A COMPARISON OF ALTERNATIVE ESTIMATORS OF MACROECONOMIC MODELS

Ray C. Fair

ECONOMETRIC RESEARCH PROGRAM
Research Memorandum No. 121

March 1971

The research described in this paper was supported by NSF Grant GS-1840 and the computer work by NSF Grant GJ-34.

Econometric Research Program
PRINCETON UNIVERSITY
207 Dickinson Hall
Princeton, New Jersey

A COMPARISON OF ALTERNATIVE ESTIMATORS OF MACROECONOMIC MODELS

Ray C. Fair

ECONOMETRIC RESEARCH PROGRAM
Research Memorandum No. 121

March 1971

The research described in this paper was supported by NSF Grant GS-1840 and the computer work by NSF Grant GJ-34.

Econometric Research Program
PRINCETON UNIVERSITY
207 Dickinson Hall
Princeton, New Jersey

A COMPARISON OF ALTERNATIVE ESTIMATORS OF MACROECONOMIC MODELS

Ray C. Fair
Department of Economics, Princeton University

I. INTRODUCTION

Three properties that characterize most macroeconomic models are the simultaneous determination of the endogenous variables, the inclusion of lagged endogenous variables among the predetermined variables, and serial correlation of the error terms. The purpose of this paper is to provide a comparison of alternative estimators of such models. Ten basic estimators are considered. Each estimator is first used to estimate the structural equations of an eight-equation macroeconomic model. The reduced form of the model is then solved for each set of estimates, and ex post predictions of the eight endogenous variables of the model are generated over the sample period. Both static (one-period-ahead) predictions and dynamic predictions are generated. The estimators are compared in terms of the accuracy of the ex post predictions. The ten estimators are discussed in Section II, the eight-equation model is discussed in Section III, and the results are presented in Section IV.

It should be noted at the outset that this study is not a Monte Carlo study. An actual model of the United States economy is used to perform the tests, and so the true structure and coefficients of the model are not known.[1] This study is rather based on the premise that the basic properties of macroeconomic models are similar enough so that conclusions obtained from the use of one model can be generalized to other

---

[1]The methodology of this study is thus similar to that of Klein's study [9], where four sets of estimates of the Klein-Goldberger model are compared.

models. To the extent that this premise is valid and to the extent that the model used in this study is an adequate representative of macroeconomic models, the results should give a good indication of the relative usefulness of the various estimators.

II. THE TEN ESTIMATORS

The General Model

The general model to be estimated is

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

where Y is an hxT matrix of endogenous variables, X is a kxT matrix of predetermined (both exogenous and lagged endogenous) variables, U is an hxT matrix of error terms, and A and B are hxh and hxk coefficient matrices respectively. T is the number of observations. The i^{th} equation of the model will be written as

$$y_i = -A_iY_i - B_iX_i + u_i \quad , \; i = 1,2,\ldots,h \quad , \tag{2}$$

where y_i is a 1xT vector of values of y_{it}, Y_i is an h_ixT matrix of endogenous variables (other than y_i) included in the i^{th} equation, X_i is a k_ixT matrix of predetermined variables included in the i^{th} equation, u_i is a 1xT vector of error terms, and A_i and B_i are 1xh_i and 1xk_i vectors of coefficients corresponding to the relevant elements of A and B respectively.

The error terms in U are assumed to follow an n^{th} order autoregressive process:

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

where the R matrices are hxh coefficient matrices, E is an hxT 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 value, to be contemporaneously correlated but not serially correlated, and to be uncorrelated in the limit with the predetermined, lagged predetermined, and lagged endogenous variables.

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).

Ordinary Least Squares plus Lagged Dependent Variables (OLSLDV)

The second estimator considered is ordinary least squares applied to each equation of (2), where each equation of (2) is expanded, when necessary, to include the lagged dependent variable ($y_{i_{-1}}$) among the predetermined variables. In the specification of most macroeconomic models, lagged dependent variables are not included in every equation, but including the lagged dependent variable in an equation can be considered to be a crude way of accounting for first order serial correlation.

Two Stage Least Squares (TSLS)

The third 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 or if there are no lagged endogenous variables in X_i; otherwise not. With a large enough sample, all of the variables in X should be used as instruments 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 instruments. 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 instrumental variables. 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 included predetermined variables. For the work below, therefore, the variables in X_i were always included in the set of instrumental variables when the i^{th} equation of (2) was estimated by TSLS. The other instruments that were used will be discussed in Section III.

Two Stage Least Squares plus Lagged Dependent Variables (TSLSLDV)

The fourth estimator considered is two stage least squares applied to each equation of (2), where each equation of (2) is expanded, when necessary, to include the lagged dependent variable among the predetermined variables. This is the two stage least squares analogue of OLSLDV.

Ordinary Least Squares plus First Order Serial Correlation (OLSAUTO1)

The fifth estimator considered accounts for first order serial correlation of the error term u_i in (2), but not for simultaneous equation bias. The estimator is based on the assumption that the error term in each equation is first order serially correlated:

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

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

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

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. The technique used for the work below was the Cochrane-Orcutt technique.

Two Stage Least Squares plus First Order Serial Correlation (TSLSAUT01)

The sixth estimator considered is two stage least squares applied to each equation of (5). This estimator accounts for both first order serial correlation and simultaneous equation bias and produces consistent estimates if $R^{(1)}$ is diagonal and the other R matrices are zero in (3). This estimator is discussed in Fair [6], where it is shown that the following variables must be included as instruments in the first stage regression 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 instrumental variables. Any standard nonlinear technique can be used to estimate the second stage regression of equation (5), and the technique used in this study was the Cochrane-Orcutt technique.

Ordinary Least Squares plus First and Second Order Serial Correlation (OLSAUTO2)

The seventh estimator considered accounts for first and second order serial correlation of the error term u_i in (2), but not for simultaneous equation bias. The estimator is based on the assumption that the error term in each equation is determined as:

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

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

$$(7) \qquad y_i = r_{ii}^{(1)} y_{i_{-1}} + r_{ii}^{(2)} y_{i_{-2}} - A_i Y_i + r_{ii}^{(1)} A_i Y_{i_{-1}} + r_{ii}^{(2)} A_i Y_{i_{-2}}$$

$$- B_i X_i + r_{ii}^{(1)} B_i X_{i_{-1}} + r_{ii}^{(2)} B_i X_{i_{-2}} + e_i \quad , \quad i = 1,2,\ldots,h \quad .$$

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 technique used in this study is an extension of the Cochrane-Orcutt technique to the second order case and is briefly described in the Appendix.

Two Stage Least Squares plus First and Second Order Serial Correlation (TSLSAUTO2

The eighth estimator considered is two stage least squares applied to each equation of (7). This estimator is an extension of the estimator discussed in Fair [6] to the second order case and produces consistent estimates if $R^{(1)}$ and $R^{(2)}$ are diagonal and the other R matrices are zero in (3). The estimator is discussed in the Appendix, where it is shown that the following variables must be included as instruments in the first stage regression 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 instrumental variables. The nonlinear technique used in the second stage regression was the extension of the Cochrane-Orcutt technique to the second order case.

Full Information Maximum Likelihood plus First Order Serial Correlation (FIMLAUTO1)

The ninth estimator considered is full information maximum likelihood applied to (1), where $R^{(1)}$ in (3) is assumed to be a diagonal matrix and where the other R matrices in (3) are assumed to be zero. The technique for estimating the general model, (1) and (3), by full information maximum likelihood is developed in Chow and Fair [2]. This estimator uses all of the information in the system and produces consistent estimates. In the present use of the technique, consistent estimates are produced if $R^{(1)}$ is diagonal and the other R matrices are zero in (3).

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 endogenous variables are known. This assumption 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) $$y_{1t} = \alpha_o + \alpha_1 y_{1t-1} + \alpha_2 x_{1t} + \epsilon_{1t} \quad , \quad t = 1,2,\ldots,T \quad ,$$

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) $$y_{1t} = (\alpha_o + \alpha_1 \alpha_o) + \alpha_1^2 y_{1t-2} + \alpha_2 x_{1t} + \alpha_1 \alpha_2 x_{1t-1} + \epsilon_{1t} \quad , \quad t = 1,2,\ldots,T \quad ,$$

and choosing those values of α_o, α_1, and α_2 that minimize the sum of squared errors:[2]

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

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 for 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, although for large values of n 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

---

[2] The suggestion that it might be useful to estimate dynamic models by minimizing the sum of multiperiod forecast errors can be found in Klein [8], p. 56.

though the error term in equation (8) may not be serially correlated, the error term in equations like (9) will be serially correlated. This is not of direct concern for the DYN procedure, however, since the procedure is concerned only with minimizing the sum of squared errors and not with accounting for the particular properties that the error terms may have. It is purely a best-fitting procedure.

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

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

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

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

Unfortunately, there appears to be no easy way to estimate the coefficient matrices A, B* and C of (12) by generalized least squares,[3] and there thus appears to be no practical way to use the DYN procedure to estimate more than one equation at a time.

For the work below, the DYN procedure was used to estimate each equation of the model under the assumption of first order serial

---

[3]The concept of generalized least squares is discussed in Chow [1].

correlation of the error term. For the two-period-ahead estimates, for example, equation (5) was solved, as above, so that $y_{i_{-1}}$ was not among the explanatory variables,[4] 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 and B_i were estimated by minimizing the sum of squared errors of this equation. Four- and five-period-ahead estimates were also made in a similar manner. Because of the high degree of first order serial correlation in some of the equations, it seemed best to choose (5) as the basic equation determining y_i and to apply the DYN procedure to this basic equation. In practice, all the assumption of first order serial correlation does is to increase the complexity of the nonlinear equation being estimated.[5]

The single-equation DYN procedure assumes that the variables in $Y_{i_{-1}}$ in equation (5) are known and that the variables in Y_i are not endogenous. There is thus no guarantee that the DYN estimates will yield better results when used within the context of the overall model.

---

[4]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.

[5]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.

Conclusion

This concludes the discussion of the ten estimators. Since on theoretical grounds no one estimator can be considered as dominating all of the rest, it is largely an empirical question as to how well each set of estimates will do in terms of predicting the endogenous variables. The accuracy of each of the estimators will depend on the characteristics of the model being estimated. If second order serial correlation is important, then OLSAUTO2 and TSLSAUTO2 should do well relative to the other estimators; if first, but not second, order serial correlation is important, then OLSAUTO1, TSLSAUTO1, and FIMLAUTO1 should do well; if the simultaneous nature of the model is important, then FIMLAUTO1 and the two stage least squares estimators should do well; and if the dynamic nature of the model is important, then DYN may do well. It is very likely that the OLS and OLSLDV estimators will be dominated by the other estimators, but the ordinary least squares estimators were included in the analysis in an attempt to determine the quantitative importance of each of the other estimators relative to ordinary least squares.

III. THE EIGHT-EQUATION MODEL

The model used for the tests in this study is the simultaneous part of the forecasting model developed in Fair [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, non-durable consumption, service consumption, plant and equipment investment, 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

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

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 =

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

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 be 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 nondurable 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 nondurable 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.[6]

For the FIMLAUTO1 estimator, the sample period was taken to be 1960II-1970III, for a total of 42 observations. For the other estimators,

---

[6] 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.

the sample period was taken to be 1956I-1970III, for a total of 59 observations, except for the nondurable consumption and housing investment equations, where the shorter 1960II-1970III period was retained.[7] 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 ex post predictions. There may thus be a bias in favor of the FIMLAUTO1 estimates relative to the others, since only information within the 1960II-1970III period was used to compute the estimates, 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 at the bottom of Tables 2-1 through 2-7 in the next section. For all of the equations 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 nondurable consumption. Because the MOOD variable is important in determining durable and nondurable consumption, extra lagged values of this variable were used

---

[7]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.

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.[8] Some of the work in Fair [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 regression, 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 a basic instrument rather than G+EX is because of the pronounced effect that the dock strike in 1968 and 1969 had on exports.

As mentioned above, the Cochrane-Orcutt iterative technique was used for OLSAUT01 and TSLSAUT01 and a technique like the Cochrane-Orcutt technique was used for OLSAUT02 and TSLSAUT02. For FIMLAUT01, counting the dummy variable coefficients, there were 40 coefficients to be estimated simultaneously -- 33 coefficients in A and B and 7 coefficients in $R^{(1)}$. The technique used to obtain the FIMLAUT01 estimates is also a technique like the Cochrane-Orcutt technique, only in this case applied to the full scale model simultaneously. For the DYN estimates, the quadratic hill climbing technique of Goldfeld, Quandt, and Trotter [7] was used. The technique worked quite well for this purpose, and no serious problems were encountered in minimizing the sum of squared errors for any of the equations.

---

[8] 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.

IV. THE RESULTS

The Coefficient Estimates

The results of estimating the seven stochastic equations of the model are presented in Tables 2-1 and 2-7. The results are fairly self explanatory and require only a limited discussion here. For the serial correlation estimators, the SE is the estimated standard deviation of the errors in E, not of the errors in U. For the DYN estimator, the SE is the estimated standard deviation of the error being minimized, and so for the two-quarter-ahead estimates, for example, the SE is based on the premise that the one-quarter-lagged value of the dependent variable is not known. Standard errors for the individual equations were not computed for the FIMLAUTO1 estimates. The standard errors are in billions of dollars.

Some of the noteable features of the results in Table 2-1 for the durable consumption equation are the poorer fits for OLS and TSLS, the different estimates for OLSLDV and TSLSLDV, and the closeness of the estimates for OLSAUTO1, TSLSAUTO1, OLSAUTO2, and TSLSAUTO2. The FIMLAUTO1 estimate of $r_{11}^{(1)}$ is also noteably smaller than the corresponding TSLSAUTO1 (42 obs.) estimate. The noteable feature of the results in Table 2-2 for nondurable consumption is the moderate degree of closeness of all of the estimates. The same holds true for the estimates of the service consumption equation in Table 2-3. The estimates of the plant and equipment investment equation in Table 2-4 are, on the other hand, somewhat less close. Note that the five-quarter-ahead DYN estimate of $r_{44}^{(1)}$ is negative. This has little effect on the final results, however, since only the fifth power of $r_{44}^{(1)}$ affects the final results. The estimates of the housing investment equation in Table 2-5 are again moderately close, except for the OLSLDV and TSLSLDV estimates and the five-quarter-ahead DYN estimate of $r_{55}^{(1)}$.

TABLE 2-1: Parameter estimates of the CD equation.

Estimator	Const.	GNP	$MOOD_{-1}$	$MOOD_{-2}$	D644	D651	CD_{-1}	$r_{11}^{(1)}$	$r_{11}^{(2)}$	SE	No.of obs.
		Coefficient Estimates for									
OLS	-24.93	.1026	.1158	.0828	-2.30	2.51	-	-	-	1.556	59
OLSLDV	-14.29	.0532	.1741	-.0604	-2.93	4.09	.4961	-	-	1.290	59
TSLS	-24.89	.1026	.1154	.0830	-2.30	2.51	-	-	-	1.556	59
TSLSLDV	-13.04	.0476	.1800	-.0761	-2.99	4.27	.5517	-	-	1.294	59
OLSAUTO1	-25.33	.1033	.1367	.0604	-2.60	2.43	-	.6053	-	1.210	59
TSLSAUTO1	-25.06	.1031	.1362	.0596	-2.53	2.59	-	.6039	-	1.211	59
OLSAUTO2	-26.76	.1044	.1478	.0560	-2.62	2.33	-	.4342	.2398	1.171	59
TSLSAUTO2	-26.47	.1042	.1474	.0552	-2.57	2.53	-	.4347	.2358	1.172	59
DYN: Two Quarters	-26.35	.1043	.1498	.0512	-2.52	2.23	-	.7148	-	1.294	59
DYN: Three Quarters	-27.22	.1040	.1383	.0740	-2.47	2.16	-	.7257	-	1.409	59
DYN: Four Quarters	-27.10	.1037	.1399	.0734	-2.36	2.24	-	.7179	-	1.494	59
DYN: Five Quarters	-28.89	.1045	.1605	.0658	-2.64	3.50	-	.8054	-	1.456	59
FIMLAUTO1	-41.28	.1090	.2176	.1021	-2.06	3.26	-	.2396	-	-	42
TSLSAUTO1	-34.62	.1085	.1830	.0695	-2.32	2.66	-	.3862	-	1.231	42

Instrumental variables used in addition to those in the basic set.

TSLS : $MOOD_{-1}$, D644, D651

TSLSLDV : $MOOD_{-1}$, D644, D651

TSLSAUTO1: $MOOD_{-1}$, $MOOD_{-3}$, D644, D651, $D644_{-1}$

TSLSAUTO2: CD_{-2}, GNP_{-2}, $MOOD_{-1}$, $MOOD_{-3}$, $MOOD_{-4}$, D644, D651, $D644_{-1}$, $D644_{-2}$

TABLE 2-2: Parameter estimates for CN equation.

Estimator	GNP	$MOOD_{-2}$	CN_{-1}	$r_{22}^{(1)}$	$r_{22}^{(2)}$	SE	No. of obs.
	Coefficient Estimates for						
OLS	.0517	.0552	.7987	-	-	1.592	42
OLSLDV	(same as OLS)						
TSLS	.0536	.0578	.7906	-	-	1.593	42
TSLSLDV	(same as TSLS)						
OLSAUTO1	.0439	.0432	.8331	-.2919	-	1.528	42
TSLSAUTO1	.0446	.0443	.8297	-.2896	-	1.528	42
OLSAUTO2	.0452	.0448	.8276	-.2551	.1102	1.519	42
TSLSAUTO2	.0477	.0484	.8165	-.2451	.1167	1.520	42
DYN: Two Quarters	.0468	.0477	.8200	-.0759	-	1.775	42
Three Quarters	.0508	.0565	.8010	-.3966	-	2.006	42
Four Quarters	.0528	.0617	.7914	-.3447	-	2.053	42
Five Quarters	.0536	.0646	.7868	-.4199	-	2.086	42
FIMLAUTO1	.0550	.0706	.7794	-.2691	-	-	42

Instrumental variables used in addition to those in the basic set

TSLS : none

TSLSAUTO1: CN_{-2}, $MOOD_{-3}$

TSLSAUTO2: CN_{-2}, CN_{-3}, GNP_{-2}, $MOOD_{-3}$, $MOOD_{-4}$

TABLE 2-3: Parameter estimates for CS equation.

Estimator	GNP	$MOOD_{-2}$	CS_{-1}	$r_{33}^{(1)}$	$r_{33}^{(2)}$	SE	No. of obs.
	Coefficient Estimates for						
OLS	.0228	-.0217	.9402	-	-	.422	59
OLSLDV	(same as OLS)						
TSLS	.0219	-.0213	.9434	-	-	.422	59
TSLSLDV	(same as TSLS)						
OLSAUTO1	.0227	-.0217	.9403	-.0085	-	.422	59
TSLSAUTO1	.0221	-.0214	.9427	-.0092	-	.422	59
OLSAUTO2	.0227	-.0217	.9404	-.0089	-.0096	.422	59
TSLSAUTO2	.0221	-.0214	.9427	-.0096	-.0107	.422	59
DYN: Two Quarters	.0239	-.0218	.9356	-.0461	-	.590	59
Three Quarters	.0238	-.0220	.9364	-.2386	-	.691	59
Four Quarters	.0240	-.0221	.9356	-.2870	-	.745	59
Five Quarters	.0244	-.0222	.9340	-.3734	-	.784	59
FIMLAUTO1	.0294	-.0235	.9149	-.1249	-	-	42
TSLSAUTO1	.0212	-.0233	.9471	.0139	-	.418	42

Instrumental variables used in addition to those in the basic set

TSLS : none

TSLSAUTO1: CS_{-2}, $MOOD_{-3}$

TSLSAUTO2: CS_{-2}, CS_{-3}, GNP_{-2}, $MOOD_{-3}$, $MOOD_{-4}$

TABLE 2-4: Parameter estimates for IP equation.

Estimator	Const.	GNP	PE2	IP_{-1}	$r_{44}^{(1)}$	$r_{44}^{(2)}$	SE	No. of obs.
	Coefficient Estimates for							
OLS	-5.96	.0515	.7831	-	-	-	1.430	59
OLSLDV	-3.78	.0322	.3852	.4612	-	-	1.303	59
TSLS	-5.89	.0507	.7921	-	-	-	1.430	59
TSLSLDV	-3.44	.0289	.3547	.5117	-	-	1.305	59
OLSAUTO1	-8.21	.0768	.4827	-	.8035	-	1.038	59
TSLSAUTO1	-7.21	.0707	.5451	-	.7685	-	1.042	59
OLSAUTO2	-8.23	.0786	.4597	-	.8873	-.0875	1.034	59
TSLSAUTO2	-7.35	.0728	.5202	-	.8373	-.0678	1.037	59
DYN: Two Quarters	-6.48	.0587	.6947	-	.6595	-	1.328	59
DYN: Three Quarters	-6.18	.0549	.7411	-	.6102	-	1.413	59
DYN: Four Quarters	-6.02	.0522	.7752	-	.5127	-	1.440	59
DYN: Five Quarters	-5.88	.0508	.7913	-	-.5864	-	1.440	59
FIMLAUTO1	-7.71	.0732	.5139	-	.8873	-	-	42
TSLSAUTO1	-9.74	.0801	.4625	-	.8135	-	1.122	42

Instrumental variables used in addition to those in the basic set

TSLS : none

TSLSLDV : IP_{-1}

TSLSAUTO1: IP_{-1}, $PE2_{-1}$

TSLSAUTO2: IP_{-1}, IP_{-2}, GNP_{-2}, $PE2_{-1}$, $PE2_{-2}$

TABLE 2-5: Parameter estimates for IH equation.

Estimator	Const.	GNP	HSQ	HSQ_{-1}	HSQ_{-2}	IH_{-1}	$r_{55}^{(1)}$	$r_{55}^{(2)}$	SE	No.of obs.
	Coefficient Estimates for									
OLS	-2.53	.0141	.0636	.0946	.0059	-	-	-	.591	42
OLSLDV	-1.96	.0097	.0665	.0722	-.0248	.3171	-	-	.566	42
TSLS	-2.54	.0142	.0636	.0946	.0059	-	-	-	.591	42
TSLSLDV	-1.97	.0098	.0665	.0723	-.0246	.3149	-	-	.566	42
OLSAUTO1	-2.91	.0141	.0660	.0868	.0147	-	.3833	-	.549	42
TSLSAUTO1	-2.92	.0141	.0660	.0869	.0146	-	.3829	-	.549	42
OLSAUTO2	-3.51	.0139	.0672	.0875	.0187	-	.2994	.3041	.525	42
TSLSAUTO2	-3.52	.0139	.0672	.0875	.0187	-	.2993	.3038	.525	42
DYN: Two Quarters	-3.23	.0140	.0 63	.0895	.0149	-	.6471	-	.568	40
DYN: Three Quarters	-3.09	.0142	.0648	.0953	.0079	-	.5759	-	.602	40
DYN: Four Quarters	-3.23	.0143	.0666	.0938	.0084	-	.6710	-	.601	40
DYN: Five Quarters	-2.52	.0143	.0656	.0923	.0052	-	-.6252	-	.609	40
FIMLAUTO1	-2.47	.0144	.0558	.0877	.0183	-	.4457	-	-	42

<u>Instrumental variables used in addition to those in the basic set</u>

TSLS : HSQ, HSQ_{-1}, HSQ_{-2}

TSLSLDV : HSQ, HSQ_{-1}, HSQ_{-2}, IH_{-1}

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}

TABLE 2-6: Parameter estimates for $V-V_{-1}$ equation.

Coefficient Estimates for

Estimator	Const.	CD+CN	CD_{-1} + CN_{-1}	V_{-1}	D593	D594	D644	D651	V_{-2}	$r_{66}^{(1)}$	$r_{66}^{(2)}$	SE	No. of obs.
OLS	-32.63	.0540	.1677	-.0984	-4.61	.95	.60	4.18	-	-	-	4.078	59
OLSLDV	-36.95	-.1398	.3798	.5286	-8.18	2.44	.30	3.84	-.6482	-	-	3.115	59
TSLS	-34.35	.6971	-.4705	-.1091	-4.85	2.50	2.89	1.03	-	-	-	4.381	59
TSLSLDV	-37.38	.0842	.1568	.5021	-8.13	2.91	1.08	2.79	-.6245	-	-	3.164	59
OLSAUTO1	-91.97	-.2004	.8100	-.3221	-8.89	-5.66	-1.73	5.09	-	.8846	-	2.623	59
TSLSAUTO1	-87.28	-.2506	.8281	-.3014	-8.32	-4.48	-.67	7.74	-	.8711	-	2.666	59
OLSAUTO2	-99.42	-.2084	.8609	-.3382	-9.72	-6.88	-1.73	5.29	-	1.1050	-.2642	2.547	59
TSLSAUTO2	-109.17	-.1374	.8510	-.3730	-9.36	-6.39	-1.46	5.20	-	1.1147	-.2593	2.554	59
DYN: Two Quarters	-103.00	-.0152	.6883	-.3522	-7.84	-2.96	-1.69	4.73	-	.8525	-	5.201	59
DYN: Three Quarters	-130.95	.1249	.7198	-.4423	-7.18	-1.52	-3.28	3.51	-	.8505	-	7.227	59
DYN: Four Quarters	-110.81	.1267	.5870	-.3649	-5.50	.97	-3.92	3.35	-	.8066	-	8.426	59
DYN: Five Quarters	-107.43	.1181	.5723	-.3489	-2.84	2.55	-2.84	.09	-	.7976	-	9.239	59
FIMLAUTO1	-165.82	.5593	1.0258	-.5187	-	-	.59	-1.31	-	.7299	-	-	42
TSLSAUTO1	-87.69	-.2373	.8296	-.3198	-	-	-1.27	6.47	-	.9101	-	2.652	42

<u>Instrumental variables used in addition to those in the basic set</u>

TSLS : D593, D594, D644, D651, $MOOD_{-1}$

TSLSLDV : D593, D594, D644, D651, V_{-2}, $MOOD_{-1}$

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}, $MOOD_{-1}$, $MOOD_{-3}$, $MOOD_{-4}$

---

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

TABLE 2-7: Parameter estimates for IMP equation.

Estimator	Const.	GNP	D644	D651	D684	D691	D692	IMP_{-1}	$r_{77}^{(1)}$	$r_{77}^{(2)}$	SE	No. of obs.
		Coefficient Estimates for										
OLS	-10.85	.0673	-2.76	-4.75	.37	-3.72	4.60	-	-	-	2.133	59
OLSLDV	-2.30	.0124	-.16	-2.02	-1.21	-4.16	7.57	.8427	-	-	.777	59
TSLS	-10.89	.0674	-2.76	-4.75	.35	-3.74	4.58	-	-	-	2.133	59
TSLSLDV	-2.24	.0121	-.15	-2.01	-1.21	-4.15	7.59	.8473	-	-	.777	59
OLSAUTO1	-19.23	.0780	.17	-2.35	-1.66	-6.07	1.93	-	.9507	-	.657	59
TSLSAUTO1	-25.98	.0851	.40	-1.97	-1.73	-6.22	1.71	-	.9638	-	.665	59
OLSAUTO2	-24.54	.0828	.18	-2.23	-1.52	-5.98	2.01	-	.7917	.1702	.650	59
TSLSAUTO2	-49.52	.0982	.26	-2.17	-1.48	-5.98	1.98	-	.7469	.2356	.661	59
DYN: Two Quarters	-34.99	.0904	-.09	-1.70	-.95	-6.05	2.39	-	.9799	-	.814	59
DYN: Three Quarters	-84.09	.1033	.47	-1.75	-1.52	-5.52	1.58	-	.9938	-	.922	59
DYN: Four Quarters	-310.35	.1092	.40	-1.62	-.83	-6.38	1.91	-	.9986	-	1.002	59
DYN: Five Quarters	271.17	.1138	.28	-1.77	-.93	-5.86	2.11	-	1.0014	-	1.028	59
FIMLAUTO1	-24.06	.0837	.23	-2.25	-1.74	-6.31	1.59	-	.8749	-	-	42
TSLSAUTO1	-25.96	.0859	.55	-1.68	-1.74	-6.24	1.66	-	.8931	-	.647	42

Instrumental variables used in addition to those in the basic set

TSLS : D644, D651, D684, D691, D692

TSLSLDV : D644, D651, D684, D691, D692, IMP_{-1}

TSLSAUTO1: 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}

The estimates of the inventory investment equation in Table 2-6 are quite different. In particular, the FIMLAUTO1 estimates are considerably different from the rest. Most of this difference cannot be attributed to the use of the shorter sample period since the TSLSAUTO1 estimates are quite close for the two periods. From the theoretical model in Chapter 6 in [5], the coefficient estimate of CD+CN is expected to lie between minus one and zero. The estimate does lie in this range for the OLSAUTO1, TSLSAUTO1, OLSAUTO2, and TSLSAUTO2 results, but it clearly does not for the FIMLAUTO1 results. The theoretical model in [5] is based on a very simple assumption about how sales expectations are formed, however, and it can be seen that changing this assumption slightly no longer requires that the coefficient of CD+CN be negative.[9] The FIMLAUTO1 results are thus not necessarily unreasonable on a priori grounds. Note that first order serial correlation appears to be very pronounced in the inventory equation, which in this case causes the OLS and TSLS estimates to be much different from the OLSAUTO1 and TSLSAUTO1 estimates.

The estimates of the import equation in Table 2-7 are also fairly different, although the import equation is of much less importance in the model than is the inventory equation. Again, first order serial correlation is quite pronounced in the equation. The large DYN constant term estimates are caused by the estimates of $r_{77}^{(1)}$ being close to one. If $r_{77}^{(1)}$ is equal

---

[9]The assumption about sales expectations in [5], equation (6.10), is that

(i) $SALES_t^e = SALES_{t-1} + \bar{S}$,

where $\bar{S}$ is a constant. If this assumption is changed to

(ii) $SALES_t^e = \delta SALES_t + (1-\delta)SALES_{t-1}, \quad 0 < \delta < 1$,

then it can be seen that the coefficient of $SALES_t$ (or in the present case the coefficient of CD_t+CN_t) need not be negative.

to one, the constant term cannot be estimated, and in the present case values of $r_{77}^{(1)}$ near one have caused large constant term estimates. These large estimates have little effect on the final results, however, because of the closeness of the $r_{77}^{(1)}$ estimates to one. Unlike the case for the inventory equation, the FIMLAUTO1 and TSLSAUTO1 estimates of the import equation are quite close.

The Ex Post Predictions

Given a set of estimates of the A, B, and R matrices in (1) and (3), one can solve for the reduced form for Y. Assume, for example, that $R^{(3)}$ through $R^{(n)}$ are zero in (3). Then (1) and (3) imply:

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

For each of the sets of estimates, ex post 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.[10] 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

---

[10]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 in Tables 2-1 through 2-7 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. The two-quarter-ahead predictions were generated by first generating the one-quarter-ahead predictions and then generating the two-quarter-ahead predictions using the generated values of the lagged endogenous variables. Similar procedures were followed for the other predictions. It should be emphasized 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.

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.

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

$$MAE = \sum_{t=1}^{T} |y_t - \hat{y}_t| \quad ,$$

and the mean absolute error in terms of changes,

$$MAE\Delta = \sum_{t=1}^{T} |(y_t - y_{t-1}) - (\hat{y}_t - \hat{y}_{t-1})| \quad ,$$

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. MAEΔ is a measure of how well the model has explained the change in the endogenous variable. For the one-quarter-ahead prediction, MAE and MAEΔ are the same.

The mean absolute errors for the GNP variable are presented in Table 3-1 for each set of estimates. The errors are in billions of dollars. For the MAE results, the most striking feature is the increased accuracy obtained from the FIMLAUT01 estimates for the one- through four-quarter-ahead predictions. This increased accuracy does not appear attributable to the use of the shorter sample period to compute the FIMLAUT01 estimates since the results from the two sets of TSLSAUT01 estimates are quite close. For the five-quarter-ahead prediction and the prediction over the whole sample period, FIMLAUT01 performs well, but it does not completely dominate the field as it does for the other predictions. The DYN estimates perform

TABLE 3-1: Mean absolute errors for GNP.

Estimator	MAE: One Quarter ahead 42obs	MAE: Two Quarters ahead 41obs	MAE: Three Quarters ahead 40obs	MAE: Four Quarters ahead 39obs	MAE: Five Quarters ahead 38obs	MAE: Entire Sample Period 42obs	MAE△: Two Quarters ahead 41obs	MAE△: Three Quarters ahead 40obs	MAE△: Four Quarters ahead 39obs	MAE△: Five Quarters ahead 38obs	MAE△: Entire Sample Period 42obs
OLS	4.57	5.11	5.06	5.53	5.58	6.08	2.96	3.06	3.02	3.02	3.07
OLSLDV	2.78	3.92	4.59	4.77	4.79	6.64	2.62	2.58	2.56	2.66	2.83
TSLS	5.25	4.92	4.77	5.34	5.40	5.95	3.69	3.33	3.32	3.32	3.37
TSLSLDV	2.71	3.78	4.38	4.51	4.61	6.49	2.57	2.58	2.54	2.65	2.78
OLSAUTO1	2.49	3.81	4.52	5.13	5.45	6.16	2.52	2.55	2.55	2.49	2.70
TSLSAUTO1	2.57	3.85	4.35	4.88	5.04	5.56	2.54	2.61	2.55	2.58	2.78
OLSAUTO2	2.40	3.54	4.14	4.77	5.13	5.98	2.44	2.47	2.43	2.46	2.69
TSLSAUTO2	2.36	3.41	4.02	4.55	4.79	5.22	2.42	2.46	2.41	2.46	2.67
DYN	a	3.58	3.68	4.23	4.11	b	2.44	2.62	2.72	2.77	b
FIMLAUTO1	1.69	2.25	2.97	3.77	4.72	5.27	2.36	2.65	2.73	2.71	2.74
TSLSAUTO1, 42 obs.	2.46	3.82	4.50	4.96	5.21	5.77	2.46	2.56	2.49	2.50	2.65

[a] Same as OLSAUTO1.

[b] Not relevant since DYN estimates only computed up to five-quarter -ahead predictions.

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. The two-stage least squares estimators perform on average better than their ordinary least squares counterparts. TSLSAUTO2 is always better than OLSAUTO2, TSLSAUTO1 is better than OLSAUTO1 for the three-quarter-ahead prediction and beyond, TSLSLDV is always better than OLSLDV, and TSLS is better than OLS for the two-quarter-ahead prediction and beyond. The AUTO2 estimators always perform better than their AUTO1 counterparts, TSLSAUTO1 always perform better than TSLS, and OLSAUTO1 performs better than OLS except for the prediction over the whole sample period. OLSLDV and TSLSLDV tend to perform fairly well relative to the others except for the prediction over the entire sample period.

The following conclusions can be drawn from the MAE results in Table 3-1. (1) For the one- and two-quarter-ahead predictions, it is quite important quantitatively to account for first order serial correlation. For the three-quarter-ahead prediction and beyond it is still beneficial to account for first order serial correlation, but of somewhat less quantitative importance. (2) It is usually beneficial to use a two stage least squares estimator as opposed to its ordinary least squares counterpart, but the use of a two stage least squares estimator relative to its ordinary least squares counterpart is quantitatively less important than accounting for first order serial correlation. (3) It is beneficial to account for second order serial correlation, but the gain in going from first to second order serial correlation is less than the gain in going from zero to first order serial correlation. (4) The DYN procedure is beneficial and is of quantiative importance for the three-through five-quarter-ahead predictions. (5) Accounting for all of the information in the system is beneficial and is of considerable quantitative importance at least up through the four-quarter-ahead predictions.[11]

For the MAE∆ error measure, the results of the various estimators are closer, and the FIMLAUTO1 estimator no longer dominates the other estimators. TSLSAUTO2 gives the best overall results, but the TSLSAUTO2 and OLSAUTO2 results are quite close. The OLSAUTO1 and TSLSAUTO1 results are slightly poorer than their AUTO2 counterparts, and in the AUTO1 case, the ordinary least squares estimator is slightly better than the two-stage least squares estimator. The OLS and TSLS estimators perform noteably worse than the others, which again indicates the importance of accounting for first order serial correlation. Otherwise, there is not too much to choose among the various estimators.

The closeness of the MAE∆ 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 FIMLAUTO1 estimator, for example, maximizes 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 MAE results do not carry over completely to the MAE∆ results.

---

[11]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 the one-period-ahead prediction, but considerably poorer results for the prediction over the whole sample 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.

In Table 3-2 the mean absolute 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, however, that for the MAE results the FIMLAUTO1 estimator does not dominate the other estimators as much for inventory investment as it did for GNP. This general pattern was true for the other six GNP components as well. The large gain from using the FIMLAUTO1 estimator comes in terms of predicting GNP and not in terms of predicting the individual components of GNP. This result is consistent with the fact that the FIMLAUTO1 estimator is the only estimator that uses information about the GNP identity directly in computing the estimates. Because of this, one would expect that much of the gain from using FIMLAUTO1 would come in terms of predicting GNP, which appears to be the case.

There were no unusual features of the results for the other components of GNP, other than the one just cited about the FIMLAUTO1 estimator, and so these results will not be presented here.

Conclusion

Most of the results in this section are consistent with what one would expect on theoretical grounds. For the MAE results for GNP, there is a close correlation between the rankings of the various estimators on theoretical grounds and the rankings of the estimators according to the empirical results. In addition, the results in this section give an indication of the quantitative importance of the various estimators.

TABLE 3-2: Mean absolute errors for $V-V_{-1}$.

Estimator	MAE One Quarter ahead 42obs	MAE Two Quarters ahead 41obs	MAE Three Quarters ahead 40obs	MAE Four Quarters ahead 39obs	MAE Five Quarters ahead 38obs	MAE Entire Sample Period 42obs	MAE△ Two Quarters ahead 41obs	MAE△ Three Quarters ahead 40obs	MAE△ Four Quarters ahead 39obs	MAE△ Five Quarters ahead 38obs	MAE△ Entire Sample Period 42obs
OLS	2.61	2.67	2.66	2.55	2.42	2.88	2.67	2.80	2.72	2.78	2.69
OLSLDV	2.24	2.62	2.78	2.63	2.48	3.46	2.66	2.80	2.83	2.98	2.90
TSLS	2.96	2.55	2.51	2.39	2.30	2.65	3.22	2.94	2.89	2.94	2.84
TSLSLDV	2.19	2.56	2.68	2.49	2.37	3.26	2.63	2.77	2.79	2.94	2.87
OLSAUTO1	1.90	2.54	2.56	2.51	2.38	3.19	2.46	2.76	2.77	2.80	2.76
TSLSAUTO1	1.93	2.57	2.55	2.51	2.38	3.21	2.41	2.75	2.74	2.80	2.76
OLSAUTO2	1.89	2.59	2.59	2.49	2.34	3.29	2.48	2.76	2.82	2.81	2.84
TSLSAUTO2	1.88	2.56	2.55	2.45	2.31	3.17	2.47	2.76	2.80	2.80	2.82
DYN	a	2.47	2.48	2.41	2.36	b	2.49	2.87	2.91	3.01	b
FIMLAUTO1	1.93	2.13	2.26	2.22	2.17	2.59	2.63	2.88	2.84	2.90	2.80
TSLSAUTO1, 42 obs.	1.92	2.54	2.58	2.54	2.41	3.09	2.40	2.73	2.73	2.78	2.75

[a] Same as OLSAUTO1.

[b] Not relevant since DYN estimates only computed up to five-quarter-ahead predictions.

The conclusions of this study are, of course, 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.[12] One of the important features of the model used in this study is its linearity, and it is an open question as to how restrictive this property is in allowing one to generalize the results of this study to nonlinear models. It is also an open question as to whether different results would be obtained if one used a much larger and more disaggregated model. 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 certainly indicate that serious attempts should be made to estimate models by full information methods and that work should proceed on trying to develop more sophisticated ways of accounting for the dynamic nature of models.

---

[12] The mean absolute errors presented in Table 3-1 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.

APPENDIX

Iteration Technique for OLSAUTO2

Assume that equation (2) is to be estimated, where u_i in (2) follows the second order autoregressive process in (6) and where the endogeneity of Y_i in (2) is ignored. Then the following iterative technique can be used to estimate the parameters A_i, B_i, $r_{ii}^{(1)}$, and $r_{ii}^{(2)}$. Choose initial values of $r_{ii}^{(1)}$ and $r_{ii}^{(2)}$, say $r_{ii}^{o(1)}$ and $r_{ii}^{o(2)}$, and estimate the parameters A_i and B_i of the following equation by ordinary least squares:

$$\text{(A1)} \qquad y_i - r_{ii}^{o(1)} y_{i_{-1}} - r_{ii}^{o(2)} y_{i_{-2}} = -A_i(Y_i - r_{ii}^{o(1)} Y_{i_{-1}} - r_{ii}^{o(2)} Y_{i_{-2}}) - B_i(X_i - r_{ii}^{o(1)} X_{i_{-1}} - r_{ii}^{o(2)} X_{i_{-2}}) + v_i \quad .$$

From these estimates of A_i and B_i, say A_i^o and B_i^o, compute estimates of u_i, from equation (2). Then estimate $r_{ii}^{(1)}$ and $r_{ii}^{(2)}$ from the following equation by ordinary least squares:

$$\text{(A2)} \qquad u_i^o = r_{ii}^{(1)} u_{i_{-1}}^o + r_{ii}^{(2)} u_{i_{-2}}^o + w_i \quad .$$

From these new estimates of $r_{ii}^{(1)}$ and $r_{ii}^{(2)}$, estimate new values of A_i and B_i from equation (A1), from which new estimates of $r_{ii}^{(1)}$ and $r_{ii}^{(2)}$ can be computed from equation (A2), and so on. This process can be repeated until successive estimates of $r_{ii}^{(1)}$ and $r_{ii}^{(2)}$ are within some prescribed tolerance level. For the work in this study, this technique converged quite rapidly. It can be proven, in a manner similar to that done by Cooper [3] for the first order serial correlation case, that if this technique converges, it converges to at least a local minimum.

TSLSAUTO2

For the two stage least squares analogue of OLSAUTO2, Y_i in equation (A1) is replaced by $\hat{Y}_i$, where $\hat{Y}_i$ is the matrix of fitted values of Y_i obtained from regressing each row of Y_i on a set of instrumental variables. Let $\hat{V}_i = Y_i - \hat{Y}_i$. Then the equation to be estimated in the second stage regression is:

$$(A3) \qquad y_i = r_{ii}^{(1)} y_{i_{-1}} + r_{ii}^{(2)} y_{i_{-2}} - A_i \hat{Y}_i + r_{ii}^{(1)} A_i Y_{i_{-1}} + r_{ii}^{(2)} A_i Y_{i_{-2}} - B_i X_i + r_{ii}^{(1)} B_i X_{i_{-1}} + r_{ii}^{(2)} B_i X_{i_{-2}} + (e_i - A_i \hat{V}_i) \quad ,$$

where e_i is given in equation (6). The general estimation technique now consists in choosing estimates of $r_{ii}^{(1)}$, $r_{ii}^{(2)}$, A_i, and B_i such that the sum of squared errors in (A3) is at a minimum. On the assumption that the error term e_i in (A3) is uncorrelated with the instrumental variables used in the first stage regression, e_i is uncorrelated with all of the explanatory variables in (A3). In order for consistent estimates to be produced, $A_i \hat{V}_i$ in (A3) also has to be uncorrelated with the explanatory variables, and the only way to insure that this will be the case is to include $y_{i_{-1}}$, $y_{i_{-2}}$, $Y_{i_{-1}}$, $Y_{i_{-2}}$, X_i, $X_{i_{-1}}$, and $X_{i_{-2}}$ in the set of instrumental variables. If this is done, then by the property of least squares $\hat{V}_i$ will be orthogonal to all of the explanatory variables in (A3).

It can be shown, in a manner similar to that done in Fair [6], footnote 5, for the first order serial correlation case, that the iterative technique described above can be used to estimate equation (A3).

[1] Chow, Gregory C.: "A Comparison of Alternative Estimators for Simultaneous Equations," Econometrica XXXII (October, 1964) 532-553.

[2] Chow, Gregory C. and Ray C. Fair: "Maximum Likelihood Estimation of Linear Equation Systems with Auto-Regressive Residuals," Research Memorandum No. 118, Econometric Research Program, Princeton University, November 1970.

[3] Cooper, J. P.: "Notes on the Use of Procedures for Linear Regression in the Presence of First Order Serially Correlated Disturbances," mimeographed.

[4] Evans, Michael K., Joel 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] Fair, Ray C.: "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] Klein, L. R.: "Estimation of Interdependent Systems in Macro-econometrics," Econometrica XXXVII (April, 1969), 171-192.