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

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

Volume Publisher: NBER

Volume ISBN: 0-870-14232-1

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

Publication Date: 1972

Chapter Title: The Predictive Performance of Quarterly Econometric Models of the United States

Chapter Author: Ronald L. Cooper

Chapter URL: http://www.nber.org/chapters/c2789

Chapter pages in book: (p. 813 - 947)

THE PREDICTIVE PERFORMANCE OF QUARTERLY ECONOMETRIC MODELS OF THE UNITED STATES

RONALD L. COOPER · University of California

1 THEORY AND METHODOLOGY

A. INTRODUCTION

The general plan of this research is to compare and evaluate the available quarterly econometric models of the United States with respect to their performance in prediction. In carrying out this research, we have attempted to answer two basic questions: (1) Which quarterly econometric model is most satisfactory as a predictive device? (2) Is any presently available econometric model superior, as a predictive device, to a purely mechanical forecasting scheme? A minimal standard for performance of an econometric model is that it must forecast more accurately than a purely mechanical scheme which incorporates no economic information whatever. Such a mechanical scheme provides a standard for assessing the validity of the information incorporated into an econometric model. The models included in the study are: (1) Friend-Taubman [24], (2) Fromm [26], (3) Klein [42], (4) Liu [48], (5) OBE [47], (6) Wharton-EFU [44], and (7) Goldfeld [31]. All of these models were fitted to data in the postwar period and are designed specifically to predict the expenditure components of the gross national product, although some are more comprehensive in scope.

The work in the present study concentrates exclusively on quarterly econometric models. Jon Cunnyngham, at the University of

NOTE: I cannot adequately convey my appreciation to all those who have helped me with the work necessary to complete this research. I wish especially to thank the following people, without whom, I am sure, I could not have completed it at all: Dale W. Jorgenson, James M. Brundy, Marjorie L. Flint, and Jan Seibert. Finally, I wish to express thanks to the Federal Bank of San Francisco for providing resources necessary to carry out the research project. Needless to say, views expressed in this paper do not necessarily reflect those of the Federal Reserve Bank of San Francisco.

Ohio, is engaged in a similar study, in which he is attempting to evaluate the predictive performance of the available annual models of the United States economy.

The outline of the present paper is as follows: the rest of this section is devoted to a description of both the structure and economic content of the models; a brief discussion of the methodology of evaluating forecasts; and last, the estimation techniques developed for the comparative study across models, and the assumptions that lie behind the chosen estimation techniques. In Section 2, we present empirical results, which are used to compare the econometric models in the reduced form over the period of fit. We also compare the models to mechanical forecasting schemes over the period of fit, in order to provide a minimum standard of performance for the econometric models. In Section 3, we compare the econometric models and the mechanical schemes over the period of forecast. In Section 4, we test each model for structural change. Finally, in Section 5, we summarize our findings on the predictive performance of the available quarterly econometric models and attempt to give some indication of the direction that future research should take to improve forecasting ability.

B. EVALUATION OF ECONOMETRIC MODELS

The evaluation of econometric models on the basis of predictive performance has been widely discussed. One line of argument, due essentially to Friedman [22, 23], is that goodness-of-fit statistics for a fixed period reflect not only the performance of the model but the persistence of the investigator. Only tests of alternative models against bodies of data not used in the specification of these models can provide an appropriate basis for comparison. None of the quarterly econometric models included in our study has been fitted to the same set of data as any other model. Simply fitting each of the models to a common set of data will provide important new information for evaluation of these models in empirical work. Forecasting observations not actually used in specifying or fitting the models will provide additional information for this purpose.

The only previous evaluation of quarterly econometric models of the United States are two studies: one undertaken by Fromm [27]; and the other, by Stekler [53].[1] Both studies are based on comparisons of forecasts published for each model with realized values of the economic variables. The models were estimated over different periods, each model using a different set of data. Second, the length and time periods of the forecasts varied among the models. Third, the forecast jointly dependent variables in each model are conditional upon a different set of predetermined variables. To illustrate the third point, suppose that we are attempting to compare the forecasts of *GNP* between two econometric models, the first model explaining net exports, while the second assumes that net exports are exogenous. Obviously, the second model has an advantage over the first in predicting *GNP*; hence, it would be inappropriate to compare the models in forecasting *GNP* without taking into account their differing properties. The present study attempts to overcome these deficiencies in order to provide a valid basis for comparison between the models: (1) by estimating all models from the same set of data; (2) by forecasting the economic variables for the same period; and (3), by making sure that the conditional forecasts based on different sets of predetermined variables are made known, so as not to invalidate the comparison of forecasts across models.

An earlier study similar to the one undertaken here was carried out by Carl Christ [6]; he analyzed Klein Model III and, subsequently, the first two versions of the annual Klein-Goldberger Model [7]. He reestimated Klein Model III for the separate periods 1929–1950 and 1929–1952, and concluded that structural change occurred in the model between the two periods. For versions one and two of the K-G Model, Christ compared their ex post predictions to those based on several naive models. The results favored the naive models. In another study, Suits [57] compared a later version of the K-G Model to several naive models and found that its performance in prediction was superior to that of the naive models used by Christ. Re-

[1] However, there has been a great deal of work done by Theil [58, 59] in evaluating forecasts of the Dutch economy. Theil has performed a very interesting study of econometric forecasting made by the Central Planning Bureau over the years 1953–62. The forecasting model used by the Central Planning Bureau is based on the Tinbergen Econometric Model.

cently, Jon Cunnyngham[2] compared the predictive performance of the K-G Model with the performance of those based on autoregressive schemes, and he concludes that the results favor the mechanical schemes. All of the studies just mentioned are confined to annual models.

In a recent paper by Adams and Evans [1], the record of quarterly forecasts based on the Wharton-EFU Model is discussed. According to Adams and Evans, the ex ante forecasting record of the Wharton-EFU Model has been more accurate than several judgmental forecasts made by leading business economists and forecasters over the last five years. Also, the predictive performance for business plant and equipment expenditures is found to be superior to that based on the McGraw-Hill and OBE-SEC investment surveys. The authors stress that good econometric forecasting cannot be "mechanistic," since a good deal of judgment is involved in preparing an ex ante forecast based on an econometric model. Adams and Evans are quite correct; the human element is required in making ex ante forecasts based on an econometric model. It is possible, however—as in this study—to hold the judgmental element constant by preparing ex post forecasts of an econometric model and comparing the accuracy of the ex post forecasts with those obtained from alternative econometric models and with purely mechanical models. In this sense, the test of the relative predictive performance of an econometric model is purely mechanistic—yet quite meaningful. Once the relative performance of the various models is evaluated using ex post forecasts, one can proceed to make ex ante forecasts with the model that predicts best ex post.

Another study, by Jorgenson and Nadiri [37, 38], comparing alternative models of quarterly investment behavior is also similar to the current study. Jorgenson and Nadiri first test two versions of an investment model, based on the neoclassical theory of the firm, against three alternative investment models put forth by Anderson; Eisner; and Meyer and Glauber. All of the models are reestimated, using a consistent set of data over the periods 1949-I through 1960-IV, and 1961-I through 1964-IV. Goodness-of-fit statistics are compared between alternative models of investment behavior over each period. In addition, the goodness-of-fit statistics based on the econometric in-

[2] Unfortunately, Cunnyngham's results have not yet been published.

vestment models are compared with those based on a four-period autoregressive scheme, to provide a minimum standard of performance for each of the four investment models. The results favor one version of the neoclassical investment model over all others, including the mechanical forecasting schemes.

Second, each model is tested for structural change over the two periods of estimation, and the results indicate that the same version of the neoclassical investment model exhibits the most structural stability over time.

C. METHODS OF ESTIMATION

The estimation methods used by the authors of the models included in the study differ significantly. For example, in the Klein, Liu, and Goldfeld Models, the authors experiment with both ordinary least squares (OLSQ) and two-stage least squares (2SLS) estimation. Both the OBE and Wharton-EFU Models are estimated by 2SLS, where for each model an arbitrary subset of the available instrumental variables is chosen for estimating the unconstrained reduced-form. The Friend-Taubman Model was also fitted by 2SLS. But only in the Friend-Taubman Model are the structural equations estimated by a 2SLS method which achieves limited information efficiency (LIE).

To achieve LIE for the estimated structural coefficients of a simultaneous-equations model, a consistent estimator of the reduced form must be used in the first stage of 2SLS, or limited information maximum likelihood (LIML)—if we assume that the disturbance terms in the structural equations are asymptotically normal, with zero expectation and finite variance. In this study we make the usual normality assumption about the disturbance terms so that asymptotically consistent sample statistics may be derived from the estimators in the models for hypothesis testing.

Since the methods of estimation differ, the models cannot be compared directly. To make a valid comparison across models, we must use a method of estimation which is statistically equivalent across models. The criterion chosen is the use of estimation techniques for all models which are asymptotically equivalent to 2SLS or LIML, under the assumption that the error terms are asymptotically normal with mean zero and finite variance. Obviously a more efficient estimator exists for the models, viz., three-stage least squares or full-information maximum likelihood. The full-information methods are not used in this study mainly for two reasons. First, the project of evaluating the predictive performance of quarterly econometric models would have taken a great deal longer, even if full-information methods were feasible for the models, owing to the problems connected with programing the algorithms necessary for full-information methods. Second, it is most probable that the included models each have at least one structural equation which is mis-specified, and a full-information method of estimation would carry this mis-specification to all parts of the model. This would make it difficult to locate the specification error and make it even more difficult to provide a valid comparison across models for the individual components of the gross national product. Thus, for these and other reasons, only limited-information methods are used to estimate the structural equations of the models.

It is relatively simple to obtain an estimator which achieves limited-information efficiency in an econometric model with relatively few equations and predetermined variables, but it is difficult (and sometimes impossible) to obtain an appropriate estimator for large econometric models with large numbers of predetermined variables, because: (1) very high multicollinearity exists between the instrumental variables; or (2) there are too many instrumental variables relative to the number of available data points. In several of the models listed above, the authors chose subsets of available instrumental variables by various arbitrary methods. In each of these cases the resulting estimators do not achieve limited-information efficiency. Before discussing the estimation techniques used for the quarterly econometric models, we shall first discuss several methods of choosing instrumental variables which have been proposed in the literature by various investigators.

Several econometricians have explored the problem of choosing an appropriate set of instrumental variables. Kloek and Mennes [46] determined such a set in an econometric model by choosing the set with the least amount of multicollinearity. More recently, Amemiya [2] also investigated the theoretical implications of using smaller sets of predetermined variables than the total number available in 2SLS estimation, and how to determine such sets. Amemiya chose to find the set of instrumental variables which minimize mean-squared error over the period of fit. Under the assumption of exact sampling theory, according to Amemiya, his 2SLS estimator may be more efficient than the ordinary estimator, since his uses fewer degrees of freedom. However, if being unbiased is at all important in estimating the reduced form, then Amemiya's method is less efficient than ordinary 2SLS. Further, if we generalize the discussion to include asymptotic sampling theory, then the consistency property of the reduced-form estimator becomes important in determining the efficiency of the 2SLS estimator. Clearly, if Amemiya's method results in a choice of a subset of instruments which is strictly less than the total number available, then Amemiya's method again becomes deficient.

A further difficulty with Amemiya's method (and, for that matter, with Kloek and Mennes' method) is that it is not clear that one does not lose degrees of freedom in choosing the appropriate set of instrumental variables, since the choice of variables involves the explicit or implicit testing of several hypotheses. Therefore, it is difficult to determine the usefulness of Amemiya's method.

Franklin Fisher [10] criticized Kloek and Mennes' method of choosing instrumental variables by pointing out that it supplies no way of preserving causal information once multicollinearity is eliminated. A similar criticism can be made about Amemiya's method, for there is no assurance that the set of instrumental variables which minimize mean-squared error in the fitted period will do the same in the forecast period: the low MSE in the fitted period may be due to spurious correlation.

Fisher's criteria for permitting an instrumental variable to be used in estimating a particular structural equation of an econometric model are: (1) that the instrumental variable should directly or indirectly causally influence the variable to be estimated in a way independent of the other instrumental variables; (2) that the more direct such influence is, the better. In general, an instrumental variable should be known to cause the included variables in the equation, at least indirectly. No problem in estimation arises if (*a*) the predetermined variables are adequate in number to identify the system; and (*b*) the variance-covariance matrix of the predetermined variables is nonsingular. However, according to Fisher, the latter condition is almost never satisfied in large-scale econometric models.

Fisher develops a method for choosing a set of instrumental variables to estimate a particular equation in a system which consists of two steps. First, a lexicographic ordering of the predetermined variables is obtained, based on the structure of the model. The predetermined variables which directly affect the right-hand jointly dependent variables in an equation are called "first-order" instrumental variables. The variables in the order of importance which indirectly determine the right-hand jointly dependent variables are referred to as "second-order," "third-order," and so on—causal instrumental variables. By this procedure, a complete causal-ordering, say r, of the predetermined variables is obtained. Then to each predetermined variable is assigned an r-component vector, where the predetermined variables of first causal order are given lower numbers than those of higher causal order. This allows the elements of the r-component vectors to be ordered lexicographically, the step being based completely on a priori information contained in the model.

The second step in Fisher's method, given the lexicographic ordering of the predetermined variables, is to use a posteriori information to choose a set of instruments for the zero-order right-hand jointly dependent variables in each structural equation. Suppose that there are n observations in the sample; Fisher regresses the zero-causal-order endogenous variables on the first n-2 instruments, preserving one degree of freedom. Next, the least preferred of these instruments is dropped. It can then be seen whether or not the coefficient of multiple determination ($\bar{R}^2$)—corrected for degrees of freedom—drops significantly. If $\bar{R}^2$ drops significantly, then the instrument is retained; if it does not, the instrument is eliminated from the set of possible instruments. The same procedure is applied to the next preferred instrument, and so forth, until all orders of predetermined variables are tested.

The two steps just described must be performed for every right-hand jointly dependent variable determined in the structural equations of the model, requiring a great deal of time and effort on the investigator's part.[3] Also, as is true of Amemiya's and Kloek and Mennes' methods, every regression involves the implicit testing of a hypothesis, so that many degrees of freedom may be lost in obtaining the appropriate set of instrumental variables. Thus, it is not clear that the disadvantages of obtaining the best-fitting set of instrumental variables, constrained by their lexicographic ordering, would outweigh the efficiency of choosing an arbitrary set. Also, Fisher's lexicographic ordering is not valid, for a predetermined variable which affects a given left-hand jointly dependent variable only indirectly may have as large (or larger) an effect on that dependent variable as a predetermined variable which has a direct effect.

The methods of estimation used in this study—which satisfy the requirements of limited-information efficiency—are based on three important observations. First, we observe that all of the available quarterly econometric models, as we have specified them, are decomposable into block-recursive systems, so that the matrix of coefficients (Γ) associated with the jointly dependent variables for each model can be expressed as:

$$\Gamma = \begin{array}{|l|l}
\Gamma_{11} 0 \ldots\ldots\ldots\ldots 0 & 1 \\ \hline
\Gamma_{21} \Gamma_{22} 0 \ldots\ldots\ldots 0 & 2 \\ \hline
\vdots \qquad \ddots \qquad \vdots & \vdots \\ \hline
\Gamma_{r1} \Gamma_{r2} \ldots\ldots\ldots \Gamma_{rr} & r
\end{array}$$

Γ_{11}, the lowest-order block, can be estimated independently of the other r-1 blocks; the jointly dependent variables in this block are a function solely of the predetermined variables contained in the block. Fitted values from endogenous variables in Block 1 which appear in Block 2 can be treated as predetermined to Block 2, with their fitted values substituted for the variable where they appear in Block 2 so that the coefficients in Block 2 (Γ_{21} Γ_{22}) are a function of the predetermined

[3] Recently, Fisher's method has been computerized by Bridger Mitchell at Stanford.

variables in Block 2 and the fitted values from Block 1 (which are taken as predetermined in Block 2). The procedure of feeding fitted values from lower-order blocks to higher-order blocks is continued until all blocks are estimated.

Besides the Γ matrix being upper-block triangular for each of the models included in the study, it also happens that for these models the matrix of parameters associated with the predetermined variables (B) for each model is either *nearly* upper-block triangular or nearly block diagonal. The properties of these two matrices have important implications for the elements of the reduced-form matrix, obtained from the product ($\Gamma^{-1}B$), since the product of a matrix which is block triangular and one which is near block triangular (diagonal) is itself a matrix which is *near* upper-block triangular. The property of the reduced-form matrix of being near upper-block triangular has an important implication for the choice of instrumental variables for purposes of estimating large-scale econometric systems. Since the reduced-form matrix is near upper-block triangular, the jth jointly dependent variable ($j = 1,2, \ldots, r$) can be solved for a set of predetermined variables which are *strictly less* than the total number of instrumental variables available in the whole system.

A more general implication of the property of upper-block triangularity of the reduced-form matrix is that it contains restrictions in the form of *known coefficients.* The known coefficients discussed above take the form of zero restrictions on the reduced-form matrix (Π). In this analysis we have also found known coefficients in the form of collapse restrictions, where a collapse restriction corresponds to two or more columns of Π, being linear combinations of one another, and allows these instrumental variables to be collapsed into a single predetermined variable. Known coefficients in the reduced form were used to reduce the number of instrumental variables used in the estimation of the Klein, OBE, and Goldfeld Models.

The advantage of taking into account a priori restrictions on the reduced form in a structural model is that the number of available instruments for estimation of the model can be reduced substantially without any loss of asymptotic efficiency.

The second important observation is that we can obtain an estimator which is asymptotically equivalent to ordinary two-stage least squares, based solely on repeated least squares techniques—even in the case where the available set of instrumental variables exceeds the number of data points.[4] That is, even if the number of instrumental variables available to a model is greater than the total number of observations, we can still obtain an estimator which achieves LIE. To verify this very important property concerning repeated least squares, we observe, first, that the efficiency of the two-stage least squares estimator depends only on the consistency of the unconstrained least squares estimator, say p, of the reduced form Π. Hence, the problem of finding a consistent estimator of the structural form which has the property of LIE is solved if—and only if—we can find a consistent estimator of the reduced form. Suppose now that we choose an arbitrary estimator for the unconstrained reduced form, selecting an arbitrary subset of the available instrumental variables in the model. (For example, it turned out that in the Wharton-EFU Model the number of instruments in the model exceeded the number of observations; so for this model, we selected an arbitrary subset of instruments from it to estimate the unconstrained reduced form.) The set of instrumental variables that we selected included only those purely exogenous variables (lagged dependent variables were excluded) that appeared in the structural equations of the simultaneous block.

After computing the unconstrained estimates for the jointly dependent variables determined in the structural equations, we feed these first-stage estimates into the identities of the large simultaneous blocks occurring in the model and solve them out, obtaining two-stage estimates for the jointly dependent variables explained in the identities. Finally, using the unconstrained structural estimates and the two-stage estimates from the identities, we estimate the structural equations for each simultaneous block, using 2SLS. If we are able to choose all of the available instrumental variables in the first stage, and thus obtain consistent estimates of the unconstrained reduced form, then the method just described achieves LIE and the estimation is completed. However, if the first-stage estimator inconsistently estimates the reduced form because we used fewer than the total number of available instrumental variables in the estimation, then the repeated least squares process is carried one step further.

In carrying out this further step, we first compute the constrained reduced form from the structural estimates, which are prepared by

[4] This method was suggested to me by Dale W. Jorgenson.

some two-stage method utilizing fewer than the total number of instruments available in the model. We observe that this reduced-form estimator, say $\hat{\Pi}$, is a consistent estimator of the reduced form, since plim $\hat{\Pi}$ = plim $(-\hat{B}\hat{T}^{-1})$ = (−plim $\hat{B}$)(plim $\hat{T}^{-1}$) $= -BT^{-1}$, since $(\hat{B},\hat{T})$ are consistent estimators of (B,T); but by definition $\Pi = -BT^{-1}$, so plim $\hat{\Pi} = \Pi$. Q.E.D.

Thus, we now have a consistent estimator of the reduced form. The final parts of this further step are, initially, computation of reduced-form fitted values, using the constrained estimator of the reduced form, as follows: $Y_j = X_j\hat{\Pi}$, where X_j is the matrix of predetermined variables in the jth simultaneous block. The last part consists of computing revised two-stage least estimates, using the constrained reduced-form fitted values as the first stage in the two-stage process. We now have an estimator of the structural form for each simultaneous block of a model which achieves LIE, and we call this estimator the repeated, reduced-form estimator (RR).[5]

It is important to note that the method of estimation just discussed can be used with any choice of first-stage regressions, so long as identification for each structural equation is achieved. Identification for a particular structural equation, say the jth, in a set of f linear structural equations

$$y_j = Y_j\gamma_j + X_j\beta_j + \epsilon_j \qquad (j = 1, \ldots, f) \tag{C-1}$$

is achieved if the rank of the matrix of instruments for the jth structural equation is at least as great as the rank of the included right-hand jointly dependent, and predetermined, variables in the equation. Thus, the smallest number of instruments that can be chosen in estimating

[5] Dhrymes has objected that this estimator is not, in fact, asymptotically equivalent to the ordinary two-stage least-squares estimator unless a certain orthogonality condition is satisfied. In general, the orthogonality condition is not satisfied for the RR estimator. However, Jorgenson has shown that Dhrymes' objection can be overcome by adding one additional step to the estimation process. This step consists of again applying least squares to the structural form of the model, using fitted values of the endogenous variables of the structural form based on the RR estimator. Jorgenson refers to the method of estimating by repeated least squares to achieve LIE as the "limited information repeated structural" method. For a detailed discussion of the orthogonality problem, see *Notes on the Selection of Instruments for Two Stage Least Squares and K Class Type Estimators of Large Models*, by Michael D. McCarthy. Discussion Paper No. 125, Wharton School of Finance, University of Pennsylvania, September, 1969.

the jth structural equation is equal to the sum of the included right-hand variables in the structural equation.

As long as the above rank condition is satisfied, the RR method can be used in combination with any of the methods for choosing instrumental variables that we have discussed above: (1) Fisher's method, (2) some form of principal components, or (3) the rule of choosing an arbitrary subset of instrumental variables which includes no lagged dependent variables. None of these methods for choosing the initial set of instrumental variables is necessarily preferable to any of the other methods. Consequently, there is a certain amount of arbitrariness in obtaining the RR estimator for finite samples. However, in spite of this difficulty, our method does, at least, take care of one of the degrees-of-freedom problems that arise in applying repeated least squares estimation to large econometric systems—viz., the number of degrees of freedom required for the initial estimate of the reduced form. We can, by using the method of RR estimation, make the loss in degrees of freedom as small as possible, subject to the constraint that each structural equation be identified.

The final significant observation that we make—which permits the use of repeated least squares techniques, and is quite independent of whether or not the reduced form can be estimated on the first pass—is that all quarterly econometric models can be made linear in the variables of the structural equations. This is accomplished by transferring all nonlinear (and linear nonexclusion-type) restrictions appearing in the structural equations to the identities, leaving behind only exclusion-type restrictions. Specifying the structural equations in this way, reduces the order condition for identification to the simple case—the number of variables excluded from any structural equation must be at least as great as the number of included jointly dependent variables, less one. The rank condition for identification of each structural equation requires that the Jacobian corresponding to the excluded variables be nonsingular for every structural equation. An example of the linearization process is given by the following model

$$C_t = a + bY_t + \epsilon_{1t} \tag{C-2}$$

$$I_t = c + d\left(\frac{K}{Y}\right)_t + \epsilon_{2t} \tag{C-3}$$

(C-4) $$Y_t = c + I_t + G_t$$

(C-5) $$K_t = (1 - \delta)K_{t-1} + I_t$$

Since the model is nonlinear [equation (C-3)] in two endogenous variables—both of which appear separately elsewhere in the model—we are unable, in its present form, to apply repeated least squares. Equation (C-3) can be made linear in the variables by adding an additional identity, as follows

(C-6) $$I_t = c + dZ_t + \epsilon_{2t}$$

(C-7) $$Z_t = \left(\frac{K}{Y}\right)_t$$

With the structural equations linearized in the variables, and their rank and order conditions satisfied in each quarterly econometric model, their parameters can be estimated by simple repeated least squares techniques, under the following *maintained hypotheses:*

(C-8) X is a stochastic matrix, distributed independently of ϵ

(C-9) X has full rank with probability one

(C-10) $$E(n^{-1}X'X) = \Sigma_{X'X}$$

(C-11) $$\text{plim } (n^{-1}X'X) = \Sigma_{X'X}$$

(C-12) $E(\epsilon) = 0$, where ϵ is E stacked

(C-13) $V(\epsilon) = \Sigma \otimes I$, Σ positive definite

(C-14) $$\text{plim } (n^{-1}X'E) = 0$$

(C-15) $$\text{plim } (n^{-1}E'E) = \Sigma$$

If we add the additional assumption

(C-16) $$n^{-1/2}[I \otimes X']\epsilon \underset{n \to \infty}{\sim} N(O, \Sigma \otimes \Sigma_{X'X})$$

then hypothesis testing can be carried out. We assume (C-8) through (C-16) to be valid for the seven econometric models described in this study.

D. EVALUATING PREDICTIVE PERFORMANCE

There are many ways to evaluate the predictive performance of an econometric model. The view taken here is that the predictive performance of an econometric model should be measured relative to its reduced form, rather than relative to its structural form. This has been made possible only recently by the development of Holt's SIMULATE program [34]. Most of the econometric models reported in the literature contain little or no information on the goodness-of-fit of the reduced form. Goodness-of-fit of the structural form cannot be used as a test of the over-all predictive performance of an econometric model. In Section 2, goodness-of-fit in the reduced form of the models is compared in the mean-squared error (MSE) over the period of fit, where the MSE is defined as

$$\text{(C-17)} \quad MSE = 1/m \sum_{i=1}^{m} (y_{ij}^{p} - y_{ij})^2 \qquad (i = l, \ldots, m; j = 1, \ldots, p)$$

where y_{ij}^{p} and y_{ij} are the predicted and actual values, respectively, of the jth dependent variable in the set of p jointly dependent variables.

Theil [58] presented several variants of this statistic for evaluating forecasts of various types of models, but they are shown by Jorgenson and Nadiri [36] to be invalid when applied to the linear statistical regression model.

More powerful evidence of the performance of an econometric model can be obtained by examining reduced-form MSE's over a period different from the period used to fit the model. This type of evaluation can be made rigorous by devising a statistical test for structural change. In Section 3, the reduced form MSE's are reported; and in Section 4, the structural stability of each of the alternative models is given. The statistic used to test the models for structural change is derived as shown below.

To test the independent equations for structural change in each quarterly econometric model, we must consider the asymptotic distribution of the statistic

$$\text{(D-1)} \qquad \frac{(m^{1/2}\hat{v}_{0j}'\hat{v}_{0j}m^{1/2})/m}{(\hat{v}_j'\hat{v}_j)/n} \qquad (j = 1, \ldots, r)$$

where the numerator is the sum of squared residuals in the jth reduced-form equation over the forecast period, and the denominator is the reduced-form MSE over the fitted period. The numerator of (D-1) is equal to plim $(v_{oj}'v_{oj})$, since plim $(\hat{v}_{oj}'\hat{v}_{oj}) =$ plim $(X_o\Pi_j + v_{oj} - X_o\hat{\Pi}_j)'$ $(X_o\Pi_j + v_{oj} - X_o\hat{\Pi}_j) =$ plim $(v_{oj}'v_{oj})$. Plim $(v_{oj}'v_{oj})$ converges in probability to a random variable distributed χ^2. This follows from the assumption that $m^{1/2}v_{oj}$ is asymptotically normal with mean zero and variance ω_j. This implies that $v_{oj}'v_{oj}/\omega_j$ has an approximate chi-square distribution. The denominator of (D-1) converges in probability to a constant (ω_j). Thus it is reasonable to assume that (D-1) has an approximate χ^2 distribution with 20 degrees of freedom (the number of quarters in the forecast). Interpreting (D-1) as a likelihood ratio statistic, we reject the null hypothesis of no structural change for large values of the statistic. This implies that a one-tail test is appropriate; for the critical region, we choose 5 per cent.

A multivariate test for structural change, more powerful than the univariate test, could be devised, taking the following form

$$\hat{v}_o'[I \otimes \hat{\Omega}]\hat{v}_o \tag{D-2}$$

where $\hat{\Omega}$ is the reduced-form variance-covariance matrix of errors in the fitted period, and $\hat{v}_o$ is the vector of reduced-form errors in the forecast period. This statistic (D-2) has an approximate χ^2 distribution with 20 degrees of freedom. Due to the enormous computational work involved, the multivariate test is not used in the current study.

In addition to the above tests, it is useful to provide a minimum standard of performance for the models by comparing their goodness-of-fit and predictions to those based on auto-regressive time series. Comparing an econometric to a naive method of forecasting supplies a technique for assessing the economic information contained in an econometric model. The defining characteristic of a "naive" forecasting method is that it depends exclusively on purely statistical properties of economic time series, such as trend, past levels, or past changes. A naive method does not incorporate any economic information, such as relationships between consumption and income, or prices and the quantity of money. Forecasts made by naive methods are then compared with forecasts made by other methods. Forecasting methods that cannot do better than a purely mechanical one should be discarded.

Two popular mechanical forecasting methods, previously used to evaluate forecasting performance, are the "no change" and "same change" models. In the no-change method, the naive forecast is simply that each economic time series will equal its own present level. In the same-change method, the naive forecast is that each economic forecast will continue to change in the same direction and by the same amount.

In early work on econometric models, the no-change and same-change models did provide, at least for a time, quite stringent standards for their forecasting performance. However, these models have become progressively less important for two reasons. First, econometric work has improved in quality, so that over-all standards of performance have gone up. Second, recent econometric work has emphasized empirical relationships involving lagged values of dependent variables as "explanatory" variables. An empirical relationship of this type incorporates past levels and past changes of economic variables as explanatory variables. But these are precisely the variables used for predictive purposes in the no-change and same-change methods; therefore, econometric forecasting methods can be expected to do better than these simple mechanical devices.

In this study, a different type of mechanical forecasting method is employed as an alternative to econometric forecasting methods: the auto-regressive scheme. An auto-regressive scheme is simply a regression of a variable on its own past values. The basic idea is to express an economic time series as a weighted sum of its own past values. The weights in the sum are determined so as to achieve the greatest possible predictive power. The auto-regressive scheme is a naive method of forecasting, since it incorporates no economic information. Forecasting on the basis of auto-regressive schemes is completely mechanical and depends exclusively on the purely statistical properties of an economic time series.

No-change and same-change forecasting models provide a better yardstick for measuring forecasting performance than simply postulating that an economic time series is constant. Similarly, an auto-regressive scheme provides a more stringent test of a method of forecasting than either the no-change or the same-change method. The no-change model is simply an auto-regressive scheme with weights equal to one for the most recent past value of the series, and zero for all other past values. The same-change model is an auto-regressive scheme with weights equal to two for the most recent past value, minus one for the next most recent value, and zero for all other past values. Thus, if either of these methods of forecasting is superior to other auto-regressive schemes, weights can be chosen from among the other possibilities. In practice, the weights that maximize predictive power are not usually those associated with either the no-change or the same-change methods.

Auto-regressive schemes include a number of other purely mechanical methods of forecasting as special cases, in addition to the no-change and the same-change methods. The same-change model, itself, may be interpreted as a linear trend; that is, the sum of a constant plus a term that is proportional to time. If a linear trend is a satisfactory representation of an economic time series, the weights of the auto-regressive scheme can be chosen to make forecasts by trend extrapolation. As a practical matter, such a representation is not usually particularly satisfactory. Similarly, an auto-regressive scheme can be chosen to make forecasts by extrapolation of a geometric trend; that is, by assuming a constant rate of growth. In the auto-regressive scheme, a weight equal to one plus the rate of growth is assigned to the immediate past value, and weights equal to zero are assigned to all other past values.

We conclude that an auto-regressive scheme, though purely mechanical and depending exclusively on purely statistical properties of an economic time series, includes a wide variety of naive forecasting schemes. If the auto-regressive scheme is designed to maximize predictive power, the resulting yardstick for evaluating forecasting performance is at least as good as the no-change and same-change methods of forecasting, as good as linear or geometric trends—or any combination of same. Accordingly, matching the forecasting performance of an econometric model against the best possible auto-regressive scheme provides a stringent test of the econometric model. Similarly, matching the forecasting performance of a judgmental method against an auto-regressive scheme provides a stringent test of the judgmental method. For these reasons, we present the forecasting performance of both an auto-regressive scheme and each econometric model. By comparing the results of forecasting by these two methods, the value of the economic information contained in the econometric models can be assessed.[6]

E. THE ALTERNATIVE MODELS

Scope, economic theory, and block-recursive structure in each of the seven quarterly econometric models will be discussed at this point.

E-1. Scope

Coverage of the quarterly econometric models is presented in Table 1.1. Note that the scope of the models differs greatly. For example, the Friend-Taubman Model determines no financial variables, while the Goldfeld Model explains nineteen. As another example, the income side of the national income and product accounts is totally lacking, or grossly incomplete, in the Friend-Taubman, Liu, and Goldfeld Models, but is explained completely in the Klein, OBE, and Wharton-EFU Models. Fromm [27] and Nerlove [50] give a more detailed description of the coverage of quarterly econometric models in tabular form.

Not only do the models differ in coverage, but sometimes what is endogenous in one model is exogenous in another. This differing property of the models poses a serious problem in comparing predictive performance; one model may attempt to explain what another assumes to be given. This problem is easily handled if what is endogenous in one model and exogenous in another appears in the former in recursive blocks. For example, suppose that net exports are solved in a recursive block in the Klein Model, and appear as an exogenous variable in the Friend-Taubman Model. Further, suppose that we wish to compare these two models in predicting *GNP*. In this case, the procedure is to

[6] The order of the auto-regressive scheme used in this study is determined for each endogenous variable by the scheme which has the smallest residual variance. However, we set an arbitrary cutoff point of eight quarters.

TABLE 1.1

Variables Determined in Quarterly Econometric Models of the United States

Variable	Friend-Taubman	Fromm	Liu	Klein	OBE	Goldfeld	Wharton-EFU
Product Side							
Consumption	1	2	3	3	4	2	3
Investment	1	1	2	1	1	1	2
Housing	1	1	1	1	1	x	1
Inventories	1	1	1	1	1	1	2
Foreign trade	x	1	x	3	2	3	x
Government	x	x	x	x	x	x	x
Income Side							
GNP	x	1	1	1	1	x	1
NNP	x	2	1	x	1	x	4
National income	x	2	x	2	3	x	2
Personal income	x	7	3	4	4	x	5
Labor Force							
Employment	x	x	x	1	1	x	5
Unemployment	x	1	x	1	1	x	1
Financial Data							
Bank reserves	x	x	x	x	x	6	x
Loans	x	x	x	x	x	3	x
Demand deposits	x	x	2	1	1	3	x
Time deposits	x	x	1	x	x	3	x
Interest rates	x	x	3	2	3	4	2
Total	4	19	18	21	24	23	31

drop the net exports equation from the Klein Model by changing net exports from an endogenous to an exogenous variable. This then provides a valid comparison of the two models in predicting *GNP*, since the basic structure of the Klein Model remains unchanged after dropping the exports equation. However, suppose that we change the example by assuming that net exports in the Klein Model are solved in a simultaneous block. We cannot drop the net exports equation now, because it is solved simultaneously with the other equations in the block; dropping this equation would change the structure of the model. Hence, for this case, the comparison of the two models in predicting *GNP* has one-way validity: only if the Klein Model predicts *GNP* best. If the Friend-Taubman Model gave the best *GNP* prediction, we would be unable to tell whether the outcome was due to net exports being exogenous or to other factors in the two models.

E-2. Economic Theory

Friend-Taubman Model. This model is similar in structure to the original underemployment model presented by Keynes. The model contains no income side and no monetary sector, so Friend-Taubman assume that interest rates, the money supply, and the employment of labor and capital always adjust immediately to support any given level of aggregate demand. Monetary policy has no place at all. Every aggregate expenditure variable in the Friend-Taubman Model is determined in real terms, and no explanation of the price level is given. The Friend-Taubman Model differs from the original Keynesian underemployment model in that it includes several anticipations variables. Both the business investment and residential construction equations contain the OBE-SEC second investment anticipations, and the inventory equation contains a sales anticipations variable.

The Friend-Taubman Model, established on a semiannual basis, has been changed to a quarterly basis by making the sum of two quarterly observations equal to the semiannual observations. In a few situations, aribitrary approximations were necessary. Finally, the model was originally formulated in terms of flows rather than levels. The authors suspected the presence of serial correlation, and first-differencing the data is a method commonly employed to eliminate this problem. However, all of the other models of the study are written in terms of levels. For this reason, the variables of the Friend-Taubman Model were converted to levels to facilitate the comparison across models. This change required adding trend terms to the behavioral equations.

Fromm Model. The Fromm Model is similar to the original Keynesian Model, although it differs in one significant respect. Whereas the original Keynesian Model says nothing directly about capacity, the Fromm Model explains the movement of three paramount variables in his system—the *GNP* price deflator, residential structures, and unemployment—as the difference between actual and capacity output. Fromm follows the policy of the Council of Economic Advisers by assuming that capacity output is independent of actual output, so that capacity output can be taken as a predetermined variable in the model. The Fromm Model has no monetary sector, so monetary policy is assumed to be a powerless tool in the world represented by Fromm's system. Personal and corporate income, indirect business taxes, and several transfer items are endogenous to the model, and a production function relating unemployed workers to the difference between potential and actual output is included. Some expenditure variables are determined in real dollars and others are explained in current dollars, with no obvious pattern or rationale. The only price deflator explained in the model is the *GNP* price deflator. Finally, the complete income side of the *GNP* accounts is determined in the model, including corporate profits and dividends.

Liu Model. The Liu Model is basically neoclassical in structure. It takes production of output as given and assumes that the labor market is always in equilibrium at full employment. However, the Liu Model does differ from the original neoclassical model with respect to financial variables, since it includes the real balance, or wealth effect on cash balances and near-money assets, which, in turn, has an effect on expenditure variables in the system.

Expenditures on consumer durables, residential and nonresidential structures, and producers' durable equipment are explained by their respective capital stocks, as well as by distributed lags in investment expenditures. In the monetary sector of the model, the long-term interest rate—itself an endogenous variable, and standing for the cost of capital services—has great importance in determining residential and nonresidential construction, as well as business inventory investment. Inventory investment largely determines changes in the price level, and is, itself, explained by the demand for cash balances and the long-term interest rate. The long-term interest rate is determined jointly—in accord with the expectations theory governing interest rates and the price-level growth rate. Government expenditures, taxes, and the money supply are exogenous to the model, and all variables on the income side of the *GNP* accounts are explained, with the exception of personal income. This last is due to the fact that tax items are exogenous to the model.

Klein Model. The Klein Model was constructed in the tradition of Keynes and Tinbergen, and represents a quarterly version of the earlier annual Klein-Goldberger Model [45]. The basic characteristics of the Klein Model are as follows: the consumption and investment equations are similar to those found in the earlier Klein-Goldberger system, with the notable exception that the quarterly consumption and investment equations contain anticipations data in the form of the Michigan Survey Index of Consumer Attitudes and OBE-SEC first investment anticipations. A production function, using labor and capital services as inputs, determines real private *GNP*; current *GNP* is determined in the usual way, as the sum of the current expenditure components of aggregate demand. The only exogenous expenditure components are government purchases of goods and services, and the price deflator for imports. The *GNP* deflator is explained as the ratio of real *GNP*, determined from the production function, and current *GNP*, determined on the demand side. In addition, implicit price deflators for consumer durables, nondurables, and services; plant and equipment expenditures; residential structures; and exports are determined in the model.

The model has a small, primitive monetary sector, in which the ratio of liquid assets to *GNP* is explained by the money supply and the long-term interest rate; the money supply is exogenous to the system. Taxes and transfer payments are explained in the model, so that the complete income side of the *GNP* accounts is determined.

OBE Model. The OBE Model is a variant of the Klein Model. Thus, both models are similarly structured, with only minor differences throughout. For example, a minor difference between the two is that the OBE Model has a production function which is constrained to be Cobb-Douglas, with coefficients summing to unity, while the Klein Model has a general linear form for its production function. Another change in the OBE from the Klein Model is that the former has disaggregated consumer durables expenditures into automobile and nonautomobile durables, while the Klein Model explains only total consumer durables. A further difference is that the OBE Model has added an equation to the financial sector explaining the yield on twenty-five-year-insured homes. Most of the remaining changes are slight modifications of the individual variables appearing in the structural equations.

Wharton-EFU Model. The 1965 version of the Wharton-EFU

Model [44], like the earlier Klein-type models, has a small monetary sector. The short-term interest rate is determined by the long-term interest rate—both current and lagged one period—and the long-term interest rate is determined by the ratio of excess to required reserves, an exogenous variable assumed to be under the control of the monetary authority. The Wharton-EFU Model is far more disaggregated in the expenditure, capital stock, and inventory variables than the Klein-NBER Model, but basically determines the expanded set of endogenous variables in a way similar to the original model. By way of disaggregation, both plant and equipment expenditures and inventory investment are separated into manufacturing and nonmanufacturing industries, while real gross product originating, determined from production functions, is explained for the following variables: manufacturing, nonmanufacturing, and nonfarm residential structures. Another difference between the Wharton-EFU and Klein Models is that the services and nondurables components of consumption in the former model have been combined and expressed as relative, rather than absolute, terms, due to the observed long-run constancy of the average propensity to consume services and nondurables. A final difference between the two models is that the Wharton-EFU has the consumption and investment equations specified in two different ways. The first specification, designed for short-term forecasting, includes anticipatory variables in the consumption and investment equations. The limit on the short-term forecast is, of course, set by the number of quarters in the future which are covered by the anticipatory variables. Alternative consumption and investment equations, which exclude the anticipatory variables, are used in the model for making forecasts beyond the coverage of the anticipatory projections. In the current study, we have chosen the first, rather than the second, specification of the consumption and investment equations, since we are solely concerned with making one-period forecasts, which are included in the coverage of the anticipatory variables.

Goldfeld Model. The Goldfeld Model provides an interesting comparison with the others, because though it is primarily a financial model of the commercial banking sector, it includes a few expenditure equations explaining the broad aggregates: consumption, gross fixed investment, inventory investment, and gross national product. All expenditure variables are measured in current dollars and no explanation of the price level is given. The supply of money and short- and long-term interest rates are endogenous in the Goldfeld Model, allowing a more complete tracing out of the effects of changes in monetary policy than the other models offer. However, the Goldfeld Model does not permit a complete general equilibrium solution to changes in monetary policy, since it has no income side. Thus, feedbacks are not possible between changes in expenditures and changes in income. In the Goldfeld Model, in addition to the endogenous money supply, several portfolio assets of the banking sector are determined, such as member bank short- and long-term holdings of federal government and municipal securities, and commercial loans. Most of these variables are determined for both country and city member banks.

E-3. Block-Recursive Structure

In Table 1.2 we present a breakdown of the simultaneous and recursive blocks within the quarterly econometric models. All of the models are block recursive; i.e., each one contains at least a single simultaneous block. The OBE and Goldfeld Models both have two large and two small simultaneous blocks, whereas the rest of the models have only one such simultaneous block. The Fromm Model comes closest to being fully recursive, since it has only one simultaneous block containing five equations.

The Friend-Taubman, Liu, and Wharton-EFU Models, with well over half of their equations in simultaneous blocks, are closest to being completely simultaneous. The Klein and OBE Models, with about one-half of their equations in simultaneous blocks, are intermediate between the recursive and simultaneous systems.

Table 1.3 contains a detailed breakdown of the contents of the simultaneous blocks across quarterly econometric models. A comparison of the simultaneous blocks reveals, first, that excepting the Fromm Model, the expenditure variables on the demand side of the *GNP* accounts are solved in the simultaneous blocks of the models. Second, for models containing production functions, the gross output variables are solved in simultaneous blocks. The same is also true of the em-

TABLE 1.2

Breakdown of Simultaneous and Recursive Blocks Within Quarterly Models

Econometric Model	Number of Simultaneous Blocks	Number of Variables in Simultaneous Blocks	Number of Recursive Blocks
Friend-Taubman	1	6	5
Fromm	1	5	45
Liu	1	26	35
Klein	1	52	25
OBE	2	1–18	46
		2–40	
Wharton-EFU	1	92	23
Goldfeld	2	1–63	19
		2–4	

ployment variables for all of the econometric models, other than the OBE. In the OBE Model, gross output and employment are determined in the small simultaneous block, ahead of the large simultaneous block which contains the expenditure variables. Next, with the exception of the Fromm Model, the income side of the *GNP* accounts—along with the implicit price deflators—is solved in the simultaneous blocks of the quarterly econometric models.

With respect to financial variables, we notice first that in the Liu Model the short-term rate of interest is determined ahead of the large simultaneous block by a bank reserves variable and the Federal Reserve discount rate; all other interest rates and financial wealth variables are solved in the simultaneous block. In the Klein, OBE, and Wharton-EFU Models both the short- and long-term rates of interest are solved recursively, ahead of the large simultaneous blocks.

In the Goldfeld Model, the short-term rate of interest is solved in the simultaneous block of the model, along with the real variables. Bank reserves and the money supply are considered endogenous. Clearly, the Goldfeld Model is, financially, the most sophisticated of all the quarterly econometric models.

TABLE 1.3

Breakdown of Contents of Simultaneous Blocks Across Quarterly Models

Econometric Models	Contents of Simultaneous Blocks	
Friend-Taubman Block	*Expenditure Side*	*Income Side*
	Real consumer expenditures Real residential structures Real plant and equipment expenditures Real gross national product	
Fromm Block	*Expenditure Side*	*Income Side*
		Federal tax payments State and local tax payments Personal income Personal disposable income
Liu Block	*Expenditure Side*	*Income Side*
	Real business structures Real producer's durable equipment Real consumer durables Real consumer nondurables Real consumer services Real nonfarm inventory investment Real final sales of goods Real gross national product	Real disposable income Real corporate profits Real dividends Real depreciation
	Financial	*Implicit Price Deflators*
	Moody's corporate AAA bond rate Yield on time deposits and savings shares Real business liquid assets Real consumer holdings of currency Real consumer holdings of time deposits and savings shares	Rate of growth in *GNP* deflator

(continued)

TABLE 1.3 (*continued*)

Econometric Models	Contents of Simultaneous Blocks	
Klein Block	*Expenditure Side*	*Employment*
	Real consumption (durables, non-durables, services)	Unemployed workers
	Real gross national product	Wage and salary employment
	Current gross national product	Total civilian labor force
	Real private *GNP*	Hours worked per week
	Real nonfarm inventory investment	Private civilian employment
	Real nonfarm residential structures	
	Current private *GNP*	
	Income Side	*Implicit Price Deflators*
	Current retained earnings	*GNP*
	Personal income	Exports
	Nonlabor personal income	Nonfarm residential structures
	Real retained earnings	Plant and equipment
	Real private wages and salaries	Consumer services
	Corporate profits	Consumer nondurables
	Current wages and salaries	Consumer durables
	Personal disposable income	
	Taxes: indirect business, corporate, and personal	*Financial*
	Average private annual wage rate	None
Goldfeld Block (1)	*Expenditure Side*	*Income Side*
	Current gross fixed investment	Personal disposable income
	Current inventory investment	
	Current consumer nondurables and services	
	Current gross national product	
	Financial	
	Excess reserves, city	
	Excess reserves, country	
	Borrowings, city	
	Borrowings, country	
	Demand deposits, city	
	Demand deposits, country	
	Time deposits, city	
	Time deposits, country	

TABLE 1.3 (*continued*)

Econometric Models	Contents of Simultaneous Blocks	
Goldfield Block (1) (*continued*)	*Financial* (*continued*)	
	Commercial loans, member	
	Potential demand deposits	
	Change in class average reserve requirement, city	
	Money supply	
	High-powered money	
	Commercial loan rate	
	Long-term government rate	
	Treasury bill rate	
	Commercial loans, city	
	Commercial loans, country	
Goldfeld Block (2)	*Financial*	
	Intermediate	
	Government rate	
OBE Block (1)	*Production Function*	*Employment*
	Real private *GNP* excluding housing	Ratio of civilian labor force to population (ages 18–64)
	Real private *GNP* at full capacity	Civilian wage and salary employment
		Private man-hours
		Civilian labor force
OBE Block (2)	*Expenditure Side*	*Income Side*
	Real consumer automobile expenditures	Corporate profits
	Real consumer nonautomobile expenditures	Rate of growth private wage rate
	Real consumer services excluding housing	Dividends
	Real inventory investment	Personal income
	Real consumer expenditures excluding housing	Personal disposable income, current wages and salaries
		Real personal disposable income
		Current private wages and salaries

(*continued*)

TABLE 1.3 (*continued*)

Econometric Models	Contents of Simultaneous Blocks	
OBE Block (2) (*continued*)	*Expenditure Side (continued)*	*Implicit Price Deflators*
	Current private *GNP* excluding housing	Private *GNP* excluding housing
	Real total consumer expenditures	Consumer nonautomobile durables
	Current consumer expenditures excluding housing	Consumer nondurables deflator
		Consumer services excluding housing
		Consumer expenditures deflator
	Financial	
	None	
Wharton-EFU Block	*Expenditure Side*	*Income Side*
	Ratio of real consumer nondurables and services to real disposable income	Depreciation, manufacturing, nonmanufacturing, and nonfarm residential structures
	Real consumer durables excluding automobiles	Taxes: indirect business, corporate, and personal
	Real consumer automobiles and parts	Government and business transfer payments
	Real plant and equipment expenditures, manufacturing and nonmanufacturing	Business income of unincorporated proprietors
	Nonfarm residential structures	Dividends
	Real inventory investment, manufacturing and nonmanufacturing	Inventory valuation adjustment
		National income
		Personal income
	Production Function	*Implicit Price Deflators*
	Real full capacity output in manufacturing	Consumer nonautomobile durables
	Real gross output originating in residential structures	Consumer automobiles
	Real gross output originating in manufacturing	Plant and equipment
	Wharton School capacity index	Nonfarm residential structures
	Real gross output originating in nonmanufacturing	Exports
		GNP
		Consumption
		Manufacturing deflator

TABLE 1.3 (*concluded*)

Econometric Models	Contents of Simultaneous Blocks	
Wharton-EFU Block (*continued*)	*Employment*	*Financial*
	Man-hours in manufacturing	None
	Average hours worked in manufacturing	
	Man-hours in nonmanufacturing	
	Average hours worked in nonmanufacturing	
	Employment in manufacturing	
	Employment in nonmanufacturing	
	Expenditure Side	*Income Side*
	Real imports of crude food and materials	Corporate profits
	Real imports of semifinished goods and services	Real disposable income
	Real exports	
	Current *GNP*	
	Real *GNP*	
	Real consumer expenditures	
	Current consumer expenditures	
	Real plant and equipment expenditures	
	Real nonfarm inventory investment	
	Real consumer durables	
	Real nondurables and services	

2 A COMPARISON OF QUARTERLY ECONOMETRIC MODELS OVER THE PERIOD OF FIT: 1949-I THROUGH 1960-IV

THE acid test of any econometric model is how well it predicts in the reduced form. Reduced-form predictions over the period of estimation are made for the jointly dependent variables of the quarterly econometric models. Using the reduced-form predictions and the realized values of the jointly dependent variables, mean-squared errors are computed for this period.[7]

It must be pointed out that some comparisons given below are valid only one way, since in several of the models, exogenous price deflators are added to differentiate the current values of economic variables from their real values. In the rankings below, asterisks appear after those econometric models in which exogenous deflators have been added. Thus, for example, a comparison between the Fromm and the OBE Models in predicting real consumer expenditures is valid only one way, because the OBE system explains the consumption deflator, while the Fromm Model does not.

A. CONSUMPTION EXPENDITURES

Mean-squared errors for consumption expenditures and its components over the period of estimation appear in Table 2.1.[8] Considering

[7] The sample period for the econometric models includes the Korean War period. Some of the models were originally estimated for a period which includes the Korean War years, while the rest were not. Those models that were originally estimated over the Korean War period included dummy variables in several of their structural equations. For the models that were originally estimated excluding the Korean War years, we decided not to add dummy variables to any of their structural equations. We do not know whether or not these model-builders would have added dummy variables to some of their structural equations if they had included the Korean War in their sample periods. However, dummy variables typically add only a very small amount to the explained variance of the left-hand variable. We do not feel that adding dummy variables to the equations of those models that were not estimated over the Korean War period would have significantly improved their performance. In fact, there is evidence from some research which I have done recently that the addition of dummy variables produces worse ex post forecasts than would be obtained by excluding them.

[8] Due to the magnitude of this study, every effort was made to eliminate errors, both in the computation and in the tabulation of the statistics presented in the tables below.

TABLE 2.1

Mean-Squared Errors Over Fitted Period for Consumption Expenditures

Variables	Econometric Models: Friend-Taubman	Fromm	Liu	Klein	OBE	Wharton-EFU	Goldfeld	Autoregressive Schemes
Real total consumer expenditures	13.5	10.1*	8.899	17.02	11.11	25.44	39.14*	7.11
Current total consumer expenditures	11.6*	7.86	7.714*	50.98	33.93	20.92	14.87	5.791
Real consumer durables	+	+	3.526	8.65	5.308	9.313	10.99*	3.711
Real consumer automobile expenditures	+	+	+	+	2.787	6.410	+	1.841
Real nonautomobile durables	+	+	+	+	0.7594	1.080	+	0.657
Current consumer durables	+	+	3.040*	9.03	4.580	7.07	9.67	2.992
Real consumer nondurables	+	+	2.094	2.50	2.403	+	+	1.369
Current consumer nondurables	+	+	1.860*	10.80	10.59	+	+	1.436
Real consumer services	+	+	0.275	0.336	0.3421	+	+	0.2176
Current consumer services	+	0.571	0.215*	3.23	2.660	+	+	0.1532
Real consumer nondurables and services	+	+	2.643	2.503	2.403	12.58	10.34*	2.195
Current consumer nondurables and services	+	+	2.322*	3.230	20.41	9.042	8.21	1.847

Note: In this and following tables, asterisks indicate addition of exogenous deflators. Pluses signify that the variable is not explained in this model.

in detail the results given in Table 2.1, we observe, first, that both real and current total consumer expenditures are best predicted by the naive models. The best performing econometric model for both variables over the fitted period is the Liu Model: worst performing are the Klein, for current consumer expenditures; and the Goldfeld, for real consumer expenditures.

In predicting real and current consumer durables expenditures, both the Liu and naive models perform well (see Table 2.1). For real consumer durables, the Liu Model performs slightly better than the naive model; but for current consumer durables, the naive model predicts slightly better than the Liu. Both are significantly superior to the other quarterly econometric models in predicting durables. The poorest performance for real and current consumer durables is given by the Klein, Wharton-EFU, and Goldfeld Models.

Real consumer automobile and nonautomobile expenditures are explained in the Wharton-EFU and OBE Models. The naive model is superior in predictive performance to both econometric models, while the OBE Model is superior to the Wharton-EFU Model for both automobile and nonautomobile expenditures.

For real and current consumer nondurables and services, the naive model registers a predictive performance superior to all of the quarterly econometric models over the period of fit. In predicting real nondurables and services separately, the Liu Model outperforms all other econometric models. Both the Goldfeld and Wharton-EFU Models explain nondurables and services together. For this variable we find, in Table 2.1, that both the Klein and OBE Models slightly outperform the Liu Model. The Wharton-EFU is the least good of all the econometric models discussed here in predicting the nondurables-services variable. We conclude that although in all cases the mechanical schemes outperform the econometric models, no ranking among the latter is possible in predicting nondurables and services—except for the Wharton-EFU, which ranks last among all the quarterly econometric models being evaluated. The rank order of the predictive performance of the components of consumption expenditures for the models is summarized in the table which follows.

Variables	Rank Order of Predictive Performance Over Fitted Period							
	Naive	Friend-Taubman	Fromm	Liu	Klein	OBE	Wharton-EFU	Goldfeld
Consumer expenditures								
Real	1	5	3*	2	6	4	7	8*
Current	1	4*	3	2*	8	7	6	5
Consumer durables								
Real	2			1	4	3	5	6*
Current	1			2*	5	3	4	6
Consumer nondurables								
Real	1			2	4	3		
Current	1			2*	4	3		
Consumer services								
Real	1			2	3	4		
Current	1		3	2*	4	5		
Consumer nondurables including services								
Real	1			4	3	2	6	5*
Current	1			2*	3	6	5	4

Note: Asterisks indicate addition of exogenous deflators.

B. BUSINESS FIXED INVESTMENT

A comparison of mean-squared errors for both real and current plant and equipment expenditures over the fitted period reveals that the Liu Model has a smaller prediction error than all the other quarterly econometric models and the naive model (*see* Table 2.2). However, in predicting current plant and equipment expenditures, a comparison of the Liu and Friend-Taubman Models with the other models is valid only one way, because neither of these two models determines the plant and equipment deflator as an endogenous variable. Also, a comparison of the Fromm Model with the other models in predicting real investment is valid only one way, since the former determines real investment from current investment through an exogenous deflator. For current plant and equipment expenditures, the Liu Model again registers a predictive performance superior to the other econometric models

TABLE 2.2

Mean-Squared Errors Over Fitted Period for Business Fixed Investment

	Econometric Models							
Variables	Friend-Taub-man	Fromm	Liu	Klein	OBE	Wharton-EFU	Gold-feld	Auto-regressive Schemes
Real plant and equipment expenditures	2.065	1.661*	0.942	3.932	2.005	98.07	+	1.70
Real nonresidential structures	+	+	0.0999	+	+	+	+	0.1224
Real producers' durable equipment	+	+	0.714	+	+	+	+	1.363
Current plant and equipment expenditures	1.469*	1.152	0.693*	6.520	1.922	26.74	+	1.389
Current nonresidential structures	+	+	0.0849	+	+	+	+	0.1024
Current producers' durable equipment	+	+	0.516	+	+	+	+	1.107
Real plant and equipment expenditures in manufacturing	+	+	+	+	+	0.6001	+	0.2154
Real plant and equipment expenditures in nonmanufacturing	+	+	+	+	+	89.13	+	1.085
Real gross private domestic investment	4.083	4.803	2.162	6.689	2.908	97.88	+	17.64
Current gross private domestic investment	3.201*	3.683	0.9969*	28.06	1.922	23.70	8.258	2.106

and the naive model. Ranking of the models in predicting plant and equipment expenditures based on alternative quarterly models is shown here.

Variables	Rank Order of Predictive Performance Over Fitted Period						
	Naive	Friend-Taub-man	Fromm	Liu	Klein	OBE	Whar-ton-EFU
Real plant and equipment expenditures	3	5	2*	1	6	4	7
Current plant and equipment expenditures	3	5*	2	1*	6	4	7

Note: Asterisks indicate addition of exogenous deflators.

Liu's is the only econometric model included in the study which explains separately the nonresidential structures and producers' durables components of business fixed investment. For both of these variables, the predictive performance of the Liu Model is superior to the naive models over the period of estimation. The Liu Model also predicts current nonresidential structures and producers' durable equipment better than do the mechanical models.

Although the Wharton-EFU Model does not separate plant from equipment expenditures, it does disaggregate business fixed investment into manufacturing and nonmanufacturing investment. For investment in both manufacturing and nonmanufacturing, the Wharton-EFU Model registers a predictive performance inferior to that of the naive models.

The Goldfeld Model explains only gross fixed investment. In comparing the mean-squared errors across models for both real and current gross fixed investment, we see that the Liu Model again registers a predictive performance superior to that of the other models. All of the remaining econometric models are inferior to purely mechanical forecasting schemes in predicting gross fixed investment over the period of fit.

TABLE 2.3

Mean-Squared Errors Over Fitted Period for Residential Structures

Variables	Econometric Models							Auto-regressive Schemes
	Friend-Taub-man	Fromm	Liu	Klein	OBE	Wharton-EFU	Goldfeld	
Real total residential structures	1.455	2.060	0.2683	1.688	1.730	0.2147	+	0.3425
Real nonfarm residential structures	1.455	2.060	0.2683	1.688	1.730	0.2147	+	0.3565
Real new nonfarm residential structures	+	+	+	+	1.730	+	+	71.59
Current total residential structures	1.269*	1.773	0.2360*	1.640	1.363	.2328	+	0.4086
Current nonfarm residential structures	1.269*	1.773	0.2360*	1.640	1.363	.2328	+	0.4058

C. RESIDENTIAL STRUCTURES

Table 2.3 presents mean-squared errors for residential structures in the reduced forms of the quarterly models over the fitted period. Considering these results in detail, we observe that the Wharton-EFU and Liu Models outperform all other quarterly models in predicting real total and nonfarm residential structures. The mean prediction error for the Wharton-EFU Model is slightly smaller than that of the Liu Model.

In predicting current residential structures a comparison among all quarterly econometric models is valid only one way because the Klein, OBE, and Wharton-EFU Models determine the housing deflator as an endogenous variable, while the other econometric models treat the housing deflator as an exogenous variable. Rankings of the quarterly models in predicting real and current residential structures follow.

Variables	Rank Order of Predictive Performance Over Fitted Period						
	Naive	Friend-Taub-man	Fromm	Liu	Klein	OBE	Wharton-EFU
Real residential structures	3	4	7	2	5	6	1
Current residential structures	3	4	7	2	6	5	1

D. IMPORTS AND EXPORTS

Mean-squared errors over the fitted period, based on reduced-form equations for imports and exports over this period, are presented in Table 2.4. From the results shown in this table, we observe, first, that in predicting real imports and exports over the period of fit, the naive model outperforms all of the quarterly econometric models. Among the econometric models, the Fromm Model registers the best predictive performance for real imports, and the Wharton-EFU Model is superior to the Klein Model in predicting exports.

TABLE 2.4

Mean-Squared Errors Over Fitted Period for Imports and Exports

Variables	Econometric Models: Fromm	Klein	OBE	Wharton-EFU	Auto-regressive Schemes
Real total imports	0.555	1.043	1.652	0.6051	0.3209
Real imports of crude materials and foodstuffs	+	0.023	0.024	0.0218	0.0191
Real imports of semifinished and finished goods and services	+	0.874	1.634	0.4851	0.2437
Current total imports	1.774	0.982	1.652	0.5854	0.4057
Real total exports	+	2.35	+	1.8149	1.02
Current total exports	+	2.798	+	1.578	0.9893

In predicting real imports of crude materials and foodstuffs, and real imports of semifinished and finished goods and services over the fitted period, the naive model registers a predictive performance superior to that of the quarterly econometric models. In predicting current-dollar imports, the naive model is again superior to all of the quarterly econometric models. However, the Fromm Model now moves from first place among the econometric models in predicting real imports to last place in predicting current imports. Rankings of the quarterly models for imports and exports follow.

Variables	Rank Order of Predictive Performance Over Fitted Period: Naive	Fromm	Klein	OBE	Wharton-EFU
Real imports	1	2	4	5	3
Current imports	1	4	3		2
Real imports of crude materials and foodstuffs	1		3	4	2
Real imports of semifinished and finished goods and services	1		3	4	2
Real exports	1		3		2

In summary, a comparison of prediction results between the econometric and naive models is unmistakably clear over the fitted period. The mechanical schemes for every imports and exports variable register a superior predictive performance when compared with that of the quarterly econometric models. The prediction results among the econometric models themselves are somewhat mixed. The Wharton-EFU Model does predict disaggregated imports and real exports better than does any other econometric model.

E. INVENTORIES, ORDERS, AND SHIPMENTS

Table 2.5 presents mean-squared errors based on reduced-form predictions for inventories, orders, and shipments for the quarterly econometric models over the fitted period. We observe first that the Liu Model outperforms all other quarterly econometric models, as well as the naive model, in predicting constant-dollar inventory investment. Next in line in predicting real inventory investment is the Wharton-EFU Model, which performs slightly better than the Fromm Model. Both econometric models outperform the naive model over the fitted period. In predicting current-dollar inventory investment, the Liu Model again outperforms all other quarterly models. Rankings of the models for real and current inventory investment are given below.

Variables	Rank Order of Predictive Performance Over Fitted Period							
	Naive	Friend-Taub-man	Fromm	Liu	Klein	OBE	Whar-ton-EFU	Gold-feld
Real inventory investment	5	8	3	1	6	4	2	
Current inventory investment	6	8	4	1	7	3	2	5

Note that in predicting both real and current inventory investment, the Klein Model registers the worst predictive performance of all the quarterly models under consideration.

TABLE 2.5

Mean-Squared Errors Over the Fitted Period for Inventory Investment, Orders, and Shipments

	Econometric Models							
Variables	Friend-Taub-man	Fromm	Liu	Klein	OBE	Whar-ton-EFU	Gold-feld	Auto-regressive Schemes
Real nonfarm inventory investment	19.65	7.211	6.464	15.26	8.385	6.954	10.18	10.06
Real total inventory investment	19.65	7.211	6.464	15.26	8.385	6.954	10.18	10.51
Current nonfarm inventory investment	15.36	6.759	6.149	13.68	6.634	5.913	8.280	8.75
Current total inventory investment	15.36	6.759	6.149	13.68	6.634	5.913	8.280	9.325
Real stock of nonfarm business inventories	+	+	0.6418	+	+	+	+	0.6418
Real unfilled orders in manufacturing	+	7.919	+	+	+	+	+	0.0078
Real new orders in manufacturing	+	+	+	0.0052	4.649	+	+	0.0019
Real shipments of durable goods in manufacturing	+	+	+	+	1.270	+	+	0.5509
Real unfilled orders of durable goods in manufacturing	+	+	+	0.0104	20.40	+	+	7.241

F. REDUCED-FORM COMPARISONS FOR GROSS NATIONAL PRODUCT

Table 2.6 presents mean-squared errors for real and current *GNP*, determined on the demand side for the quarterly econometric and naive models over the period of fit. The mean-squared errors for *GNP* based on the econometric models are derived from reduced-form predictions. Detailed consideration of these results shows that in predicting both real and current *GNP* over the period of fit, the Fromm Model ranks first, followed by the Liu and naive models. The naive model is slightly better than the Liu system for predicting real *GNP*, with positions reversed for current *GNP*. Poorest prediction performances over the fitted period are registered by the Goldfeld Model for real *GNP*; and by the Klein for current *GNP*, as shown by the following ranking.

Variables	Rank Order of Predictive Performance Over Fitted Period							
	Naive	Friend-Taub-man	Fromm	Liu	Klein	OBE	Whar-ton-EFU	Gold-feld
Real *GNP*	2	7	1	3	5	4	6	8*
Current *GNP*	3	5*	1	2*	8	4	7	6

Note: Asterisks indicate addition of exogenous deflators.

This ranking of quarterly econometric models is not completely valid, since the endogenous variables are conditional on different sets of predetermined variables. However, further inspection of the results

TABLE 2.6

Mean-Squared Errors Over Fitted Period for Gross National Product Determined on Demand Side

Variables	Econometric Models							Auto-regres-sive Schemes
	Friend-Taub-man	Fromm	Liu	Klein	OBE	Whar-ton-EFU	Gold-feld	
Real *GNP*	64.44	18.16	25.95	44.48	31.53	178.5	452.8*	25.50
Current *GNP*	51.67*	18.75	21.82*	135.2	35.92	129.2	108.3	27.71

TABLE 2.7

Mean-Squared Errors Over the Fitted Period for Gross Product Originating

Variables	Econometric Models: Klein	OBE	Wharton-EFU	Auto-regressive Schemes
Real private *GNP*	14.36	101.1	178.6	24.41
Real private *GNP* at full capacity	19.40	136.2	+	62.16
Real *GNP* originating in manufacturing	+	+	189.1	6.844
Real *GNP* originating in nonmanufacturing	+	+	23.58	14.05
Real *GNP* originating in residential structures	+	+	0.608	0.7146
Real full capacity output in manufacturing	+	+	456.1	15.26

indicates that the above ranking would hold even after allowing the models to be conditional on different sets of predetermined variables.

G. GROSS PRODUCT ORIGINATING

Reduced-form mean-squared errors for gross product originating variables in the quarterly econometric models over the period of fit are presented in Table 2.7.[9] In predicting both real actual and full capacity private *GNP*, the Klein Model turns in the best predictive performance, followed by the naive model. Both are superior to the OBE and Wharton-EFU Models, which rank third and fourth, respectively.

Considering the predictive performance over the fitted period for the more disaggregated gross product variables, we observe that the naive model outperforms the Wharton-EFU Model in predicting actual

[9] The reader may wonder how these models can determine more than one value for *GNP*. These models are all overdetermined systems, so that it is possible to determine *GNP* both from the demand side and from the aggregate production function. Consequently, more than one value of *GNP* can be obtained.

and full capacity output in both manufacturing and nonmanufacturing sectors.

H. CAPITAL CONSUMPTION ALLOWANCES

Reduced-form mean-squared errors for depreciation over the fitted period are presented in Table 2.8. We find, with one exception, that the mechanical models provide the best prediction for both real and accounting depreciation: the OBE model provides the best forecast of the real stock of plant and equipment.

I. TAX AND TRANSFER ITEMS

In Table 2.9, we present mean-squared errors based on reduced-form tax and transfer equations over the period of fit. Beginning with indirect business tax and nontax liability, we observe this category to be best predicted by the naive model. The best-performing econometric model is the Wharton-EFU, although the others perform almost as well.

In predicting corporate profits tax liability over the fitted period, the naive model outperforms all of the quarterly econometric models. The Klein Model registers the best predictive performance over the fitted period for personal tax and nontax liability, although the naive and OBE models perform almost as well. Rankings of the models for these variables are as shown.

Variables	Rank Order of Predictive Performance Over Fitted Period				
	Naive	Fromm	Klein	OBE	Wharton-EFU
Indirect business taxes	1	3	5	4	2
Corporate profits taxes	1	3	4	2	5
Personal taxes	2	4	1	3	5

The Fromm Model explains federal, and state and local, personal taxes separately; for these two variables, the predictive performance of the naive models is superior to that of the Fromm Model.

TABLE 2.8

Mean-Squared Errors Over Fitted Period for Capital Consumption Allowances

	Econometric Models					
Variables	Fromm	Liu	Klein	OBE	Wharton-EFU	Autoregressive Schemes
Real accounting depreciation	+	0.6103	+	+	+	0.0755
Current accounting depreciation	0.1206	0.5353	+	+	0.6891	0.0434
Real depreciation on nonfarm capital stock	+	+	0.2265	.0017	+	0.2218
Current accounting depreciation in manufacturing	+	+	+	+	0.1671	0.0069
Current accounting depreciation in nonmanufacturing	+	+	+	+	0.3069	0.0585
Current accounting depreciation in nonfarm residential structures	+	+	+	+	0.0311	0.0038
Corporate accounting depreciation	0.0511	+	+	+	+	0.0433

TABLE 2.9

Mean-Squared Errors Over the Fitted Period for Tax and Transfer Items

Variables	Econometric Models				Autoregressive Schemes
	Fromm	Klein	OBE	Wharton-EFU	
Indirect business tax and nontax liability	1.046	2.391	1.061	0.9839	0.2671
Corporate profits tax liability	6.123	10.35	4.580	17.48	2.637
Personal tax and nontax liability	8.912	4.362	5.236	10.95	0.4989
Federal personal tax and nontax liability	8.662	+	+	+	0.4481
State and local personal tax and nontax liability	0.106	+	+	+	0.0052
State unemployment insurance benefits	0.1474	+	0.8875	+	0.6458
OASI and veterans' benefits	0.100	+	+	+	0.100
Relief payments and other transfers	1.540	+	+	+	1.733
Business and government transfer payments	+	+	+	7.725	2.450

J. NONLABOR INCOME, CORPORATE PROFITS, DIVIDENDS, AND RETAINED EARNINGS

Mean-squared errors for nonlabor income and other related items over the fitted period are given in Table 2.10. Considering these results in detail, we observe, first, that the Wharton-EFU Model is superior to all other quarterly models in predicting nonlabor personal income over the fitted period, while the naive model outperforms the other econometric models.

In predicting before-tax corporate profits, the Liu Model is superior to all other quarterly models, while the naive model also outperforms them.

In predicting after-tax corporate profits, the naive model outperforms all econometric models (excluding the Liu Model). In second place is the OBE Model, which outperforms the remaining econometric models.

For predicting current retained earnings for the fitted period, the naive model outperforms all others; the Liu performs best of the quarterly econometric models.

In predicting constant dollar retained earnings, the Klein outperforms the naive model, but the naive performs better than the Wharton-EFU in predicting retained earnings in manufacturing.

In predicting current dividends, the naive model narrowly outdoes the econometric models, all of which perform about equally well, with the OBE Model showing a slight superiority. Rankings of the models for current nonlabor personal income and its components are shown below.

Variables	Rank Order of Predictive Performance Over Fitted Period					
	Naive	Fromm	Liu	Klein	OBE	Wharton-EFU
Nonlabor personal income	2			4	3	1
Before-tax corporate profits	2	4	1	5	3	6
After-tax corporate profits	1	4		3	2	5
Current retained earnings	1	4	2	3	5	6
Current dividends	1	3	4	6	2	5

TABLE 2.10

Mean-Squared Errors Over Fitted Period for Nonlabor Personal Income, Corporate Profits, Retained Earnings, and Dividends

Variables	Econometric Models: Fromm	Liu	Klein	OBE	Wharton-EFU	Auto-regressive Schemes
Nonlabor personal income	+	+	9.644	2.181	0.6104	1.082
Corporate profits before taxes, current dollars	13.66	5.453	23.16	7.548	255.6	6.169
Corporate profits before taxes, constant dollars	+	6.229	+	+	+	6.868
Corporate profits after taxes, current dollars	7.273	+	6.462	4.199	120.4	3.625
Retained earnings (including *IVA*), deflated by plant and equipment deflator	+	+	5.374	+	+	8.09
Current retained earnings	6.932	4.562	5.864	7.512	115.0	1.829
Retained earnings in manufacturing	+	+	+	+	4.158	1.092
Real dividends	+	0.2278	+	+	+	0.1995
Current dividends	0.1355	0.1762	0.7614	0.1337	0.2437	0.1235

Finally, in predicting real dividends, the naive model outperforms the Liu Model.

K. EMPLOYMENT, HOURS, AND WAGES

In Table 2.11 we present the predictive results over the fitted period, based on the reduced forms of the various models explaining hours and wages. Considering these results in detail, we observe, first, that regarding employment, the Fromm Model best forecasts the number of unemployed workers. In second place is the naive model, which outperforms all other econometric models.

Next, we observe for the total civilian labor force, and for total employment, that the naive model has a smaller prediction error than either the OBE or Klein Model. In predicting total employment, the Klein outperforms the OBE Model; but for the total civilian labor force, the OBE Model performs better than the Klein.

For average hours worked per week in the private sector, the OBE Model has a smaller prediction error than the Klein Model over the period of fit.

In predicting current-dollar total civilian and private wages and salaries, the naive model outperforms all of the quarterly econometric models. In second place is the Klein Model, which outperforms both the Wharton-EFU and OBE Models. Rankings for labor force, and wages and salaries, are as shown.

Variables	Rank Order of Predictive Performance Over Fitted Period				
	Naive	Fromm	Klein	OBE	Wharton-EFU
Unemployment	2	1	4	5	3
Employment	1		2	4	3
Labor force	1		3	2	
Wages and salaries	1		2	4	3

TABLE 2.11

Mean-Squared Errors Over Fitted Period for Employment Hours and Wages

Variables	Econometric Models: Fromm	Klein	OBE	Wharton-EFU	Autoregressive Schemes
Unemployed workers	0.0737	1.905	2.177	3.190	0.106
Total employment	+	0.9058	2.907	1.311	0.223
Employment in manufacturing	+	+	+	2.778	0.0485
Employment in nonmanufacturing	+	+	+	0.6353	0.0338
Total civilian labor force	+	0.8476	0.283	+	0.2739
Total civilian wages and salaries, current	+	30.36	103.3	31.23	4.791
Private wages and salaries, current	+	30.36	103.3	31.23	3.785
Real private wages and salaries	+	3.760	+	+	63.93
Wages and salaries in manufacturing	+	+	+	26.91	4.659
Wages and salaries in nonmanufacturing	+	+	+	6.476	0.6369
Average hours worked per week	+	0.1872	0.3249[a]	+	0.2377[a]
Average hours worked per week in manufacturing	+	+	+	0.0017	0.7811[a]
Average hours worked per week in nonmanufacturing	+	+	+	0.0005	0.241[a]
Private average hourly wage rate	+	0.529[b]	0.0092	+	0.1678[b]
Private hourly wage rate for manufacturing	+	+	+	0.0075	0.0027
Private hourly wage rate for nonmanufacturing	+	+	+	0.0027	0.0007

[a] E-04.
[b] E-05.

In predicting wages and salaries, both in manufacturing and nonmanufacturing, the naive outperforms the Wharton-EFU Model. For constant-dollar private wages and salaries, the Klein demonstrates a predictive performance superior to the naive model for the fitted period.

With respect to the total average private wage rate, the naive model has a smaller prediction error than either the Klein or OBE Model. For the private wage rate, in both manufacturing and nonmanufacturing, the naive model registers a predictive performance superior to that of the Wharton-EFU Model over the period of fit.

L. IMPLICIT PRICE DEFLATORS

Given in Table 2.12 are mean-squared errors based on reduced-form predictions for implicit price deflators for *GNP* and its components over the fitted period. In predicting the implicit price deflator for *GNP* over the fitted period, the naive is the best performer of all the models; the Liu, best of the econometric models. Both the OBE and Wharton-EFU Models determine the private *GNP* deflator. For this variable, the Wharton-EFU Model registers a performance superior to that of the OBE and naive models, the OBE outperforming the naive model. The total consumption deflator is available only for the OBE and Wharton-EFU Models. For this deflator, both econometric models outperform the naive model over the fitted period. However, in predicting each of the following deflators over the fitted period, the naive model outperforms the econometric models: consumer services, consumer nondurables, consumer durables, nonauto durables. Over the fitted period, the plant and equipment deflator is predicted best by the naive model. This model again outperforms all econometric models over the fitted period in predicting the nonfarm residential structures deflator; the Wharton-EFU outperforms the OBE and Klein Models. Rankings of some of these deflators are shown here.

Variables	Rank Order of Predictive Performance Over Fitted Period					
	Naive	Fromm	Liu	Klein	OBE	Wharton-EFU
GNP price deflator	1	3	2	4		5
Plant and equipment deflator	1			4	3	2
Nonfarm residential structures deflator	1			3	4	2
Consumer services deflator	1			2	3	
Consumer nondurables deflator	1			3	2	
Consumer durables deflator	1			3		2
Nonauto durables deflator	1				3	2

The exports deflator is determined by the Klein and Wharton-EFU Models; but for this variable, over the fitted period the naive model again outperforms both econometric models. The Wharton-EFU Model is the poorest-performing of the econometric models.

M. MISCELLANEOUS ITEMS

Reduced-form mean-squared errors for net interest paid by government and consumer, and inventory valuation adjustment over the fitted period, are presented in Table 2.13. These results show that in predicting both *IVA* and net interest the naive models perform best.

N. INTEREST RATES

Table 2.14 presents reduced-form mean-squared errors for interest rates over the period of estimation. Detailed consideration of

TABLE 2.12

Mean-Squared Errors Over Fitted Period for Implicit Price Deflators

Variables	Econometric Models					Auto-regressive Schemes
	Fromm	Liu	Klein	OBE	Whar-ton-EFU	
GNP price deflator	0.3136[a]	0.1936[a]	0.906[b]	0.929[d]	0.1089[d]	0.1824[a]
Private *GNP* price deflator	+	+	+	0.185[a]	0.1388[a]	0.2019[a]
Total consumption deflator	+	+	+	0.2[b]	0.160[a]	0.2268[a]
Consumer services deflator	+	+	0.2403[b]	0.5476[b]	+	0.3919[c]
Consumer nondurables deflator	+	+	0.404[b]	0.3648[b]	+	0.3997[a]
Consumer durables deflator	+	+	0.4754[b]	+	0.188[b]	0.9746[a]
Consumer nonauto durables deflator	+	+	+	0.219[b]	0.9409[a]	0.3307[a]
Plant and equipment deflator	+	+	0.1756[d]	0.1464[b]	0.5625[a]	0.3965[c]
Nonfarm residential structures deflator	+	+	0.5018[b]	0.3764[b]	0.8649[a]	0.6243[a]
Exports deflator	+	+	0.3133[b]	+	0.5244[b]	0.1679[b]

[a] E-04.
[b] E-03.
[c] E-05.
[d] E-02.

TABLE 2.13

Mean-Squared Errors Over Fitted Period for Miscellaneous Items

Variables	Econometric Models: Wharton-EFU	Econometric Models: Fromm	Autoregressive Schemes
Net interest paid by government and consumers	+	0.0265	0.008
Inventory valuation adjustment	3.00	+	1.71

these results shows that over the fitted period, all of the econometric models yield better predictions of the 4- to 6-month prime commercial paper rate than does the naive model, with the Liu Model registering the best predictive performance of all quarterly models. For Moody's AAA corporate bond rate, the Liu Model again outperforms all other quarterly models. In second place for this variable is the naive model, which outperforms the remaining econometric models. Rankings of the quarterly models for these two interest rates are listed here.

Variables	Rank Order of Predictive Performance Over Fitted Period: Naive	Liu	Klein	OBE	Wharton-EFU
4- to 6-month prime commercial paper rate	4	1	2	2	3
Moody's AAA corporate bond yield	2	1	2	2	3

In predicting the weighted average yield on time deposits and savings shares, the naive model outperforms the Liu Model. Also, the naive model predicts the mortgage yield better than the OBE Model. Finally, in the fitted period, the naive models outperform the Goldfeld Model in predicting the bill, intermediate-term, and long-term government rates.

TABLE 2.14

Mean-Squared Errors Over Fitted Period for Interest Rates

Variables	Econometric Models: Liu	Klein	OBE	Wharton-EFU	Goldfeld	Auto-regressive Schemes
4- to 6-month prime commercial paper rate	0.0424	0.0475	0.0480	0.049	+	0.0646
Moody's AAA corporate bond yield	0.0078	0.0147	0.0147	0.0177	+	0.0125
Weighted average yield on time deposits and savings shares	0.300	+	+	+	+	0.0009
FHA mortgage yield on insured homes	+	+	0.0161	+	+	0.1409
Commercial loan rate	+	+	+	+	1.431	0.0187
Treasury bill rate	+	+	+	+	1.167	0.0675
Intermediate government bond rate (3- to 5-years)	+	+	+	+	1.0349	0.0619
Long-term government bond rate	+	+	+	+	0.4223	0.0108

To summarize, for the commercial paper rate and private bond rate the Liu Model outperforms all quarterly models. For the commercial paper rate, all of the econometric models outperform the naive model. With the exception of the Liu Model, the naive model registers a performance superior to that of the econometric models in predicting the private long-term bond rate. Over the fitted period, the naive models outperform the econometric models in predicting all other interest rates.

O. OTHER FINANCIAL VARIABLES

Table 2.15 presents reduced-form mean-squared errors over the fitted period for financial variables other than interest rates. Results considered in detail show that the Liu outperforms the naive model in predicting real business liquid assets, and both real and current holdings of consumer demand deposits. The naive outperforms the Liu Model in predicting current business liquid assets and constant and current dollar holdings of consumer time deposits and savings shares.

According to Table 2.15, the Liu Model predicts constant-dollar total consumer liquid assets better than does the naive model. However, in predicting current-dollar total consumer liquid assets, the Liu is outperformed by the naive model, being followed, in order, by the Klein and OBE Models.

The remaining mean-squared errors in Table 2.15 are based on financial series unique to the Goldfeld Model. Detailed consideration of these results shows that the naive model outperforms the Goldfeld Model in predicting city and country member bank excess reserves, Federal Reserve borrowings, city member bank long-term securities, and country bank municipal securities. The Goldfeld outperforms the naive model in predicting city and country member bank holdings of government securities, country member bank long-term securities, and city bank municipal securities.

The prediction results for most of the deposit variables largely favor the naive models. Over the fitted period, they outperform the Goldfeld Model in predicting total demand and time deposits, city and country member bank time deposits, commercial loans, and both total

TABLE. 2.15

Mean-Squared Errors Over Fitted Period for Financial Variables Other Than Interest Rates

Variables	Econometric Models: Liu	Klein	OBE	Goldfeld	Auto-regressive Schemes
Real holdings of business demand and time deposits	0.6469	+	+	+	1.984
Current holdings of business demand and time deposits	0.6089	+	+	+	0.2162
Real holdings of consumer demand deposits	0.6947	+	+	+	0.8636
Current holdings of consumer demand deposits	0.4588	+	+	+	0.5941
Real holdings of consumer time deposits and savings shares	1.215	+	+	+	0.5379
Current holdings of consumer time deposits and savings shares	1.189	+	+	+	0.6827
Real total liquid assets of consumers	2.304	+	+	+	3.208
Current total liquid assets	1.939	411.2	46.22	+	1.606
Excess reserves of city member banks	+	+	+	0.0042	0.0028
Excess reserves of country member banks	+	+	+	0.0064	0.0019
City member bank borrowings	+	+	+	0.1447	0.0421

Country member bank borrowings	+	+	+	0.0085	0.0011
City member bank holdings of short-term securities	+	+	+	1.133	2.57
Country member bank holdings of short-term securities	+	+	+	0.1851	0.7068
City member bank holdings of long-term securities	+	+	+	1.672	1.344
Country member bank holdings of long-term securities	+	+	+	0.4046	0.5729
Currency component of money supply	+	+	+	0.0314	0.0203
Demand deposits component of money supply	+	+	+	0.4294	0.4685
Time deposits city member banks	+	+	+	0.1966	0.1299
Time deposits country member banks	+	+	+	0.1875	0.0394
Time deposits member banks	+	+	+	1.1825	0.2297
Demand deposits city member banks	+	+	+	0.1669	0.7268
Demand deposits country member banks	+	+	+	0.0428	0.1233
Commercial loans member banks	+	+	+	2.573	2.206
Municipal securities city member banks	+	+	+	0.0228	0.0361
Municipal securities country member banks	+	+	+	0.0051	0.0039
City member bank commercial loans	+	+	+	1.646	1.859
Country member bank commercial loans	+	+	+	0.1033	0.0998
Change in class average reserve requirements on demand and time deposits, city member banks	+	+	+	0.36E-06	18.13E-06
Change in class average reserve requirements on demand and time deposits, country member banks	+	+	+	0.16E-06	13.28E-06

member and country member bank loans. The Goldfeld Model outperforms the naive models in predicting city and country member bank demand deposits, and city member bank commercial loans. It is distinctly superior in predicting changes in class average reserve requirements on city and country member bank deposits.

To summarize, except for short- and long-term government security holdings, the Goldfeld Model is generally inferior—where the Klein and OBE Models are definitely inferior—to the mechanical forecasting schemes in predicting noninterest rate financial variables. The Liu Model is generally superior to the naive models for real financial variables; but for current-dollar financial variables, it is slightly inferior.

P. PERSONAL AND DISPOSABLE INCOME

All of the econometric models, except the Friend-Taubman, determine personal disposable income—with the Liu Model explaining both real and current. The variable of personal income, however, is determined only in the Klein, OBE, and Wharton-EFU Models.

The mean-squared errors based on reduced forms for both personal and disposable personal income over the fitted period are given in Table 2.16. Considering these results for both personal and disposable personal income, we find that the naive models register predictive performances superior to those of all of the quarterly econometric models—of which, the Fromm Model performs best. A ranking of performance over the period of fit for both variables follows.

Variables	Rank Order of Predictive Performance Over Fitted Period						
	Naive	Fromm	Liu	Klein	OBE	Wharton-EFU	Goldfeld
Personal income	1	2		4	5	3	
Personal disposable income	1	2	3	4	6	5	7

Q. SUMMARY

This concludes the goodness-of-fit comparisons across econometric models in the reduced form. However, as pointed out by Friedman, the problem with comparing the goodness-of-fit of econometric models

TABLE 2.16

Mean-Squared Errors Over Fitted Period for Personal and Disposable Income

	Econometric Models						
Variables	Fromm	Liu	Klein	OBE	Wharton-EFU	Goldfeld	Autoregressive Schemes
Personal income	15.56	+	54.80	117.2	18.44	+	1.083
Personal disposable income	11.08	15.29	54.67	88.26	14.37	108.6	6.935
Real personal disposable income	+	19.94	+	+	+	+	9.10

over a fixed period is that they may reflect not only the performance of the model but the persistence of the investigator. More conclusive evidence as to the predictive performance of the econometric models can be obtained by comparing them over a period not used in fitting the models. To this, we now turn.

3 A COMPARISON OF QUARTERLY ECONOMETRIC MODELS OVER THE PERIOD OF FORECAST: 1961-I THROUGH 1965-IV

WITH the econometric models and mechanical forecasting schemes ranked for the period of estimation, we may now evaluate the predictive performance of the quarterly econometric models over the forecast period: 1961-I through 1965-IV.

A. CONSUMER EXPENDITURES

Mean-squared errors for consumer expenditures over the forecast period are presented in Table 3.1. Considering these results in detail, we observe that for real total consumer expenditures the naive model forecasts better than all of the quarterly econometric models; the best of which are the Klein and OBE, and the worst, the Friend-Taubman.

Current consumer expenditures, in contrast, are forecast best by the OBE, second-best by the naive, and most poorly by the Friend-Taubman Model.

In forecasting real consumer durables, the Liu Model outperforms all other quarterly models, the naive model forecasting better than the remaining econometric models. Current consumer durables are forecast best by the Klein Model, next best by the naive model. Real consumer durables are separated into auto and nonauto durables in the OBE and Wharton-EFU Models, both of which forecast these variables less satisfactorily than does the naive model. In forecasting real automobile expenditures, the OBE outperforms the Wharton-EFU Model, while the reverse occurs in forecasting nonauto durables. Real consumer nondurables are forecast best by the Klein Model, according to the results presented in Table 3.1; the naive model outperforms all other econometric models. The naive model outperforms all econometric models in forecasting current consumer nondurables. Both the Wharton-EFU and Goldfeld Models determine nondurables and services together; for this combined variable, the naive model registers the best forecast. Rankings for consumption and its components are given below in tabular form.

Variables	Rank Order of Predictive Performance Over Forecast Period							
	Naive	Friend-Taubman	Liu	Klein	Fromm	OBE	Wharton-EFU	Goldfeld
Real consumer expenditures	1	7	5	3	4*	2	8	6*
Current consumer expenditures	2	8*	4*	6	3	1	7	5
Real consumer durables	2		1	3		4	5	*
Current consumer durables	2		3*	1		5	6	4
Real consumer nondurables	2		3	1		4		
Current consumer nondurables	1		2*	3			4	
Real services	1		2	3		4		
Current services	1		2*	5	3	4		
Real nondurables (including services)	1		4	2		5	3	
Current nondurables (including services)	1		3*	5		5	4	2

Note: Asterisks indicate addition of exogenous deflators.

TABLE 3.1

Mean-Squared Errors Over Forecast Period for Consumption Expenditures

	Econometric Models							
Variables	Friend-Taubman	Fromm	Liu	Klein	OBE	Wharton-EFU	Goldfeld	Auto-regressive Schemes
Real total consumer expenditures	181.3	24.25*	48.91	21.89	21.06	156.9	71.75	15.46
Current total consumer expenditures	212.2*	28.10	52.37*	126.1	7.414	175.5	105.9	16.76
Real consumer durables	+	+	8.93	11.98	64.33	102.13	20.05	9.820
Real consumer automobile expenditures	+	+	+	+	8.777	40.24	+	4.366
Real nonautomobile durables	+	+	+	+	102.8	14.78	+	2.364
Current consumer durables	+	+	8.969*	4.073	61.41	96.96	20.06	7.957
Real consumer nondurables	+	+	20.19	1.522	21.28	+	+	2.691
Current consumer nondurables	+	+	22.23*	24.18	40.41	+	+	3.019
Real consumer services	+	+	0.5233	1.404	2.554	+	+	0.3233
Current consumer services	+	0.9532	0.637*	30.22	4.931	+	+	0.4625
Real consumer nondurables and services	+	+	24.83	2.844	36.32	14.81	10.54*	3.874
Current consumer nondurabies and services	+	+	27.67*	105.9	60.14	28.32	23.19	4.525

B. BUSINESS FIXED INVESTMENT

Reduced-form forecasts for business fixed investment are given in Table 3.2. Considering these results, we observe that in forecasting real plant and equipment expenditures the Fromm Model outperforms all other quarterly models, with the Liu Model in second place. Current plant and equipment expenditures are also forecast best by the Fromm Model. In second place this time, however, is the naive model, which forecasts current business investment better than all of the quarterly models. The Liu Model explains business structures and producers' durable equipment separately. For these variables, expressed in constant dollars, the Liu Model registers a better forecast for equipment than does the naive model, but the reverse occurs for structures. The naive models forecast current business structures and producers' durable equipment better than the Liu Model results. The Wharton-EFU Model disaggregates plant and equipment expenditures into manufacturing and nonmanufacturing. For these variables, expressed in real terms, the naive models project better forecasts than the Wharton-EFU Model. The Goldfeld Model determines gross fixed investment. Comparing the mean-squared errors across the quarterly models for the forecast period, we find that the naive model forecasts gross fixed investment better than any of the econometric models. In the table below, the models are ranked by investment class.

Variables	Rank Order of Predictive Performance Over Forecast Period							
	Naive	Friend-Taub-man	Fromm	Liu	Klein	OBE	Whar-ton-EFU	Gold-feld
Real plant and equipment	5	4	1*	2	6	3	7	
Current plant and equipment	2	4*	1	3*	6	5	7	

Note: Asterisks indicate addition of exogenous deflators.

C. RESIDENTIAL STRUCTURES AND HOUSING STARTS

Total and nonfarm residential structures are forecast best by the naive models (Table 3.3). In second place in these categories is the

TABLE 3.2

Mean-Squared Errors Over Forecast Period for Business Fixed Investment

Variables	Econometric Models							
	Friend-Taubman	Fromm	Liu	Klein	OBE	Wharton-EFU	Goldfeld	Autoregressive Schemes
Real plant and equipment expenditures	14.26	1.021*	4.677	113.6	10.55	185.3	+	84.83
Real nonresidential structures	+	+	1.802	+	+	+	+	0.6461
Real producers' durable equipment	+	+	3.614	+	+	+	+	5.161
Current plant and equipment expenditures	16.22*	1.131	5.161*	422.75	4.97	62.25	+	2.346
Current nonresidential structures	+	+	2.222	+	+	+	+	0.7893
Current producers durable equipment	+	+	3.855	+	+	+	+	1.752
Real plant and equipment expenditures in manufacturing	+	+	+	+	+	2.961	+	0.567
Real plant and equipment expenditures in nonmanufacturing	+	+	+	+	+	142.1	+	1.276
Real gross private domestic investment	43.45	79.31	38.80	93.21	144.3	198.1	445.2	26.78
Current gross private domestic investment	21.58*	93.28	43.84*	6.648	75.30	56.62	10.55	3.936

TABLE 3.3

Mean-Squared Errors Over Forecast Period for Residential Structures

Variables	Econometric Models						Auto-regressive Schemes
	Friend-Taubman	Fromm	Liu	Klein	OBE	Wharton-EFU	
Real total residential structures	1.892	36.62	6.876	23.24	10.54	0.4283	0.4252
Real nonfarm residential structures	1.892	36.62	6.876	23.24	10.54	0.4283	3.4184
Real new nonfarm residential structures	+	+	+	+	574.3	+	88.80
Current total residential structures	2.319*	45.83	8.762*	35.16	13.07	0.7529	0.484
Current nonfarm residential structures	2.319*	45.83	8.762*	35.16	13.07	0.7529	0.4658

Wharton-EFU Model, which performs almost as well as the naive models and substantially better than the other econometric models. Ranked for real and current residential structures over the period of forecast, the models appear in the following order.

Variables	Rank Order of Predictive Performance Over Forecast Period						
	Naive	Friend-Taub-man	Fromm	Liu	Klein	OBE	Whar-ton-EFU
Real residential structures	1	3	7	4	6	5	2
Current residential structures	1	3*	7	4*	6	5	2

Note: Asterisks indicate addition of exogenous deflators.

D. IMPORTS AND EXPORTS

Mean-squared errors for imports and exports based on reduced forms are presented in Table 3.4. In forecasting both real and current imports, the Fromm Model outperforms all other quarterly models. In forecasting real imports of crude materials and foodstuffs, all of the econometric models outperform the naive model, with the Wharton-

TABLE 3.4

Mean-Squared Errors Over Forecast Period for Imports and Exports

Variables	Econometric Models				Auto-regressive Schemes
	Fromm	Klein	OBE	Wharton-EFU	
Real total imports	0.5985	2.691	2.150	0.9181	1.013
Real imports of crude materials and foodstuffs	+	0.0121	0.0147	0.0104	0.0192
Real imports of semifinished and finished goods and services	+	2.707	2.158	0.8709	0.9033
Current total imports	0.6191	2.740	2.178	0.9921	1.245
Real total exports	+	5.427	+	1.382	3.996
Current total exports	+	2.927	+	127.2	3.318

EFU Model showing the best performance. The Wharton-EFU Model also surpasses all other quarterly models in forecasting real imports of semifinished and finished goods and services. In second place in forecasting this variable is the naive model. Rankings of the models for each category of imports are presented below.

Variables	Rank Order of Predictive Performance Over Forecast Period				
	Naive	Fromm	Klein	OBE	Wharton-EFU
Total imports	3	1	5	4	2
Real imports of crude materials and foodstuffs	4		2	3	1
Real imports of semifinished and finished goods and services	2		4	3	1

E. INVENTORIES, ORDERS, AND SHIPMENTS

Reduced-form mean-squared errors for inventories, orders, and shipments are agiven in Table 3.5. Considering these results for inventory investment first, we observe that the Wharton-EFU Model provides the best forecast for real, total, and nonfarm inventory investment. For real total inventory investment, the naive model is in second place, slightly outperforming the Liu Model. In forecasting real nonfarm inventory investment, the Liu Model moves to second place. The Wharton-EFU Model again provides the best forecasts for current nonfarm and total inventory investment. Both the Liu and Fromm Models outperform the naive models in forecasting current total and nonfarm inventory investment. A performance ranking of the models in predicting the classes of inventory investment over the forecast period follows.

The Liu Model explains the real stock of nonfarm business inventories, and in forecasting this variable, it registers a predictive performance superior to that of the naive model. Unfilled orders in manufacturing are explained in the Fromm, Klein, and OBE Models.

Variables	Rank Order of Predictive Performance Over Forecast Period							
	Naive	Friend-Taub-man	Fromm	Liu	Klein	OBE	Whar-ton-EFU	Gold-feld
Real nonfarm inventory investment	3	7	5	4	6	8	1	2
Real total inventory investment	4	7	5	3	6	8	1	2
Current inventory investment	4	7	3	2	6	8	1	5

For both unfilled and new orders in manufacturing, the naive model outperforms all others over the forecast period. However, the OBE outperforms the naive model in forecasting unfilled orders in durable manufacturing.

F. GNP DETERMINED ON DEMAND SIDE

Real gross national product determined on the expenditure side of national income and product accounts is forecast best by the Fromm Model, as shown in Table 3.6. Current *GNP*, on the other hand, is forecast best by the naive model. Rankings of the models in forecasting real and current *GNP* are shown here.

Variables	Rank Order of Predictive Performance Over Forecast Period							
	Naive	Friend-Taub-man	Fromm	Liu	Klein	OBE	Whar-ton-EFU	Gold-feld
Real *GNP*	2	7	1	4	6	5	8	3*
Current *GNP*	1	7*	3	6*	2	5	8	4

Note: Asterisks indicate addition of exogenous deflators.

G. GROSS PRODUCT ORIGINATING

Reduced-form mean-squared errors for gross output originating by various sectors are presented in Table 3.7. Starting with real private

TABLE 3.5

Mean-Squared Errors Over Forecast Period for Inventory Investment, Orders, and Shipments

Variables	Econometric Models							Autoregressive Schemes
	Friend-Taubman	Fromm	Liu	Klein	OBE	Wharton-EFU	Goldfeld	
Real nonfarm inventory investment	16.48	9.266	8.727	13.87	65.72	6.005	7.393	7.808
Real total inventory investment	16.48	9.266	8.727	13.87	65.72	6.005	7.393	9.056
Current nonfarm inventory investment	19.8	9.693	9.084	12.47	79.01	5.765	10.70	9.751
Current total inventory investment	19.8	9.693	9.084	12.47	79.01	5.765	10.70	10.18
Real stock of nonfarm business inventories	+	+	0.5391	+	+	+	+	0.7358
Real unfilled orders in manufacturing	+	1.822	+	+	+	+	+	0.167[a]
Real new orders in manufacturing	+	+	+	0.0107	3.832	+	+	0.232[a]
Real shipments of durable goods in manufacturing	+	+	+	+	1.5341	+	+	0.8285
Real unfilled orders of durable goods in manufacturing	+	+	+	0.483	1.411	+	+	1.974

[a] E-03.

TABLE 3.6

Mean-Squared Errors Over Forecast Period for GNP Determined on Demand Side

	Econometric Models							
Variables	Friend-Taub-man	Fromm	Liu	Klein	OBE	Whar-ton-EFU	Gold-feld	Auto-regressive Schemes
Real *GNP*	386.1	20.87	149.6	279.3	255.2	2380	127.6*	50.6
Current *GNP*	467.9*	136.8	259.0*	65.69	229.0	883.9	224.7	32.67

TABLE 3.7

Mean-Squared Errors Over Forecast Period for Gross Product Originating

	Econometric Models			
Variables	Klein	OBE	Whar-ton-EFU	Auto-regressive Schemes
Real private *GNP*	379.4	486.7	484.9	44.46
Real private *GNP* at full capacity	501.6	692.6	+	94.69
Real *GNP* originating in manufacturing	+	+	82.02	16.25
Real *GNP* originating in nonmanufacturing	+	+	229.3	27.40
Real *GNP* originating in residential structures	+	+	18.28	3.519
Real full capacity output in manufacturing	+	+	1860	37.43

GNP, we observe that the naive model forecasts this variable better than do the quarterly econometric models, of which the Klein Model performs best. A ranking for real private *GNP* can be enumerated as follows: (1) naive; (2) Klein; (3) Wharton-EFU; and (4) OBE. Real gross product is explained in the Wharton-EFU Model in the manufacturing, nonmanufacturing, and residential construction sectors. For all three variables, the naive models exhibit a superior predictive performance in the forecast period when compared with the Wharton-EFU Model.

H. CAPITAL CONSUMPTION ALLOWANCES

Reduced-form mean-squared errors for capital consumption allowances are given in Table 3.8. Current accounting depreciation allowances are forecast best by the Fromm Model. In order of performance on current accounting depreciation, the models rank as follows: (1) Fromm; (2) Liu; (3) naive; and (4) Wharton-EFU.

The Liu Model also explains constant-dollar accounting depreciation, but in forecasting this variable, it is outperformed by the naive model. In the Fromm Model, current corporate accounting depreciation is explained and is better forecast than by the naive model. Finally, although current accounting depreciation in manufacturing, nonmanufacturing, and nonfarm residential structures is determined in the Wharton-EFU Model, the naive models perform better in their forecast.

I. TAX AND TRANSFER ITEMS

Reduced-form mean-squared errors for tax and transfer items are given in Table 3.9. Considering these results in detail, we observe, first, that for indirect business tax and nontax liability, the naive model provides the best forecast for indirect taxes, the Klein Model being the best performing econometric model. Corporate profits tax liability is forecast best by the OBE Model. In second place is the naive model. The naive model also provides the best forecast for personal tax and nontax liability, according to Table 3.9. Of the econometric models, the OBE Model provides the best forecast of this variable. The rankings for these variables follow.

Variables	Rank Order of Predictive Performance Over Forecast Period				
	Naive	Fromm	Klein	OBE	Wharton-EFU
Business tax and nontax liability	1	3	2	5	4
Corporate profits tax liability	2	4	3	1	5
Personal tax liability	1	5	3	2	4

The Fromm Model explains both federal, and state and local, personal tax and nontax liability. For both variables, the naive models provide better forecasts than the Fromm Model.

State unemployment insurance benefits are determined in the Fromm and OBE Models. In forecasting this variable, the naive model outperforms both econometric models. The Fromm Model also determines two other transfer items: OASI and veterans' benefits, and relief payments and other transfers. The naive models forecast both of these variables better than the Fromm Model. Business and government transfer payments are explained in the Wharton-EFU Model. In forecasting this variable, the naive model outperforms the econometric model.

J. NONLABOR PERSONAL INCOME, CORPORATE PROFITS, RETAINED EARNINGS, AND DIVIDENDS

In Table 3.10 reduced-form mean-squared errors are presented for nonlabor personal income and several of its components. Beginning with nonlabor personal income, we observe that the naive model registers a predictive performance superior to that of the econometric models which explain this variable. Considering the predictive results for components of nonlabor personal income, we observe that the

TABLE 3.8

Mean-Squared Errors Over Forecast Period for Capital Consumption Allowances

	Econometric Models					
Variables	Fromm	Liu	Klein	OBE	Wharton-EFU	Autoregressive Schemes
Real accounting depreciation	+	0.7048	+	+	+	0.4572
Current accounting depreciation	0.3518	0.533	+	+	12.26	0.553
Real depreciation on nonfarm capital stock	+	+	24.23	0.0056	+	25.74
Current accounting depreciation in manufacturing	+	+	+	+	5.285	0.1108
Current accounting depreciation in nonmanufacturing	+	+	+	+	0.4382	0.1165
Current accounting depreciation nonfarm residential structures	+	+	+	+	0.5648	0.0041
Corporate accounting depreciation	0.2989	+	+	+	+	0.398

TABLE 3.9

Mean-Squared Errors Over the Forecast Period for Tax and Transfer Items

Variables	Econometric Models				Autoregressive Schemes
	Fromm	Klein	OBE	Wharton-EFU	
Indirect business tax and nontax liability	7.508	6.915	31.18	29.03	0.1493
Corporate profits tax liability	8.713	3.998	1.263	119.0	2.954
Personal tax and nontax liability	56.95	26.27	19.27	32.38	2.726
Federal personal tax and nontax liability	88.28	+	+	+	2.734
State and local personal tax and nontax liability	3.698	+	+	+	0.0430
State unemployment insurance benefits	0.414	+	3.377	+	0.1899
OASI and veterans' benefits	0.9103	+	+	+	0.909
Relief payments and other transfers	9.151	+	+	+	1.021
Business and government transfer payments	+	+	+	39.99	1.342

TABLE 3.10

Mean-Squared Errors Over Forecast Period for Nonlabor Personal Income, Corporate Profits, Retained Earnings, and Dividends

Variables	Econometric Models					Autoregressive Schemes
	Fromm	Liu	Klein	OBE	Wharton-EFU	
Nonlabor personal income	+	+	12.36	10.44	4.4867	1.131
Corporate profits before taxes, current dollars	9.802	176.7	110.7	62.93	662.1	16.08
Corporate profits before taxes, constant dollars	+	145.0	+	+	+	23.69
Corporate profits after taxes, current dollars	6.607	+	74.77	57.29	243.7	9.178
Retained earnings (including *IVA*): deflated by plant and equipment price deflator	+	+	112.9	+	+	5.379
Current retained earnings	6.019	149.9	12.36	54.97	225.6	18.94
Retained earnings in manufacturing	+	+	+	+	4.440	2.657
Real dividends	+	0.7603	+	+	+	0.2444
Current dividends	0.0603	0.9919	0.5448	0.1492	0.4793	0.1070

Fromm Model provides the best forecast of current-dollar corporate profits (before and after tax), retained earnings, and dividends. Retained earnings in manufacturing, which are explained in the Wharton-EFU Model, are forecast best by the naive model. Finally, real dividends in the Liu system and real retained earnings in the Klein are both forecast best by the naive models. Here we present the models ranked according to predictive performance for nonlabor personal income and its components over the forecast period.

Variables	Rank Order of Predictive Performance Over Forecast Period					
	Naive	Fromm	Liu	Klein	OBE	Wharton-EFU
Nonlabor personal income	1			4	3	2
Before-tax corporate profits	2	1	5	4	3	6
After-tax corporate profits	2	1		4	3	5
Dividends	2	1	6	5	3	4
Retained earnings	2	1		3	4	5

K. EMPLOYMENT, HOURS, AND WAGES

Table 3.11 presents reduced-form mean-squared errors over the forecast period for employment, hours, and wages. Considering these results, first, for the employment variables, we observe that for unemployed workers the naive model forecasts better than all of the econometric models. Among the latter, the Fromm Model performs best. The naive model also outperforms all other models in predicting total employment. Next comes the Wharton-EFU, which registers the best predictive performance of all the econometric models. Employed workers in manufacturing and nonmanufacturing are determined separately in the Wharton-EFU Model but are forecast better by the naive models. Over the forecast period, current wages and salaries are predicted best by the naive model. The best performing econometric model for this variable is the Wharton-EFU. Rankings of the quarterly models in forecasting these variables are given here.

Variables	Rank Order of Predictive Performance Over Forecast Period				
	Naive	Fromm	Klein	OBE	Wharton-EFU
Unemployment	1	2	5	4	3
Total employment	1		3	4	2
Current wages and salaries	1		3	4	2

Real private wages and salaries are explained in the Klein Model, and for this variable, the model registers a forecast superior to that of the naive model. Current wages and salaries are explained in the Wharton-EFU Model for both manufacturing and nonmanufacturing. In forecasting these variables, the naive models outperform the Wharton-EFU Model.

Average hours worked per week in the private sector are explained in the Klein and OBE Models, while average hours worked per week in manufacturing and nonmanufacturing are explained in the Wharton-EFU Model. In forecasting these variables, the naive outperform the econometric models.

The private hourly wage rate is determined in the Klein and OBE Models, the private hourly wage rates in manufacturing and nonmanufacturing being explained in the Wharton-EFU Model. Over the forecast period, the predictive performance of the naive models for these variables is again superior to that of the econometric models.

L. IMPLICIT PRICE DEFLATORS

Table 3.12 presents reduced-form mean-squared errors for implicit price deflators contained in the national income and product accounts. Beginning with the *GNP* price deflator, we observe that the Fromm Model registers the best predictive performance of the quarterly models. In second place is the naive model. In order of performance for the *GNP* deflator over the forecast period, the models are: (1) Fromm; (2) naive; (3) Wharton-EFU; (4) Liu; (5) OBE; and (6)

TABLE 3.11

Mean-Squared Errors Over Forecast Period for Employment, Hours, and Wages

	Econometric Models				
Variables	Fromm	Klein	OBE	Wharton-EFU	Auto-regressive Schemes
Unemployed workers	0.7093	33.79	11.68	6.319	0.02013
Total employment	+	6.716	9.250	3.370	0.5024
Employment in manufacturing	+	+	+	5.975	0.06141
Employment in nonmanufacturing	+	+	+	0.8938	0.05675
Total civilian labor force	+	10.78	0.4896	+	0.113
Total civilian wages and salaries, current	+	293.2	447.2	103.9	8.318
Private wages and salaries, current	+	293.2	447.2	103.9	7.724
Real private wages and salaries	+	9.766	+	+	1731
Wages and salaries in manufacturing	+	+	+	157.1	2.274
Wages and salaries in nonmanufacturing	+	+	+	23.40	1.414
Average hours worked per week	+	1.803	0.90[a]	+	0.1487
Average hours worked per week in manufacturing	+	+	+	0.001	0.426[b]
Average hours worked per week in nonmanufacturing	+	+	+	0.0002	0.3157[b]
Private average hourly wage rate	+	0.36[b]	0.0191	+	0.2534[a]
Private hourly wage rate for manufacturing	+	+	+	0.003	0.00171
Private hourly wage rate for nonmanufacturing	+	+	+	0.0043	0.00179

[a] E-05.
[b] E-04.

TABLE 3.12

Mean-Squared Errors Over Forecast Period for Implicit Price Deflators

	Econometric Models					
Variables	Fromm	Liu	Klein	OBE	Wharton-EFU	Auto-regressive Schemes
GNP price deflator	0.529E-5	0.4356E-4	0.0028	0.001209	0.1225E-4	0.6399E-5
Private *GNP* deflator	+	+	+	0.0001	0.035	0.1003
Total consumption deflator	+	+	+	0.0002	0.1296E-4	0.2013E-2
Consumer services deflator	+	+	0.0021	0.0010	+	0.6475E-5
Consumer nondurables deflator	+	+	0.0008	0.0002	+	0.1031E-4
Consumer durables deflator	+	+	0.0039	+	0.1999E-4	0.1976E-4
Consumer nonauto durables deflator	+	+	+	0.0001	0.2304E-4	0.2340E-4
Plant and equipment deflator	+	+	0.0130	0.0001	0.3249E-4	0.2072E-4
Nonfarm residential structures deflator	+	+	0.0004	0.0001	0.0001	0.4665E-4
Exports deflator	+	+	0.0007	+	0.0001	0.9165E-4

Klein. The private *GNP* deflator is forecast best by the naive model, which outperforms both the OBE and Wharton-EFU Models.

Mean-squared errors for the total consumption deflator are available in the OBE and Wharton-EFU Models. In forecasting this variable, the Wharton-EFU outperforms the OBE and the naive models. These last models provide the best forecasts of the consumer nondurables, durables, and services deflators. The implicit price deflator for consumer nonauto durables is explained by the OBE and Wharton-EFU Models, but is forecast best by the naive model. The naive models also furnish the best forecasts for the implicit price deflator corresponding to the following expenditure items: business investment in plant and equipment, nonfarm residential structures, and exports of goods and services.

M. MISCELLANEOUS ITEMS

Table 3.13 presents mean-squared errors for net interest paid by government and consumers, and *IVA*. The former variable is forecast best by the naive models; and the latter, by Wharton-EFU.

N. INTEREST RATES

Table 3.14 presents mean-squared errors for interest rates based on reduced-form forecasts. For the econometric models which explain the 4- to 6-month prime commercial paper rate, we observe that the

TABLE 3.13

Mean-Squared Errors Over Forecast Period for Miscellaneous Items

	Econometric Models		
Variables	Wharton-EFU	Fromm	Auto-regressive Schemes
Net interest paid by government and consumers	+	.0301	.0281
Inventory valuation adjustment	0.5697	+	1.153

TABLE 3.14

Mean-Squared Errors Over Forecast Period for Interest Rates

Variables	Econometric Models: Liu	Klein	OBE	Wharton-EFU	Gold-feld	Auto-regressive Schemes
4- to 6-month prime commercial paper rate	0.0754	0.0674	0.0682	0.0603	+	0.03159
Moody's AAA corporate bond rate	0.0231	0.0128	0.0128	0.0045	+	0.02735
Weighted average of yield on time deposits and savings shares	0.2785	+	+	+	+	0.004735
FHA mortgage yield	+	+	0.2107	+	+	0.02482
Commercial loan rate	+	+	+	+	1.433	0.01327
Treasury bill rate	+	+	+	+	4.177	0.03992
Intermediate rate on government bonds (3 to 5 years)	+	+	+	+	2.394	0.05626
Long-term government bond rate	+	+	+	+	1.241	0.01839

naive model provides the best forecast. In contrast, Moody's AAA corporate bond rate is best explained by the Klein and OBE Models. All of the econometric models forecast this variable better than does the naive. Rankings of the quarterly models in predicting these two interest rates over the forecast period are given below.

Variables	Rank Order of Predictive Performance Over Forecast Period			
	Naive	Liu	Klein and OBE	Wharton-EFU
4- to 6-month prime commercial paper rate	1	4	2	3
Moody's AAA corporate bond yield	4	3	2	1

The naive models outperform the Liu and OBE Models in forecasting the yield on time deposits and savings shares, and the FHA mortgage yield, respectively. The naive models outperform the Goldfeld Model in forecasting the following interest rates: the commercial loan rate, the Treasury bill rate, the intermediate government rate, and the long-term government rate.

O. FINANCIAL VARIABLES OTHER THAN INTEREST RATES

Reduced-form mean-squared errors for financial variables other than interest rates are given in Table 3.15. The forecast results strongly favor the naive models. Of the financial variables explained in the Liu Model, the naive models provide the best forecasts for all but real holdings of business demand and time deposits.

Current total consumer liquid assets are determined in the Liu, Klein, and OBE Models. The naive model registers the best performance in predicting this variable, as the ranking shows: (1) naive; (2) Liu; (3) OBE; and (4) Klein.

According to Table 3.15, the naive model outperforms the Goldfeld Model in forecasting: city member bank holdings of short-term

TABLE 3.15

Mean-Squared Errors Over Forecast Period for Financial Variables Other Than Interest Rates

Variables	Econometric Models				Auto-regressive Schemes
	Liu	Klein	OBE	Goldfeld	
Real holdings of business demand and time deposits	6.029	+	+	+	6.352
Current holdings of business demand and time deposits	5.967	+	+	+	3.233
Real holdings of consumer demand deposits	8.403	+	+	+	2.266
Current holdings of consumer demand deposits	7.906	+	+	+	5.205
Real holdings of consumer time deposits and savings shares	27.30	+	+	+	3.13
Current holdings of consumer time deposits and savings shares	19.88	+	+	+	5.122
Real total liquid assets of consumers	62.13	+	+	+	10.73
Current total liquid assets	44.76	12141	1376	+	9.872
Excess reserves of member banks	+	+	+	0.0033	0.0013
Excess reserves country member banks	+	+	+	0.0246	0.0020
Borrowings city member banks	+	+	+	0.3107	0.01311
Borrowings country member banks	+	+	+	0.0253	0.0006759

Short-term securities city member banks	+	+	+	2.469	1.436
Short-term securities country member banks	+	+	+	0.194	0.5298
Long-term securities city member banks	+	+	+	10.01	0.4656
Long-term securities country member banks	+	+	+	1.424	0.1492
Municipal securities city member banks	+	+	+	0.0903	0.5462
Municipal securities country member banks	+	+	+	0.0267	0.0746
Member bank commercial loans	+	+	+	5.481	2.584
Commercial loans city member banks	+	+	+	3.104	2.299
Commercial loans country member banks	+	+	+	0.3364	0.0294
Currency component of money supply	+	+	+	46.36	0.0808
Demand deposit component of money supply	+	+	+	1.612	1.355
Demand deposits city member banks	+	+	+	0.4817	0.375
Demand deposits country member banks	+	+	+	0.201	0.4103
Time deposits member banks	+	+	+	1.061	1.805
Time deposits city member banks	+	+	+	1.061	1.805
Time deposits country member banks	+	+	+	0.6777	0.0814
Potential deposits member banks	+	+	+	34.85	3.192
Change in class average reserve requirement city member bank deposits	+	+	+	0.490[a]	0.1607[b]
Change in class average reserve requirement country member banks	+	+	+	0.16[a]	0.1429[b]

[a] E-06.
[b] E-05.

securities; city and country member bank holdings of long-term securities; commercial loans, and total time and demand deposits; city and country member bank time deposits; and member bank potential deposits. The Goldfeld Model outperforms the naive model in forecasting: country member bank holdings of short-term securities; city and country municipal securities; country member bank time deposits; and changes in class average reserve requirements.

P. PERSONAL AND DISPOSABLE PERSONAL INCOME

In forecasting both of these variables, the naive models perform best of all, with the Fromm best of the econometric models (as shown in Table 3.16). Rankings of the models with respect to their performance in prediction over the forecast period are listed below.

Variables	Rank Order of Predictive Performance Over Forecast Period						
	Naive	Fromm	Liu	Klein	OBE	Wharton-EFU	Goldfeld
Personal income	1	2		4	5	3	
Disposable personal income	1	2	6	5	7	3	4

The naive model outperforms the Liu Model in forecasting real disposable income (*see* Table 3.16).

Q. SUMMARY

Beginning with the expenditures side of the income and product accounts, we observe that the Fromm Model provides the best forecast of real *GNP*, but the naive produces the best forecast of current *GNP*. The naive model forecasts total consumption better than all of the econometric models. However, for the durables component of consumption, several of the econometric models outperform it. The Fromm

TABLE 3.16

Mean-Squared Errors Over Forecast Period for Personal and Personal Disposable Income

Variables	Econometric Models						Auto-regressive Schemes
	Fromm	Liu	Klein	OBE	Whar-ton-EFU	Gold-feld	
Personal income	6.807	+	335.8	717.8	24.23	+	1.265
Personal disposable income	51.67	399.6	335.3	585.3	66.54	224.7	17.98
Real personal disposable income	+	306.9	+	+	+	+	10.74

Model provides the best forecast of plant and equipment expenditures. Residential structures are predicted best by the naive model. Finally, imports, exports, and inventory investment are forecast best by the Wharton-EFU Model.

The income side of the income and product accounts is dominated almost entirely by the mechanical forecasting schemes. The naive models provide the best forecasts of gross product originating, current wages and salaries, employment, and most tax and transfer items. However, the Fromm Model does outperform the naive models in forecasting several components of nonlabor personal income. The implicit price deflators and financial variables are consistently forecast best by the mechanical models.

We conclude that no one model predominates in forecasting the components of *GNP*. However, the mechanical models forecast about 30 per cent of the expenditure side—and about 90 per cent of the income side—of *GNP* accounts better than the econometric models. Further, almost 100 per cent of the price deflators and financial variables are forecast best by the naive models. Thus, as a means of forecasting economic variables, the naive models stand up well when confronted with econometric models.

4 UNIVARIATE TEST FOR STRUCTURAL CHANGE

UNIVARIATE test statistics for structural change in the reduced-form equations of each quarterly econometric model are found in Tables 4.1 through 4.14.

Considering these test results by sector, we observe from Table 4.1, for consumption expenditures, that structural change has occurred in most of the equations. Such change has occurred in all of the consumption equations in the Friend-Taubman, Liu, OBE, and Goldfeld Models. In the Fromm Model, the per capita services over per capita income equation shows no evidence of structural change. In both the consumer durables and nondurables equations in the Klein Model, the null hypothesis of no structural change is accepted. Finally, the null hypothesis of no structural change is accepted for the ratio of consumer nondurables and services to disposable income equation in the Wharton-EFU Model. A ranking of the models with respect to structural stability appears below.

Econometric Model	Number of Equations	Per Cent Showing Structural Change
Klein	3	33
Fromm	2	50
Wharton-EFU	3	66
Friend-Taubman	1	100
Liu	3	100
OBE	4	100
Goldfeld	2	100

Test statistics for business fixed investment are given in Table 4.2. Out of all the investment equations in the econometric models, only the current plant and equipment equation in the Fromm Model, and the current gross private domestic investment equation in the Goldfeld Model, are structurally stable. Ranked for structural stability, the models appear as follows.

Econometric Model	Number of Equations	Per Cent Showing Structural Change
Fromm	1	0
Goldfeld	1	0
Friend-Taubman	1	100
Liu	2	100
Klein	1	100
OBE	1	100
Wharton-EFU	2	100

Structural change statistics for investment in residential structures are presented in Table 4.3. The results here show that the null hypothesis of no structural change is accepted only for the real total residential structures equation in the Friend-Taubman Model. Also included in Table 4.3 are single-family housing starts in the OBE Model. For this equation, the null hypothesis of no structural change is rejected. A ranking of the models with respect to structural stability is listed here.

Econometric Model	Number of Equations	Per Cent Showing Structural Change
Friend-Taubman	1	100
Fromm	1	0
Liu	1	0
Klein	1	0
OBE	2	0
Wharton-EFU	1	0

Structural change statistics for imports and exports of goods and services appear in Table 4.4. Consideration of these results shows that the total current imports equation in the Fromm Model gives no evidence of structural change. The null hypothesis of no structural change is accepted for real imports of crude food and materials in the Klein, OBE, and Wharton-EFU Models. In the OBE Model, the semi-finished imports equation does not show evidence of structural change,

TABLE 4.1

Consumption Expenditures: Test for Structural Change
($\alpha = 0.05$ critical region)

Econometric Model	Variable	$\hat{\chi}^2_{20}(20)$	$\chi^2_{20}(20)$
Friend-Taubman	Real consumer expenditures	267.8	31.4
Fromm	Current per capita consumer expenditures	53.3	31.4
	Current per capita services over per capita disposable income	1.1	31.4
Liu	Real consumer durables	50.6	31.4
	Real consumer nondurables	192.8	31.4
	Real consumer services	38.1	31.4
Klein	Real consumer durables	27.7	31.4
	Real consumer nondurables	12.2	31.4
	Real consumer services	83.49	31.4
OBE	Real consumer automobile expenditures	63.0	31.4
	Real consumer nonauto durables	2,707.9	31.4
	Real consumer nondurables	177.0	31.4
	Real consumer services (excluding housing)	151.3	31.4
Wharton-EFU	Real consumer automobile expenditures	125.5	31.4
	Real consumer nonauto durables	273.6	31.4
	Ratio of real consumer nondurables and services to real disposable income	20.0	31.4
Goldfeld	Current consumer durables	41.5	31.4
	Current consumer nondurables and services	56.4	31.4

TABLE 4.2

Business Fixed Investment: Test for Structural Change
($\alpha = 0.05$ critical region)

Econometric Model	Variable	$\hat{\chi}^2_{20}(20)$	$\chi^2_{20}(20)$
Friend-Taubman	Real plant and equipment expenditures	138.1	31.4
Fromm	Current plant and equipment expenditures	19.6	31.4
Liu	Real nonresidential structures	360.8	31.4
	Real producers' durable equipment	101.3	31.4
Klein	Real plant and equipment expenditures	577.5	31.4
OBE	Real plant and equipment expenditures	105.2	31.4
Wharton-EFU	Real plant and equipment expenditures in manufacturing	98.7	31.4
	Real plant and equipment expenditures in nonmanufacturing	31.9	31.4
Goldfeld	Current gross private domestic investment	25.56	31.4

TABLE 4.3

Residential Structures and Housing Starts: Test for Structural Change
($\alpha = 0.05$ critical region)

Econometric Model	Variable	$\hat{\chi}^2_{20}(20)$	$\chi^2_{20}(20)$
Friend-Taubman	Real total residential structures	26.0	31.4
Fromm	Current nonfarm residential structures	516.9	31.4
Liu	Real total residential structures	512.6	31.4
Klein	Real nonfarm residential structures	280.7	31.4
OBE	Real nonfarm residential structures excluding additions and alterations	121.8	31.4
	Single-family housing starts	199.2	31.4
Wharton-EFU	Real nonfarm residential structures	39.9	31.4

TABLE 4.4

Imports and Exports: Test for Structural Change
($\alpha = 0.05$ critical region)

Econometric Model	Variable	$\hat{\chi}^2_{20}(20)$	$\chi^2_{20}(20)$
Klein	Real imports of crude food and materials	10.5	31.4
	Real imports of semifinished and finished goods and services	62.0	31.4
	Real exports	46.2	31.4
OBE	Real imports of crude food and materials	12.2	31.4
	Real imports of semifinished and finished goods and services	26.4	31.4
Wharton-EFU	Real imports of semifinished and finished goods and services	35.9	31.4
	Real imports of crude foodstuffs and materials	9.54	31.4
	Real exports	15.2	31.4
Fromm	Current imports of goods and services	23.8	31.4

while structural change occurs for this equation in the Klein and Wharton-EFU Models. Finally, the null hypothesis of no structural change for the exports equation is rejected in the Klein Model but accepted in the Wharton-EFU. Here is a ranking of the econometric models on the basis of structural stability.

Econometric Model	Number of Equations	Per Cent Showing Structural Change
Fromm	1	0
Wharton-EFU	3	33
Klein	3	66
OBE	2	100

In Table 4.5 are presented structural change statistics for inventories, orders, and shipments. Considering the results for inventory investment first, we observe that the null hypothesis of no structural change is accepted for the real total inventory investment equations in the Friend-Taubman, Klein, and OBE Models. Structural change is rejected in the real nonfarm inventory investment equations in the Fromm and Liu Models, and in the real manufacturing and nonmanufacturing inventory investment equation in the Wharton-EFU Model. Current inventory investment in the Goldfeld Model accepts the null hypothesis.

TABLE 4.5

Inventories, Orders, and Shipments: Test for Structural Change ($\alpha = 0.05$ critical region)

Econometric Model	Variable	$\hat{\chi}^2_{20}(20)$	$\chi^2_{20}(20)$
Friend-Taubman	Real total inventory investment	16.8	31.4
Fromm	Real nonfarm inventory investment	25.7	31.4
	Real change in unfilled orders	6.7	31.4
Liu	Real change in nonfarm inventory investment	27.0	31.4
Klein	Real total inventory investment	14.9	31.4
	Real total manufacturing new orders	41.2	31.4
	Real total manufacturing unfilled orders	92.9	31.4
OBE	Real total inventory investment	156.7	31.4
	Real shipments of manufacturing durable goods	24.2	31.4
	Real new orders manufacturing durable goods	16.5	31.4
	Real unfilled orders manufacturing durable goods	1.4	31.4
Goldfeld	Current inventory investment	25.9	31.4
Wharton-EFU	Real inventory investment in manufacturing	11.6	31.4
	Real inventory investment in nonmanufacturing	11.6	31.4
	Changes in unfilled orders manufacturing	5.596	31.4

The equation explaining the real change in unfilled orders is structurally stable between the forecast and fitted period; the real unfilled orders equation also shows no evidence of structural change. In the Klein Model, both the new and unfilled orders equations exhibit structural change. In the OBE Model, the null hypothesis of no structural change is accepted for the real new orders equation. A ranking based on structural stability is given here for inventories, orders, and shipments appearing in the quarterly econometric models.

Econometric Model	Number of Equations	Per Cent Showing Structural Change
Goldfeld	1	0
Friend-Taubman	1	0
Fromm	2	0
Liu	1	0
OBE	4	50
Klein	3	66

Considering the test for structural change for gross product originating, we observe from Table 4.6 that the reduced-form equations determining real private *GNP* in the Klein Model, and the log of private *GNP* in the OBE Model, both reject the null hypothesis of no structural change. In the Wharton-EFU Model, the equations for full capacity, actual output in manufacturing, and output of residential structures all show evidence of structural change.

Test results for structural change in the depreciation equations are presented in Table 4.7. From these results, we observe that with the exception of current depreciation in manufacturing in the Wharton-EFU Model, the null hypothesis of no structural change is rejected for all the depreciation equations in the econometric models.

Structural change statistics for tax and transfer items are given in Table 4.8. With the exception of corporate profits and tax equations, all of the tax and transfer equations show evidence of structural change. These results are not surprising, since the tax and transfer rates have changed at least three times between 1949 and 1965.

TABLE 4.6

Gross Product Originating: Test for Structural Change
($\alpha = 0.05$ critical region)

Econometric Model	Variable	$\hat{\chi}^2_{20}(20)$	$\chi^2_{20}(20)$
Klein	Real private gross national product	528.3	31.4
OBE	Log real private *GNP* at capacity less log of function of civilian labor force	85.0	31.4
Wharton-EFU	Real full capacity output originating in manufacturing	81.6	31.4
	Real output originating in manufacturing	69.6	31.4
	Real output originating in residential structures	601.2	31.4

TABLE 4.7

Capital Consumption Allowances: Test for Structural Change
($\alpha = 0.05$ critical region)

Econometric Model	Variable	$\hat{\chi}^2_{20}(20)$	$\chi^2_{20}(20)$
Fromm	Current accounting depreciation	58.3	31.4
	Corporate accounting depreciation	117.0	31.4
Klein	Real depreciation on nonfarm capital stock	65.9	31.4
Wharton-EFU	Current depreciation in manufacturing	632.5	31.4
	Current depreciation in nonmanufacturing	28.6	31.4
	Current depreciation in nonfarm residential construction	363.2	31.4

TABLE 4.8

Tax and Transfer Items: Test for Structural Change
($\alpha = 0.05$ critical region)

Econometric Model	Variable	$\hat{\chi}^2_{20}(20)$	$\chi^2_{20}(20)$
Fromm	Indirect business tax and nontax liability	143.5	31.4
	Personal contributions for social insurance	1,534	31.4
	Corporate profits tax liability	28.5	31.4
	Personal tax and nontax federal payments	203.8	31.4
	Tax and nontax payments to state and local governments	695.7	31.4
	OASI and veterans' benefits	182.1	31.4
	Unemployment benefits	56.2	31.4
	Relief payments and other transfer payments	118.8	31.4
Klein	Indirect business tax and nontax liability	57.8	31.4
	Corporate profits tax liability	7.7	31.4
	Personal tax and nontax liability	120.5	31.4
OBE	Personal tax and nontax payments	73.6	31.4
	Corporate profits tax liability	5.5	31.4
	Indirect business tax and nontax liability	587.6	31.4
	State unemployment insurance benefits	183.5	31.4
Wharton-EFU	Indirect business taxes and business transfer payments	590.2	31.4
	Corporate income taxes	136.2	31.4
	Government and business transfer payments	103.8	31.4
	Personal income taxes	59.2	31.4

Table 4.9 presents structural change statistics for corporate profits, dividends, retained earnings, and nonlabor income. From these results we observe that structural change is found in the corporate profits equations in the Liu, Klein, and OBE Models. The corporate profits equation in the Fromm Model shows no evidence of structural change. Structural change occurs for the dividends equations in the Liu and Wharton-EFU Models. The dividend equations in the Fromm and OBE Models show no evidence of structural change. The hypothesis of no structural change is rejected for the nonlabor personal income in the OBE Model. Finally, structural change occurs in the Wharton-

TABLE 4.9

Corporate Profits, Dividends, and Retained Earnings, and Nonlabor Income: Test for Structural Change
($\alpha = 0.05$ critical region)

Econometric Model	Variable	$\hat{\chi}^2_{20}(20)$	$\chi^2_{20}(20)$
Liu	Real before-tax corporate profits including *IVA*	465.6	31.4
	Real dividends	66.8	31.4
Fromm	Current before-tax corporate profits, excluding *IVA*	14.3	31.4
	Current dividends	0.197	31.4
Klein	Retained earnings (including *IVA*): deflated by plant and equipment price deflator	328.5	31.4
	Corporate profits before taxes, current dollars	95.6	31.4
OBE	Corporate profits before taxes, current dollars	166.7	31.4
	Current dividends	22.3	31.4
	Nonlabor personal income	554.0	31.4
Wharton-EFU	Business income of unincorporated enterprises	80.9	31.4
	Current dividends	39.3	31.4
	Retained earnings in manufacturing	20.95	31.4

TABLE 4.10

Employment, Hours, and Wages: Test for Structural Change
($\alpha = 0.05$ critical region)

Econometric Model	Variable	$\hat{\chi}^2_{20}(20)$	$\chi^2_{20}(20)$
Fromm	Unemployed workers	192.5	31.4
Klein	Real private wages and salaries	52.0	31.4
	Private average hourly wage rate	135.8	31.4
	Average hours worked per week	192.6	31.4
	Total civilian labor force	254.4	31.4
OBE	Ratio of labor force to population	28.2	31.4
	Private average hourly wage rate	2.0	31.4
	Average weekly hours index for private employees	0.56	31.4
	Ratio of total private man-hours to private *GNP* at full capacity	6.8	31.4
Wharton-EFU	Private hourly wage rate for manufacturing	8.0	31.4
	Private hourly wage rate for nonmanufacturing	31.9	31.4
	Man-hours worked in manufacturing	46.1	31.4
	Average hours worked per week in manufacturing	11.8	31.4
	Man-hours worked in nonmanufacturing	40.8	31.4
	Average hours worked per week in nonmanufacturing	8.0	31.4
	Ratio civilian labor force to civilian population	65.9	31.4

EFU equations determining retained earnings in manufacturing and business income of unincorporated enterprises.

The results of testing the employment, hours, and wages equations for structural change are presented in Table 4.10. Beginning with the Fromm Model, we observe that the null hypothesis showing no structural change is rejected for unemployed workers. In the Klein Model, structural change occurs in all of the equations in the employment, hours, and wages sector. In the OBE Model, as well, these equations all show evidence of structural change. In the Wharton-EFU Model, structural change occurs in the nonmanufacturing wage rate equation but not in the manufacturing wage equation. No structural change occurs in the equations determining average hours worked per week in manufacturing, but it does occur in the two man-hours equations. The null hypothesis is rejected for equations in the Wharton-EFU Model explaining the ratio of the civilian labor force to the civilian population.

Table 4.11 presents statistics testing the implicit price deflator equations for structural change. The null hypothesis of no structural change is accepted for the *GNP*, and rate of growth in *GNP*, equations in the Fromm and Liu Models. The equation determining the private *GNP* deflator is structurally stable in the OBE Model. In the Klein Model, structural change occurs in all of the price equations except the consumer durables and residential structures price deflators. In the OBE Model, on the other hand, only the equation explaining the consumer services deflator exhibits structural change. In the Wharton-EFU Model, all of the price equations are structurally stable between the fitted and forecast periods. A ranking of the models on the basis of structural stability is given here.

Econometric Model	Number of Equations	Per Cent Showing Structural Change
Klein	6	33
OBE	6	83
Liu	1	100
Fromm	1	100
Wharton-EFU	7	100

Test statistics for the miscellaneous economic variables are given in Table 4.12. From these results, we observe that the equations explaining net interest paid by government and consumers in the Fromm Model, and inventory valuation adjustment in the Wharton-EFU, do not show evidence of structural change. But structural change does occur for the equation determining the rent and interest component of national income in the Wharton-EFU Model.

TABLE 4.11

Implicit Price Deflators: Test for Structural Change
($\alpha = 0.05$ critical region)

Econometric Model	Variable	$\hat{\chi}^2_{20}(20)$	$\chi^2_{20}(20)$
Fromm	Implicit price deflator for gross national product	0.34	31.4
Liu	Implicit price deflator for gross national product	30.0	31.4
Klein	Implicit price deflator for:		
	Consumer durables	0.70	31.4
	Consumer nondurables	40.0	31.4
	Consumer services	210.0	31.4
	Plant and equipment expenditures	144.4	31.4
	Nonfarm residential structures	16.0	31.4
	Exports	46.7	31.4
OBE	Implicit Price deflator for:		
	Private *GNP* (excluding housing services)	108.7	31.4
	Consumer nonautomobile durables	9.13	31.4
	Consumer nondurables	11.1	31.4
	Consumer services (excluding housing)	40.0	31.4
	Nonfarm residential structures	5.0	31.4
	Nonresidential fixed investment	20.0	31.4
Wharton-EFU	Implicit price deflator for manufacturing (wholesale price index)	2.5	31.4
	Change in implicit price deflator for:		
	Nondurables and services	9.9	31.4
	Nonautomobile durables	4.6	31.4
	Automobiles	3.3	31.4
	Fixed business investment	6.5	31.4
	Nonfarm residential structures	20.0	31.4
	Exports	4.0	31.4

TABLE 4.12

Miscellaneous Variables: Test for Structural Change
($\alpha = 0.05$ critical region)

Econometric Model	Variable	$\hat{\chi}^2_{20}(20)$	$\chi^2_{20}(20)$
Fromm	Net interest paid by government and consumers	22.7	31.4
Wharton-EFU	Rent and interest component of national income, current	79.9	31.4

The results of testing the null hypothesis of no structural change for interest rates are given in Table 4.13. Considering these results, we observe for the Liu Model that all of the interest rate equations exhibit structural change. In contrast, in the Klein, OBE, and Wharton-EFU Models the interest rate equations determining the short and long rates are structurally stable between the fitted and forecast periods. In the Goldfeld Model, structural change occurs in all of the interest rate equations: the commercial loan rate, the spread between the long-term bond rate and the bill rate, and the difference between the intermediate rate and the bill rate. Econometric models rank as shown for interest rates on the basis of structural stability.

Econometric Model	Number of Equations	Per Cent Showing Structural Change
Klein	2	0
Wharton-EFU	2	0
OBE	3	66
Liu	3	100
Goldfeld	3	100

Table 4.14 presents statistics which test the null hypothesis of no structural change for financial variables other than interest rates. Considering these results, we observe that most of the noninterest-rate financial variables in the econometric models show evidence of

TABLE 4.13

Interest Rates: Test for Structural Change
($\alpha = 0.05$ critical region)

Econometric Model	Variable	$\hat{\chi}^2_{20}(20)$	$\chi^2_{20}(20)$
Liu	Moody's AAA corporate bond rate	59.2	31.4
	Short-term rate (4- to 6-month prime commercial paper)	35.6	31.4
	Interest rate on time deposits and savings shares	185.7	31.4
Klein	Average yield, corporate bonds, per cent	17.4	31.4
	Average yield, 90-day commercial paper	28.4	31.4
OBE	Moody's AAA corporate bond rate	17.4	31.4
	Rate on 4- to 6-month prime commercial paper, per cent	28.4	31.4
	Per cent yield, secondary market, FHA insured new homes	261.7	31.4
Wharton-EFU	Average yield on 4- to 6-month prime commercial paper, per cent	24.6	31.4
	Moody's AAA corporate bond rate	5.1	31.4
	Commercial loan rate	200.6	31.4
	Long-term bond rate less treasury bill rate	76.8	31.4
	Intermediate government rate less treasury bill rate	92.8	31.4

structural change between the fitted and forecast periods. In the Liu Model, the equations explaining real business liquid assets, real consumer currency and demand deposits, and consumer holdings of time deposits and savings shares, all show evidence of structural change. In the Klein Model, structural change occurs in the equation determining the ratio of consumer liquid assets to *GNP*; it also occurs in the equation in the OBE Model explaining consumer liquid assets. Finally, out of all the noninterest-rate financial equations, only three are found to be structurally stable: excess reserves at city member banks, short-term securities, and municipals held by country member banks.

TABLE 4.14

Financial Variables Other Than Interest Rates:
Test for Structural Change
(α = 0.05 critical region)

Econometric Model	Variable	$\hat{\chi}^2_{20}(20)$	$\chi^2_{20}(20)$
Liu	Real business liquid assets	186.4	31.4
	Real personal holdings of currency and demand deposits	241.9	31.4
	Real personal holdings of time deposits and savings shares	449.5	31.4
Klein	End-of-quarter cash balance divided by gross national product	116.9	31.4
OBE	End-of-quarter liquid assets of households	595.6	31.4
Goldfeld	Excess reserves, city member banks	15.7	31.4
	Excess reserves, country member banks	76.9	31.4
	Borrowings, city member banks	42.9	31.4
	Borrowings, country member banks	59.5	31.4
	Short-term securities, city member banks	43.6	31.4
	Short-term securities, country member banks	21.0	31.4
	Long-term securities, city member banks	119.7	31.4
	Long-term securities, country member banks	70.4	31.4
	Municipals, city member banks	79.2	31.4
	Municipals, country member banks	4.7	31.4
	Currency component of money supply	29,530.3	31.4
	Demand deposits	75.1	31.4
	Time deposits	85.0	31.4
	Commercial loans, city and country member banks	42.6	31.4

5 SUMMARY AND CONCLUSIONS

THE predictive performances of the quarterly econometric models—for selected variables and over both fitted and forecast periods—are summarized in Table 5.1. Comparing the results reveals that the Fromm Model provides the best prediction, on the average, for real *GNP* in both fitted and forecast periods. The naive model outpredicts the econometric models for real consumer expenditures in both periods, and for current consumer expenditures during the fitted period. For current consumer expenditures in the forecast period, however, the OBE Model predicts best. For the components of consumer expenditures, the Liu Model registers the best prediction for real consumer durables in both fitted and forecast periods, while for current consumer expenditures, the naive model performs best during the fitted period—the Klein excelling it during the forecast period. The naive model gives the best prediction for both real and current nondurables and services during fitted and forecast periods.

The Liu Model predicts plant and equipment expenditures better than the other models during the period of fit, but the Fromm Model is superior to them in predicting this variable during the forecast period. The Wharton-EFU Model provides the best prediction for residential structures in the fitted period; and the naive model, in the forecast period. Total expenditures on imports are predicted best in the fitted period by the naive model, but in the forecast period the Wharton-EFU Model is superior. Best predictions for the two components of imports are turned in by the naive models during the fitted period, and by the Wharton-EFU Model during the forecast period.

For both periods, econometric models outperform the naive models in predicting inventory investment: the Liu Model in the fitted period; and the Wharton-EFU in the forecast period. For *GNP* originating in both periods, however, the naive models register the best predictive performance.

For capital consumption allowances, the naive model predicts best over the fitted period, while the Fromm performs best over the forecast period.

In the forecast period, corporate profits, dividends, and retained earnings are predicted best by the naive models. In the fitted period,

TABLE 5.1

Predictive Performance of Quarterly Models for Selected Econometric Variables

Variable	Best-Performing Model: Fitted Period	Best-Performing Model: Forecast Period
Real *GNP* determined on demand side	Fromm	Fromm
Current *GNP* determined on demand side	Fromm	Naive
Real consumer expenditures	Naive	Naive
Current consumer expenditures	Naive	OBE
Real consumer durables	Liu	Liu
Current consumer durables	Naive	Klein
Consumer nondurables and services, current and real	Naive	Naive
Plant and equipment expenditures	Liu	Fromm
Residential structures	Wharton-EFU	Naive
Total imports	Naive	Fromm
Real imports of crude materials and foodstuffs	Naive	Wharton-EFU
Real imports of semifinished and finished goods and services	Naive	Wharton-EFU
Inventory investment	Liu	Wharton-EFU
Gross product originating—private *GNP*	Naive	Naive
Before-tax corporate profits	Liu	Fromm
Retained earnings	Naive	Fromm
Dividends	Naive	Naive
Unemployment	Fromm	Naive
Total employment	Naive	Naive
Wages and salaries	Naive	Naive
Indirect business taxes	Naive	Naive
Corporate profits taxes	Naive	OBE
Personal taxes	Klein	Naive
GNP price deflator	Naive	Fromm
Plant and equipment deflator	Naive	Naive
Nonfarm residential structures deflator	Naive	Naive
Exports deflator	Naive	Naive
4- to 6-month prime commercial paper rate	Liu	Naive
Moody's AAA corporate bond rate	Liu	Wharton-EFU
Consumer liquid assets	Naive	Naive
Treasury bill rate	Naive	Naive
Other interest rates (Goldfeld)	Naive	Naive
Capital consumption allowances	Naive	Fromm

corporate profits are best predicted by the Liu Model; retained earnings and dividends best by the naive models.

Best predictive performances are registered by the naive models for total labor force and employment over the forecast period; for total employment in the fitted period; and for wages and salaries in both fitted and forecast periods. Over the fitted period, best performance for unemployed workers is turned in by the Fromm Model.

The naive prediction is superior for the *GNP* deflator over the fitted period, but the Fromm Model excels during the forecast period. The econometric models are inferior to the naive models in predicting all other price deflators in both the fitted and forecast periods.

All of the interest rate variables and consumer liquid assets are predicted best by the naive models, except for Moody's corporate AAA bond rate. For this variable, the econometric models are best: the Liu Model in the fitted period, and the Wharton-EFU Model in the forecast period.

Conclusions drawn from these results are, first, that no one econometric model surpasses its counterparts in predicting the components of the national income accounts. However, for three of the broad expenditure aggregates—real *GNP*, plant and equipment expenditures, and imports—the Fromm Model's predictive performance is better than that of all other quarterly models. Except for current consumer expenditures and consumer durables expenditures, the naive models outperform the econometric models in predicting the components of consumer expenditures. Second, the income side of the national income and product accounts is almost completely dominated by the naive models in both the fitted and forecast periods. Finally, in almost all cases, the naive models outperform the econometric models in predicting implicit price deflators, interest rates, and other financial variables.

We can conclude from the univariate tests of the models for structural change, that except the inventory investment equations, almost all of the structural equations on the demand side of the *GNP* accounts show evidence of such change. Exceptions—other than the inventory equations—are real consumer durables in the Klein Model; per capita services and plant and equipment expenditures in the Fromm Model.

Second, most of the employment and earnings equations show evidence of structural change. Exceptions here are a few equations explaining average hours worked per week and the private annual wage rate. Third, all but two of the tax equations show structural change. This, of course, may happen every time tax rates change—and they have changed several times during the postwar period.[10] Fourth, most of the equations determining the price deflators show structural stability over time. Exceptions here are the nondurables, services, and investment deflator equations in the Klein Model. Finally, almost all the interest rate and other financial equations in the econometric models show evidence of structural change.

If the reduced-form equations in which no structural change occurs are matched with the accuracy of predictions based on these equations, it appears that for several reasons there is no clear-cut relationship between the univariate test for structural change of a given reduced-form equation and its ability to predict economic variables. First, a univariate test ignores the interdependence between equations. Thus, a multivariate test for structural change, more powerful than the univariate test, could easily be devised. Second, the coefficients in any given structural equation may so change over time that they provide better reduced-form forecasts than would have been made without structural change. In this case, structural change is desirable, since econometric forecasts are improved. Third, the test for structural change may have failed in some cases; i.e., some Type 1 errors were made.

A number of general conclusions may be drawn from the research undertaken here. First, no single quarterly econometric model included in this study is overwhelmingly superior to all of the other quarterly

[10] Two things should be pointed out concerning the assumption of constant tax rates. First, in performing actual ex ante forecasting, the model-builders would most likely have allowed for changes in tax rates. Second, we assume that effective, rather than statutory, tax rates remain constant. Although statutory tax rates have changed twice over the period of this study, effective rates have remained relatively constant. For example, the corporate effective tax rate ranged from about .41 to .51 over the period 1949 through 1965, with an average of about .46. The personal effective tax rate over this period ranged from approximately 11.3 to 15.4, with a mean of 12.5 billion. The largest variations came during the two statutory tax cuts, which occurred in 1954 and 1964. The range of values of the effective tax rates appears to be small enough so as not to distort significantly the results or the conclusions of this study.

models in predicting the components of the national income and product accounts.[11] Second, the econometric models are not, in general, superior to purely mechanical methods of forecasting. However, there are modules of the econometric models which are definitely superior to purely mechanical models. Third, the econometric models are, in general, structurally unstable.

A. CURRENT AND PAST RESEARCH COMPARED

The study undertaken here parallels that done many years ago by Carl Christ [6]. The general conclusions of the Christ project and the current study are basically the same: existing econometric models show evidence of structural change, and are not, in general, superior to purely mechanical models in predicting the components of the national income and product accounts. The results from both studies imply that econometric model-building has not been highly successful. It is as true now as it was at the time of Christ's study that mechanical forecasting models can be constructed which predict economic variables about as well as econometric models.

The research presented here has two possible intepretations. First, it can be considered simply as determining those models which serve as the best predictive devices. Second, the evaluation of the predictive performance of the alternative econometric models can be thought of as an experiment designed to test the underlying theory. Most of the text here has been devoted to comparing the predictive performance of alternative models. Further study and analysis are required to evaluate the underlying theory.

B. THE DIRECTION OF FUTURE RESEARCH

Since econometric models generally forecast no better than autoregressive schemes—and since they are structurally unstable—the information contained in the structures of these models has little pre-

[11] We should mention that revisions have occurred in two of the econometric models included in this study—the OBE and the Wharton-EFU. In addition, several other econometric models have become available since this study was undertaken. The most noteworthy of these models are the Brookings-SSRC, the Michigan Quarterly, and the FRB-MIT.

dictive value. There is a natural question to ask on the basis of these results: "How can forecasting performance be improved?" One way that we might improve forecasting performance is by combining the instrumental variables from the unconstrained reduced forms of the econometric models with the auto-regressive schemes. This research remains to be completed.

BIBLIOGRAPHY

1. Adams, F. Gerard, and Evans, Michael K., *Econometric Forecasting with the Wharton-EFU Model, the Five Year Forecasting Record.* Discussion Paper No. 25, Wharton School of Finance, University of Pennsylvania, 1967.
2. Amemiya, Takeshi, "On the Use of Principal Components of Independent Variables in Two-Stage Least-Squares Estimation," *International Economic Review,* September, 1966.
3. Ando, Albert, and Goldfeld, Stephen M., "An Econometric Model for Evaluation of Stabilization Policies," in Albert Ando, E. Cary Brown, and Ann F. Friedlaender (eds.), *Studies in Economic Stabilization.* Washington, D.C., The Brookings Institution, National Committee on Government Finance, 1968.
4. Bentzel. R., and Hansen, B., "On Recursiveness and Interdependency in Econometric Models," *Review of Economic Studies,* Vol. 22, No. 3, 1954–55.
5. Chow, Gregory C., "Tests for Equality Between Sets of Coefficients in Two Linear Regressions," *Econometrica,* July, 1960.
6. Christ, Carl F., "A Test of an Econometric Model of the United States, 1921–1947," in *Conference on Business Cycles.* New York, National Bureau of Economic Research, 1951.
7. ———, "Aggregate Econometric Models," in R. A. Gordon and L. R. Klein (eds.), *Readings in Business Cycles,* Vol. X. Homewood, Ill., Richard D. Irwin, Inc., 1965.
8. de Leeuw, Frank, and Gramlich, Edward, "The Federal Reserve–MIT Econometric Model," in *The Economic Outlook.* Ann Arbor, University of Michigan, 1968.

9. ———, "The FRB-MIT Econometric Model," *Federal Reserve Bulletin,* January, 1968.
10. Duesenberry, James S., et al. (eds.), *The Brookings Econometric Model of the United States.* Chicago, Rand McNally & Company, 1965.
11. ———, Eckstein, Otto, and Fromm, Gary, "A Simulation of the United States Economy in Recession," *Econometrica,* October, 1960.
12. Durbin, J., "Testing for Serial Correlation in Systems of Simultaneous Regression Equations," *Biometrika,* December, 1957.
13. ———, and Watson, G. S., "Testing for Serial Correlation in Least Squares Regression, I," *Biometrika,* Vol. 37, 1950.
14. ———, "Testing for Serial Correlation in Least Squares Regression, II," *Biometrika,* Vol. 38, 1951.
15. Einzig, Robert S., *Forecasting the Economic Outlook for Investment Purposes.* San Francisco, Mitchell, Hutchins & Company, 1966.
16. Evans, Michael K., *Multiplier Analysis of a Postwar Quarterly U.S. Model and a Comparison with Several Other Models.* Discussion Paper No. 11, Wharton School of Finance, University of Pennsylvania, 1963.
17. ———, and Klein, L. R., *The Wharton Econometric Forecasting Model* (2nd, enlarged ed.). Philadelphia, University of Pennsylvania, 1968.
18. Fisher, Franklin M., "Identifiability Criteria in Nonlinear Systems," *Econometrica,* October, 1961.
19. ———, "The Choice of Instrumental Variables in the Estimation of Economy-Wide Econometric Models," *International Economic Review,* September, 1965.
20. ———, *The Identification Problem in Econometrics.* New York, McGraw-Hill Book Company, 1966.
21. ———, "Uncorrelated Disturbances and Identifiability Criteria," *International Economic Review,* May, 1963.
22. Friedman, Milton, Review of article on business cycles in the United States of America, 1919–32, *American Economic Review,* September, 1940.
23. ———, Comment on "A Test of an Econometric Model for the

United States, 1921–1947," *Conference on Business Cycles*. New York, National Bureau of Economic Research, 1951.

24. Friend, Irwin, and Jones, Robert C., "Short-Run Forecasting Models Incorporating Anticipatory Data," in *Models of Income Determination*. New York, National Bureau of Economic Research, 1964.
25. ———, and Taubman, Paul, "A Short Term Forecasting Model," *The Review of Economics and Statistics*, August, 1964.
26. Fromm, Gary, "Inventories, Business Cycles, and Economic Stabilization," in *Inventory Fluctuations and Economic Stabilization*, Part III, Joint Economic Committee, 87 Congress, first session. Washington, D.C., Government Printing Office, 1961.
27. ———, "Forecasting, Policy Simulation, and Structural Analysis: Some Comparative Results of Alternative Models." Proceedings of the Business and Economic Statistics Section, American Statistical Association, December, 1964.
28. ———, and Taubman, Paul, *Policy Simulations with an Econometric Model*. Washington, D.C., The Brookings Institution, 1968.
29. Goldberger, Arthur S., *Econometric Theory*. New York, John Wiley and Sons, Inc., 1964.
30. ———, *Impact Multipliers and Dynamic Properties of the Klein-Goldberger Model*. Amsterdam, North-Holland Publishing Company, 1959.
31. Goldfeld, S. M., *Commercial Bank Behavior and Economic Activity*. Amsterdam, North-Holland Publishing Company, 1966.
32. Goldsmith, Raymond W., *The National Wealth of the United States in the Postwar Period*. Princeton, Princeton University Press, 1962.
33. Hildebrand, F. B., *Introduction to Numerical Analysis*. New York, McGraw-Hill Book Company, 1956.
34. Holt, Charles C.; Shirey, Robert; Steward, Donald; Schrank, William E.; Palit, Dawn; Midler, Joseph L.; and Stroud, Arthur H., *Program SIMULATE, AB, A User's and Programmer's Manual*. Madison, Social Systems Research Institute, The University of Wisconsin, 1965.
35. Hood, William C., and Koopmans, Tjalling C., *Studies in Econometric Method*. New York, John Wiley & Sons, Inc., 1953.

36. Jorgenson, Dale W., "Minimum Variance, Linear Unbiased Seasonal Adjustment of Economic Time Series," *Journal of the American Statistical Association,* September, 1964.
37. ———, Hunter, Jerald, and Nadiri, M. I., "A Comparison of Alternative Econometric Models of Quarterly Investment Behavior," *Econometrica,* March, 1970.
38. ———, "The Predictive Performance of Econometric Models of Quarterly Investment Behavior," *Econometrica,* March, 1970.
39. ———, and Siebert, C. D., "A Comparison of Alternative Theories of Corporate Investment Behavior," *American Economic Review,* September, 1968.
40. ———, "Optimal Capital Accumulation and Corporate Investment Behavior," *Journal of Political Economy,* December, 1968.
41. Klein, L. R., *Economic Fluctuations in the United States, 1921–1941,* Cowles Commission Monograph 11. New York, John Wiley and Sons, Inc., 1950.
42. ———, "A Postwar Quarterly Model: Description and Applications," in *Models of Income Determination,* Studies in Income and Wealth, Vol. 28. Princeton, Princeton University Press for the NBER, 1964.
43. ———, "The Social Science Research Council Econometric Model of the United States," *Econometric Analysis for National Economic Planning,* Colston Papers, No. 16, P. E. Hart, G. Mills, and J. K. Whitaker (eds.). London, Butterworths, 1964.
44. ———, Unpublished Wharton School Model. Wharton School of Finance, University of Pennsylvania.
45. ———, and Goldberger, Arthur S., *An Econometric Model of the United States, 1929–1952.* Amsterdam, North-Holland Publishing Company, 1954.
46. Kloek, T., and Mennes, L. B. M., "Simultaneous Equations Estimations Based on Principal Components of Predetermined Variables," *Econometrica,* January, 1960.
47. Liebenberg, Maurice, Hirsh, Albert A., and Popkin, Joel, "A Quarterly Econometric Model of the United States: A Progress Report," *Survey of Current Business,* May, 1966.
48. Liu, T. C., "An Exploratory Quarterly Econometric Model of Effective Demand in the Postwar U.S. Economy," *Econometrica,* July, 1963.

49. Malinvaud, E., *Statistical Methods of Econometrics.* Chicago, Rand McNally & Company, 1966.
50. Nerlove, Marc A., "A Tabular Survey of Macro-Econometric Models," *International Economic Review,* May, 1966.
51. Rasche, Robert H., and Shapiro, Harold T., "The FRB-MIT Econometric Model: Its Special Features," *AER Papers and Proceedings,* May, 1968.
52. Scheffé, Henry, *The Analysis of Variance.* New York, John Wiley and Sons, Inc., 1959.
53. Stekler, H. O., "Forecasting with Econometric Models: An Evaluation," *Econometrica,* October, 1968.
54. ———, "Data Revisions and Economic Forecasting," *Journal of American Statistical Association,* June, 1967.
55. Strotz, Robert H., "Interdependence as a Specification Error," *Econometrica,* April, 1960.
56. ———, and Wold, H. O. A., "Recursive vs. Nonrecursive Systems: An Attempt at Analysis," *Econometrica,* April, 1960.
57. Suits, Daniel B., "Forecasting and Analysis with an Econometric Model," in R. A. Gordon and L. R. Klein (eds.), *Readings in Business Cycles,* Vol. X. Homewood, Ill., Richard D. Irwin, Inc., 1965.
58. Theil, H., *Economic Forecasts and Policy.* Amsterdam, North-Holland Publishing Company, 1958.
59. ———, *Applied Economic Forecasting.* Amsterdam, North-Holland Publishing Company, 1960.
60. Tinbergen, Jan, *An Econometric Approach to Business Cycle Problems.* Paris, Hermann et cie, 1937.
61. ———, *Statistical Testing of Business Cycle Theories, I: A Method and Its Application to Investment Activity.* Geneva, League of Nations, Economic Intelligence Service, 1939.
62. ———, *Statistical Testing of Business Cycle Theories, II: Business Cycles in the United States of America, 1919–1932.* Geneva, League of Nations, Economic Intelligence Service, 1939.
63. Wold, H. O. A., "A Generalization of Causal Chain Models," *Econometrica,* April, 1960.
64. ———, "Causality and Econometrics," *Econometrica,* April, 1954.
65. Zarnowitz, Victor, *An Appraisal of Short-Term Economic Fore-*

casts. New York, National Bureau of Economic Research, 1967.
66. *Flow of Funds Accounts, 1945–62: 1963 Supplement*. Washington, D.C., Board of Governors of the Federal Reserve System, 1963.
67. *The National Income and Product Accounts of the United States, 1929–1965*, supplement to *Survey of Current Business*. Washington, D.C., Office of Business Economics, U.S. Department of Commerce, 1966.
68. *Survey of Current Business*. Washington, D.C., Office of Business Economics, U.S. Department of Commerce (*various issues*).

DISCUSSION

STEPHEN M. GOLDFELD

PRINCETON UNIVERSITY

Mr. Cooper has produced an ambitious paper in which he attempts to run a fair prediction contest between seven quarterly econometric models. The models are put on more of an equal footing by reestimating them for the same sample period – an effort clearly prodigious in scope. The "rabbit" in this prediction race is a mechanical auto-regressive scheme which has a nasty tendency of outdistancing virtually all of the competitors in the race – sometimes monumentally. In short, Mr. Cooper is driven to the conclusion that "econometric models are not, in general, superior to mechanical methods of forecasting." While at first blush, the general findings of this paper may be disheartening to model-builders, I believe that more careful thought leads one to temper Mr. Cooper's findings significantly. Let me turn directly to some of the relevant issues.

NATURE OF THE GAME

As indicated above, Cooper finds that an auto-regressive scheme for each endogenous variable generally outforecasts the econometric models under examination. It is important, however, to emphasize that in each instance Cooper is comparing one-period forecasts. In other words, for each quarter he assumes as known the lagged values of all endogenous variables. Given that realistic forecasting situations rarely involve only one-period forecasts, his conclusions are not terribly disturbing, nor as interesting as they might be.[1] More particularly, while continual re-initialization helps an econometric model, in that the root mean-square errors of one-period forecasts are significantly less than corresponding errors over longer forecast periods,[2] it is clear that this procedure helps an auto-regressive scheme even more. As is well known, such schemes can deteriorate badly if used to forecast more than one period ahead. For example, using a second order auto-regressive equation for real *GNP* yields the following average absolute errors for one-period forecasts through six-period forecasts: 3.60, 6.93, 9.55, 11.39, 13.17, 14.43. The corresponding results for the OBE Model are as follows: 2.35, 3.58, 4.47, 4.92, 5.42, 5.82. In other words, the auto-regressive equation produces a six-period forecast error which is four times the one-period error, while the OBE Model yields a six-period error which is only 2.5 times as large as the one-period error.[3] In summary, part of the reason for Cooper's pessimistic findings stems from the limited nature of the comparison he has chosen to make.

However, even if one wished to play the game according to Cooper's rules, one additional modification would seem desirable. To be specific, Cooper took no account, either in estimation or in forecasting, of the serial properties of the error terms. Several papers at this Conference have reported significant improvement in forecasting accuracy – *especially* for one-period forecasts – when account is taken of the most recent residuals. For example, the OBE evidence indicates that the one-period forecast error can be reduced by about 25 per cent. This suggests that even within the rather restrictive context of

[1] This is not meant to suggest that existing econometric models have produced satisfactory forecasts, but that is a separate matter.

[2] For example, see the evidence presented in L. R. Klein, *An Essay on the Theory of Economic Prediction* (Helsinki, 1968), and G. R. Green, "Short- and Long-term Simulations with the OBE Econometric Model" (*this conference*).

[3] The results cited in the text are for the period 1955-I to 1966-IV and are in billions of dollars. They are based on the auto-regression equation found in Green, *op. cit.* The author kindly provided me with both sets of numbers.

Cooper's comparisons, proper handling of the models would have reduced the discrepancies between the models and the naive standard.[4]

MODEL COMPARISONS

I turn now from contrasting the naive model with the econometric models to some remarks about the relative performance of the various models. As we shall see, however, these remarks have implications for the absolute performance of the models, as well.

As indicated above, Cooper reestimated all of the models over one sample period: 1949-I to 1960-IV. While this was done to provide a more equal treatment of the different models, the way in which it was done may, in some respects, muddy the waters. There are at least two reasons for this. First, with the exception of the OBE and Wharton efforts, all of the models were originally fitted to data available prior to 1966. In 1966, there was, however, a major historical revision of U.S. National Income statistics. From various econometric efforts, there is ample evidence that many of the relationships fitted to prerevision data are no longer satisfactory when reestimated.[5] This is an interesting finding in its own right and may be worthy of further study. Nevertheless, the evidence calls into question the procedure of mechanically reestimating the models for prediction purposes, especially since it takes only one "bad" expenditure function to deteriorate the forecasting ability of a model significantly. (This mechanical procedure is hardly the sort of thing that the user of an econometric model is likely to do.)

A related difficulty stems from the fact that some of the models were originally estimated excluding the Korean War period. Typically, this was done in the belief that structural shifts had taken place. Indeed, if the sample originally included the early 1950's, the ubiquitous "shift dummy" is often in evidence. This suggests that mechanically including

[4] Also see M. K. Evans, Y. Haitovsky, and G. I. Treyz, "An Analysis of the Forecasting Properties of U.S. Econometric Models" (*this conference*). Of course, one should make the same corrections for the auto-regressive schemes. However, the basic ingredient in the technique is the serial correlation coefficient. Since this is biased toward zero by the auto-regressive schemes, correcting both models and naive standards should improve the performance of the former relative to the latter.

[5] It is not exactly clear what weight to place on this observation, since the two best-performing models were fit to prerevision data.

the Korean period, if it was not originally used, may introduce unwanted differences into the model comparisons.[6]

Aside from problems of this sort, Cooper experienced some other difficulties in reestimating the models. Given the scope of his task, and the occasional imprecisions of model-builders in describing such things as the measurement of their variables, this is certainly understandable. A few examples may help. In the translation of my own model, for instance, Cooper reestimates all of my expenditure functions in current dollars, although they were originally estimated in real terms. Similarly, the investment category used does not correspond to the originally specified category.[7] I am, of course, not as familiar with the details of the other models but some similar questions come to mind. For example, the OBE Model is overdetermined in the sense that it endogenously determines the statistical discrepancy; this appears to have been ignored. It is also unclear how Cooper treated endogenous revenue equations (e.g., in the Wharton Model) which have tax-rate parameters that appear as fixed coefficients.

Aside from these types of difficulty, Cooper explicitly changed the character of some of the models. For instance, the Friend-Taubman Model, which originally was a semiannual first-difference model, was turned into a quarterly model estimated in level form with trend terms! My own model was originally estimated with data unadjusted for seasonal variation but this format was not preserved. While these types of change are an attempt to put the models on a comparable basis, I believe that they point up the fact that the models used are not always fully consistent with their original versions. This "ownership" difficulty is compounded by the fact that a number of these models (e.g., OBE, Wharton) have undergone substantial changes over time.

Passing on from the above types of difficulty, there is a more fundamental problem (which Cooper is well aware of) in comparing the ex post forecasting performance of different econometric models. In

[6] It may be noted that if one is interested in forecasting under conditions of relatively low unemployment rates, it could be a mistake to ignore the Korean period in estimation, for it provides, aside from recent history, the only other observations in that range.

[7] See S. M. Goldfeld, *Commercial Bank Behavior and Economic Activity* (Amsterdam, North-Holland Publishing Company, 1966), pp. 165–66, for the statement about real vs. current dollars. Cooper's erroneous choice for the investment variable is more understandable, due to my use of some imprecise language, but see p. 202.

particular, the different models are not based on the same set of predetermined variables. There is ample evidence that one can change the forecasting characteristics of a given model by changing the exogenous-endogenous status of different variables. Most of this evidence is based on forecasts of longer than one period, but the same is true in the one-period case. For example, I generated a series of one-period forecasts from one version of the FRB-MIT Model (for 1958-I to 1968-IV), which produced a mean-square error of $12.2 billion for real *GNP*. However, throwing out the stock market and the currency equation produced a mean-square error of $9.9 billion. For a shorter period (1963-I to 1968-IV), the corresponding results were $9.6 and $12.4 billion.

In other words, making certain variables exogenous in the shorter period deteriorated performance, whereas it had helped in the longer period. If one is willing to make more things exogenous (e.g., prices), even more dramatic shifts could be reported. In short, if the forecasting ability of a given model varies significantly when different sectors are made exogenous, what are we to make of different models which have different sets of exogenous variables?[8]

SOME ECONOMETRIC DETAILS

Cooper's concern with putting all of the models on a comparable basis extends to his method of estimation as well. Although, as suggested before, the treatment of autocorrelation might have been a worthwhile endeavor, Cooper concentrates his efforts on the simultaneous-equations problem, and on securing comparably efficient estimates for the different models. His basic estimation method is two-stage least squares, and the major problem he encounters is that in some of the models, the number of predetermined variables exceeds the number of observations. In treating these models, rather than resorting to principal components or to some other selection procedure for the first-stage regressors, Cooper utilizes a method he calls repeated reduced-

[8] As another illustration, consider the treatment of strikes. In ex ante forecasting, the Wharton Model, for example, typically deals with this by intercept adjustments. Other models may have strike dummy variables. Clearly, this is not a comparable treatment of the two types of models.

form estimation (RR).[9] For linear models, it appears that this method is more efficient than any of these selection procedures and than the two-stage least squares approach itself if the error terms in different equations are uncorrelated. If the residuals have nonzero covariances, however, it appears that the RR method can be less efficient than these other methods.[10] Of course, these statements concern large samples. In small samples, not much is known about comparable properties of the two methods; this may further confound the nature of the comparison between models.[11]

Aside from this, however, one remark directed at the repeated reduced-form method itself is necessary: namely, this technique will not in general have the same desirable large-sample properties in the context of nonlinear models. First, consider the linear case. The method involves calculating consistent two-stage least-squares estimates (using some subset of the predetermined variables). Employing standard notation, one gets estimates: $\hat{B}, \hat{\Gamma}$ in $Y\hat{\Gamma} + X\hat{B} = U$. Next set $U = 0$, and solve for $Y = -X\hat{B}\hat{\Gamma}^{-1}$. This indirect estimate of the endogenous variables gives us new "corrected" values for these variables, which can be used to compute a second two-stage least-squares estimate. This estimate is also consistent but should improve in efficiency, since we have a consistent estimate of the reduced form on which to base our final estimates. If the number of predetermined variables exceeds the number of observations, selecting a subset of the predetermined variables will not produce such a consistent estimate for the first stage.

Now consider the nonlinear case. Consistent estimates of the structural parameters are still available.[12] However, we can no longer get a consistent estimator of the reduced-form coefficients by setting the error term equal to zero and solving for the endogenous variables. In-

[9] The method has a variety of other names; e.g., in Evans, Haitovsky, and Treyz, *op. cit.*, it is called "regression on predicted values."

[10] These findings were communicated to me by Phoebus Dhrymes.

[11] There is a limited bit of evidence in Evans, Haitovsky, and Treyz, *op. cit.*, that the RR method works better for one-period forecasts than it does for longer spans. Given the nature of the comparison discussed above, this may introduce additional complications. Another problem with small-sample comparisons is that the definition of "large" sample is not invariant to the specification of the model. For some evidence on this point, see S. M. Goldfeld and R. E. Quandt, "Nonlinear Simultaneous Equation," *International Economic Review* (February, 1968), pp. 113–36.

[12] See, for example, H. Kelejian, "Two-Stage Least Squares and Nonlinear Systems," *Journal of the American Statistical Association (forthcoming).*

deed, the very notion of the reduced form needs to be carefully reexamined in the case of nonlinearities. Suppose the model is $Y\Gamma_1 + F(Y)\Gamma_2 + XB = U$, where $F(Y)$ is a set of nonlinear functions in Y. Setting $U = 0$ gives us one "solution" for Y as a function of X, say $y = f_S(X)$. However, the conditional expectation of Y given X, i.e., $E(Y/X)$, will in general not be equal to $f_S(X)$; rather, it will be $f_T(X)$, say. This means that Y can be written as $Y = f_S(X) + [f_T(X) - f_S(X)] + V$, where V is an error term such that $E(V/X) = 0$.[13] Consequently, simply using $\hat{Y} = f_S(X)$, as in the RR method, does not provide a consistent estimate of the reduced form.[14] Furthermore, since it throws a term $f_T(X) - f_S(X)$ into the error term of the second two-stage regression, and since this term will be correlated in general (even asymptotically) with X, the RR method for nonlinear models may not even give consistent structural estimates.[15] In short, while the method seems to have much to recommend it for linear models, further investigations are needed to ascertain its suitability for nonlinear models.

CONCLUSION

From the above discussion, it should be clear that despite the monumental proportions of Cooper's study, it is of only limited relevance for the rather broad questions to which it is nominally addressed. Cooper's pessimistic conclusions on the performance of models relative to a "naive" standard stem from both the mechanical reestimation

[13] For a more detailed discussion of this development, see E. P. Howrey and H. H. Kelejian, "Dynamic Econometric Models: Simulation vs. Analytical Solution," in T. H. Naylor, ed., *The Design of Computer Simulation Experiments*. Durham, Duke University Press, 1969, pp. 207–231.

[14] This same point is made in Klein, *op. cit.*

[15] Perhaps a concrete example (taken from Howrey and Kelejian, *op. cit.*) will help. Consider the model

$$y_{1t} = b_1 X_t + u_{1t}$$

$$y_{2t} = b_2 y_{1t-1} + b_3 \exp(y_{1t}) + u_{2t}$$

One can write $y_{2t} = b_2 y_{1t-1} + b_3 \exp(b_1 X_t) \exp(u_{1t}) + u_{2t}$. Now, $E(\exp(u_{1t})/X_t, y_{1t-1}) = e^{\sigma_1^2/2}$ where σ_1^2 is the variance of u_{1t}. Consequently, one can write $u_{1t} = e^{\sigma^2/2} + u_{3t}$ where $E(u_{3t}/X_t, y_{1t-1}) = 0$. Hence, $y_{2t} = b_2 y_{1t-1} + (b_3 e^{\sigma_1^2/2}) \exp(b_1 X_t) + (u_{2t} + u_{3t})$. Thus, simply setting $u_{1t} = u_{2t} = 0$, and solving, would lead to a reduced form which ignored the term $e^{\sigma_1^2/2}$ a term which does not vanish.

of all the models and the use of one-period forecasts. More recent evidence, such as the FRB-MIT results cited earlier and the results with the OBE Model presented at this Conference, suggests that even for one-period forecasts, carefully estimated large econometric models outperform the auto-regressive standards.[16]

As for picking the "best" model, given that different models employ different sets of predetermined variables, it is difficult to do this by ex post analysis alone. Such types of analysis give undue credit to models for hard-to-forecast exogenous variables. Clearly, what is needed for each model is a notion of the sensitivity of prediction errors to errors in forecasting the exogenous variables. Furthermore, the evidence on ex ante forecasting needs to be examined as it accumulates.[17] Finally, we need more systematic procedures for mixing and matching the best parts of various econometric models so that rather than focusing on picking a "winner" out of a fixed set of models, we can more generally improve our set of forecasting tools.

MICHAEL D. McCARTHY

WHARTON SCHOOL

The paper under discussion represents a report on the results of an obviously massive study. The author is to be congratulated for having had the courage to undertake such a task. It is clear that he has thought through many of the problems associated with testing the predictive performance of competing models (hypotheses). This alone insures that the study has merit as a commendable first effort. Whether or not it has further merit hinges on whether its striking conclusion, that econometric models really cannot be expected to be better forecasting tools than "naive" auto-regressive schemes, is credible. If it is, econometricians are obligated to reconsider their goals.

[16] Green, *op. cit.*, reports a root mean-square error for real *GNP* of $3.1 billion for the OBE Model, and of about $4.6 billion for an auto-regressive scheme over the same period. For the FRB-MIT Model, the numbers cited above are clearly better than either this $4.6 billion or the evidence given by Cooper.

[17] This is the thrust of Evans, Haitovsky, and Treyz, *op. cit.*, where it is argued that ex ante forecasting, with all its ad hoc constant adjustments, often outperforms ex post forecasting with known values of the exogenous variables.

The credibility of the conclusions clearly depends on the quality of evidence given as support; in deciding on credibility, we seek evidence that the author did a convincingly good job of testing the competing models. Specifically, we seek evidence that (1) taking the model structure as given, appropriate estimation techniques have been used; (2) in estimating the models, adequate use has been made of prior knowledge; and (3) the tests of predictive performance are good tests. Under Category 3, we require evidence that efficient use was made of the models in performing the prediction experiments. All of these categories of evidence are important. Failure in any one area is sufficient to cast doubt on credibility. In the case of the study being discussed here, it is most tempting to begin by considering the evidence in Category 2.

In testing competing hypotheses, it is clear that the tests must be applied to a consistent body of data. The study appears to have satisfied this requirement – but much more is required. For instance, in estimating a model, the sample observations must be drawn from the same population over time. If observations are drawn from different populations, the estimates (and forecasts) can be expected to be seriously biased. One thing we know well is that owing to factors such as technical change, strikes, government economic policy actions, and wars, the structure of economic models does shift over time. An uncritical pooling of time-series data will surely involve sampling from different populations. Accordingly, under Category 2 we seek evidence that the author has taken adequate account of prior knowledge concerning shifts in the economic structure. The quarterly sample used in the study covered 1949 through 1960. During this period a war occurred, and tax laws changed. Moreover, it is known that due to strikes and technical progress, production functions – and hence, investment demand functions – and price markup relations probably shifted. These shifts need not be of a simple exponential sort. The shifts in the tax laws can also be expected to affect functions other than the equations for government tax receipts. Depreciation equations shift, and investment incentives are affected. This is only the beginning of a very long list of complications due to structural shifts.

It should be clear that building a good forecasting model involves a great deal of hard work. The economist begins by specifying an initial structural hypothesis – a prior or null hypothesis. This prior hypothesis involves a statement (perhaps probabilistic) about the form of the structural equations, a statement of what knowledge is available concerning structural shifts, a statement of all prior knowledge about coefficient signs and magnitudes, and some statement about the error properties of the model. (Typically, the prior hypothesis implicitly involves a statement that some of the equations have stable structures.) The model is then estimated, using some appropriate method, and some effort is made to improve on it. This involves testing the prior hypotheses against competing hypotheses suggested by the initial estimates. Under Category 2, we seek evidence that such testing was undertaken. If the original state of knowledge dictated that a coefficient was positive, and a significantly negative estimate was obtained, using an "appropriate" estimator, there are grounds for rejecting (revising) some part of the prior hypothesis. If the prior hypothesis specified that the errors of the structural equations were non-autocorrelated, and the calculated residuals of the estimated equations show strong systematic behavior with time, this too is justification for rejecting (revising) the prior hypothesis. Rejection of the prior hypothesis involves rejection of one or more of its component statements. Sometimes attention is focused on the form of the equations; variables may have been omitted or included in an inappropriate fashion. New knowledge (additional data and information) may suggest a ready answer. In reconsidering the prior hypotheses, attention is often focused on structural shifts not originally hypothesized. Here, too, new knowledge (information not available in the initial data set) is required.

Under Category 2, the most striking thing about the Cooper paper is the lack of evidence that anything was done other than estimating the various models, using a consistent body of data. In spite of the evidence of structural change presented in analyzing the forecasting properties of the various models – we shall not refer to them as the Fromm Model, or the Liu Model, or whatever – no evidence is presented of any effort to take into account the obvious sources of structural shift over the period 1949–60. There is no evidence that the prior hypotheses were ever tested against competing hypotheses. In view of the fact that the body of data used for the study differed from the bodies of data used by the original authors, such tests are certainly called for. What is lacking is evidence that, in constructing his models, the author gave the same loving care to each one that the authors of the earlier studies gave when constructing theirs. At this point, we merely note that if such care is lacking, there is a presumption that the models tested will yield seriously biased ex post and ex ante forecasts. (In light of the fact that the author's largest model does not seem to fare too well in forecasting in comparison with his more aggregative models, the following seems worth noting: large models, because they make a great many explicit statements about structure, can be expected to be more adversely affected by a failure to take great care in equation specification than are small models. Errors due to the sources discussed above may well be swamped by the aggregation of the small models.)

Turning now to the evidence under Category 1, I concur with Professor Goldfeld's criticism that because of the treatment of nonlinearity used by the author, his estimators do not have the property of consistency. This problem can be circumvented if the author wishes to undertake the study a second time.

Next, let us consider the evidence under Category 3. The author performed a series of one-quarter reduced-form forecasts over the sample period, and over the span 1961 through 1965. Mean-squared errors were then computed and compared equation by equation. Included in this comparison were the mean-squared errors obtained from the naive auto-regressive models. Because of the bias problems raised in earlier paragraphs, one might expect the "structural" models studied by the author to do relatively poorly in comparison with the auto-regressive models. This is, in fact, true for the sample period forecasts and for the 1961–65 forecasts.

Actually, the comparisons in the sample period seem irrelevant. It is well known that as far as one-period forecasts are concerned, least-squares auto-regressions provide a convenient device for obtaining as close a fit to the sample as desired. Moreover, the fit of the one-period forecasts will be exactly the same as the least-squares fit. To prove that one can obtain a close auto-regressive fit to the sample proves nothing.

With respect to the forecasts outside the sample, there is an additional irritation. Model forecasters are well aware of the problem caused by structural shifts in forecasting outside the sample period, and exert much effort on measures designed to cope with it. Models are frequently updated and appropriate tests of structural stability performed. When shifts are found, corrective measures are taken. In a slightly different vein, it is also to be noted that the errors of econometric models often exhibit auto-regressive behavior (in spite of the model-builders' intuition which suggests that a perfectly specified model should not have this property). However, there are also well-known techniques for taking account of the auto-regressive errors in forecasting.

Not surprisingly, the effect of the corrective measures is a significant improvement in one-quarter forecasts. One striking thing about this study is the lack of evidence that in performing the one-quarter forecasts, the author took any such measures. In fact, it appears that the model used to obtain the forecast for 1961-I was the same as the model used to forecast 1965-IV. Apparently, there was not even an attempt to take account of the most obvious of all the structural shifts that took place in the forecast period: the tax law changes. This particular structural shift was one that the forecaster would have been aware of some time before it actually took place; corrective measures would have been taken before the fact.

It should also be stressed that a one-period forecast comparison of the models puts the auto-regressive models in an unduly favorable light. There is much reason to believe that they would have fared much worse had the author chosen a four-quarter or eight-quarter forecast. In fact, for the sample period (in which the naive models inevitably appear in a favorable light, regardless of their structural significance), the auto-regressive models appear to perform worse than do the structural models. For instance, consider the following auto-regressive model of constant price GNP fit to quarterly data from the period 1948–64.

$$\underset{}{GNP^{58}} = \underset{(.1139)}{1.443GNP^{58}_{-1}} - \underset{(.1150)}{.4413GNP^{58}_{-2}} + \underset{(4.0264)}{1.6837}$$

$$SE = \text{Standard Deviation} = 5.091$$

The numbers in parentheses are standard errors. This second-order scheme was the best auto-regression by the same least-squares standards as were used in Cooper's paper. One-quarter forecasts of the Wharton Model over this period yielded a standard deviation for GNP^{58} of 7.343. On the other hand, when the two models were used to forecast GNP^{58} for the entire period 1948 through 1964, using only the initial information available at the start of the period, the standard deviation for the Wharton Model was 15.3 billion; and for the auto-regressive scheme, 17.5. Moreover, the auto-regressive model showed no cyclical sensitivity over the period, while the Wharton Model GNP forecasts closely followed the direction of the observed cyclical movements. If the auto-regressive scheme cannot pick out cycles in the sample period, why should we expect it to do so outside the sample?

The relevance of the one-quarter forecast should also be questioned. Very few institutions in this society have a planning period as short as one quarter. Four to eight quarters is a more representative span.

In summary, the present study leaves too many questions unanswered (or answered inadequately) to pass the credibility test. The model builder's attitude is still the following: a model built according to the principles outlined above will consistently beat an auto-regressive model in one-period forecasts and, especially, in forecasts for two, three, and four periods. Cooper's study does not provide a basis for rejecting such a hypothesis.

COMMENT

GEORGE R. GREEN
MAURICE LIEBENBERG
ALBERT A. HIRSCH

Mr. Cooper undertook a rather ambitious task. Knowing the enormous work involved in estimating and processing one model, one can only admire him for taking on the estimation and processing of seven models. On the other hand, one can question whether Mr. Cooper undertook too much. First of all, the size of the task apparently made it impossible for him to develop an adequate text. We find, for example, that his summary accounts of the various models are quite inadequate, and it is doubtful whether any reader unacquainted with

NOTE: All of the contributors listed above are staff members of the Office of Business Economics, U.S. Department of Commerce.

the models will be able to grasp their essential features. Moreover, one can find instances of ambiguous, and sometimes incorrect, statements. For example, we find it somewhat puzzling that output is taken as given in the Liu Model. Cooper also states that there are only minor differences between the Klein and OBE Models. Although the OBE Model can be regarded as a variant of the Klein Model, the differences are by no means minor.

A more serious difficulty pertains to the absence of an adequate statement or explanation of the precise procedures followed. To a reader who is made aware of the rather startling mean-square errors found by Mr. Cooper, there is particular interest in a detailed account of his procedures so that results can be appraised in the light of the methods used. There are a number of instances where such additional information would be of value. As an example of this, one can mention the anomalous result that for the OBE Model, the mean-square error for constant-dollar *GNP* on the demand side was $31.5 billion; while for real private *GNP*, it is somewhat over $101 billion. It is odd that such results are possible, since the differences can only be due to exogenous real government product, which is assumed known. We think that this result requires something by way of explanation.

Again, referring to the OBE Model, no account is given—either in the text or in communication with the author—of how the statistical discrepancy is treated. In that model, we have included equations for all major income magnitudes instead of adopting the more common procedure of leaving one income item—usually profits—as a residual. We avoided the problem of having more equations than unknowns by defining the statistical discrepancy as a variable which we constrain to vary only within prescribed limits. This requires that we make adjustments to selected income items and then re-solve the entire model to ensure that the income and product identity holds within the prescribed limits. If Cooper solved the model without regard to the discrepancy, he must, in effect, have removed the importantly constraining income-product identity. Then, a correct solution is only obtained fortuitously. In particular, since Cooper did not take into account serial correlation, we can expect rather wide swings in the discrepancy if indeed he ignored the problem, as he has apparently done.

Another problem is the rather mechanical refitting of models to a common sample period. Some of the models were originally fitted to a sample period which included the Korean War, while others were not. Because of the unique conditions which prevailed, dummy variables were used in some instances. In the Cooper study, the war was also included in the sample period, but Cooper gives no indication that he introduced such variables in estimating those models which originally did not cover the war period. Such a step would have demanded careful and detailed work on Cooper's part, and we have no inkling of what was actually done. If this factor was also ignored, efforts to compare the various models would hardly be worthwhile; we would naturally expect differences both in the errors obtained and in the estimates of the structural parameters.

This brings us directly to the question of whether it is proper to reestimate a model using a sample period other than that initially used by the model-builders. To scientific investigators who profess to have captured basic structure, such a transposition is indeed desirable. To say that a model is only valid for the period to which it is initially fitted, would put its worth seriously in question. But we must also admit that structure can indeed change; our hope is that it changes slowly enough so that we may obtain useful equations over the short run. Our tolerance for slow structural change does not mean, however, that we hold our equations to be valid over periods that include wars, which obviously bring into play very special factors. Mr. Cooper apparently ignores the model-builder's preference in this regard. It is impossible to determine the impact of this point on the results which he obtained, and we can only register uneasiness. As a minimum, we can say that it is not surprising that Mr. Cooper finds evidence of structural change. We would expect it; especially in the case of the OBE Model, since we quite intentionally confined our observations to the post–Korean War period.

Mr. Cooper uses as a standard of comparison a simple autocorrelative function and concludes that a "minimal standard for performance of an econometric model is that it must forecast more accurately than a purely mechanical scheme which incorporates no economic information whatever." Mr. Cooper finds that, by and large, the naive autocorrelative function outperforms the model – a rather disappointing result from the point of view of model-builders, at least with respect to the use of models in forecasting. In light of some of the points raised above, there is some feeling of uncertainty about whether his conclusions are indeed valid. But, assuming that they are correct for single-period predictions, one would have preferred that the examination be carried out for more than one quarter, since there are reasons to expect deterioration of the autocorrelative form in multiperiod predictions.

Both during informal meetings at the Conference and later, the above finding of Cooper's was mentioned frequently. We noted that the superiority of the autocorrelative form was frequently quoted, particularly by persons skeptical of model-building. The statement apparently confirmed their own suspicions that models consist mainly of elaborate facades with nothing of substance behind them. Once out of the parental environment of the model-builders and exposed to scientific scrutiny, they seemed to emerge inferior to even the most primitive substitute.

It is because of this reaction to the Cooper paper that we include below a table presenting some results which were obtained using a later version of the OBE Model—one used in the paper presented by Mr. Green—which underlies the analysis included in some of the other papers in this volume. The table gives average absolute errors in real *GNP* obtained over the sample period 1955-I through 1966-IV for one- through six-period forecasts. Figures are given for a second-order auto-regressive equation, for the model without any adjustment, and for strictly mechanical adjustments based on considerations of first-order serial correlation.

Average Absolute Errors in Real GNP Over the 40 Quarters: 1955-I Through 1966-IV (billions of 1958 dollars)

		OBE Model	
	Auto-regressive Equation	No Constant Adjustments	Automatic Constant Adjustments
First quarter forecasts	3.60	3.09	2.35
Second quarter forecasts	6.93	4.28	3.58
Third quarter forecasts	9.55	4.98	4.47
Fourth quarter forecasts	11.39	5.43	4.92
Fifth quarter forecasts	13.17	5.81	5.42
Sixth quarter forecasts	14.43	5.98	5.82

It can be seen that the OBE Model outperforms the naive autoregressive form in every instance; and that the improvement is more apparent, the larger the horizon over which the forecast is made.

Of course, the results given do not negate, per se, Cooper's findings with an earlier version of the model, but they should serve to make model critics somewhat hesitant to generalize from the Cooper findings.

Some final remarks can be made regarding the nature of the Cooper study, taken as a whole. We think that tests of model performance should indeed be made. We also insist that model performance must be appraised in terms of post-sample predictions, with exogenous variables taken as given. Denying the worth of such an exercise leaves one deservedly open to criticism regarding the degree to which models have captured underlying structure. From this point of view, the usefulness of models in actual forecasting situations—when in the hands of capable investigators—is not directly relevant.

There is a question, however, whether the research carried out by Mr. Cooper contributes much to the problem of intermodel comparisons. Some of the reasons for this reservation have been outlined above and others have been included in various comments. What is required, apparently, is that efforts be expended prior to the building of various models to achieve a maximum degree of comparability in regard to such factors as: sample time-period; degree of endogenicity; use of dummy variables; and the extent to which serial correlation is taken into account. It is evident that something can be done to lay down basic ground rules to be observed by all model-builders. Among those rules, one would include the meticulous observance of whatever parameter and solution constraints are desired on the part of the model-builders.

This program, of course, far transcends anything attempted by Mr. Cooper, who took the models as given and attempted to satisfy the requirement of identical time-period by undertaking the estimation himself. We feel that the required task would be made easier if it were undertaken as a project by an organization such as the National Bureau of Economic Research. This project would require considerable discussion among model-builders prior to beginning the actual task of estimation and testing.

REPLY

COOPER

Owing to limitations of space, I shall comment only on those criticisms of my study which I consider to be the most important: (1) My study, which was carried out on an ex post rather than ex ante basis, does not adequately reflect the predictive performance of the alternative econometric models, since using a priori information (e.g., anticipated strikes) improves forecasts. Also, not enough attention was devoted to allowing for factors that the original authors would have taken into account had they used their respective models for forecasting (e.g., correcting for serial correlation). (2) I did not handle the statistical discrepancy correctly in the OBE Econometric Model. I made the discrepancy exogenous, while the OBE econometric staff assumed it endogenous. (3) I should have added dummy variables to those models that were not originally estimated over periods including the Korean War, since most of the models that did include it in their original estimation periods also included dummy variables in some of the structural equations. (4) Single-period forecasts are not a sufficiently stringent test for the econometric models, because those models which do make forecasts do so for more than one period. (5) The asymptotic properties of the repeated reduced-form (RR) estimator used in the Wharton-EFU Model are unknown, since the model is nonlinear in the variables.

(1) We must be extremely careful not to confuse two quite separate issues. One is the predictive performance of econometric forecasting, and the other is the predictive performance of econometric models. This study is concerned not with measuring the performance of econometric forecasting, but with measuring that of econometric models. In particular, my study is an attempt to test the specification of the econometric models. One way to do this is to put all of the econometric models on a comparable basis, as I have attempted to do, and to compare the ex post predictive performance of the models with that of naive models, which contain no economic information whatever. This provides a very stringent test of the economic information contained

in the econometric models, or alternatively, the form in which the economic information is entered into the models.

Another way to test the specification of econometric models is to test them for structural change. This test is equivalent to determining whether the same specification used in fitting the model for a given period holds for a different period. Both of these tests were carried out in my study.

To my knowlege, there are no ways other than the tests mentioned above for evaluating the specification of an econometric model. Saying that the authors would have respecified their models, had structural change occurred, may be good hindsight, but it offers little—if any—support for making accurate forecasts in the future with econometric models which have been fitted to historical data. A model that can properly reflect the structure of the economy during a given period, but not for another period, cannot be expected to give reliable forecasts outside the original period.

It is quite possible that the authors would have tried to take into account a priori information in forecasting, but I had no possible way of evaluating how the authors would have adjusted their models in making forecasts. This is a good reason for evaluating the econometric models on an ex post, rather than on an ex ante, basis.

(2) As it stands, the OBE Model is an overdetermined system. In solving their model, the OBE econometric staff attempts to get around this problem by treating the statistical discrepancy in the national accounts as an endogenous variable. The statistical discrepancy is prevented from exceeding a certain level by distributing its forecast error to the other variables on the income side of the national accounts.

Since I had no way of determining how the OBE econometric staff would have distributed the errors of the statistical discrepancy to other variables in the national accounts, I decided to make the discrepancy exogenous. It is not possible to tell what effect this problem has on the over-all solution of the OBE Model but I suspect that it is fairly minor.

(3) Although the answer to this question cannot really be put forward without additional testing, I honestly do not feel that the performance of any model that did not originally include dummy variables for the Korean War period would have been any better with them than

without them. There are two reasons for this. First, the dummy variables in the models that include them explain only a very small percentage of the variance of the left-hand variables (usually less than 2 per cent), so that their inclusion over the fitted period probably does not significantly improve the performance of the model during this period. Second, the inclusion of the dummy variables during the fitted period may cause predictions to be biased outside the sample period, in which case it will produce worse forecasts than would be obtained without it. I have much evidence, from a piece of unpublished research which I have done, that this is, in fact, the case. Therefore, I do not consider the dummy variable problem as serious as some of the model-builders whose work was included in this study seem to think.

(4) As a consequence of the magnitude of my study, it was necessary for me to place certain limitations on the research. One such limitation was to make single- rather than multiple-period forecasts. We feel that the one-period forecasts are a reasonable choice, since it is most likely that a necessary condition for making accurate multiple-period forecasts is the making of accurate one-period forecasts. Furthermore, we should point out that although the forecasts are one-period in the sense that actual values of the lagged endogenous variables are used, the forecasts over twenty quarters (1961-I through 1965-IV) are based on the same set of coefficients (1949-I through 1960-IV). A twenty-quarter forecast based on one set of coefficients certainly represents a stringent test of an econometric model's predictive power, even though forecasts are made on a one-period basis.

To buttress this argument, the following table presents *MSE*'s and performance rankings for the Friend-Taubman and naive models, using both single- and multiple-period forecasts. The mean-squared errors are computed over the first eight quarters of the forecast period – 1961-I through 1962-IV.

Considering these results in detail, we notice, first, that with one exception, the mean-squared errors of the *GNP* components for both the Friend-Taubman and naive models are larger for the multiple-period forecasts than they are for the single-period forecasts. Second, we observe that, with one exception, the performance rankings of the Friend-Taubman and naive models do not change for the *GNP* components in going from the single- to the multiple-period forecasts.

Comparison of Mean-Squared Errors of Friend-Taubman and Naive Models Over Period 1961-I Through 1962-IV (billions of 1958 dollars)

	Single Period		Multiple Period	
	Friend-Taubman	Naive	Friend-Taubman	Naive
Real *GNP*	49.97	27.01	2893.0	646.9
Performance rankings	2	1	2	1
Real consumer expenditures	16.66	3.229	2146.0	10.04
Performance rankings	2	1	2	1
Real residential structures	1.503	0.1458	7.747	2.598
Performance rankings	2	1	2	1
Real nonresidential structures	1.762	1.318	21.08	10.03
Performance rankings	2	1	2	1
Real inventory investment	7.010	5.318	4.811	9.294
Performance rankings	2	1	1	2

Of course, these results are not conclusive but only suggestive, since we have carried out a multiple-period forecasting test for just one of the seven econometric models included in the study.

In their comment on my paper, Green, Liebenberg, and Hirsch have presented some results which show that the OBE Model is slightly superior to that of an auto-regressive scheme in predicting *GNP* over both single- and multiple-periods. However, they fail to point out that these predictive results are obtained *within* the sample period. As I, and others, have pointed out, this is not an adequate test of the model, since the specification of an econometric model typically reflects the persistence of the investigator as well as the underlying theory. What they should have done was compare their model's forecasts with those based on naive models *outside* the sample period. I might add, also, that should they decide to do this, they should be careful to keep the comparison of the econometric with the naive model a fair one. For example, in making multiple-period ex post forecasts, if they correct any of their econometric equations for serial correlation, they should do this for the corresponding auto-regressive schemes as well. For practical reasons, it is probably best to make no adjustments to either the econometric or the naive equations when carrying out the comparative ex post forecasting test.

(5) It is difficult to evaluate the validity of this criticism. If an econometric model is nonlinear in the variables, then it is true that the asymptotic properties of the RR estimator are unknown. However, it is also true that the asymptotic properties of any repeated least-squares estimator are unknown (including the ordinary 2SLS estimator). As Goldfeld correctly indicates, more research is needed to determine the asymptotic properties of limited information least-squares estimators when the system is nonlinear in the variables.