INTERNATIONAL ECONOMIC REVIEW

June, 1973
Vol. 14, No. 2

# A COMPARISON OF ALTERNATIVE ESTIMATORS OF MACROECONOMIC MODELS*

BY RAY C. FAIR

## 1. INTRODUCTION

ALTHOUGH THE DEVELOPMENT OF MACROECONOMIC MODELS has reached a stage where a number of models are now being used on a regular basis for forecasting purposes, many models do not appear to be capable as yet of producing accurate forecasts without the extensive use of constant-term adjustments and other "fine-tuning" procedures.[1] Part of the need for constant-term adjustments in many models appears to be due to serial correlation of the error terms in the models. Adjusting the constant term in an equation to make the last estimated residual zero can be considered to be a crude way of accounting for serial corrlation. There are, of course, more formal ways of accounting for serial correlation than by adjusting constant terms, and when a more formal approach was taken for the model developed in [5], the model appeared capable of generating good forecasts in a mechanical way without the need for any constant-term adjustments or other fine-tuning procedures. In the estimation of the model in [5], account was taken of both simultaneous-equations bias and first-order serial correlation of the error terms, and the estimates of the serial correlation coefficients were used in the generation of the forecasts from the model.

The results obtained in [5] thus indicate that considerable gain in forecasting accuracy may be achieved by the use of more advanced estimation techniques, but no formal comparison of the forecasting accuracy of alternative estimators was made in [5]. The purpose of this paper is to make such a comparison. Ten estimators are considered. Each estimator is first used to estimate the seven stochastic expenditure equations of the model developed in [5]. The reduced form of the model is then solved for each set of estimates, and within-sample predictions (both static and dynamic) of the endogenous variables of the model are generated over the sample period. The estimators are compared in terms of the accuracy of the within-sample predictions. Seven of the estimators account for first-order serial correlation of the error terms, three of the estima-

* Manuscript received March 30, 1972; revised July 19, 1972.

[1] Evans, Haitovsky, and Treyz [4] in their study of the Wharton and OBE models, for example, concluded that: "From the previous results it should be obvious that econometric models cannot generate good forecasts if they are used only in a mechanical fashion. The art of forecasting still requires that a great deal of fine tuning be used with any econometric model presently in existence."

tors also account for second-order serial correlation, three of the estimators are two-stage least squares estimators, three of the estimators are full-information maximum likelihood estimators, and one of the estimators attempts to account for the fact that values of the lagged endogenous variables are not known beyond the one-period-ahead forecast.

Parts of this study would not have been possible without the recent advances that have been made in numerical analysis and computer technology. Work by Goldfeld, Quandt, and Trotter [7] and others on nonlinear estimation techniques, for example, has greatly increased the feasibility of estimating equations that are nonlinear in coefficients, and recent work by Chow [2] and Chow and Fair [3] has increased the feasibility of estimating linear economic models by full-information maximum likelihood. Advances in computer technology have increased the feasibility of handling larger models and more complicated problems. For the work in this study, the technique described in Chow [2] was used to obtain the full-information maximum likelihood estimates under the assumption of no serial correlation, and the technique described in Chow and Fair [3] was used to obtain the full-information maximum likelihood estimates under the assumptions of first- and second-order serial correlation. The quadratic hill-climbing technique of Goldfeld, Quandt, and Trotter [7] was used to obtain the estimates that accounted for the fact that values of the lagged endogenous variables are not known beyond the one-period-ahead forecast.

It should be noted at the outset that this study is not a comparison of estimators in terms of the standard properties of unbiasedness, efficiency, and consistency. Rather, this study is an attempt to determine, using an actual model of the United States economy, which estimators lead to the most accurate multi-period predictions. The methodology of the study is thus similar to that of Klein's study [9], where four sets of estimates of the Klein-Goldberger model were compared in terms of the accuracy of the within-sample predictions from the model and in terms of the accuracy of one-year-ahead outside-sample predictions. In the present study only within-sample predictions are considered.

This study is based on the premise that the basic properties of macroeconomic models are similar enough so that the conclusions obtained from the use of one model can be generalized to other models. The model used in this study is small, linear, and was designed primarily for short-run forecasting purposes. Three of the variables treated as exogenous in the model—an index of consumer sentiment, plant and equipment investment expectations, and housing starts—are really not exogenous beyond about the two-quarter-ahead predictions. It is an open question as to how restrictive the linearity property is in allowing one to generalize the results of this study to nonlinear models, and it is also an open question as to whether different results would be obtained if one used a much larger and more disaggregated model that relied less heavily on exogenous expectational variables. Until more work and experimentation has been done on large-scale nonlinear models, the results in this paper are put forth as indicating what might be the case for such models. The results do indicate that serious attempts should be made to estimate models by techniques other than simple ordinary least squares or two-stage least squares.

## 2. THE TEN ESTIMATORS

2.1. *The general model.* The general model to be estimated is

$$AY + BX = U, \tag{1}$$

where $Y$ is an $h \times T$ matrix of endogenous variables, $X$ is $k \times T$ matrix of predetermined (both exogenous and lagged endogenous) variables, $U$ is an $h \times T$ matrix of error terms, and $A$ and $B$ are $h \times h$ and $h \times k$ coefficient matrices respectively. $T$ is the number of observations. The $i$th equation of the model will be written as

$$y_i = -A_i Y_i - B_i X_i + u_i, \qquad i = 1, 2, \ldots, h, \tag{2}$$

where $y_i$ is a $1 \times T$ vector of values of $y_{it}$, $Y_i$ is an $h_i \times T$ matrix of endogenous variables (other than $y_i$) included in the $i$-th equation, $X_i$ is a $k_i \times T$ matrix of predetermined variables included in the $i$-th equation, $u_i$ is a $1 \times T$ vector of error terms, and $A_i$ and $B_i$ are $1 \times h_i$ and $1 \times k_i$ vectors of coefficients corresponding to the relevant elements of $A$ and $B$ respectively.

The error terms in $U$ are assumed to follow a second-order auto-regressive process:[2]

$$U = R^{(1)} U_{-1} + R^{(2)} U_{-2} + E, \tag{3}$$

where the $R$ matrices are $h \times h$ coefficient matrices, $E$ is an $h \times T$ matrix of error terms, and the subscripts denote lagged values of the terms of $U$. The error terms in $E$ are assumed to have zero expected values, to be contemporaneously correlated but not serially correlated, and to be uncorrelated in the limit with the predetermined, lagged predetermined, and lagged endogenous variables.

2.2. *Ordinary least squares* (*OLS*). The first estimator considered is ordinary least squares applied to each equation of (2). Ordinary least squares does not, of course, produce consistent estimates of the coefficients of the model. The estimates are inconsistent both because of the correlation between $u_i$ and $Y_i$ in (2) and because of the correlation between $u_i$ and the lagged endogenous variables in $X_i$ in (2).[3]

2.3. *Two-stage least squares* (*TSLS*). The second estimator considered is two-stage least squares applied to each equation of (2). Two-stage least squares produces consistent estimates if the error term $u_i$ in (2) is not serially correlated

[2] The process in (3) can easily be generalized to higher-order processes, but that will not be done here since only processes up to second order are considered in the empirical work.

[3] Although, as mentioned above, the estimators considered in this study are not compared in terms of properties like consistency and efficiency, some of these properties will be noted in the discussion of the estimators. This is done in part so that the ranking of the estimators in terms of forecasting accuracy can be compared with the ranking of the estimators in terms of properties like consistency and efficiency.

or if there are no lagged endogenous variables in $X$; otherwise not. With a large enough sample, all of the variables in $X$ should be used as regressors in the first-stage regression for each equation. In practice, however, it is usually necessary to use only a subset of variables in $X$ as regressors or to use only certain linear combinations of all of the variables in $X$ as regressors. A necessary condition for TSLS to produce consistent estimates is that the included predetermined variables in the equation being estimated be in the set of regressors. Otherwise there is no guarantee that TSLS will produce consistent estimates even if the error term is not serially correlated or if there are no lagged endogenous variables among the predetermined variables.[4] For the work below, therefore, the variables in $X_i$ were always included in the set of regressors when the $i$th equation of (2) was estimated by TSLS. The other regressors that were used will be discussed in Section 3.

2.4. *Ordinary least squares plus first-order serial correlation* (*OLSAUTO1*). The third estimator considered accounts for first-order serial correlation of the error term $u_i$ in (2), but not for simultaneous-equations bias. The estimator is based on the assumption that the error term in each equation is first-order serially correlated:

$$u_i = r_{ii}^{(1)} u_{i_{-1}} + e_i , \qquad i = 1, 2, \ldots, h, \tag{4}$$

which means that $R^{(1)}$ in (3) is assumed to be a diagonal matrix and $R^{(2)}$ in (3) to be zero. Under this assumption, equations (2) and (4) can be combined to yield:

$$y_i = r_{ii}^{(1)} y_{i_{-1}} - A_i Y_i + r_{ii}^{(1)} A_i Y_{i_{-1}} - B_i X_i + r_{ii}^{(1)} B_i X_{i_{-1}} + e_i , \qquad i = 1, 2, \ldots, h. \tag{5}$$

Ignoring the fact that $Y_i$ and $e_i$ are correlated, equation (5) is a simple nonlinear equation in the coefficients $r_{ii}^{(1)}$, $A_i$, and $B_i$ and can be estimated by a variety of techniques. Two of the most common techniques are the Cochrane-Orcutt iterative technique and the Hildreth-Lu scanning technique, but any standard technique for estimating nonlinear equations can be used.[5] The technique used for the work below was the Cochrane-Orcutt technique.

2.5. *Two-stage least squares plus first-order serial correlation* (*TSLSAUTO1*). The fourth estimator considered is two-stage least squares applied to each equation of (5). This estimator accounts for both first-order serial correlation and simultaneous-equations bias and produces consistent estimates if $R^{(1)}$ is diagonal

[4] Two-stages least squares is taken here to mean the technique of replacing in the second-stage regressions the actual values of the right-hand-side variables with their predicted values from the first-stage regressions. The above discussion of the necessity of including certain variables as regressors in the first-stage regressions is not relevant for the instrumental-variable estimator in which the predicted values from the first stage regressions merely serve as instruments for the right-hand-side endogenous variables.

[5] In particular, the quadratic hill-climbing technique could be used. It is the author's experience, however, that for the serial-correlation problem the Cochrane-Orcutt technique converges faster than does the quadratic hill-climbing technique. Sargan [10] has shown that the Cochrane-Orcutt technique converges to a stationary value.

and $R^{(2)}$ is zero in (3). This estimator is discussed in [6], where it is shown that the following variables must be included as regressors in the first-stage regressions in order to insure consistent estimates of equation (5): $y_{i_{-1}}$, $Y_{i_{-1}}$, $X_i$, and $X_{i_{-1}}$. For the work below, these variables were always included in the set of regressors. Any standard nonlinear technique can be used for the second-stage regression of equation (5), and the technique used in this study was the Cochrane-Orcutt technique.

2.6. *Ordinary least squares plus first- and second-order serial correlation* (*OLSAUTO2*). The fifth estimator considered accounts for first- and second-order serial correlation of the error term $u_i$ in (2), but not for simultaneous-equations bias. The estimator is based on the assumption that the error term in each equation is determined as:

$$u_i = r_{ii}^{(1)} u_{i_{-1}} + r_{ii}^{(2)} u_{i_{-2}} + e_i\,, \qquad i = 1, 2, \ldots, h, \tag{6}$$

which means that $R^{(1)}$ and $R^{(2)}$ in (3) are assumed to be diagonal matrices. Under this assumption, equations (2) and (6) can be combined to yield:

$$\begin{aligned} y_i = {} & r_{ii}^{(1)} y_{i_{-1}} + r_{ii}^{(2)} y_{i_{-2}} - A_i Y_i + r_{ii}^{(1)} A_i Y_{i_{-1}} + r_{ii}^{(2)} A_i Y_{i_{-2}} \\ & - B_i X_i + r_{ii}^{(1)} B_i X_{i_{-1}} + r_{ii}^{(2)} B_i X_{i_{-2}} + e_i\,, \qquad i = 1, 2, \ldots, h. \end{aligned} \tag{7}$$

Again, ignoring the fact that $Y_i$ and $e_i$ are correlated, equation (7) is a simple nonlinear equation in the coefficients $r_{ii}^{(1)}$, $r_{ii}^{(2)}$, $A_i$, and $B_i$ and can be estimated by a variety of techniques. The Cochrane-Orcutt technique can be extended in an obvious way to the second-order case, and the extended Cochrane-Orcutt technique was the one used in this study. The technique converged quite rapidly in almost all cases.

2.7. *Two-stage least squares plus first- and second-order serial correlation* (*TSLSAUTO2*). The sixth estimator considered is two-stage least squares applied to each equation of (7). This estimator is an extension of the estimator discussed in [6] to the second-order case and produces consistent estimates if $R^{(1)}$ and $R^{(2)}$ are diagonal in (3). It is easy to show, following the analysis in [6], that the following variables must be included as regressors in the first-stage regressions in order to insure consistent estimates of equation (7): $y_{i_{-1}}$, $y_{i_{-2}}$, $Y_{i_{-1}}$, $Y_{i_{-2}}$, $X_i$, $X_{i_{-1}}$, and $X_{i_{-2}}$. For the work below, these variables were always included in the set of regressors. The nonlinear technique used for the second-stage regressions was the extension of the Cochrane-Orcutt technique to the second-order case.

2.8. *Full-information maximum likelihood* (*FIML*). The seventh estimator considered is full-information maximum likelihood applied to (1), where $R^{(1)}$ and $R^{(2)}$ in (3) are assumed to be zero. This estimator takes into account the contemporaneous correlations of the error terms $E$ in (3), and if $R^{(1)}$ and $R^{(2)}$ are zero in (3), it produces consistent and asymptotically efficient estimates.

2.9. *Full-information maximum likelihood plus first-order serial correlation* (*FIMLAUTO1*). The eighth estimator considered is full-information maximum likelihood applied to (1), where $R^{(1)}$ in (3) is assumed to be a diagonal matrix and where $R^{(2)}$ in (3) is assumed to be zero. If $R^{(1)}$ is diagonal and $R^{(2)}$ is zero in (3), this estimator produces consistent and asymptotically efficient estimates.

2.10. *Full-information maximum likelihood plus first- and second-order serial correlation (FIMLAUTO2).* The ninth estimator considered is full information maximum likelihood applied to (1), where $R^{(1)}$ and $R^{(2)}$ in (3) are assumed to be diagonal matrices. If $R^{(1)}$ and $R^{(2)}$ are diagonal matrices, this estimator produces consistent and asymptotically efficient estimates.

A practical method of estimating the model (1) by full-information maximum likelihood under the assumption of no serial correlation has been developed by Chow [**2**]. This method was extended in Chow and Fair [3] to cover the case of first- and higher-order auto-regressive properties of the error terms. This basic method was the one that was used to obtain the FIML, FIMLAUTO1, and FIMLAUTO2 estimates in the present study. For FIML there were 32 coefficients in $A$ and $B$ to be estimated simultaneously; for FIMLAUTO1 there were 40 coefficients to be estimated simultaneously—33 coefficients in $A$ and $B$ and 7 coefficients in $R^{(1)}$; and for FIMLAUTO2 there were 47 coefficients to be estimated simultaneously—33 coefficients in $A$ and $B$, 7 coefficients in $R^{(1)}$, and 7 coefficients in $R^{(2)}$. The method worked quite well in these three cases, and in none of the cases were any problems of convergence encountered. The method even succeeded in converging when all of the intial coefficient values were set equal to zero.

2.11. *Accounting for the dynamic nature of the model (DYN).* All of the estimators considered so far are based on the assumption that the values of the lagged endogeneous variables are known. This assumptions is, of course, not true in an actual forecasting situation, where values of the lagged endogenous variables are only known for the one-period-ahead forecast. After the one-period-ahead forecast, generated values of the lagged endogenous variables must be used. One way of trying to account for this dynamic nature of the model in the estimation of the coefficients of each equation is the following. Assume that the equation to be estimated is

$$(8) \qquad y_{1t} = \alpha_0 + \alpha_1 y_{1t-1} + \alpha_2 x_{1t} + \varepsilon_{1t}\,, \qquad t = 1, 2, \ldots, T,$$

and assume that one is interested in minimizing the two-period-ahead forecast error. Then the two-period-ahead forecast error can be minimized by solving for $y_{1t}$ in terms of $y_{1t-2}$ and the exogenous $x_1$ variable:

$$(9) \qquad y_{1t} = (\alpha_0 + \alpha_1\alpha_0) + \alpha_1^2 y_{1t-2} + \alpha_2 x_{1t} + \alpha_1\alpha_2 x_{1t-1} + \varepsilon_{1t} + \alpha_1\varepsilon_{1t-1}\,, \qquad t = 1, 2, \ldots, T,$$

and choosing those values of $\alpha_0$, $\alpha_1$, and $\alpha_2$ that minimize the sum of squared errors:[6]

[6] The suggestion that it might be useful to estimate dynamic models by minimizing the sum of multi-period forecast errors can be found in Klein [**8**, (56)].

$$(10) \qquad SSE = \sum_{t=1}^{T} [y_{1t} - (\alpha_0 + \alpha_1\alpha_0) - \alpha_1^2 y_{1t-2} - \alpha_2 x_{1t} - \alpha_1\alpha_2 x_{1t-1}]^2 .$$

Since equation (9) is nonlinear in the coefficients, the minimization of (10) requires that a nonlinear technique be used.

In a similar manner, if one is interested in minimizing the three-period-ahead forecast error, then $y_{1t}$ can be solved in terms of $y_{1t-3}$ and the exogenous $x_1$ variable and the resulting sum of squared errors minimized. This procedure can be followed for any $n$-period-ahead forecast error. For large values of $n$, of course, the equation for $y_{1t}$ becomes somewhat cumbersome. It is also necessary to have extra observations at the beginning of the sample period in order to use this procedure or else to shorten the sample size accordingly. It also should be noted that even though the error term in equation (8) may not be serially correlated, the error term in equation (9) is likely to be serially correlated. This is not of direct concern for the DYN procedure, however, since the procedure is concerned only with minimizing the $n$-period-ahead forecast error and not with accounting for the particular properties that the error terms may have.

The DYN procedure as just described is a single-equation procedure and has not accounted for the fact that other endogenous or lagged endogenous variables may be included as explanatory variables in equation (8). It is easy to solve the general model in (1) so that, for example, no one-period-lagged endogenous variables are included among the predetermined variables. This can be seen by rewriting (1) as

$$(11) \qquad AY + B^*X^* + CY_{-1} = U ,$$

where $X^*$ includes only exogenous variables, and then solving for $AY$ in terms of $Y_{-2}$ and the exogenous variables:

$$(12) \qquad AY + B^*X^* - CA^{-1}B^*X^*_{-1} - CA^{-1}CY_{-2} = U - CA^{-1}U_{-1} .$$

The generalization of the DYN procedure to the estimation of (12) would be to estimate the coefficient matrices $A$, $B^*$, and $C$ by the method of generalized least squares as discussed in Chow [**1**]. The coefficients $C$ and $A^{-1}$ in the error term would be ignored, and one would minimize $|VV'|/|AYY'A|$, where $V = AY + B^*X^* - CA^{-1}B^*X^*_{-1} - CA^{-1}CY_{-2}$. Unfortunately, there appears to be no practical way to carry out this minimization, and so at the present time this approach does not appear to be particularly fruitful. For the work below, therefore, the DYN procedure was merely used to estimate one equation at a time, with no account being taken of right-hand-side endogenous variables nor of lagged endogenous variables other than the lagged dependent variable in the equation.

It should be stressed that the single-equation DYN procedure is tested here not with the idea that the procedure should be used in practice, but to see if further work on this type of an estimator is warranted. If the results using the single-equation DYN procedure are good (it will be seen below that the results are in fact quite good), then considerable gain in multi-period forecasting accuracy may be made if practical ways can be found to estimate models using *full-information* DYN procedures.

The DYN procedure was actually used in this study to estimate each equation under the assumption of first-order serial correlation. For the two-period-ahead estimates, for example, equation (5) was solved, as in equation (9), so that $y_{i_{-1}}$ was not among the explanatory variables,[7] and the coefficients $r_{ii}^{(1)}$, $A_i$, and $B_i$ were estimated by minimizing the sum of squared errors of this equation. Similarly, for the three-period-ahead estimates, equation (5) was solved so that $y_{i_{-1}}$ and $y_{i_{-2}}$ were not among the explanatory variables, and the coefficients $r_{ii}^{(1)}$, $A_i$ $B_i$ were estimated by minimizing the sum of squared errors of this equation. Four- and five-period-ahead estimates were obtained in a similar manner. Because of the high degree of first-order serial correlation in some of the equations, it seemed best to estimate each equation under the serial-correlation assumption. In practice, all the assumption of first-order serial correlation does is to increase the complexity of the nonlinear equation being estimated.[8]

For the DYN estimates, the quadratic hill-climbing technique of Goldfeld, Quandt, and Trotter [7] was used. Even though the four- and five-quarter-ahead equations were quite nonlinear in the coefficients, no serious problems of any kind were encountered in minimizing the sum of squared errors of any of the equations. It turned out to be quite inexpensive to estimate the model by the DYN procedure using the quadratic hill-climbing technique.

2.12. *Conclusion.* This concludes the discussion of the ten estimators. Other estimators could have been considered, but in order to limit the size and cost of this study, the above ten estimators were chosen as some of the more important ones to consider. The performance of each of the estimators will, of course, depend on the characteristics of the data to which they are applied, and it is hoped that the results of this study will give an indication of the *quantitative* importance of each of the estimators when applied to the estimation of quarterly macroeconomic models of the United States economy.

## 3. THE EIGHT-EQUATION MODEL

The model used for the tests in this study is the simultaneous part of the forecasting model developed in [5]. The model is quarterly and consists of eight equations—seven equations explaining seven components of current dollar GNP and a GNP identity. The seven components are durable consumption, nondurable consumption, service consumption, plant and equipment investment,

[7] Remember that $y_{i_{-1}}$ may be included in the $X_i$ matrix in addition to its being included directly in the equation because of the serial-correlation assumption.

[8] If $y_{i_{-1}}$ is not in $X_i$ in equation (5) (i.e., if $y_{i_{-1}}$ enters as an explanatory variable in equation (5) only because of the serial-correlation assumption), then for the two- and four-period-ahead estimates, the estimate of $r_{ii}^{(1)}$ is not identified between $r_{ii}^{(1)}$ and $-r_{ii}^{(1)}$. In other words, the procedure determines only $(r_{ii}^{(1)})^2$ or $(r_{ii}^{(1)})^4$ and does not determine whether $r_{ii}^{(1)}$ is positive or negative. In practice, the positive value was used if the OLSAUTO1 estimate of $r_{ii}^{(1)}$ was positive and negative if the OLSAUTO1 estimate was negative.

TABLE 1

THE EIGHT-EQUATION MODEL

| Dependent Variable | Explanatory Variables | | | | |
|---|---|---|---|---|---|
| 1) CD | Const. | GNP | $MOOD_{-1}$ | $MOOD_{-2}$ | |
| 2) CN | | GNP | $MOOD_{-2}$ | $CN_{-1}$ | |
| 3) CS | | GNP | $MOOD_{-2}$ | $CS_{-1}$ | |
| 4) IP | Const. | GNP | PE2 | | |
| 5) IH | Const. | GNP | HSQ | $HSQ_{-1}$ | $HSQ_{-2}$ |
| 6) $V-V_{-1}$ | Const. | CD+CN | $CD_{-1}+CN_{-1}$ | $V_{-1}$ | |
| 7) IMP | Const. | GNP | | | |

8) $GNP = CD + CN + CS + IP + IH + V - V_{-1} - IMP + EX + G$

Notation:

CD = Durable-Consumption Expenditures
CN = Non-Durable-Consumption Expenditures
CS = Service-Consumption Expenditures
IP = Plant and Equipment Investment
IH = Nonfarm Housing Investment
$V - V_{-1}$ = Change in Total Business Inventories
IMP = Imports
GNP = Gross National Product
EX = Exports
G = Government Expenditures plus Farm Housing Investment
MOOD = Michigan Survey Research Center Index of Consumer Sentiment
PE2 = Two-quarter-ahead Expectation of Plant and Equipment Investment
HSQ = Quarterly Nonfarm Housing Starts
V = Stock of Total Business Inventories (arbitrary base period value of zero in 1953IV)

Note: The Subscript $-1$ or $-2$ after a variable denotes the one-quarter or two-quarter lagged value of the variable.

Basic set of instrumental variables =

$\{Const., MOOD_{-2}, PE2, G, CD_{-1}, CN_{-1}, CS_{-1}, V_{-1}, GNP_{-1}\}$.

Dummy variables D644 and D651 added to CD equation; dummy variables D593, D594, D644, and D651 added to $V - V_{-1}$ equation, and dummy variables D644, D651, D684, D691 and D692 added to IMP equation.

nonfarm housing investment, inventory investment, and imports. Government spending, exports and farm housing investment are taken to be exogenous. The model is presented in Table 1. There are four lagged endogenous variables among the predetermined variables: lagged durable consumption enters the inventory equation, lagged non-durable consumption enters the non-durable-consumption equation and the inventory equation, lagged service consumption enters the service-consumption equation, and the lagged stock of inventories enters the inventory equation. A detailed description of the eight-equation model is presented in [5], along with a description of the overall forecasting model, and this description will not repeated here.

For the work in [5], the model was estimated by TSLSAUTO1 for the basic 1956I–1969IV sample period. Observations for 1959III, 1959IV and 1960I were omitted from the sample period because of the steel strike, and observations for

1964IV, 1965I and 1965II were omitted from the sample period because of the automobile strike. Observations for 1968IV, 1969I, 1969II, and 1969III were also omitted from the sample period for the import equation because of a dock strike. In addition, for the non-durable-consumption and housing-investment equations, the beginning of the sample period was taken to be 1960II rather than 1956I because of an apparent structural shift in the non-durable-consumption equation around 1960 and because of lack of good data on housing starts before 1959. For the work here, no observations were omitted from the sample period because of strikes, but rather dummy variables were used in those equations most affected by the strikes. Two dummy variables—D644 and D651—were added to the durable-consumption equation; four dummy variables—D593, D594, D644, and D651—were added to the inventory equation; and five dummy variables—D644, D651, D684, D691, and D692—were added to the import equation.[9]

For the FIML, FIMLAUTO1 and FIMLAUTO2 estimators, the sample period was taken to be 1960II–1970III, for a total of 42 observations. For the other estimators, the sample period for all of the equations except the non-durable-consumption and housing-investment equations was taken to be 1956I–1970III, for a total of 59 observations. For the non-durable-consumption and housing-investment equations the shorter 1960II–1970III period was retained.[10] In addition, TSLSAUTO1 estimates were obtained for all of the equations for the 1960II–1970III period to allow a direct comparison between these estimates and the FIMLAUTO1 estimates to be made.

The 1960II–1970III period was used for the predictions. Since only information within the 1960II–1970III period was used to compute the FIML, FIMLAUTO1, and FIMLAUTO2 estimates, there may be a bias in favor of these estimates relative to the others, but it will be seen below that this bias does not appear to be very large. The probable size of the bias can be determined by comparing the results obtained from the TSLSAUTO1 estimates based on the 1956I–1970III period with the results obtained from the TSLSAUTO1 estimates based only on the 1960II–1970III period.

The basic set of instrumental variables used for the two-stage least squares estimators is presented at the bottom of Table 1. In addition, as mentioned above, other variables were added to this basic set for each equation when their addition was a necessary condition for the estimates to be consistent. The variables that were added for each equation are listed in Table 2. For all of the stochastic equations of the model except the inventory equation, the endogenous variable on the right-hand-side of the equation is the GNP variable. For the inventory equation, the endogenous variable is the sum of durable and non-

[9] D593 denotes a variable that takes on a value of one in the third quarter of 1959 and zero otherwise; D594 denotes a variable that takes on a value of one in the fourth quarter of 1959 and zero otherwise; and so on.

[10] For the DYN estimates of the housing-investment equation, the sample period was taken to be 1960IV–1970III, for a total of 40 observations. This was done because of lack of data on housing starts before 1959, data that would have been needed for the four- and five-quarter-ahead DYN estimates had the sample period begun in 1960II.

TABLE 2

INSTRUMENTAL VARABLES USED IN EACH EQUATION IN ADDITION TO THOSE IN THE BASIC SET

| Dependent Variable | Estimator | Additional Instrumental Variables |
|---|---|---|
| CD | TSLS | MOOD$_{-1}$, D644, D651 |
| | TSLSAUTO1 | MOOD$_{-1}$ MOOD$_{-3}$, D644, D651, D644$_{-1}$ |
| | TSLSAUTO2 | CD$_{-2}$, GNP$_{-2}$, MOOD$_{-1}$, MOOD$_{-3}$, MOOD$_{-4}$, D644, D651, D644$_{-1}$, D644$_{-2}$ |
| CN | TSLS | none |
| | TSLSAUTO1 | CN$_{-2}$, MOOD$_{-3}$ |
| | TSLSAUTO2 | CN$_{-2}$, CN$_{-3}$, GNP$_{-2}$, MOOD$_{-3}$, MOOD$_{-4}$ |
| CS | TSLS | none |
| | TSLSAUTO1 | CS$_{-2}$, MOOD$_{-3}$ |
| | TSLSAUTO2 | CS$_{-2}$, CS$_{-3}$, GNP$_{-2}$, MOOD$_{-3}$, MOOD$_{-4}$ |
| IP | TSLS | none |
| | TSLSAUTO1 | IP$_{-1}$,PE2$_{-1}$ |
| | TSLSAUTO2 | IP$_{-1}$, IP$_{-2}$, GNP$_{-2}$, PE2$_{-1}$, PE2$_{-2}$ |
| IH | TSLS | HSQ, HSQ$_{-1}$, HSQ$_{-2}$ |
| | TSLSAUTO1 | HSQ, HSQ$_{-1}$, HSQ$_{-2}$, HSQ$_{-3}$, IH$_{-1}$ |
| | TSLSAUTO2 | HSQ, HSQ$_{-1}$, HSQ$_{-2}$, HSQ$_{-3}$, HSQ$_{-4}$, IH$_{-1}$, IH$_{-2}$, GNP$_{-2}$ |
| V-V$_{-1}$ | TSLS | D593, D594, D644, D651, MOOD$_{-1}$ |
| | TSLSAUTO1 | D593$^{a}$, D594$^{a}$, D644, D651, D593$^{a}_{-1}$, D644$_{-1}$, V$_{-2}$, CD$_{-2}$, CN$_{-2}$, MOOD$_{-1}$, MOOD$_{-3}$ |
| | TSLSAUTO2 | D593, D594, D644, D651, D593$_{-1}$, D593$_{-2}$, D644$_{-1}$, D644$_{-2}$, V$_{-2}$, V$_{-3}$, CD$_{-2}$, CD$_{-3}$, CN$_{-2}$, CN$_{-3}$, MOOD$_{-1}$, MOOD$_{-3}$, MOOD$_{-4}$ |
| IMP | TSLS | D644, D651, D684, D691, D692 |
| | TSLSAUTO2 | D644, D651, D684, D691, D692, D644$_{-1}$, D684$_{-1}$, IMP$_{-1}$ |
| | TSLSAUTO2 | D644, D651, D684, D691, D692, D644$_{-1}$, D644,$_{-2}$ D684$_{-1}$, D684$_{-2}$, IMP$_{-1}$, IMP$_{-2}$, GNP$_{-2}$ |

[a] Not used as an instrument for TSLSAUTO1, 42 observations.

durable consumption. Because the MOOD variable is important in determining durable and non-durable consumption, extra lagged values of this variable were used as instruments in the estimation of the inventory equation. Otherwise, the only variables added to the basic set of instruments for each equation were the ones necessary for consistency.[11] Some of the work in [6] indicates that the small sample properties of two-stage least squares estimators may be adversely affected by the use of a large number of instrumental variables in the first stage regressions, and thus an attempt was made in this study to keep the basic set of instrumental variables fairly small. The reason that G in Table 1 was used as

[11] It should be noted when examining the additional instrumental variables used that, for example, D644 and D651$_{-1}$ are the same variable and so are obviously not both included in the list of instruments.

a basic instrument rather than G + EX is because of the pronounced effect that the dock strike in 1968 and 1969 has on exports.

## 4. THE RESULTS

4.1. *The coefficient estimates.* The results of estimating the seven stochastic equations of the model are available from the author on request. First-order serial correlation tended to be fairly pronounced in most of the equations, whereas second-order serial correlation tended to be much less pronounced. The FIMLAUTO1 and FIMLAUTO2 estimates tended to differ from the other estimates more than the other estimates tended to differ from each other. The service-consumption equation was the equation in which the estimates differed the least from each other, and the inventory-investment equation was the equation in which the estimates differed the most from each other.

4.2. *The within-sample predictions.* Given a set of estimates of the $A$, $B$, and $R$ matrices in (1) and (2), one can solve for the reduced form for $Y$. From (1) and (3) the reduced form for $Y$ is

$$Y = -A^{-1}BX + A^{-1}R^{(1)}AY_{-1} + A^{-1}R^{(1)}BX_{-1} + A^{-1}R^{(2)}AY_{-2} + A^{-1}R^{(2)}BX_{-2} + A^{-1}E. \tag{13}$$

For each of the sets of estimates, within-sample predictions of the eight endogenous variables in $Y$ were generated for the 1960II–1970III period using (13) and the assumption that $E$ is zero. One- through five-quarter-ahead predictions were generated, as well as one prediction over the whole sample period.[13] For each variable, there were 42 one-quarter-ahead predictions generated, 41 two-quarter-ahead predictions, 40 three-quarter-ahead predictions, 39 four-quarter-ahead predictions, and 38 five-quarter-ahead predictions. The one prediction over the whole sample period began in 1960II and consisted of 42 observations. For all of the predictions, generated values of the lagged endogenous variables were used after the one-quarter-ahead prediction.[14]

Two error measures were computed for each set of predictions: the root mean square error in terms of levels,

[13] For the DYN estimates, only two- through five-quarter-ahead predictions were generated since these were the only relevant predictions for the estimates. The two-quarter-ahead DYN estimates were used for the two-quarter-ahead predictions, the three-quarter-ahead DYN estimates were used for the three-quarter-ahead predictions, and so on. It should be noted that, for example, the two-quarter-ahead predictions based on the two-quarter-ahead DYN estimates were not used as inputs for the three- quarter-ahead predictions. The latter were generated using only the three-quarter-ahead DYN estimates.

[14] For the one-quarter-ahead prediction, unrestricted reduced-from estimates of each equation in (13) would by the property of least squares yield the smallest sum of squared errors for each equation. This is not necessarily true for the two-quarter-ahead predictions and beyond, however, and in general unrestricted reduced-form estimation is not of much interest, since for models only slightly larger than the model considered in this study, there are likely to be more variables on the right-hand-side of (13) than there are observations, which would make unrestricted reduced-form estimation impossible.

TABLE 3

ROOT MEAN SQUARE ERRORS FOR GNP

| | RMSE | | | | | | RMSEΔ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Estimator | One Quarter ahead 42obs | Two Quarters ahead 41obs | Three Quarters ahead 40obs | Four Quarters ahead 39obs | Five Quarters ahead 38obs | Entire sample Period 42obs | Two Quartear ahead 41obs | Three Quarters ahead 40obs | Four Quartear ahead 39obs | Five Quarters ahead 38obs | Entire Sample Period 42obs |
| OLS | 6.54 | 6.90 | 7.18 | 7.51 | 7.69 | 8.45 | 4.10 | 4.12 | 4.13 | 4.17 | 4.14 |
| TSLS | 6.92 | 6.56 | 6.81 | 7.21 | 7.43 | 8.07 | 4.96 | 4.59 | 4.61 | 4.66 | 4.60 |
| FIML | 4.97 | 4.84 | 5.13 | 5.34 | 5.62 | 7.38 | 3.59 | 3.57 | 3.63 | 3.66 | 3.62 |
| OLSAUTO1 | 3.19 | 4.82 | 5.80 | 6.44 | 6.94 | 8.61 | 3.48 | 3.53 | 3.48 | 3.54 | 3.69 |
| TSLSAUTO1 | 3.27 | 4.95 | 5.86 | 6.42 | 6.75 | 8.27 | 3.48 | 3.54 | 3.49 | 3.60 | 3.73 |
| TSLSAUTO1[a] | 3.16 | 4.85 | 5.92 | 6.66 | 7.18 | 7.88 | 3.46 | 3.50 | 3.41 | 3.48 | 3.59 |
| FIMLAUTO1 | 2.48 | 3.10 | 3.96 | 4.79 | 5.87 | 6.55 | 3.20 | 3.45 | 3.56 | 3.56 | 3.60 |
| OLSAUTO2 | 3.08 | 4.62 | 5.54 | 6.14 | 6.61 | 8.40 | 3.36 | 3.43 | 3.38 | 3.51 | 3.70 |
| TSLSAUTO2 | 3.03 | 4.53 | 5.37 | 5.90 | 6.28 | 7.78 | 3.31 | 3.39 | 3.35 | 3.48 | 3.64 |
| FIMLAUTO2 | 2.30 | 3.01 | 4.02 | 4.95 | 6.39 | 6.99 | 3.22 | 3.55 | 3.69 | 3.70 | 3.79 |
| DYN | *b* | 4.57 | 4.92 | 5.31 | 5.43 | *c* | 3.37 | 3.58 | 3.67 | 3.65 | *c* |

[a] Based on 42 observations only.
[b] Same as OLSAUTO1.
[c] Not relevant since DYN estimates only computed up to five-quarter-ahead predictions.

$$\text{RMSE} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(y_t - \hat{y}_t)^2}$$

and the root mean square error in terms of changes,

$$\text{RMSE}\Delta = \sqrt{\frac{1}{T}\sum_{t=1}^{T}[(y_t - y_{t-1}) - (\hat{y}_t - \hat{y}_{t-1})]^2}$$

where $y_t$ denotes the actual value of $y$ for quarter $t$ and $\hat{y}_t$ denotes the predicted value of $y$ for quarter $t$. RMSEΔ is a measure of how well the model has explained the change in the endogenous variable. For the one-quarter-ahead prediction, RMSE and RMSEΔ are the same.

The root mean square errors for the GNP variable are presented in Table 3 for each set of estimates. The errors are in billions of dollars. The results in the table are fairly self-explanatory, and only a brief description of them will be presented here. The most striking feature of the RMSE results is perhaps the increased accuracy obtained from the FIMLAUTO1 and FIMLAUTO2 estimates for all but the five-quarter-ahead prediction. This increased accuracy does not appear attributable to the use of the shorter sample period to compute the FIMLAUTO1 and FIMLAUTO2 estimates since the results from the two sets of TSLSAUTO1 estimates are quite close. For the five-quarter-ahead prediction, FIMLAUTO1 and FIMLAUTO2 perform well, but they do not completely dominate the field as they do for the other predictions.

For the RMSE results the DYN estimates perform consistently well and always do better than the OLSAUTO1 estimates, to which they are comparable in terms of accounting for first-order serial correlation but not for the simultaneous nature of the model. For the three- and four-quarter-ahead predictions, the DYN estimator outperforms all estimators except FIMLAUTO1 and FIMLAUTO2, and for the five-quarter-ahead prediction, the DYN estimator outperforms even FIMLAUTO1 and FIMLAUTO2.

The RMSE results in Table 3 also show that the two-stage least squares estimators perform on average better than their ordinary least squares counterparts, that the full-information maximum likelihood estimators perform on average better than their two-stage least squares counterparts, that the AUTO1 estimators perform on average better than their no-serial-correlation counterparts, and that the AUTO2 estimators perform on average better than their AUTO1 counterparts: TSLS is better than OLS except for the one-period-ahead prediction, TSLSAUTO1 is better than OLSAUTO1 for the four-quarter-ahead prediction and beyond, and TSLSAUTO2 is always better than OLSAUTO2; FIML is always better than TSLS, FIMLAUTO1 is always better than TSLSAUTO1, and FIMLAUTO2 is better than TSLSAUTO2 except for the five-quarter-ahead prediction; OLSAUTO1 is better than OLS except for the prediction over the whole sample period, TSLSAUTO1 is better than TSLS except for the prediction over the whole sample period, and FIMLAUTO1 is better than FIML except for the five-quarter-ahead prediction; OLSAUTO2 is always better than OLSAUTO1, TSLSAUTO2 is always better than TSLSAUTO1, and FIMLAUTO2 is better than FIMLAUTO1 for the one- and two-quarter-ahead predictions.

With respect to the quantitative importance of the various estimators, the gain in going from an ordinary least squares estimator to a two-stage least squares estimator appears to be less than the gain in going from a two-stage least squares estimator to a full-information maximum likelihood estimator, and the gain in going from an AUTO1 estimator to an AUTO2 estimator appears to be less than the gain in going from a no-serial-correlation estimator to an AUTO1 estimator. Also, the gain in going from OLS to TSLS appears to be less than the gain in going from OLS to OLSAUTO1, so that if one had a choice between TSLS and OLSAUTO1, but not TSLSAUTO1, then OLSAUTO1 would appear to be the estimator to use. The gain in going from a no-serial-correlation estimator to an AUTO1 or AUTO2 estimator generally lessens as the period to be predicted moves further away from the starting point. Finally, the gain in using the full-information maximum likelihood estimators or the DYN estimator appears to be quite large.[15]

[15] Klein, in his study [9] using the (annual) Klein-Goldberger model, compared the accuracy of ordinary least squares, two-stage least squares using four and eight principal components, and full-information maximum likelihood. None of the estimators accounted for serial correlation. Klein found that the full-information estimator gave on average slightly better results for

(*Continued on next page*)

TABLE 4

ROOT MEAN SQUARE ERRORS FOR $V - V_{-1}$

| | RMSE | | | | | | RMSEΔ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Estimator | One Quarter ahead 42obs | Two Quarters ahead 41obs | Three Quarters ahead 40obs | Four Quarters ahead 39bos | Five Quarters ahead 38obs | Entire Sample Period 42obs | Two Quarters ahead 41obs | Three Quarters ahead 40obs | Four Quarters ahead 39obs | Five Quarters ahead 38obs | Entire Sample Period 42obs |
| OLS | 3.82 | 3.78 | 3.75 | 3.63 | 3.50 | 4.09 | 3.63 | 3.73 | 3.68 | 3.72 | 3.63 |
| TSLS | 4.12 | 3.64 | 3.58 | 3.49 | 3.40 | 3.85 | 4.22 | 3.90 | 3.87 | 3.91 | 3.79 |
| FIML | 3.89 | 3.63 | 3.60 | 3.47 | 3.34 | 3.90 | 3.70 | 3.79 | 3.75 | 3.79 | 3.69 |
| OLSAUTO1 | 2.59 | 3.34 | 3.49 | 3.40 | 3.28 | 4.17 | 3.49 | 3.67 | 3.69 | 3.68 | 3.66 |
| TSLSAUTO1 | 2.62 | 3.38 | 3.53 | 3.43 | 3.31 | 4.23 | 3.46 | 3.65 | 3.67 | 3.65 | 3.65 |
| TSLSAUTO1[a] | 2.62 | 3.31 | 3.48 | 3.45 | 3.36 | 4.10 | 3.46 | 3.67 | 3.67 | 3.67 | 3.65 |
| FIMLAUTO1 | 2.77 | 3.00 | 3.10 | 3.11 | 3.18 | 3.54 | 3.56 | 3.77 | 3.79 | 3.85 | 3.72 |
| OLSAUTO2 | 2.65 | 3.34 | 3.49 | 3.38 | 3.23 | 4.26 | 3.48 | 3.70 | 3.73 | 3.69 | 3.73 |
| TSLSAUTO2 | 2.64 | 3.30 | 3.44 | 3.34 | 3.18 | 4.13 | 3.46 | 3.70 | 3.72 | 3.69 | 3.71 |
| FIMLAUTO2 | 2.74 | 2.90 | 3.07 | 3.10 | 3.26 | 3.61 | 3.66 | 3.81 | 3.82 | 3.91 | 3.79 |
| DYN | *b* | 3.28 | 3.38 | 3.30 | 3.24 | *c* | 3.49 | 3.81 | 3.86 | 3.90 | *c* |

[a] Based on 42 observations only.
[b] Same as OLSAUTO1.
[c] Not relevant since DYN estimates only computed up to five-quarter-ahead predictions.

For the RMSEΔ error measure in Table 3, the results of the various estimators are closer. The OLS and TSLS estimators continue to perform poorly relative to the others, but for the other eight estimators the RMSEΔ results are quite close, and no one estimator can be considered as dominating all of the rest. The closeness of the RMSEΔ results can probably be explained by the fact that none of the estimators is explicitly designed to minimize the errors in terms of changes. The full-information maximum likelihood estimators, for example, maximize the likelihood function in terms of the levels of the endogenous variables and not in terms of their changes. It is thus not too surprising that the conclusions reached from examining the RMSE results do not carry over completely to the RMSEΔ results.

In Table 4 the root mean square errors for inventory investment are presented for each set of estimates. Generally, the basic conclusions reached for the GNP results also hold for the inventory results, although the inventory results tend to be somewhat closer. It is interesting to note that for the RMSE results the full-

(*Continued*)
the one-period-ahead prediction, but considerably poorer results for the prediction over the whole period. The results in the present study are thus much more optimistic than Klein's results regarding the gain that can be achieved using full-information estimators. I have been informed, however, that at least part of the poor results for the prediction over the whole sample period from the Klein-Goldberger model was due to a computing error.

information maximum likelihood estimators do not dominate their two-stage least squares counterparts as much for inventory investment as they did for GNP. This general pattern was true for the other six GNP components as well. The large gain from using the full-information maximum likelihood estimators comes in terms of predicting GNP and not in terms of predicting the individual components of GNP. This result can probably be explained by the fact that the full-information maximum likelihood estimators are the only estimators that take into account the covariance of the error terms in the $E$ matrix in (3). Since GNP is merely the sum of its individual components, one would expect that the full-information maximum likelihood estimators, by not merely minimizing the sum of the individual error variances, would perform best in terms of predicting GNP.

There were no unusual features of the results for the other components of GNP, other than the one just cited about the full-information maximum likelihood estimators, and so these results will not be presented here.

4.3. *Conclusion.* The results in this section indicate that considerable gain in forecasting accuracy can be achieved by the use of more advanced estimation techniques. Certainly, accounting for first-order serial correlation is important, and even more gain appears possible by accounting for second-order serial correlation and by using a two-stage least squares technique as opposed to its ordinary least squares counterpart. The results also indicate that considerable gain can be achieved by using full-information maximum likelihood estimators and by accounting for the fact that values of lagged endogenous variables are not known after the one-quarter-ahead forecast. Given the success of the single-equation DYN procedure and of the full-information maximum likelihood estimators, if practical ways can be found to implement full-information DYN procedures, a significant increase in forecasting accuracy may result.

As mentioned in the Introduction, the conclusions of this study are based heavily on the premise that the basic properties of macroeconomic models are similar and that the particular model used is a good representative of macroeconomic models.[16] To the extent that this premise is valid, the results give an indication of the relative usefulness of the various estimators for multi-period forecasting purposes. More work on larger models is needed, however, before too much force should be put on the detailed conclusions of this study. Further work is also needed to see if any of the conclusions need modification when outside-sample predictions are considered.[17] In general, however, the results do indicate that considerable forecasting accuracy can be achieved by using more complicated techniques than simple ordinary least squares or two-stage least squares to estimate macroeconomic models. Because of the increased feasibility

[16] The root mean square errors presented in Table 3 are quite low relative to the results from previous models (see, for example, Evans, Haitovsky, and Treyz [4] or the results for the Klein-Goldberger model in Klein [9]), which in itself is encouraging but cannot be used in any rigorous way to argue that the present model is a good representative model.

[17] Some preliminary work in comparing within-sample and outside-sample predictions can be found in [5, (Chapters 11-13)].

of using more advanced estimation techniques, there is now less need for model builders to limit themselves to simpler techniques if more advanced techniques lead to improved results.

*Princeton University, U. S. A.*

REFERENCES

[1] CHOW, GREGORY C., "A Comparison of Alternative Estimators for Simultaneous Equations," *Econometrica*, XXXII (October, 1964), 532–553.

[2] ———, "Two Methods of Computing Full-Information Maximum Likelihood Estimates in Simultaneous Stochastic Equations," *International Economic Review*, IX (February, 1968), 100–112.

[3] ———, AND RAY C. FAIR, "Maximum Likelihood Estimation of Linear Equation Systems with Auto-Regressive Residuals," *Annals of Economic and Social Measurement*, II (January, 1973), 17–28.

[4] EVANS, MICHAEL K., YOEL HAITOVSKY, AND GEORGE I. TREYZ, "An Analysis of the Forecasting Properties of U.S. Econometric Models," paper presented to the Conference on Research in Income and Wealth, Cambridge, Massachusetts, November, 1969.

[5] FAIR, RAY C., *A Short-Run Forecasting Model of the United States Economy* (Lexington: D. C. Heath and Company, 1971).

[6] ———, "The Estimation of Simultaneous Equation Models with Lagged Endogenous Variables and First Order Serially Correlated Errors," *Econometrica*, XXXVIII (May, 1970), 507–516.

[7] GOLDFELD, S.M., R.E. QUANDT, and H.F. TROTTER, "Maximization by Quadratic Hill-Climbing," *Econometrica*, XXXIV (July, 1966), 541–551.

[8] KLEIN, L.R., *An Essay on the Theory of Economic Prediction* (Helsinki: The Academic Book Store, 1968).

[9] ———, "Estimation of Interdependent Systems in Macroeconomics," *Econometrica*, XXXVII (April, 1969), 171–192.

[10] SARGAN, J.D., Wages and Prices in the United Kingdom: A Study in Econometric Methodology," in P. E. Hart, G. Mills, and J. K. Whitaker eds., *Econometric Analysis for National Economic Planning* (London: The Colston Research Society, 1964), 25–54.