INTERNATIONAL ECONOMIC REVIEW
Vol. 21, No. 2, June, 1980

ON PREDICTIVE EVALUATION OF ECONOMETRIC MODELS*

By Jayasri Dutta

1. INTRODUCTION

The analysis of the predictive power of an econometric model as a guide to model evaluation and selection has become fairly routine in applied econometric research. The popularity of predictive evaluation is due not merely to computational simplicity but also to the belief that extrapolation is a good guide to a model's explanatory ability. Thus, a "good" model should be able to generate "good" or accurate forecasts.[1]

The first problem is to measure the accuracy of a set of forecasts. Several techniques are available, which include overall scalar measures as well as parametric tests (cf. Granger and Newbold [1977], Dhrymes et al. [1972]). In this paper, I consider only some scalar measures, further restricting the class to those which are generated by a quadratic loss function. Thus, if e is the $N \times 1$ prediction error vector, then we can define

$$m = m(e) = e'Qe \tag{1.1}$$

as a measure of forecast accuracy, Q being an $N \times N$ matrix of weights. Clearly, m is zero only for perfect forecasts, and a higher value of m corresponds to a lower accuracy of forecasts; m is symmetric and monotonic.

The choice of weights (i.e., the matrix Q) may not be very important in the comparison of two forecast series. In fact, a multiple of the identity matrix would serve quite well. However, if m is to be used as a guide to model evaluation, it is being forced to yield additional information on the model's specification. In that case, one can expect the statistic to satisfy some requirements other than simple monotonicity. I examine the performance of four such measures against one requirement that can reasonably be set on an evaluation procedure.

I start from the premise that any evaluation procedure should enable the researcher to choose the "true" model out of any set of competing specifications. A model may be specified (and correspondingly, misspecified), in several respects. I deal with only one aspect of the specification problem — the selection of regressors in the linear regression model. The question is framed as follows: can a particular measure of forecast accuracy enable a researcher to choose the correct set of regressors over competing sets, at least on the average? If it can,

* Manuscript received February 2, 1978; revised December 27, 1978.

[1] What else a "good" model is supposed to do is not clear. The answer depends at least partly on the final use of the model. It seems generally agreed that the "true" model is the "best."

it is said to satisfy the regressor selection criterion.

It should be noted that this requirement is far from demanding. In the first place, the ability to discriminate between "true" and incorrect specifications is only one among several requirements that can be asked of an evaluative technique. Further, regressor selection is a very small part of the specification problem. The fulfillment of this criterion can be viewed as a necessary, (though by no means sufficient) condition for a good evaluative technique.

2. SOME MEASURES OF PREDICTIVE ABILITY

Four statistics for measuring the accuracy of forecasts are discussed below. Each of them was suggested with a different end in view.

Let $P_1, P_2, \ldots, P_N$, be a set of predictions on some variable of interest. After the actual values $A_1, A_2, \ldots, A_N$, become available, we have the series of prediction errors as $e_i = P_i - A_i$, $i = 1, 2, \ldots, N$. If e be the $N \times 1$ vector with e_i in the i-th place, then we write

$$m_i = e'Q^i e, \tag{2.1}$$

$i = 1, 2, 3, 4$, as the statistic, where Q^i is the corresponding weights matrix.

a. *MSE of forecasts.* Writing

$$Q^1 = \frac{1}{N} I, \tag{2.2}$$

with I as the $N \times N$ identity matrix, we have

$$m_1 = \frac{1}{N} \sum_{i=1}^{N} e_i^2, \tag{2.3}$$

as the mean squared error of forecasts, which is the most widely used variant of the quadratic loss function.

b. *Theil's forecast accuracy statistic.*

$$m_2 = \frac{e'e}{A'A}, \tag{2.4}$$

where A is the $N \times 1$ vector with A_i in the i-th place (cf. Theil [1966]). Thus,

$$Q^2 = S^{-1} I, \tag{2.5}$$

$$S = A'A. \tag{2.6}$$

For forecasts on the same series, m_1 and m_2 can be used interchangeably. The advantage of m_2 is that it is scale free, so that comparisons of forecast *procedures* are possible.

c. *The "Janus" coefficient.* This statistic was suggested by Gadd and Wold [1964] with the idea of forecasting from an estimated model. Let us define

(2.7) $$e_S = P_S - A_S,$$

as the $T \times 1$ vector of forecast errors for the sample, or "fitted" period. Then

(2.8) $$m_3 = \frac{\frac{1}{N} e'e}{\frac{1}{T} e_S' e_S} = \frac{m_1}{m_{1S}}.$$

Hence

(2.9) $$Q^3 = (m_{1S})^{-1} Q^1.$$

This coefficient is supposed to "look both ways," measuring deviations from perfect forecasting ($m_3 \simeq 0$), as well as from structural stability ($m_3 \simeq 1$).

d. *The variance ratio.* This statistic was suggested by Granger and Newbold [1977] as follows

(2.10) $$m_4 = \frac{\hat{\sigma}_e^2}{\hat{\sigma}_a^2},$$

where σ_e^2 is the variance of the series e_i, σ_a^2 the variance of the series A_i, and the numerator and denominator of m_4 are their estimates. This can be written as

(2.11) $$m_4 = \frac{\frac{1}{N}\Sigma(e_i - \bar{e})^2}{\frac{1}{N}\Sigma(A_i - \bar{A})^2},$$

where $\bar{e}$ and $\bar{A}$ are the respective arithmetic means. Thus

(2.12) $$Q^4 = (A'UA)^{-1}U,$$

with

(2.13) $$U = I - \frac{1}{N} i\, i',$$

i being the $N \times 1$ sum vector. U is symmetric and idempotent, with rank $N-1$.

3. A COMPARISON OF ALTERNATIVE STATISTICS

Let us write a linear regression model as

(3.1) $$y = x\beta + u,$$

with a competing specification

(3.2) $$y = z\partial + u.$$

x is the $T \times K$ matrix of T observations on a set of K exogenous variables; z is a $T \times P$ matrix of observations on a set of P exogenous variables. The disturbance vector u is distributed by a T dimensional normal law with zero mean and scalar dispersion matrix $\sigma^2 I$. Further, (3.3) $$x\beta \neq z\partial$$

so that only one of the models is "true." Let (3.1) be H_0, and (3.2), H_1. Consider the usual sort of expost predictions from the model estimated by OLS. To check whether the measure m_i chooses the correct set of regressors on the average, I calculate Em_{i0}, Em_{i1}, i.e., the expectations of m_i calculated from the estimated models under H_0, H_1 respectively. Now, since either H_0 is true, or H_1 is, we can calculate $E(m_{ij}|H_j)$, $i, j=0, 1$, as four different expectations. This gives rise to two interpretations of the regressor selection criterion.

Let us write the first interpretation as

(3.4) $$E(m_{i0}|H_0) < E(m_{i1}|H_0).$$

This condition follows naturally from the argument. Two competing models are set up, and the model with lower m_i (i.e., higher accuracy of forecasts) is chosen. One expects the "true" model to yield the more accurate forecasts, at least on average.

A second interpretation is

(3.5) $$E(m_{i1}|H_0) > E(m_{i1}|H_1).$$

This condition is not self-evident, and needs justification. Here, one is comparing an actual situation (H_0 "true") with a hypothetical one (H_1 "true"). In any given empirical situation, one can not compare $m_{i1}|H_0$ with $m_{i1}|H_1$. However, intuitively, one can speak of models H_0 and H_1 as having the same predictive power when "true" (thus, $E(m_{ij}|H_j)$ is the same for all j). In that case, the two conditions are the same. (In fact, if the two conditions are not equivalent, so that (say) $E(m_{i0}|H_0)<E(m_{i1}|H_1)$, then the measure m_i has the undesirable characteristics of being biased against models of type H_1). It is sometimes easier to get a definitive answer to the second condition, since one deals with the same estimators.

Let us write the matrices of values of the x and z variables in the forecast period as x_*, z_* which are $N \times K$, $N \times P$, respectively. The predictions from the two models are, respectively

(3.6) $$P_0 = \hat{y}_{*0} = x_*(x'x)^{-1}x'y$$

(3.7) $$P_1 = \hat{y}_{*1} = z_*(z'z)^{-1}z'y.$$

If H_0 is true, the actual values are generated from the model

(3.8) $$y_* = x_*\beta + u_*,$$

where u_* has the same properties as u. One can similarly define the forecast period model if H_1 is "true". The properties of the prediction error vectors, e_0, e_1, will depend on which hypothesis is true. Let us write $e_{jk}(j, k=0, 1)$, as the prediction error vector from model H_j when H_k is true. Thus,

(3.9) $$e_{00} = x_*(x'x)^{-1}x'u - u_*$$

(3.10) $$e_{10} = [z_*(z'z)^{-1}z'x - x_*]\beta + z_*(z'z)^{-1}x'u - u$$

$$e_{11} = z_*(z'z)^{-1}z'u - u_*. \tag{3.11}$$

The two conditions (3.4), (3.5) can be rewritten as

$$E(e'_{00}Q^i e_{00}) < E(e'_{10}Q^i e_{10}) \tag{3.12}$$

and

$$E(e'_{10}Q^i e_{10}) > E(e'_{11}Q^i e_{11}), \tag{3.13}$$

for each $i=1, 2, 3, 4$.

a. *M.S.E. of forecast* $(m_1 = e'e/N)$. We can write

$$\begin{aligned} E(m_{10}|H_0) &= \frac{1}{N}E(e'_{00}e_{00}) \\ &= \sigma^2 + \frac{1}{N}\sigma^2 r_1, \end{aligned} \tag{3.14}$$

where

$$r_1 = \mathrm{Tr}\,[(x'x)^{-1}x'_*x_*].$$

Further,

$$\begin{aligned} E(m_{11}|H_0) &= \frac{1}{N}E(e'_{10}e_{10}) \\ &= \frac{1}{N}\beta' D' D\beta + \sigma^2\frac{r_2}{N} + \sigma^2, \end{aligned} \tag{3.15}$$

where

$$D = z_*(z'z)^{-1}z'x - x_*;$$

$$r_2 = \mathrm{Tr}\,(z'z)^{-1}z'_*z_*;$$

and

$$\begin{aligned} E(m_{11}|H_1) &= \frac{1}{N}E(e'_{11}e_{11}) \\ &= \sigma^2 + \sigma^2\frac{r_2}{N}. \end{aligned} \tag{3.16}$$

Thus, m_1 satisfies the second condition (3.13) straightaway, as $\beta'D'D\beta$ is positive semi-definite. We can write

$$E(m_{11} - m_{10}|H_0) = \frac{\sigma_2}{N}(r_2 - r_1) + \frac{1}{N}\beta' D' D\beta. \tag{3.17}$$

The second term of the r.h.s. is clearly non-negative; the first term is of uncertain sign. If $r_2 > r_1$, then the first condition is fulfilled straightaway. Otherwise, one can not say anything about the sign. As N grows indefinitely large r_2 and r_1

tend to finite numbers[2]; so that the first term in the r.h.s. tends to become positive.

Thus, m_1 satisfies one part of the regressor selection criterion straightaway; it may not satisfy the other part, but this is likely to correct itself as N grows larger.

b. *Theil's forecast accuracy statistic* ($m_2 = e'e/A'A$). We can write

$$m_{2j} = 1 + \frac{P'_j P_j}{A'A} - \frac{2P'_j A}{A'A}, \qquad j = 0, 1. \tag{3.18}$$

We have

$$\begin{aligned} &E(m_{21} - m_{20} \mid \mathrm{H}_0) \\ &\quad = s_1 \beta'(D'D + D'x_* + x'_* D)\beta + \sigma^2(r_2 - r_1) - 2\beta' D' s_2, \end{aligned} \tag{3.19}$$

where

$$s_1 = E(y'_* y_*)^{-1}$$
$$s_2 = E[y_*(y'_* y_*)^{-1}].$$

s_1 is positive. The difference in (3.19) may be negative even if $(r_2 - r_1)$ is non-negative. Thus, the condition that $r_2 - r_1 > 0$, which is sufficient to ensure that m_1 satisfies the condition (3.4), does not ensure that m_2 satisfies it.

The exact moments s_1 and s_2 have been worked out in Appendix. A large-concentration parameter expansion of these moments shows that the first term in the expansion of the difference in (3.19) is positive if $r_2 - r_1$ is positive.

In carrying over the analysis of m_2 with respect to condition (3.5), comparisons become difficult because the concentration parameters in the distribution of $y'_* y_*$ depend on H_0 and H_1, i.e., the terms in Q^2 are stochastic and vary according to which hypothesis is true. One can get an idea of the problem from asymptotic analysis for "approximate specification."

"Approximate specification" is used to characterise a situation where the specified hypothesis is incorrect; however, the difference between the "true" and "specified" model is very small. This is illustrated as follows. Let H_1 be specified as before. In considering the situation where H_0 is "true", we assume the following:

$$x\beta = z\partial + \varepsilon q; \tag{3.20}$$

$$x_*\beta = z_*\partial + \varepsilon q_*; \tag{3.21}$$

where ε is a small scalar in the neighborhood of zero; q, q_* are non-stochastic vectors with finite elements. Clearly,

[2] If we make the assumption that

$$\lim_{T\to\infty} \frac{x'x}{T} = \lim_{N\to\infty} \frac{x'_* x_*}{N} = \Sigma_{xx'},$$

and similarly for the z, z^* matrices, then $\lim_{N,T} r_1 = K$; $\lim_{N,T} r_2 = P$.

(3.22) $$\lim_{\varepsilon \to 0} H_0 = H_1.$$

For ε small enough, H_1 is an approximate specification for true H_0 (for both sample forecast periods). Let us write

$$y_{*1} = z_* \partial + u_*;$$

when H_1 is true, $y_* = y_{*1}$. When H_1 is not true, but H_0 is, then

(3.23) $$\begin{aligned} y_* &= x_* \beta + u_* \\ &= z_* \partial + \varepsilon q_* + u_* \\ &= y_{*1} + \varepsilon q_*. \end{aligned}$$

Similarly,

(3.24) $$e_{10} = e_{11} + [z_*(z'z)^{-1} z' q - q_*].$$

Thus,

(3.25) $$m_{21} | H_0 = \frac{(e_{11} + \varepsilon b)'(e_{11} + \varepsilon b)}{(y_{*1} + \varepsilon q_*)'(y_{*1} + \varepsilon q_*)},$$

(3.26) $$m_{21} | H_1 = \frac{e_{11}' e_{11}}{y_{*1}' y_{*1}},$$

where b is appropriately defined. It follows that

(3.27) $$m_{21} | H_0 = \left(m_{21} | H_1 + 2\varepsilon \frac{e_{11}' b}{y_{*1}' y_{*1}} + \varepsilon^2 \frac{b'b}{y_{*1}' y_{*1}} \right) \left(1 + \varepsilon^2 \frac{q_*' q_*}{y_{*1}' y_{*1}} + 2\varepsilon \frac{q_*' y_{*1}}{y_{*1}' y_{*1}} \right)^{-1}.$$

Applying the geometric expansion, we can write

(3.28) $$m_{21} | H_0 = m_{21} | H_1 + 2\varepsilon \left(\frac{e_{11}' b}{y_{*1}' y_{*1}} \frac{q_*' y_{*1} e_{11}' e_{11}}{(y_{*1}' y_{*1})^2} \right) + 0(\varepsilon)^2$$

where $0(\varepsilon^2)$ refers to terms of order ε^2 and lower. To have

$$E(m_{21} | H_0) > E(m_{21} | H_1)$$

to order ε^2, we have to impose some conditions on b and q_*, and hence on q and q_*.

Thus, m_2 does not fulfill the regressor selection criterion straightaway, unless some further conditions are met. It will, of course, give the same ranking as m_1 for modeling the same series, and hence may be used interchangably.

Since much of the claimed superiority of a criterion such as m_2 over the simple M.S.E. lies in its ability to compare predictions on non-corresponding series, it should ideally fulfill a more demanding test of its ability than simple regressor selection (for example, regressand selection). The analysis for approximate specification will still be valid if two models with different regressands (and the same regressor set) are written as differing by a term of order ε.

c. *The "Janus" coefficient* ($m_3 = (e'e/N)/(e_S'e_S/T)$). For the regression model, we can write

$$m_{30} = \frac{m_{10}}{\frac{1}{T} y' M_x y} \tag{3.29}$$

$$m_{31} = \frac{m_{11}}{\frac{1}{T} y' M_z y}, \tag{3.30}$$

where

$$M_x = I - x(x'x)^{-1}x'$$
$$M_z = I - z(z'z)^{-1}z'.$$

Thus,

$$E(m_{30} \mid H_0) = E(m_{10} \mid H_0) T E((y'M_x y)^{-1} \mid H_0) \tag{3.31}$$

$$E(m_{31} \mid H_0) = E(m_{11} \mid H_0) T E((y'M_z y)^{-1} \mid H_0) \tag{3.32}$$

$$E(m_{31} \mid H_1) = E(m_{11} \mid H_1) T E((y'M_z y)^{-1} \mid H_1). \tag{3.33}$$

Recalling the normality assumption about u and hence y, it can be seen that $y'M_x y/2$ is distributed as a central chi-square with $T-K$ degrees of freedom, and $y'M_z y/2$ is a non-central chi-square with $T-P$ degrees of freedom, and non-centrality parameter

$$V = \frac{\beta' x' M_z x \beta}{2\sigma^2}$$

under H_0. Under H_1, $y'M_z y/2$ is a chi-square variate with $T-P$ degrees of freedom. Thus,

$$E((y'M_x y)^{-1} | H_0) = \frac{1}{\sigma^2} \frac{1}{T-K-2} \tag{3.34}$$

$$E((y'M_z y)^{-1} | H_0) = \frac{1}{\sigma_2} \frac{1}{T-P-2} e_1^{-v} F_1\left(\frac{T-P}{2} - 1; \frac{T-P}{2}; V\right) \tag{3.35}$$

$$E((y'M_z y)^{-1} | H_1) = \frac{1}{\sigma^2} \frac{1}{T-P-2}; \tag{3.36}$$

where ${}_1F_1(;\,;)$ is the confluent hypergeometric function. Once again, we run into the problem of noncomparability involving r_1 and r_2. Even if we take the asymptotic (large-v) expansion of the confluent hypergeometric function to write

$$E(m_{31}|H_0) - E(m_{30}|H_0) = \frac{1}{v}\left(\frac{\beta' D' D \beta}{N\sigma^2} + \frac{r_2}{N} + 1\right) \tag{3.37}$$

$$-\left(\frac{r_1}{N}+1\right)\frac{T}{T-K-2}+0\left(\frac{1}{v^2}\right);$$

where $0(1/v^2)$ refers to terms of order v^{-2} and lower; this does not yield any information as to the sign. This problem disappears when the second condition is examined.

$$E(m_{31}|H_1)=\left(\frac{r_2}{N}+1\right)\frac{T}{T-P-2}, \tag{3.38}$$

and hence

$$E(m_{31}|H_1)-E(m_{31}|H_0)=\left(\frac{r_2}{N}+1\right)\frac{T}{T-P-2}\Big[1-e_1^{-v}F_1 \left(\frac{T-P}{2}-1;\frac{T-P}{2};v\right)-\frac{\beta'D'D\beta}{2}e_1^{-v}F_1\left(\frac{T-P}{2}-1;\frac{T-P}{2};v\right)\Big]; \tag{3.39}$$

we can write

$$ {}_1F_1\left(\frac{T-P}{2}-1;\frac{T-P}{2};v\right) = 1+\frac{\frac{T-P}{2}-1}{\frac{T-P}{2}}v+\frac{\frac{T-P}{2}-1}{\frac{T-P}{2}+1}\frac{v^2}{2!}+\cdots = 1+e_1v+e_2\frac{v^2}{2!}+\cdots \tag{3.40}$$

where the successive coefficients e_i are all less than 1. Thus, the first term on the right hand side of (3.39) is positive, whereas the second one is negative. However, for large T, the second term will dominate; as long as $v/(T-P)$ is assumed to have a finite limiting value, we can write

$$e_1^{-v}F_1\left(\frac{T-P}{2}-1;\frac{T-P}{2};v\right)=\frac{T-P}{T-P-2v}+0(T^{-2}) \tag{3.41}$$

(cf. Slater [1960]), and the first term will be negative. Thus, m_3 satisfies the second regressor selection criterion at least in the asymptote. (It will probably also do so for finite T).

d. *The variance ratio* ($m_4=e'ue/A'uA$). The analysis for m_4 is very similar to that for m_2. Let J be an $N\times(N-1)$ matrix which satisfies

$$J'J=I \tag{3.42}$$

$$J\,J'=U. \tag{3.43}$$

Defining

$$P^*=J'P, \tag{3.44}$$

(3.45) $$A^* = J'A,$$

the analysis of m_2 in this section goes through, reading P^* for P, A^* for A, etc. The only major change is in degrees of freedom ($N-1$ rather than N) and the concentration parameter. As far as the fulfillment of the regressor selection criteria is concerned, the analysis is unchanged. However, one must remember that this requirement is by no means the only possible one that a good evaluative measure should satisfy.

One may get a radically different ranking if some other condition is set up. In any case, a more interesting variation of this problem would be the ranking these statistics give to a series of "untrue" models. It is rather optimistic to assume that in any situation, one of the small number of models estimated by the researcher is necessarily the "true" one.

University of Delhi, India

APPENDIX

EXACT MOMENTS AND APPROXIMATIONS

To derive the exact moments of m_2, etc., I use some results on moments of functions of noncentral chi-square variables. These are stated without proof (for proof, see e.g., Nagar and Amanullah [1973]).

Let Z be distributed by an N dimensional normal law, with mean $\bar{Z}$ and I (the identity matrix) as its dispersion matrix. Then the r.v. $w=z'z$ is distributed as a noncentral chi-square variate with N degrees of freedom and noncentrality parameter $v^*=\bar{z}'z/2$. We can write the following expectations

(A.1) $$E(w^{-1}) = \frac{1}{2}e^{-v^*}f_{-1,0};$$

(A.2) $$E(z_i w^{-1}) = \frac{1}{2}\bar{z}_i e^{-v^*}f_{0,1};$$

(A.3) $$E(z_i^2 w^{-1}) = \frac{1}{2}e^{-v^*}(\bar{z}_i^2 f_{-1,2} + f_{0,1});$$

(A.4) $$E(z_i z_j w^{-1}) = \frac{1}{2}e^{-v^*}\bar{z}_i\bar{z}_j f_{-1,2} \qquad \text{for} \quad i \neq j;$$

where $N>2$; and

$$f_{a,c} = \frac{\Gamma\left(\frac{N}{2}+a\right)}{\Gamma\left(\frac{N}{2}+c\right)}\,{}_1F_1\left(\frac{N}{2}+a;\frac{N}{2}+c;v^*\right).$$

To derive the expectation of m_2, we have to find s_1, s_2, defined in (3.19) above.

Let us write $z=y_*/\sigma$, and $v^*=(\beta' x_*' x_* \beta)/2\sigma^2$. The results of (A.1)–(A.4) are applicable. In particular,

$$s_1 = E(y_*' y^*)^{-1} = \frac{1}{\sigma^2} E(w^{-1}) \tag{A.5}$$

$$= \frac{1}{2\sigma^2} e^{-v^*} f_{-1,0}$$

Further,

$$s_2 = E[y_*(y_*' y_*)^{-1}]$$

$$= \frac{1}{\sigma} E(zw^{-1})$$

$$= \frac{1}{2\sigma^2} x_* \beta e^{-v^*} f_{0,1}.$$

If we employ a large v^* expansion, with N bounded, we can write

$$e^{-v^*} f_{a,c} \simeq \frac{1}{v^{*(c-a)}} \left[1 + \frac{\left(1 - \frac{N}{2} - a\right)}{v} + \frac{\left(1 - \frac{N}{2} - a\right)}{v} + \cdots \right] \tag{A.7}$$

for N bounded and $v^* \to \infty$. Thus

$$e^{-v^*} f_{-1,0} \simeq \frac{1}{v^*} + 0\left(\frac{1}{v^{*2}}\right) \tag{A.8}$$

$$e^{-v^*} f_{0,1} \simeq v^{*-1} + 0\left(\frac{1}{v^{*2}}\right). \tag{A.9}$$

Substituting for D in expression (3.19), we get

$$E(m_{21} - m_{20}) \,|\, \mathrm{H}_0 = \sigma^2(r_2 - r_1) - v^* e^{-v^*} f_{-1,0} + 2v^* e^{-v^*} f_{0,1} + K_1 e^{-v^*} f_{-1,0} + K_2 e^{-v^*} f_{0,1}, \tag{A.10}$$

where K_1, K_2, do not contain v^*. Combining with (A.8), (A.9), this yields

$$E(m_{21} - m_{20}) \,|\, \mathrm{H}_0 = \sigma^2(r_2 - r_1) + 1 + 0\left(\frac{1}{v^*}\right). \tag{A.11}$$

Thus, $r_2 - r_1 > 0$ is sufficient to make this difference positive to order v^{*-1}.

REFERENCES

DHRYMES, P. J. et al., "Criteria for Evaluation of Econometric Models," *Annals of Social and Economic Measurement*, 1 (1972), 291–324.

FISHER, F. M., "On the Cost of Approximate Specification in Simultaneous Equations Models," *Econometrica*, 29 (1961), 139–170.

GADD, A. AND H. WOLD, "The Janus Coefficient: A Measure for the Accuracy of Predictions," in, H. Wold, ed., *Econometric Model Building* (Amsterdam: North-Holland, 1964).

Granger, C. W. J. and Paul Newbold, *Forecasting Economic Time Series* (New York: Academic Press, 1977).

Nagar, A. L. and Aman Ullah, "Note on Approximate Skewness and Kurtosis of the Two-stage Least Squares Estimator," *Indian Economic Review*, 8 (April, 1973), 69–80.

Slater, L. J., *Confluent Hypergeometric Functions* (Cambridge: Cambridge University Press, 1970).

Theil, H., *Applied Economic Forecasting* (Amsterdam: North-Holland, 1966).