Summary Statistics and Forecasting Performance

By Don Larson*

Abstract

The simultaneous equation econometric model has recently come under increasing attack as a policy analysis and forecasting tool However, traditional methods of choosing among competing models rely heavily on the use of summary statistics It is shown, by example, that choosing a model with relatively better summary statistics does not guarantee getting the best unconditional out-of-sample forecasts Other methods of forecasting time series that allow relative evaluation of the out-of-sample forecasting ability of econometric simultaneous equation models are also examined

Keywords

Summary statistics, econometric models, autoregression, ARIMA, identification

Introduction

The simultaneous equation econometric model has recently come under increasing attack as a policy analysis and forecasting tool The criticisms are based on both theoretical results and the more pragmatic findings of *ex ante* forecast comparisons The criticisms lead to questions concerning not only the efficiency of out-of-sample forecasts but also the developmental methodology of most standard multivariate models

Despite the accumulating doubts surrounding such models, a variety of firms exist and prosper by selling the results of their econometric models—a fact which attests to the perceived usefulness of their models While there are a variety of uses to which econometric models might be put other than forecasting, the ability to replicate history is often the final measure of a model's validity

The presence of a large commercial market in econometric models and forecasts is consistent with a growing list of practical and theoretical modeling difficulties In fact, such problems may help explain the specialization of forecasting firms

Given the combination of doubt and usefulness associated with econometric models, the evaluation of competing models and methodologies becomes doubly important Persons in both the public and private sectors are required, with increasing frequency, to evaluate econometric models or economic analysis based on models or modeling techniques

The traditional method of quickly evaluating the accuracy of such models relies heavily on the use of summary statistics, such as t-scores, F-tests, R^2's, and Durbin-Watson (D-W) statistics, and these statistics represent the type of information generally requested and provided by the major economic journals One purpose of this article is to discuss the limited information provided by such statistics and to offer an illustration of those limits Another related purpose is to discuss some easily available, alternative methods which, through comparison, can provide evaluations of a model's relative forecasting abilities

I will review earlier criticisms of fixed parameter econometric models, present an *ad hoc* quarterly model of the corn sector along with estimation results from the model, discuss possible misinformation provided by initial model results and provide an alternative estimation, discuss several alternative methods of unconditional forecasting, and, finally, compare out-of-sample results.

*The author is an agricultural economist with the National Economics Division, ERS

Model Limits

There is an abundance of economic literature that discusses the gaps between modern economic theory and the structure of a simultaneous equation econometric model estimated from time series data Such problems supersede the practical difficulties of data definition and collection, along with estimation efficiency, and relate to the very logic by which economic behavior is expressed in terms of a fixed-parameter mathematical model Sims (*30*), in particular, provides an excellent review of such problems [1]

Perhaps the best known dilemma facing modelers who use time series data centers on Lucas' (*20*) argument that policy variables normally considered exogenous and independent actually determine the way in which each of the variables relate to all others His conclusion is based on Muth's (*23*) assertion that "[t]he way expectations are formed depends specifically on the structure of the relevant system describing the economy " Lucas argues that as policy changes are effectively incorporated into expectations, new policy results in new rules of response so that the reduced forms of econometric models do not have fixed parameters Prescott and Kydlund (*26*) have shown, in turn, that policy formation without regard to expectations formation can lead to what Sims has termed "Peter White" policymaking,[2] which, by necessity, must always go awry

Lucas' results lead to the dilemma that no economically or politically determined variable can truly be considered exogenous More recently, "causality" tests have been used to identify exogenous/endogenous relationships, but on a practical level, the tests themselves must be performed on a reduced form consistent with a multitude of theoretical structures See, for example, Sims (*28, 29, 31*), Granger and Newbold (*15*), Bishop (*4*), Ashley, Granger and Schmalensee (*1*), Barnett, Bessler, and Thompson (*2*)

Should the problem of exogeneity be solved, Sims (*30*) has suggested that yet another identification problem exists Sims cites the work of Hatanaka (*16*), who reevaluated Sargan's (*27*) conclusions concerning the simultaneous equation identification problem in models containing lagged dependent variables and autocorrelated residuals Hatanaka found that if the lag length and lag distribution are not known *a priori*, but rather are determined in the estimation process, then the identification rule is changed In order to identify an equation, at least one strictly exogenous variable must be located in each of the other equations of the model Repeat occurrences of the same variable with different lags in a single equation do not qualify as "strictly exogenous "

The application of Hatanaka'a criteria leaves many models underidentified Sims also notes that the exogeneity of variables is often determined by convenience While this does limit the size of the model, it further exacerbates the confusion between exogenous/endogenous variables and may limit the number of variables actually available to identify the system

Given these problems, it is not surprising that simultaneous equation models have had limited success as forecasting tools [3] Gordon (*12*) has drawn on the work of McNees (*22*) to document the failure of macroeconomic forecasters in the seventies More recently, Just and Rausser (*17*) have presented evidence that commodity price forecasts are generally better provided by the futures market than by large-scale econometric models Their findings are consistent with a growing list of works describing the consistency between futures markets and Muth's hypothesis (see (*8*), (*10*))

As already mentioned, model forecasts do exist and prosper in the market place, and it is this marketability which, in fact, documents their usefulness At the same time, there are a variety of models from which to choose and a number of potential shortcomings to avoid In the following section, I will argue that it is generally difficult to evaluate a model from the usual summary statistics provided by vendors and by most economic journals To illustrate this point, I provide and evaluate a quarterly corn model

[1]Italicized numbers in parentheses refer to items in the References at the end of the article

[2]Peter White will ne'er go right/Would you know the reason why?/He follows his nose where'er he goes/And that stands all awry —Nursery Rhyme

[3]Criticisms of particular models over particular periods are by no means a recent phenomenon (see Christ (*7*)), however, as marketed forecasts have come of age, historical evaluations of their accuracy have become available

The Model

Consider the following quarterly model

$$
\begin{aligned}
CPLANT_t &= f_1(CAL_t, CP/SMP_{t-3}, CP/SMP_{t-4}, SGAL_t) && \text{(1a)}\\
CHARV_t &= f_2(CPLANT_t) && \text{(1b)}\\
CDEMD_t &= f_3(BEEFPR_t, CHKPR_t, CP/IN_t, SMP/IN_t, WHP/IN_t, D2, D3, D4) && \text{(1c)}\\
CSTKS_t &= CSTKS_{t-1} + CPR_t + CIMP_t - CEXP_t - CDEMD_t && \text{(1d)}\\
CPR_t &= CHARV_t * CYLD_t && \text{(1e)}
\end{aligned}
$$

where variable definitions are provided in table 1

The model ignores several important structural aspects of the corn market, and prices are exogenous However, the model is similar to some large commercial agricultural forecasting models, and it serves well in illustrating several points about summary statistics

Equation (1a), (1b), and (1e) are annual equations when the model is simulated as CPLANT and CHARV and CPR are nonzero only one quarter of the year As a result, seasonal dummies have been included only in equation (1c) Production and stocks are solved by identities When performing forecasts greater than three quarters, one often takes yields for the type of model above from trend levels, so that yields become a function of time or lagged yields However, as the time horizon shortens, the U S Department of Agriculture's (USDA) forecasted yields, based on weather and/or survey data, are often used

A linear stochastic form of the three behavioral equations was estimated by use of a three-stage least squares procedure on quarterly data from USDA's T-DAM data base for the 1960-76 period Data for the 1977-79 period were retained for out-of-sample tests Table 2 shows the first stage results Third-stage results are available from the author upon request [4]

[4]From a theoretical point of view, there are several arguments for choosing one estimation process over another The production and sale of corn is a recursive process, a fact which is often evoked when one justifies a set of OLS estimates as unbiased However, reported data are often reviewed and adjusted by a USDA committee to provide consistency across aggregates Although the adjustments represent a logical necessity for data presentation, the changes potentially disrupt predetermination among the series, justifying a three stage least squares, or seemingly unrelated estimation procedure Both estimation techniques were used, and they resulted in only minor differences

Table 1—Variable definitions

Variable	Definition	Unit
Endogenous		
CPLANT	Planted acreage of corn	1,000 acres
CHARV	Harvested acreage of corn	do
CDEMD	Domestic disappearance of corn	Million bushels
CSTKS	Corn stocks	do
CPR	Corn production	do
Exogenous		
CAL	Acres allocated to corn	1,000 acres
SGAL	Acres allocated to sorghum	do
CP/SMP	Ratio of corn farm price to soy meal price	Dollars/bushel
BEEFPR	Beef production	Million pounds
CHKPR	Broiler production	do
CP/IN	Ratio of corn farm price to real disposable income	Dollars/bushel over billion 1972 dollars
SMP/IN	Ratio of soy meal price to real disposable income	do
WHP/IN	Ratio of wheat farm price to real disposable income	do
CIMP	Corn imports	Million bushels
CEXP	Corn exports	do
CYLD	Corn yield	Bushels/acre

Table 2—First-stage results: Original series

Equation

EQ1 CPLANT = 40608.12 + 0.054*CAL + 828929*CP/SMP3 − 298507*CP/SMP4 + 1.05*SGAL
(6.47) (4 93) (3.20) (−0 96) (0 32)

D-W D Statistic = 2.0

F Ratio	6782 80
Prob>F	0.0001
R-Square	0.9984

EQ2 CHARV = 978.83 + 0.84*CPLANT
(0 74) (44 1)

D-W D Statistic = 2.0

F Ratio	296368.19
Prob>F	0 0001
R-Square	0.9999

EQ3 CDEMD = 415 48 + 0 067*BEEFPR + 0 28*CHKPR − 1230345*CP/IN +
(3 07) (1 00) (2.02) (−4 26)

1422.1*SMP/IN + 561138*WHP/IN − 509 8*D2 − 43.0*D3 + 107.7*D4
(0 93) (3.83) (−10 9) (−0.94) (3 30)

D-W D Statistic = 2.07

F Ratio	61 48
Prob>F	0.0001
R-Square	0.9044

Note t-ratios are given in parentheses

For the first-stage statistics, the distinctive features of the equations as a group are that they possess remarkably high R^2s and exceedingly solid Durbin-Watson (D-W) statistics [5] All the parameter signs match *a priori* expectations, and most of the t-scores are significant at standard levels Third-stage results differ only slightly from the first-stage results

Examining the residuals in the equation for domestic disappearance, however, reveals an autocorrelation problem (table 3) Although the sample correlation coefficient at lag 1 confirms the earlier D-W statistic, problems arise at lags 2, 3, and particularly 4 Given the seasonal nature of agriculture and the fact that quarterly data are used, it is not surprising to find signs of fourth-order serial correlation Although the potential problem is well known and Wallis (*32*) has devised a measure similar to the D-W statistic to detect its presence, such information is seldom provided to purchasers or users of existing models

The Effects of Autocorrelation

It is well known that regression coefficients are less efficient, but unbiased, when estimated in the presence of an autocorrelated error structure (*18*, p 275) Moreover, the multiple correlation coefficient increases in the presence of autoregressive residuals As Bishop (*5*, p 14) notes, Granger and Newbold have been particularly critical of reporting high R^2s under such circumstances

> In time series regressions involving the levels of economic variables, one frequently sees coefficients of multiple correlation (R^2) much higher than 0 9 If these indicate anything at all, they presumably imply an extremely strong relationship between the dependent variable and the independent variables This is extremely misleading on many occasions, as comments noting poor forecast performance which sometimes follow these equations will testify In fact, the high R^2

[5]Given the structure of the model, variables which have a value only once every four quarters will, of course, always have a D W value of 2, given the formula

$$d = \frac{\Sigma (e_t - e_{t-1})^2}{\Sigma (e_t)^2}$$

Table 3—Sample autocorrelation coefficients on CDEMD residuals: Original series

Lag	Correlation	−1 9 8 7 6 5 4 3 2 1	0 1 2 3 4 5 6 7 8 9 1
0	1 00000		******************************
1	−.04088	. *	
2	−.23487	*****	.
3	−.22213	.******	
4	(.34186)	.	*********
5	−.14682	. ***	.
6	−.07209	. *	
7	−.08108	. **	
8	10161	.	** .
9	− 00398	.	
10	−.13777	. ***	.
11	−.13508	. ***	
12	.05048	.	* .
13	.15769	.	***
14	−.01194	.	.
15	− 08283	. **	
16	− 01853	.	.
17	21537	.	**** .
18	.02823	.	*
19	−.18358	. ****	.
20	−.04120	. *	.

Note *marks correlations, • marks standard error

Note this is not close to one See Kmenta (1971, 292) therefore, we should not expect to see an improvement by first differences

values could be no more than a reflection of the fact that the dependent variable is highly autocorrelated and could easily be achieved simply by regressing the variable on its own past Thus, in such circumstances, the value of R^2 says nothing at all about the strength of the relationship between the dependent and independent variable

As the multiple correlation coefficient has proven an unreliable tool in applied econometrics, emphasis has been placed on other tests of significance, primarily t-scores Bishop (*5*, p 13) notes, however

> What is also well documented in the literature, but often overlooked in practice, is that the usual tests of significance, when performed in the presence of autocorrelated errors, are biased For example, if positive first-order autocorrelation is present in the error structure and the independent variable is also autocorrelated, the estimates of the standard errors on each of the coefficients (s_b) will be biased downward in most situations When the standard error of the coefficient is underestimated, the t-statistic on that coefficient is obviously overstated as it is computed as $t = (b-b_o)/s_b$, implying greater explanatory power for that variable than actually exists This situation can easily lead to the inclusion of a statistically irrelevant variable in the final model If the error structure exhibits negative serial correlation and the independent variable is positively autocorrelated, the standard errors of the coefficients are likely to be overestimated, possibly leading to the elimination of a statistically significant variable from the model

The result of Bishop's caveat is that, unless additional information concerning residuals is provided, the relevance of any set of t-scores remains unknown Let the buyer beware

* Fortunately for the modeler, simple differencing usually greatly reduces the serial correlation problem To illustrate, the following transformation was made of the model's original data.

$$x^*_t = x_t - x_{t-4} \tag{2}$$

so that quarterly observations now represent year-to-year changes between quarters

* This is incorrect see Kmenta (1971, 289-92).

With the exception of intercept values and the error terms, the transformation does not structurally affect the expected values of the estimated coefficients. However, by reducing the serial correlation problem, the estimation process should prove more efficient Table 5 shows the first-stage results

In terms of first-order autocorrelation, the transformation appears a failure The D-W statistic, which is the statistic most often provided relating to serial correlation, dropped from 2 07 to 1 66 in the disappearance equation, which is still indeterminate at a 5-percent level of confidence Table 4, however, provides the sample autocorrelation coefficients of the residuals which, in fact, show no sign of serial correlation In addition, I performed a collective test on the errors by utilizing the sample autocorrelation coefficients across 24 lags to provide the Box-Pierce Q-statistic, which has an approximate Chi-square distribution (see Nelson (*24,* p 115)) As a result of the transformation, the Q-statistic dropped from 28 04 to 18 53 The confidence with which the series could be termed "white noise" increased from 21 percent to 76 percent

Transformation of the data requires a transformation in the form of the model A final measure of the appropriateness of the transformation is to test whether the seasonal parameters dropped are statistically different from zero when retained and estimated with transformed data

An F-test was performed to test the hypothesis that the coefficients on the seasonal dummy variables collectively equal zero The hypothesis could not be rejected for any of the equations F-score results are reported in table 6

The reduction in serial correlation had few meaningful effects on the summary statistics As already mentioned, the D-W statistic on the demand equation dropped slightly into the indeterminate range The changes on the t-scores were mixed Some scores, such as the score on the soy meal price variable in the demand equation, improved, whereas others, such as the scores on the beef and chicken production coefficient, declined

The multiple correlation coefficients all dropped dramatically, however, the two sets of numbers are

Table 4—Sample autocorrelation coefficients on CDEMD residuals: Transformed series

Lag	Correlation	−1 9 8 7 6 5 4 3 2 1	0 1 2 3 4 5 6 7 8 9 1
0	1 0000		******************************
1	1607	.	***
2	−.1535	***	
3	−.1488	. ***	.
4	−1833	. ****	
5	− 0763	. **	.
6	2305		***** .
7	.2055	.	****
8	− 1445	. ***	.
9	− 0525	*	.
10	− 1034	. **	.
11	− 0070	.	
12	−0032	.	.
13	.0582	.	*
14	.0302	.	* .
15	− 0007	.	
16	− 0313	. *	
17	.1532	.	***
18	0697	.	*
19	−.1503	. ***	
20	−.0360	. *	

Note *marks correlation, . marks standard error

Table 5—First-stage results: Transformed series

Equation	
EQ1 $\text{CPLANT} = \underset{(2.59)}{0.036*\text{CAL}} + \underset{(4.54)}{705924*\text{CP/SMP3}} - \underset{(-2.65)}{191553*\text{CP/SMP4}} + \underset{(9.27)}{0.17*\text{SGAL}}$	
D-W D Statistic = 2.0	F Ratio 32.44 Prob>F 0.0001 R-Square 0.7100
EQ2 $\text{CHARV} = \underset{(28.4)}{0.87*\text{CPLANT}}$	
D-W D Statistic = 2.0	F Ratio 809.39 Approx PR>F 0.0001 R-Square 0.9353
EQ3 $\text{CDEMD} = \underset{(0.25)}{0.016*\text{BEEFPR}} + \underset{(1.97)}{0.30*\text{CHKPR}} - \underset{(-3.43)}{770789*\text{CP/IN}} + \underset{(1.40)}{1592.1*\text{SMP/IN}} + \underset{(3.22)}{396995*\text{WHP/IN}}$	
D-W D Statistic = 1.65	F Ratio 6.81 Prob>F 0.0001 R-Square 0.3956

Note t-ratio given in parentheses

Table 6—F-scores on hypothesis that coefficients on seasonal dummies collectively equal zero

Equation	F-score	Value at 99-percent confidence level
Acres planted	0.04	7.1
Acres harvested	58	7.2
Total disappearance	13	4.2

not comparable, and they reveal very little about the relative efficiency of either model

Finally, the F-scores associated with each of the equations fell substantially with the data transformation. However, the hypothesis that the equations contain no explanatory power could still be rejected overwhelmingly

Forecasting

Although a reduction in the level of serial correlation should improve the parameter estimates for any particular model, questions remain unanswered concerning the reliability and efficiency of the model's forecasts. In this section, I describe three alternative methods of forecasting time series, each of which requires less information, and can thus be more quickly estimated and forecasted than the model in the second section. In the last section, I compare out-of-sample forecast performance for these models and for the restricted equation model

ARIMA Model

A general class of auto regressive-integrated-moving-average (ARIMA) models can supply quick and often relatively efficient forecasts for many time series (6) For the most part, ARIMA models are purely statistical and impart little information about the economic processes involved Future innovations in a series are modeled on past innovations, allowing a univariate model form and a self-contained forecast of the time series

The general form of the model, ARIMA(p,d,q) is

$$w_t = b_1 w_{t-1} + \quad + b_p w_{t-p} + u_t \\ - c_1 u_{t-1} - \quad - c_q u_{t-q} \tag{3}$$

where w is a form of the original series differenced d times, u is an error term, and b and c are fixed coefficients

Box and Jenkins provide a three-stage modeling process of (1) identifying potential models, (2) estimating one or more models, and (3) testing the restriction each form places on the model's parameters With the advent of the appropriate computer software, ARIMA models can be quickly and cheaply estimated and forecasted

Vector Regression

Sims (*30,* p 14) has noted

> Much recent theoretical work gives rigorous foundation for a rule of thumb that in high dimensional models restricted parameters can easily produce smaller forecast or projection errors than unrestricted estimators, even when restrictions are false Thus models whose self-proclaimed behavioral interpretation is widely disbelieved may nonetheless find satisfied users as tools of forecasting and policy projection

As a possible solution, Sims proposed the estimation of a reduced version of the model, treating all variables as endogenous, without imposing any *a priori* restrictions on the parameters Restrictions could then be added and tested in a more systematic manner [6] Sims termed this process vector autoregression

Although models tend to become large quickly under Sims' methodology, vector autoregression does provide a readily available standard which can be applied to more restricted model forms The functional form of a vector regression system is simply

$$X_t = P(L)X_{t-1} + Z_t \tag{4}$$

where X is a vector of endogenous variables, L represents a lag operator whose length is determined by the data, P is a matrix of estimated parameters, and Z is a vector of disturbance terms

Although the choice of variables may be guided by economic theory, the general form of the model is most likely consistent with a variety of competing theoretical models, and as such, trades the potential efficiencies of a parsimonious parameterization for the ability to test particular restrictions on a more general form

Autoregressive Open Model

A third alternative to the standard simultaneous equation system has been recently used by Lamm (*19*) for the U S food and agricultural sector The method is similar to Sims' method with the exception that the distinction between exogenous and endogenous variables is retained

This distinction helps reduce the size of the model and links the modeled sector to relevant information outside the sector

The model is written as

$$X_t = P(L)X_{t-1} + H(L)Y_t + U_t \tag{5}$$

where X is again a vector of endogenous variables, Y is a vector of exogenous variables, U is a vector of disturbance terms, L is the lag operator, and P, H are vectors of estimated parameters

The form of the model allows for checks on any set of restrictions across parameters as well as for checks on the assumption of exogeneity

Out-of-Sample Forecasts

In this section, I present the out-of-sample forecast performance for five methodologies

[6] As Malinvaud (*21*) notes, the assumptions concerning efficiency gains implicit in restricted-form equation systems, with few exceptions, have gone untested in most models since the pioneering work of Frisch and Tinbergen

1. The restricted simultaneous equation system presented in the second section, estimated with the original time series,
2. The same model estimated with seasonally differenced data as described in Section 3,
3. ARIMA models for each series,
4. A vector autoregression model, and
5. An autoregressive open model

The forecasts were performed as follows

1. Each model was estimated from a sample, and then resulting parameters were used to forecast one, two, three, and four periods ahead
2. One period was then added to the sample, and the process in the first step was repeated until the observation set had been exhausted

Because the emphasis here is on a quarterly model, only the variables for which quarterly values exist, CSTKS and CDEMD, were forecast In order not to penalize the more fully specified models, I used actual values for all endogenous variables other than CSTKS and CDEMD when forecasting these two variables

Table 7 shows the mean absolute error (MAE) associated with each methodology for each of the forecast periods when CDEMD is forecast Table 8 provides the same information for forecasts of CSTKS

The most dramatic result of the out of-sample simulation is the more than threefold reduction in the MAE in the restricted equation models due to simply differencing the data series This reduction occurred despite the fact that the differenced model had a lower R^2, a mixed set of t-scores, and a less impressive D-W statistic In short, the summary statistics had provided all the wrong signals

Once the differencing occurred, the restrictions did indeed seem to improve the model's forecasting efficiency Even in the demand equation, where the restrictions were *ad hoc,* the forecast errors were on average much smaller than under any other methodology For the stock equation, in which the restrictions produced an identity, the results were even more pronounced

Another result with practical applications is that, for some variables, very simple models requiring little in the way of data can provide relatively efficient forecasts, the best example being the effectiveness with which an ARIMA model forecasted CDEMD The vector autoregression model also provided fairly efficient forecasts, especially as the forecast horizon lengthened

Another related result is that it seems difficult to determine beforehand whether or not a method will work Although Lamm reported success with an open autoregressive model at a more aggregated

The high MAE for the restricted equations case is due to the way the model was specified If, instead, we had estimated an equation for stocks and solved for demand as a residual, we would have obtained better results

Table 7—Mean absolute error (MAE) of CDEMD forecasts

Model	Forecasting period			
	1	2	3	4
Restricted equations original series	132 9	133.0	110 7	121 9
Restricted equations differenced series	31.9	41 9	48 0	48 2
ARIMA model	47 6	60 4	64.0	68 6
Vector autoregression	60.0	63.2	53 1	60 3
Open autoregression	332.9	310 9	254.9	250 1

Table 8—Mean absolute error (MAE) for CSTKS forecasts

Model	Forecasting period			
	1	2	3	4
*Restricted equations original series	132 9	133 0	110 7	121.9
*Restricted equations differenced series	31.9	41.9	48.0	48 2
ARIMA model	158 3	262.7	351.9	385.0
Vector autoregression	418.0	331 1	509.2	433.7
Open autoregression	332.5	597.0	587.5	718 5

*Because CSTKS is determined by an identity in the restricted equation models, the MAE here is equal to the error associated with CDEMD

level, the model here grew quite large and, plagued with multicollinearity problems, failed to produce useful forecasts Although methods with the potential power to reduce the problem are available, such applications here would have detracted from the model's value as a quick alternative method to restricted equation forecasts

Conclusions

Well-constructed simultaneous restricted-equation models are useful and marketable policy tools, despite practical and theoretical problems Because of their usefulness, econometric models, forecasts, and analysis based on models have become more abundant and more comprehensive However, the types of summary statistics generally made available to model users provide limited information with which to evaluate the model or its forecasting capabilities As has been shown here, modelers who attempt to maximize the summary statistics normally reported in economic journals need not arrive at an optimal model Model users, judging two competing models on the basis of the summary statistics normally reported by economic journals, will not necessarily be able to choose the better model

Information about the structure of the error terms is essential to accurately evaluate any model and its summary statistics Such information is neither generally requested nor provided by most economic journals Furthermore, such information is generally neither requested by buyers nor provided by sellers of models or model forecasts.

Another effective method for evaluating a model and its forecasting abilities is to compare the out-of-sample forecasts of complicated models with the out-of-sample forecasts of simple models ARIMA and vector auto-regression models are two such models which provide relatively good forecasts

References

(1) Ashley, R , C W J Granger, and R Schmalensee "Advertising and Aggregate Consumption An Analysis of Causality," *Econometrica*, Vol 48, 1980, pp 1149-67

(2) Barnett, Richard C , David A Bessler, and Robert L Thompson "Agricultural Prices in the 1970s and the Quantity Theory of Money " Paper presented at American Agricultural Economics Association annual meeting, Clemson, S C , July 1980

(3) Berndt, E R , and N E Savin "Estimation and Hypothesis Testing in Singular Equation Systems with Autoregressive Disturbances," *Econometrica*, Vol 43, 1979, pp 937-57

(4) Bishop, Robert V "The Construction and Use of Causality Tests," *Agricultural Economics Research*, Vol 31, No 4, 1979, pp 1-6

(5) ________ "The Use and Misuse of Summary Statistics in Regression Analysis," *Agricultural Economics Research*, Vol 33, No 1, 1981, pp 13-18

(6) Box, G E P , and G M Jenkins *Time Series Analysis, Forecast and Control* San Francisco Holden-Day, 1970

(7) Christ, C "Aggregate Econometric Models," *American Economic Review*, Vol 46, 1956, pp 385-408

(8) Danthine, J "Information, Futures Prices, and Stabilizing Speculation," *Journal of Economic Theory*, Vol 17, 1978, pp 79-98

(9) Durbin, J "Testing for Serial Correlation in Least Squares Regression When Some of the Regressors are Lagged Dependent Variables," *Econometrica*, Vol 38, 1970, pp 410-21

(10) Gardner, B L "Futures Prices in Supply Analysis," *American Journal of Agricultural Economics*, Vol 58, 1979, pp 81-84

(11) ________ "On the Power of Macroeconomic Linkages to Explain Events in U S Agriculture," *American Journal of Agricultural Economics*, Vol 63, 1981, pp 871-78

(12) Gordon, Robert *Macroeconomics* Boston Little, Brown, and Co , 1978

(13) Granger, C W J "Investigating Causal Relationships by Econometric Models and Cross-Spectral Methods," *Econometrica*, Vol 37, 1969, pp 424-38

(14) ________, and P Newbold, "Spurious Regressions in Econometrics," *Journal of Econometrics*, Vol 2, 1974, pp 110-20

(15) ________ "The Time Series Approach to Econometric Model Building," *New Methods in Business Cycle Research Proceedings from a Conference* (ed by C Sims). Minneapolis Federal Reserve Bank of Minneapolis, 1977

(16) Hatanaka, M "On the Global Identification of the Dynamic Simultaneous Equation Model with Stationary Disturbances," *International Economic Review*, Vol 16, 1975, pp 545-54

(17) Just, R E , and G G Rausser "Commodity Price Forecasting with Large-Scale Econometric Models and the Futures Market," *American Journal of Agricultural Economics*, Vol 57, 1981, pp 197-208

(18) Kmenta, Jan *Elements of Econometrics* New York Macmillan Publishing Co , 1971

(19) Lamm, R M , Jr "Properties and Forecasting Ability of a Multivariate Model for the U S Food and Agricultural Sector " Unpublished manuscript U S Dept Agr , Econ Res Ser , 1982

(20) Lucas, Robert E , Jr "Econometric Testing of the Natural Rate Hypothesis," in *The Econometrics of Price Determination* Board of Governors, Federal Reserve System, 1972, pp 50-59

(21) Malinvaud, E "Econometrics Faced with the Needs of Macroeconomic Policy," *Econometrica*, Vol 49, 1981, pp 1363-76

(22) McNees, Stephen K "The Recent Record of Thirteen Forecasters," *New England Economic Review*, Sept /Oct 1981, pp 5-21

(23) Muth, John F "Rational Expectations and the Theory of Price Movements," *Econometrica*, Vol 29, 1979, pp 315-35

(24) Nelson, Charles *Applied Time Series for Managerial Forecasting* San Francisco Holden-Day, 1973

(25) Pierce, David A "Relationships—and the Lack Thereof—Between Economic Time Series, With Special Reference to Money, Reserves, and Interest Rates " Special Studies Papers 55 Board of Governors, Federal Reserve System, Feb 1975

(26) Prescott, E C , and F Kydlund "Rules Rather than Discretion The Inconsistency of Optimal Plans," *Journal of Political Economy*, Vol 85, 1977, pp 473-92

(27) Sargan, J D "The Maximum Likelihood Estimation of Economic Relationship with Autoregressive Residuals," *Econometrica*, Vol 29, 1961, pp 414-26

(28) Sims, Christopher A "Comment" *Journal of the American Statistical Association*, Jan 1977, pp 23-24

(29) ________ "Exogeneity, Income, and Causality," in *New Methods in Business Cycle Research Proceedings from a Conference* (ed. C. Sims) Minneapolis Federal Reserve Bank of Minneapolis, 1977

(30) ________ "Macroeconomics and Reality," *Econometrica*, Vol 48, 1980, pp 1-48

(31) ________ "Money, Income, and Causality," *American Economic Review*, Vol 62, 1972, pp 540-52

(32) Wallis, K F "Testing for Fourth-order Autocorrelation in Quarterly Regression Equations,", *Econometrica*, Vol 40, July 1972, pp 617-36