta LP_{-1} + b_2 \Delta LP_{-2} + b_3 \Delta LM1 + b_4 i_{-1} + \mu_{13}, \tag{16}$$

where all variables are as previously defined.

In the present comparison, ΔLG is the only admissible unlagged autonomous variable. Incorporating minimum lags in other autonomous variables where appropriate gives, for the M1 version,

$$\begin{aligned} \Delta Ly = a_0 &+ a_1 D_1 + a_2 D_2 + a_3 D_3 + a_4 \Delta LM1_{-1} + a_5 \Delta LG \\ &+ a_6 \Delta LG_{-1} + a_7 \Delta LG_{-2} + a_8 \Delta LG_{-3} + a_9 \Delta SH_{-1} \\ &+ a_{10} \Delta LS\&P_{-1} + a_{11} \Delta LP_{-1} + a_{12} \Delta LP_{-2} \\ &+ a_{13} i_{-1} + \mu_{12}. \end{aligned} \tag{15) M1 (}$$

For (15) M2, an additional lag is employed for ΔLM, ΔSH, $\Delta LS\&P$, and Δi, while the value of ΔLP_{-1} is obtained from

$$\Delta LP = b_0 + b_1 \Delta LP_{-1} + b_2 \Delta LP_{-2} + b_3 \Delta LM1_{-1} + b_4 i_{-1} + \mu_{13}. \tag{16) M2 (}$$

For (15) M3, a third lag is used in the four variables above, while (16) M3 lags ΔLM and i by an additional quarter and uses the value for ΔLP calculated by (16) M2.

In their discussion of the OMB model, the authors repeatedly stressed their finding that the money supply best enters their system in an unlagged form.[25] When one has successively lagged this variable by one, two, and three quarters in the tested M1, M2, and M3 versions of the model, it becomes particularly important to consider whether this practice has destructively altered the OMB system. As before, we shall focus on standard errors obtained for the equations involved. These data are presented in table 4 for various time periods in this analysis along with the original OMB results. The table shows only slight deterioration in the goodness of fit among the three tested versions of the model.

Indeed, F-tests made on the associated variances show that no significant deviation occurs in residual variation in moving from M1 to M2 to M3 for the three sample periods shown.[26] While it is some-

25. For example Laffer and Ranson, result 2, p. 257, and conclusion 2, p. 264.

26. Inspection also suggests this to be the case for the remaining 29 periods.

Table 4
Goodness of Fit (OMB Model)

Time Period	S.E. ΔLy	S.E. ΔLP
Original:		
1947.4–1969.4	...	0.0051
1952.1–1969.4	...	0.0027
1948.1–1969.4	0.0133	...
M1:		
1955.3–1962.4	0.0140	n.a.
1958.3–1965.4	0.0089	n.a.
1963.4–1970.3	0.0082	n.a.
M2:		
1955.3–1962.4	0.0127	0.0023
1958.3–1965.4	0.0096	0.0015
1963.4–1970.3	0.0085	0.0020
M3:		
1955.3–1962.4	0.0136	0.0023
1958.3–1965.4	0.0097	0.0015
1963.4–1970.3	0.0092	0.0021

NOTE.—n.a. = not applicable.

what dangerous to directly compare the M1, M2, and M3 results with the original OMB results due to differing sample sizes and possibly differing circumstances associated with the sample periods in the two cases, it is quite interesting to note that seven of the nine standard errors for the variously lagged M1, M2, M3 models are lower than the original OMB unlagged results.

As a second test of the comparability of the modified versions of the OMB model with the original, annual average ex post forecasts for 1963–70 were computed in accordance with the procedure described by the OMB authors.[27] These data are shown in table 5 along with the original OMB results. Both the forecasts and the direction of error differ noticeably between the two. However, further inspection reveals that the average absolute error is quite comparable between the models, and the total difference in the forecasts summed over the 8-year period is only 1 billion, that is, the sum total of the forecasts is virtually identical. The suggestion is that the additional time lag of the M1 model affects the timing of predicted change, but not in a systematic way and not in a way which damages the forecasting results in this expost test. Between tables 4 and 5, the general conclusion is that the imposition of the longer time lags of our tested models do no damage to the original OMB goodness-of-fit and ex post results, meaning that contrary to the conclusion of the OMB authors, the performance of the

27. Due to the overlapping 30 quarter data periods of this study, a longer comparison was not possible without piecing together ex post results from several different time bases. This would have damaged the comparison and was not undertaken.

Table 5
OMB Real GNP Predictions—Ex Post

	Annual Ex Post Forecasts, M1 Version of OMB			Original OMB Results		
	Actual	Predicted	Error	Actual	Predicted	Error
1963	551	549	2	551	561	−10
1964	583	579	4	581	583	− 2
1965	617	624	−7	618	614	4
1966	653	661	−8	658	655	3
1967	674	667	7	675	685	−10
1968	707	715	−8	707	703	4
1969	727	725	2	727	722	5
1970	724	731	−7	724	727	− 3
Average absolute error			5.6	...	...	5.1

NOTE.—OMB results are from Laffer and Ranson n. 24, above), p. 262. Actual real GNP figures differ due to our use of the Commerce Department series as of late 1970 and the apparent use by OMB of a later series.

OMB model on these two criteria appears rather insensitive to the choice of currrent or lagged money supply and other variables. This enhances the use of the OMB in our comparison and at the same time casts important doubts on earlier claims by the OMB authors that unlagged money supply is superior to lagged money supply for prediction.

In summary, the five models compared consist of my versions of (1) the Friend-Taubman small-scale, crude Keynesian structure (FT); (2) an essentially Keynesian reduced form based on key variables similar to those in the Wharton system (Evans); (3) a theory-free system based exclusively on NBER leading economic indicators (NBER); (4) the Saint Louis Federal Reserve's monetarist model, based essentially on the quantity theory framework (Saint Louis); and (5) the highly unconventional forecasting apparatus used by the Office of Management and Budget (OMB). To further aid in the comparison, forecasts from two naive autoregressive models were evaluated. These models were:

$$\text{NI}: \quad \Delta Y_t = 0 \qquad \text{(no change)},$$
$$\text{NII}: \quad \Delta Y_t = \Delta Y_{t-1} \qquad \text{(same change)}.$$

RESULTS

Equivalently revised real GNP data extended through 1970.4 at the time of this analysis. Thus, 32 (1963.1–1970.4) one-quarter-ahead, 31 two-quarters-ahead, and 30 three-quarters-ahead forecasts were made. The actual change in real GNP plus forecast errors for all models are shown in Appendix A. As a means to judge these results, average absolute errors and the original version of Theil's inequality coefficient were computed for all versions of the models.[28] These re-

28. See Henri Theil, *Applied Economic Forecasting* (Amsterdam: North-Holland Publishing Co., 1966), p. 28.

Table 6
Inequality Coefficients and Average Absolute Prediction Error in Real GNP

	First Quarter Ahead		Second Quarter Ahead		Third Quarter Ahead	
M1:						
FT	3.9	(0.92)	...	...	...	...
Evans	4.8	(1.68)	...	...	...	...
NBER	5.7	(1.98)	...	...	...	...
Saint Louis	3.6	(1.01)	...	...	...	...
OMB	7.6	(4.49)	...	...	...	...
NI	6.6	(2.41)	...	...	...	...
NII	3.8	(1.00)	...	...	...	...
M2:						
FT	3.7	(0.86)	3.2	(0.68)	...	...
Evans	4.3	(1.25)	4.1	(1.15)	...	...
NBER	6.2	(2.14)	7.1	(2.94)	...	...
Saint Louis	4.2	(1.32)	4.3	(1.47)	...	...
OMB	7.9	(3.97)	8.5	(4.18)	...	...
NI	6.6	(2.41)	6.7	(2.44)	...	...
NII	3.8	(1.00)	3.9	(1.00)	...	...
M3:						
FT	3.5	(0.87)	3.7	(0.87)	3.6	(0.82)
Evans	4.9	(1.70)	4.4	(1.33)	5.4	(1.99)
NBER	5.4	(1.84)	7.2	(2.99)	8.8	(3.90)
Saint Louis	4.6	(1.67)	4.9	(1.97)	4.5	(1.74)
OMB	7.9	(4.58)	8.1	(4.14)	8.4	(3.83)
NI	6.6	(2.41)	6.7	(2.44)	6.7	(2.42)
NII	3.8	(1.00)	3.9	(1.00)	4.0	(1.00)

NOTE.—Entries outside of parentheses are average absolute errors, while numbers in parentheses are inequality coefficients.

sults are shown in table 6. Based on these measures, one-quarter-ahead forecasts made with the M1 versions of the models are most accurately achieved by either the Friend-Taubman of the Saint Louis has a somewhat lower average absolute error due to a few disproportionately large forecasting errors by the Saint Louis model. The Evans model ranks next on both measures. The NBER and OMB systems both show up as largely ineffective at one-quarter-ahead projections in comparison with other models considered. Indeed, the average absolute forecasting error of the OMB model is twice that resulting from the "same-change" NII autoregressive scheme. In the same vein, measured against NII, the best-performing FT and Saint Louis models offer no real improvement over this autoregressive scheme for M1 one-quarter projections, although they each appear to perform about as well as the NII baseline.

Two other sets of one-quarter-ahead results are shown in table 6, involving the longer-lagged M2 and M3 versions of the models. The assumption that the timelags specified in the original model are approximately correct leads to the conclusion that M2 and M3 forecasts for one quarter ahead should progressively deteriorate. This pattern actually occurs in only the Saint Louis case (3.6, 4.2, 4.6) while in the Friend-Taubman case just the opposite occurs (3.9, 3.7, 3.5). In general, these observed forecasting results are consistent with the findings reported earlier that the goodness of fit of the models is generally insensitive to the choice of lag period.[29]

For forecasting the second-quarter ahead with the M2 version of the models, the FT model again ranks highest in terms of the inequality coefficient and also has a lower average absolute forecasting error than any other model. FT thus provides the best model for second-quarter-ahead projections relative to the other four models evaluated. In addition, the FT system improves upon the performance of the "no-change" NII baseline for two-quarter-ahead forecasts, a property not shared by an other model evaluated for that period. The next best model for two quarters ahead appears to be the Evans model, followed by the Saint Louis model. Finally, the NBER and OMB models are noticeably inferior, with the OMB framework again showing an average absolute error more than double that of the NII baseline.

For the second-quarter-ahead forecast with the M3 versions of the model, the longer lags exerted little effect on either the rankings or the size of the absolute error, although the latter did increase for four of the five models.

For forecasts three periods ahead with the M3 version of the models, the Friend-Taubman system is again a better performer in terms of inequality coefficient and average absolute error than the other models and again improves upon the performance of the no-change (NII) scheme. The Saint Louis model rates second on both criteria but is inferior to NII in both respects. As before, the NBER and OMB models are the least useful of those considered for three-quarters-ahead predictions.

As central-tendency-type statistics, both the inequality coefficient and the average absolute forecasting error are influenced by infrequent but larger forecasting errors in a manner indistinguishable from the influence of a large number of smaller forecasting errors. To enlarge on this aspect of our computed statistics, an ordinal ranking of forecasting errors for all models has been undertaken for the M1 one-quarter-ahead, M2 two-quarters-ahead, and M3 three-quarters-ahead forecasts. These results are shown in table 7. Entries in table 7 show the number of times the model listed down the side of the table had a smaller forecasting error than the model listed across the top of the

29. It is interesting to note that the M2 and M3 versions of the FT model continue to show the smallest inequality coefficient for predicting one quarter ahead and, in addition, show the smallest average absolute error. This suggests that its dominance over the other models tested is not greatly influenced by choice of time lags.

Table 7
Comparisons—Forecasting Errors for M1, M2, M3

	NII	Saint Louis	FT	Evans	OMB	NBER
NII:						
M1	...	16	20	18	21 (1)	20
M2	...	16 (1)	14	17 (1)	23 (1)	21
M3	...	13	14	16	22	23
Sum ..	...	45	48	51	66	64
Saint Louis:						
M1	15	...	19 (1)	19	20 (1)	22
M2	13 (1)	...	11 (1)	15	23	19
M3	16	...	15	16	23	24
	44	...	45	50	66	65
FT:						
M1	11	11 (1)	...	17	18	22
M2	16	18 (1)	...	15 (1)	22 (1)	23 (1)
M3	15	14	...	20	22	23
Sum ..	42	43	...	52	62	68
Evans:						
M1	13	12	14	...	18	16 (1)
M2	12 (1)	15	14 (1)	...	20	21
M3	13	13	9	...	19	18
Sum ..	38	40	37	...	57	55
OMB:						
M1	9 (1)	10 (1)	13	13	...	18
M2	6 (1)	7	7 (1)	10	...	15
M3	7	6	7	10	...	16
Sum ..	22	23	27	33	...	49
NBER:						
M1	11	9	9	14 (1)	13	...
M2	9	11	6 (1)	9	15	...
M3	6	5	6	11	13	...
Sum ..	26	25	21	34	41	...

NOTE.—The entry for each row shows the number of forecasts in which the row model (listed vertically) had a smaller forecasting error than the column model (listed horizontally). Numbers in parentheses show ties in forecasting error.

table. For example, the M1 version of the Saint Louis model had a smaller forecasting error than the Friend-Taubman model in 19 of 31 one-quarter forecasts.

The ordinal comparisons of table 7 suggest a stronger performance for the Saint Louis model relative to the FT model than the earlier results. In fact, for one-quarter-ahead forecasts, the Saint Louis model is closer to the mark more often than any other model except NII. For the combined results for M1, M2, and M3, we find the ordinal performance of the Saint Louis, Friend-Taubman, and NII no-change models virtually indistinguishable, as each of these three models outperforms the others in an approximately equal number of the 60 total comparisons.[30] At the same time, any of the three does substantially better than the OMB or NBER models on this criterion and somewhat better than the Evans model.

So, on ordinal criteria, we find little basis upon which to choose between the Friend-Taubman and Saint Louis models. This suggests that if "being close often" is more important than occasionally producing disastrous results, our outcome for one, two, and three quarters is likely to be influenced little by our choice between these two models. For only one-quarter-ahead projections, applying this ordinal criterion shows that the Saint Louis system comes closer than FT in 19 of 30 forecasts with one tie, which suggests an ordinal superiority for this system. But, since this pattern is completely reversed for two-quarter projections (11 of 29 with one tie) and further altered for three-quarter projections (15 of 29), and since the inequality coefficient for the Fried-Taubman system was earlier found to be lower in all three cases, I hesitate to put this finding forth as reflecting a systematic pattern of behavior in our sample.

Another way to view the present forecasting results is in terms of the total cumulative change predicted by the M1, M2, and M3 versions of the models compared with the actual cumulative real GNP change for one, two, and three quarters. The average absolute forecasting errors resulting from this comparison are shown in table 8. Since

Table 8

Cumulative Average Absolute Forecast Error

	Over One Quarter	Over Two Quarters	Over Three Quarters
FT	3.9	4.1	4.4
Evans	4.8	7.4	11.0
NBER	5.7	11.2	14.9
Saint Louis	3.6	6.7	9.0
OMB	7.6	10.8	14.6
NII	3.8	3.9	4.7

NOTE.—Figures are in billions of real GNP and represent the average absolute error of the sum of the M1 one-quarter-ahead forecast error, the M2 two-quarters-ahead forecast error, and the M3 three-quarters-ahead forecast error.

successive residuals are added together in table 8, the results are more of a tracking test than the M2 and M3 results of table 6. In this evaluation, the Friend-Taubman model produces almost as small a cumulative average absolute error for two- and three-quarters-ahead predictions as for one quarter. Every other model shows a clear progression toward higher absolute errors for longer prediction periods. Thus, over two- and three-quarter periods, the Friend-Taubman model shows a growing

30. There were 31 one-quarter, 30 two-quarter, and 29 three-quarter results, due to the 1970.4 cutoff point.

tracking superiority over the other models tested, reaching an absolute forecasting error of less than half the next-closest Saint Louis model in the third quarter ahead. Compared with the NII baseline, the FT is the only model providing comparable results. As in the earlier comparisons, the OMB and NBER models were distinctly inferior in their tracking capabilities.

This pattern emerges with equal distinction in ordinal comparisons. Table 9 shows these results for the cumulative three-quarter forecasting error. As seen, the Friend-Taubman model dominates this comparison, having a smaller forecasting error than the NII model in 18 of 28 cases (one tie), smaller than the Saint Louis model in 20 of 28 cases, and smaller than the Evans, NBER, and OMB models in 23, 23, and 24 of 29 cases, respectively. Confirming the table 8 data, the NBER and OMB models are seen to have no noticeable tracking capabilities. In fact, the OMB model only improves upon the "no-change" forecast in three of 29 cases.

Table 9

Comparisons—Three-Quarters Cumulative Forecast Error

	FT	NII	Saint Louis	Evans	NBER	OMB
FT	...	18 (1)	20 (1)	23	23	24
NII	10 (1)	...	18	23	23	26
Saint Louis	8 (1)	11	...	17	17	18
Evans	6	6	12	...	17	18
NBER	6	6	12	12	...	15
OMB	5	3	11	10	14	...

NOTE.—Entries in the table are the number of times the model listed vertically had a lower three-quarters cumulative forecasting error than the model listed horizontally. Numbers in parentheses are ties.

Taken together, the results now presented yield the overall conclusion that the Friend-Taubman model is somewhat superior to the Saint Louis model, particularly with respect to its tracking capabilities, and more clearly superior to others tested in the comparisons we have made. In addition, the FT model was alone in generally providing better results than the "same-change" autoregressive baseline. All other models were generally inferior to this baseline. Considering absolute errors, inequality coefficients, and our ordinal rankings, the next best forecasting mechanism was found to be the Saint Louis model, which performed best in one-quarter-ahead projections and which deteriorated steadily in average accuracy and tracking thereafter. However, our ordinal results suggest the observed deterioration in average absolute error is due in significant part to a proportionately small number of disastrous forecasts, since it betters other models' forecasts in a comparatively large number of cases. The Evans model generally ranked next, followed by the NBER and OMB systems. The latter two models repeatedly demonstrated on all criteria virtually no worth in conditional short-run predictions, suggesting they may be largely disregarded as practical forecasting mechanisms.

IMPLICATIONS

We have found the Friend-Taubman crude Keynesian structure to be slightly superior to monetarist-type and distinctly superior to other systems evaluated in short-run economic predictions. This result has several implications:

1. Evans's previously noted statement that real GNP projections could be as effectively handled by a partially reduced form incorporating key exogenous variables does not stand the quarterly test provided by this comparison. The predictions from the structure of our version of the Friend-Taubman model consistently outperformed Evans's reduced form.
2. In the comparative performance of our version of Moore's leading-indicator system, we find no support in quarterly data for his contention that a model built on the NBER variables constitutes a demanding standard of comparison. Instead, we find the FT model comes closer to comprising such a standard.
3. Contrary to claims made by the authors, the unconventional treatment of seasonality and other unique features of the OMB model are of no forecasting value over the time frame of our study as considerably better results are obtained from a "no-change" naïve model. On a relative basis, every other model in our comparison was found to be more effective in short-term forecasting, which suggests that a variety of alternative forecasting vehicles are likely to be more useful than the OMB system.
4. The relatively strong one-quarter-ahead performance of the Saint Louis model compared with the Friend-Taubman model suggests that over this period predictive results are independent of the choice of model between these two. Indeed, in terms of both the number of times each model produced closer forecasts than the other and in terms of average absolute error and inequality coefficient, we find little basis for a solid distinction between the conditional record of the two models for individual one-, two-, or three-quarter-ahead predictions.
5. The short-run tracking performance of the models as indicated by the cumulative forecast errors produces a clear distinction on this dimension of the models' performance. On all criteria, the Friend-Taubman model is superior to other models compared. The Saint Louis model ranks second among the economic models tested. In addition, the Friend-Taubman system improves somewhat upon the no-change baseline employed

in the study, and is alone among models compared on this count.

Finally, in his recent book, *Theory of Economic Prediction* (Chicago: Markham Publishing Co., 1971), Lawrence Klein states: "For some time, I have maintained the position that there is no point in trying to construct models that are purely of use in prediction and deny that such models have an existence of their own apart from structural models. I would repeat earlier assertions that best predictions will be made from best structural models" (p. 99).

The present study is by no means a definitive test of Klein's proposition, since the Friend-Taubman system is not a definitive representation of complete Keynesian or neo-Keynesian theory and so cannot be classified as a true structural model vis-à-vis this theory. The same may be said of the Saint Louis model with regard to its relationship with quantity theory. Nevertheless, the structure of both these models is explicitly inspired by Keynesian and monetarist specifications, respectively. This statement cannot be made concerning the other models tested. The NBER, Evans, and OMB systems are at best reduced forms of models with unspecified structures. Thus, our finding that the Friend-Taubman and Saint Louis structural predictions dominate the others considered is at least consistent with, if not proof for, Klein's observation.

APPENDIX A

Table A1

Residuals from Prediction (ΔY Actual $-$ ΔY Predicted)
(One-Quarter-Ahead Prediction)

Year and Quarter	Actual ΔY_T	N II	M1 FT	M1 Evans	M1 NBER	M1 Saint Louis	M1 OMB	M2 FT	M2 Evans	M2 NBER	M2 Saint Louis	M2 OMB	M3 FT	M3 Evans	M3 NBER	M3 Saint Louis	M3 OMB
1963:																	
1	2.9	−2.0	1.7	−1.8	−0.5	6.8	1.6	1.2	1.0	−0.3	6.6	5.6	−1.6	−1.2	−7.5	5.0	−10.5
2	4.8	1.9	0.6	−3.3	0.9	4.0	−3.4	2.8	−1.3	−5.3	5.4	2.9	1.4	−0.6	−1.7	7.9	−.9
3	8.7	3.9	3.0	1.9	10.5	3.6	−5.3	5.3	0.7	3.0	5.3	−3.1	5.2	2.3	−5.6	6.9	−1.7
4	7.4	−1.3	3.0	2.2	7.5	0.6	2.9	3.1	5.2	3.7	2.0	7.3	3.9	3.0	1.2	2.2	6.6
1964:																	
1	15.0	7.6	11.0	16.8	12.8	−4.1	7.6	8.5	14.4	10.0	−3.4	12.7	10.3	13.8	4.0	−2.3	13.2
2	1.5	−13.5	0.9	−8.8	−7.0	−7.4	−6.9	1.2	−6.1	−8.7	−7.3	−2.0	1.0	−5.2	−11.9	−4.6	−3.4
3	7.2	5.7	−1.8	1.5	−1.5	1.4	−5.0	−3.2	5.1	−1.8	2.5	−2.0	−5.1	6.7	−9.8	1.5	−1.3
4	2.7	−4.5	4.8	−8.7	−4.1	−2.5	1.3	5.8	−7.7	1.4	−0.3	6.4	2.4	−2.6	−8.2	−0.2	11.3
1965:																	
1	13.0	10.3	3.7	6.2	8.3	7.8	11.0	8.9	9.8	8.8	8.1	15.1	7.5	12.7	8.9	12.3	15.8
2	8.2	−4.8	9.3	2.2	13.6	−2.5	5.1	7.5	5.0	13.5	−3.0	−6.7	6.8	5.4	4.6	−3.2	−.3
3	11.0	2.8	4.1	2.3	7.6	1.4	4.8	−0.1	2.2	8.7	2.4	6.0	2.2	6.8	1.1	0.4	1.0
4	13.7	2.7	6.2	4.8	10.2	4.4	−2.0	7.6	7.4	6.7	4.9	10.7	0.6	9.2	0.7	5.9	3.2
1966:																	
1	11.0	−2.7	3.2	−0.5	−3.9	−1.2	−1.4	1.9	−0.5	−3.8	−1.1	7.7	1.0	5.0	−0.7	1.0	14.0
2	3.9	−7.1	−7.2	−12.0	−9.0	−8.2	−8.2	−7.2	−6.7	−10.8	−9.4	−16.5	−9.5	−2.6	−7.3	−10.9	−7.7
3	5.5	1.6	−4.1	−14.5	5.4	−6.3	7.3	−6.1	−8.3	2.5	−7.2	−1.7	−7.6	0.9	−9.2	−8.2	−2.3
4	6.3	0.8	2.9	−4.4	5.9	0.9	6.7	1.7	−4.0	7.3	−1.3	− .1	−0.5	3.3	−2.2	−3.8	−1.4

Table A1 (Continued)

Year and Quarter	Actual ΔY_T	N II	M1 FT	M1 Evans	M1 NBER	M1 Saint Louis	M1 OMB	M2 FT	M2 Evans	M2 NBER	M2 Saint Louis	M2 OMB	M3 FT	M3 Evans	M3 NBER	M3 Saint Louis	M3 OMB
1967:																	
1	4.6	−1.7	−2.3	−6.7	−1.8	−1.9	−1.3	−1.9	−3.4	3.0	−1.5	11.4	−3.4	−6.1	−5.4	−6.9	1.3
2	3.5	−1.1	−9.5	−1.4	4.1	7.1	−2.6	−5.3	−7.3	5.8	9.6	−1.0	−5.3	−10.9	−5.0	8.6	8.3
3	8.4	4.9	−6.1	4.8	−3.6	5.5	−.6	−3.2	5.5	−6.3	7.2	3.5	−3.1	15.6	2.0	11.4	6.0
4	4.2	−4.2	−5.4	−4.1	−4.8	−6.6	−12.9	−4.0	0.2	−7.6	−5.1	−9.2	−2.3	5.6	−6.1	−0.7	−1.7
1968:																	
1	10.9	5.7	−2.4	8.0	15.2	0.2	26.4	−3.4	3.0	9.9	−0.8	25.8	−1.5	−6.4	15.9	2.3	25.9
2	10.7	−0.2	−1.7	6.4	9.3	1.8	−2.5	−4.2	−5.3	3.8	2.9	.9	−4.2	−3.4	4.4	3.1	−2.4
3	8.9	−1.8	0.4	2.7	−3.9	−3.3	.9	−1.2	3.5	1.2	0.7	6.3	−1.4	3.3	−2.7	1.6	3.8
4	6.8	−2.1	3.1	4.0	−3.7	−3.2	−18.6	−4.8	−2.2	−10.5	−2.1	−11.5	1.5	5.4	−3.2	2.5	−9.4
1969:																	
1	3.0	−3.8	1.5	−2.7	1.8	0.7	14.4	2.1	−1.7	−1.9	−2.1	8.7	1.1	−3.1	−8.1	0.4	12.9
2	4.0	1.0	2.1	1.4	−4.8	1.5	−3.8	2.8	−1.5	−7.8	2.4	−3.2	2.7	1.1	−1.7	−1.2	1.4
3	4.8	0.8	4.0	1.0	−0.1	1.1	−.3	0.0	−1.7	4.3	0.8	.7	1.6	1.1	−1.5	1.6	−1.5
4	−1.7	−6.5	4.2	−4.7	−4.3	−4.2	−13.4	−0.2	−5.4	−6.1	−7.6	−13.1	0.1	−2.6	−12.4	−6.6	−14.2
1970:																	
1	−5.4	−3.7	−4.7	−8.0	−5.2	−4.0	4.2	−6.1	−6.9	−6.0	−5.6	3.1	−8.0	−6.1	−7.3	−9.3	−2.4
2	1.1	6.5	−1.8	1.0	2.6	−0.2	10.5	2.5	0.7	12.2	3.2	20.7	−5.7	2.2	4.9	2.0	13.8
3	2.6	1.5	−3.0	−2.3	2.8	−0.1	−17.9	−0.1	−0.6	10.4	2.0	−8.9	−0.5	−1.1	2.6	3.3	−24.9
4	−7.1	−9.7	−4.5	n.a.	n.a.	−10.0	−33.7	−4.5	n.a.	n.a.	−11.1	−19.6	−4.5	n.a.	n.a.	−8.7	−27.1

NOTE.—n.a. = not applicable.

Table A2

Residuals from Prediction (ΔY Actual − ΔY Predicted)
(Two-Quarters-Ahead and Three-Quarters-Ahead Prediction)

			Two Quarters Ahead										Three Quarters Ahead				
			M1					M2					M3				
Year and Quarter	Actual ΔY_T	N II	FT	Evans	NBER	Saint Louis	OMB	FT	Evans	NBER	Saint Louis	OMB	FT	Evans	NBER	Saint Louis	OMB
1963:																	
1	2.9	−2.0	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2	4.8	1.9	1.7	6.2	−4.0	5.1	2.6	3.1	3.8	−5.2	7.6	−2.6	...	...	...	...	...
3	8.7	3.9	2.4	1.8	0.7	6.2	−3.8	4.3	2.5	−6.6	7.4	− .7	4.3	5.2	−9.8	7.7	−2.0
4	7.4	−1.3	−1.8	−0.3	4.2	3.2	7.7	−0.9	1.2	−0.4	3.1	9.0	−1.1	1.7	−1.6	2.4	12.5
1964:																	
1	15.0	7.6	6.4	12.5	11.2	−2.8	10.9	6.9	10.3	4.2	−1.8	8.9	6.8	8.7	2.6	−0.8	9.4
2	1.5	−13.5	−7.8	0.6	−4.7	−8.2	− .2	−11.5	0.1	−10.4	−5.3	.1	−11.5	−3.4	−9.4	−5.2	.5
3	7.2	5.7	−3.7	0.2	−3.7	1.2	−1.7	−4.3	0.3	−14.6	0.8	− .3	−4.6	6.4	−11.7	1.2	2.0
4	2.7	−4.5	9.1	0.5	1.2	0.2	7.2	7.3	2.4	−13.5	0.1	10.7	7.2	−4.2	−19.2	0.3	10.9
1965:																	
1	13.0	10.3	4.3	2.2	7.8	8.1	15.3	4.0	8.8	3.3	12.2	17.1	3.7	13.0	−5.2	12.5	14.3
2	8.2	−4.8	−0.2	4.8	12.7	−0.8	−1.3	4.2	8.4	8.4	−1.5	− .7	2.5	4.5	1.1	−1.5	−2.2
3	11.0	2.8	−3.2	7.5	14.7	−0.4	5.6	−4.2	8.5	6.5	−3.4	1.5	−4.8	11.7	10.4	−2.2	− .3
4	13.7	2.7	5.8	5.0	13.1	4.2	18.5	5.0	9.4	2.8	4.6	10.9	5.3	11.1	8.4	2.4	13.5
1966:																	
1	11.0	−2.7	−3.7	4.3	0.1	1.2	11.2	−4.1	6.4	0.5	3.7	14.5	−3.9	6.8	3.3	3.2	17.7
2	3.9	−7.1	−8.7	−7.9	−10.5	−9.3	−11.7	−10.2	−2.3	−8.3	−9.6	−10.2	−9.1	−0.8	−7.9	−7.3	−7.5
3	5.5	1.6	2.1	−6.5	−5.1	−11.2	−15.5	1.4	−1.3	−11.5	−13.5	−17.8	1.2	−1.3	−12.5	−11.5	−12.1
4	6.3	0.8	6.7	−5.8	8.4	−3.4	−6.6	5.2	2.1	−10.8	−6.1	−10.5	5.2	−0.1	−13.1	−10.7	−16.3

Table A2 (Continued)

Year and Quarter	Actual ΔY_r	N II	Two Quarters Ahead										Three Quarters Ahead				
			M1					M2					M3				
			FT	Evans	NBER	Saint Louis	OMB	FT	Evans	NBER	Saint Louis	OMB	FT	Evans	NBER	Saint Louis	OMB
1967:																	
1	4.6	−1.7	−1.8	−9.6	14.1	−1.7	9.2	−2.5	−0.1	−7.2	−7.5	−1.6	−2.8	−1.6	−13.3	−11.4	−7.0
2	3.5	−1.1	−3.8	1.3	2.9	8.5	2.2	−2.4	−3.7	−14.1	5.8	9.5	−3.5	0.4	−17.2	4.6	9.3
3	8.4	4.9	0.7	−2.0	−4.0	11.0	4.8	1.8	5.5	−1.0	15.1	13.1	2.0	2.6	−7.6	11.2	13.6
4	4.2	−4.2	−1.3	1.3	−9.4	−1.8	−21.5	−0.4	11.4	−6.4	3.1	−12.2	−1.0	−9.7	−8.7	3.5	−8.5
1968:																	
1	10.9	5.7	0.2	6.4	2.0	−3.1	13.1	1.0	12.1	9.5	1.2	12.2	1.1	17.6	7.7	2.0	8.3
2	10.7	−0.2	−2.4	2.8	9.0	2.2	10.4	−4.0	−6.6	17.6	1.9	3.0	−2.8	12.0	11.4	0.5	−6.4
3	8.9	1.8	0.7	−7.3	5.5	0.4	.7	0.8	1.4	−0.8	0.6	−1.6	0.8	−8.4	12.1	0.3	−2.6
4	6.8	−2.1	1.0	1.4	−8.2	−3.7	−12.2	1.3	1.2	−3.4	0.8	−7.5	1.2	−0.7	−2.9	0.0	−6.8
1969:																	
1	3.0	−3.8	−1.6	−3.9	−6.8	−2.5	4.4	−1.5	3.1	−8.0	1.0	11.9	−2.1	−3.6	−8.5	−1.6	10.9
2	4.0	1.0	1.6	−0.4	−9.5	1.6	−2.0	1.4	−2.2	−7.1	−1.2	.6	1.5	2.1	−7.3	−0.5	−2.5
3	4.8	0.8	−2.2	−1.0	−1.9	2.6	−.8	−1.2	1.9	−2.5	1.8	−3.9	−1.1	−1.3	−10.1	1.6	−1.8
4	−1.7	−6.5	−0.4	−8.1	−3.2	−7.2	−15.3	−1.6	−5.4	−11.7	−5.8	−16.5	−1.5	−4.6	−12.5	−5.2	−16.8
1970:																	
1	−5.4	−3.7	−5.4	−8.9	−8.1	−7.7	5.1	−7.9	−6.2	−12.3	−11.3	−2.2	−7.9	−9.1	−11.7	−11.2	−3.3
2	1.1	6.5	3.6	0.2	7.5	1.0	15.8	1.7	0.7	1.1	−1.4	7.7	2.1	0.5	−8.0	−4.9	4.5
3	2.6	1.5	1.2	0.6	19.3	3.0	3.3	3.3	3.1	5.8	4.6	−1.7	2.9	2.4	0.3	1.2	−5.3
4	−7.1	−9.7	−4.5	n.a.	n.a.	−10.4	−22.4	n.a.	n.a.	n.a.	−8.0	−29.4	n.a.	n.a.	n.a.	−7.0	−24.3