E C O N O M I C S B U L L E T I N

The error of prediction for a simultaneous equation model

Alexander Gorobets
Sevastopol National Technical University

Abstract

In this paper the formulas for the matrices of the mean squared prediction error are derived for both the underfitted and the overfitted models of unrestricted reduced form of a linear simultaneous equation system.

Author is grateful to Erasmus Research Institute of Management for supporting of this research
Citation: Gorobets, Alexander, (2005) "The error of prediction for a simultaneous equation model." *Economics Bulletin,* Vol. 3, No. 17 pp. 1−7
Submitted: December 4, 2004. **Accepted:** March 25, 2005.
URL: http://www.economicsbulletin.com/2005/volume3/EB−04C30004A.pdf

1. Introduction

One of the most important functions of a simultaneous equation model is prediction the values of endogenous variables given the values of the predetermined variables and a lot of work has been done to estimate the accuracy of such predictions. Hooper and Zellner (1961) obtained the covariance matrix of the prediction error for unrestricted reduced form and Goldberger, Nagar and Odeh (1961) derived one for restricted reduced form. Properties of predictions for partially restricted reduced form have been analyzed by Amemiya (1966), Kakwani and Court (1972) and Nagar and Sahay (1978). The comparison of these estimators in the context of prediction has been carried on by Dhrymes (1973) and Park (1982). However all these derivations are made for reduced forms of correctly specified linear simultaneous equation models and they still remain unknown for the under and the over specified models.

The purpose of this paper is to derive the matrices of the mean squared prediction error for both the underfitted and the overfitted models of unrestricted reduced form of a linear simultaneous equation system.

The paper is organized as follows: Section 2 presents the basic model and its assumptions. Sections 3 and 4 derive the matrices of the mean squared prediction error for the underfitted and the overfitted models of unrestricted reduced form respectively. Section 5 gives the conclusions. An appendix contains the proofs of these derivations.

2. The model specification

A standard linear simultaneous equation system is given by

$$Y\Gamma + XB + U = 0\,, \tag{1}$$

where Y is a $N \times M$ matrix of observations on M endogenous variables, X is a $N \times K$ non-stochastic matrix of observations on K exogenous variables, U is a $N \times M$ matrix of independent structural disturbances distributed as $N(0,\Sigma)$, Γ and B are matrices of structural parameters of order $M \times M$ and $K \times M$ respectively and $|\Gamma| \neq 0$.

The reduced form of the model (1) is

$$Y = X\Pi + V\,, \tag{2}$$

where $\Pi = -B\Gamma^{-1}$ and $V = -U\Gamma^{-1}$.

It follows that $V \sim N(0,\Omega)$, where $\Omega = (\Gamma^{-1})'\Sigma(\Gamma^{-1})$ is positive definite.

Suppose that the model (2) is true, i.e. correctly specified in variables. Then a consistent estimate of the matrix of the mean squared prediction error for the true model of unrestricted reduced form is

$$\widehat{\Omega}_t = (1 + x_\tau'(X'X)^{-1}x_\tau)\widehat{\Omega}\,, \tag{3}$$

where x_τ is a $K \times 1$ vector of values for X in the prediction period τ and

$$\widehat{\Omega} = \frac{(Y - X\widehat{\Pi})'(Y - X\widehat{\Pi})}{N - K} \tag{4}$$

is an unbiased estimator of covariance matrix Ω, where $\widehat{\Pi} = (X'X)^{-1}X'Y$.

3. The underfitted model prediction

Let's consider the error of prediction for the underfitted model of unrestricted reduced form. The reduced form (2) may be partitioned as

$$Y = X\Pi + V = \begin{bmatrix} X_1 & X_2 \end{bmatrix} \begin{bmatrix} \Pi_1 \\ \Pi_2 \end{bmatrix} + V = X_1\Pi_1 + X_2\Pi_2 + V, \tag{5}$$

where X_1 is a $N \times k$ submatrix of regressors included in the underfitted model and X_2 is a $N \times (K-k)$ submatrix of omitted regressors, Π_1 and Π_2 are submatrices of parameters.
Let $x_{1\tau}$ and $x_{2\tau}$ be the vectors of observations on X_1 and X_2 at the prediction period. Then the values of the endogenous variables can be predicted by the underfitted model as

$$\widehat{y}_\tau = \widehat{P}'x_{1\tau}, \tag{6}$$

where $\widehat{P}$ is a biased estimate of parameters of order $k \times M$. The observed values of the endogenous variables in the prediction period are given by

$$y_\tau = \Pi'x_\tau + v_\tau = \Pi_1'x_{1\tau} + \Pi_2'x_{2\tau} + v_\tau, \tag{7}$$

where v_τ is a column vector of disturbances at time τ.
The error of prediction is

$$\varepsilon_\tau = y_\tau - \widehat{y}_\tau = (\Pi_1' - \widehat{P}')x_{1\tau} + \Pi_2'x_{2\tau} + v_\tau. \tag{8}$$

The first result may now be stated.

The matrix of the mean squared prediction error for the underfitted (biased) model of reduced form is

$$\Omega_b = (1 + x_{1\tau}'(X_1'X_1)^{-1}x_{1\tau})\Omega + (x_{1\tau}'C\Pi_2 - x_{2\tau}'\Pi_2)'(x_{1\tau}'C\Pi_2 - x_{2\tau}'\Pi_2), \tag{9}$$

where $C = (X_1'X_1)^{-1}X_1'X_2$.
Proof. See the appendix to this paper.

The first term of (9) is a covariance matrix of the prediction error for the underfitted model and the second one is a bias due to underfitting, which depends on both the postulated and the true models.

In practice Ω and Π_2 are unknown but instead the consistent estimates of them are available (see (4)). A consistent estimate of Ω_b is obtained by using these estimates in (9).

To compare the quality of prediction for the biased and the true models we can use the generalized error of prediction for the system, which is defined as the trace or the determinant of the matrix of the mean squared prediction error.

By partitioning X on X_1 and X_2 and then inversing the block matrix in (3) we have

$$\widehat{\Omega}_t = (1 + x'_{1\tau}(X'_1X_1)^{-1}x_{1\tau})\widehat{\Omega} + (x'_{1\tau}\widetilde{L} - x'_{2\tau})\widetilde{D}(x'_{1\tau}\widetilde{L} - x'_{2\tau})'\widehat{\Omega}\,, \tag{10}$$

where $\widetilde{L} = (X'_1X_1)^{-1}X'_1X_2 = C$, $\widetilde{D} = (X'_2\widetilde{R}X_2)^{-1}$, $\widetilde{R} = I_N - X_1(X'_1X_1)^{-1}X'_1$.

Let $g = x'_{1\tau}C - x'_{2\tau}$ and $J = (1 + x'_{1\tau}(X'_1X_1)^{-1}x_{1\tau})\widehat{\Omega}$, then the prediction from the underfitted model is superior to that from the true model if

$$g\widehat{\Pi}_2\widehat{\Pi}'_2g' \le \mathrm{tr}(g\widetilde{D}g'\widehat{\Omega}) \tag{11}$$

or

$$\left|J + \widehat{\Pi}'_2g'g\widehat{\Pi}_2\right| \le \left|J + g\widetilde{D}g'\widehat{\Omega}\right|. \tag{12}$$

The left-hand side of (11) is a scalar and hence the trace operator is left out.

4. The overfitted model prediction

In the following section, I will focus on the problem of prediction, using an overfitted model of an unrestricted reduced form. This model is given by

$$Y = X\Pi + X_A\Pi_A + V = \begin{bmatrix} X & X_A \end{bmatrix}\begin{bmatrix} \Pi \\ \Pi_A \end{bmatrix} + V = WG + V\,, \tag{13}$$

where X_A is a $N \times h$ matrix of non-relevant regressors included in the true model, Π_A is a $h \times M$ matrix of parameters, W is a $N \times (K+h)$ block matrix of all regressors in the overfitted model and, accordingly G is a $(K+h) \times M$ block matrix of all parameters. Then the prediction period values of the endogenous variables are estimated by

$$\widehat{y}_\tau = \widehat{G}'w_\tau\,, \tag{14}$$

where $\widehat{G}$ is an unbiased estimate of all parameters in the overfitted model and $w'_\tau = \begin{bmatrix} x'_\tau & x'_{A\tau} \end{bmatrix}$ is a $1 \times (K+h)$ vector of observations on W at the prediction period. The true value of Y in this prediction period is defined by (7). Then the error of prediction is

$$\varepsilon_\tau = y_\tau - \widehat{y}_\tau = \Pi'x_\tau + v_\tau - \widehat{G}'w_\tau = \begin{bmatrix} \Pi' & 0' \end{bmatrix}\begin{bmatrix} x_\tau \\ x_{A\tau} \end{bmatrix} - \widehat{G}'w_\tau + v_\tau = (\overline{G}' - \widehat{G}')w_\tau + v_\tau\,, \tag{15}$$

where $\overline{G}' = \begin{bmatrix} \Pi' & 0' \end{bmatrix}$, 0 is a $h \times M$ matrix of zeros.

The second result may now be formulated.

The matrix of the mean squared prediction error for the overfitted model of reduced form is

$$\Omega_o = (1 + x'_\tau (X'X)^{-1} x_\tau)\Omega + (x'_\tau L - x'_{A\tau}) D (x'_\tau L - x'_{A\tau})'\Omega\,, \tag{16}$$

where $L = (X'X)^{-1} X'X_A$, $D = (X'_A R X_A)^{-1}$ and $R = I_N - X(X'X)^{-1} X'$.

Proof. See the appendix to this paper.

The first term of (16) is a covariance matrix of the prediction error for the true model and the second one is a bias due to overfitting. A consistent estimate of Ω_o is obtained by using $\widehat{\Omega}$ in (16).

Comparison between (3) and (16) shows that a quality of prediction for the true model is always better than or equal to that in the overfitted model. Equality holds if X_A is orthogonal to X.

5. Conclusions

In this paper the matrices of the mean squared prediction error for both the underfitted and the overfitted models of unrestricted reduced form of a linear simultaneous equation system are obtained. It should be noted that they are the generalization of the mean squared prediction error for a single miss specified regression equation (see, e.g., Hocking, 1976; Seber, 1977).

Further it is necessary to derive the analogous matrices for the structural form of a simultaneous equation system and compare the quality of prediction between structural and reduced forms.

Appendix

A. Derivation of equation (9)

We have

$$\begin{aligned}
\Omega_b &= E(\varepsilon_\tau \varepsilon'_\tau) = E((y_\tau - \widehat{y}_\tau)(y_\tau - \widehat{y}_\tau)') = \\
&= E(((\Pi'_1 - \widehat{P}')x_{1\tau} + \Pi'_2 x_{2\tau} + v_\tau)((\Pi'_1 - \widehat{P}')x_{1\tau} + \Pi'_2 x_{2\tau} + v_\tau)') = \\
&= E((\Pi_1 - \widehat{P})' x_{1\tau} x'_{1\tau} (\Pi_1 - \widehat{P}) + (\Pi_1 - \widehat{P})' x_{1\tau} x'_{2\tau} \Pi_2 + \\
&+ (\Pi_1 - \widehat{P})' x_{1\tau} v'_\tau + \Pi'_2 x_{2\tau} x'_{1\tau} (\Pi_1 - \widehat{P}) + \Pi'_2 x_{2\tau} x'_{2\tau} \Pi_2 + \\
&+ \Pi'_2 x_{2\tau} v'_\tau + v_\tau x'_{1\tau} (\Pi_1 - \widehat{P}) + v_\tau x'_{2\tau} \Pi_2 + v_\tau v'_\tau) = \\
&= E((\Pi_1 - \widehat{P})' x_{1\tau} x'_{1\tau} (\Pi_1 - \widehat{P})) + E(\Pi_1 - \widehat{P})' x_{1\tau} x'_{2\tau} \Pi_2 + \\
&+ \Pi'_2 x_{2\tau} x'_{1\tau} E(\Pi_1 - \widehat{P}) + \Pi'_2 x_{2\tau} x'_{2\tau} \Pi_2 + \Omega
\end{aligned} \tag{17}$$

The 1st term of (17) can be evaluated by using the following facts:

$$\begin{aligned}\widehat{P} &= (X_1'X_1)^{-1}X_1'Y = (X_1'X_1)^{-1}X_1'(X_1\varPi_1 + X_2\varPi_2 + V) = \\ &= \varPi_1 + (X_1'X_1)^{-1}X_1'X_2\varPi_2 + (X_1'X_1)^{-1}X_1'V = \\ &= \varPi_1 + C\varPi_2 + (X_1'X_1)^{-1}X_1'V,\end{aligned} \tag{18}$$

$$E(\widehat{P}) = E(\varPi_1 + C\varPi_2 + (X_1'X_1)^{-1}X_1'V) = \varPi_1 + C\varPi_2. \tag{19}$$

Then

$$\begin{aligned}&E((\varPi_1 - \widehat{P})'x_{1\tau}x_{1\tau}'(\varPi_1 - \widehat{P})) = \\ &= E((C\varPi_2 + (X_1'X_1)^{-1}X_1'V)'x_{1\tau}x_{1\tau}'(C\varPi_2 + (X_1'X_1)^{-1}X_1'V)) = \\ &= E((C\varPi_2)'x_{1\tau}x_{1\tau}'C\varPi_2 + V'X_1(X_1'X_1)^{-1}x_{1\tau}x_{1\tau}'C\varPi_2 + \\ &+ (C\varPi_2)'x_{1\tau}x_{1\tau}'(X_1'X_1)^{-1}X_1'V + V'X_1(X_1'X_1)^{-1}x_{1\tau}x_{1\tau}'(X_1'X_1)^{-1}X_1'V) = \\ &= (C\varPi_2)'x_{1\tau}x_{1\tau}'C\varPi_2 + E(V'X_1(X_1'X_1)^{-1}x_{1\tau}x_{1\tau}'(X_1'X_1)^{-1}X_1'V).\end{aligned} \tag{20}$$

The last term in (20) can be simplified as follows

$$E(V'X_1(X_1'X_1)^{-1}x_{1\tau}x_{1\tau}'(X_1'X_1)^{-1}X_1'V) = x_{1\tau}'(X_1'X_1)^{-1}x_{1\tau}\varOmega. \tag{21}$$

The 2nd and 3rd terms of (17) are derived by using (19):

$$E(\varPi_1 - \widehat{P})'x_{1\tau}x_{2\tau}'\varPi_2 = -(C\varPi_2)'x_{1\tau}x_{2\tau}'\varPi_2. \tag{22}$$

Collecting terms, we obtain (9) in the text:

$$\begin{aligned}\varOmega_b &= (C\varPi_2)'x_{1\tau}x_{1\tau}'C\varPi_2 + x_{1\tau}'(X_1'X_1)^{-1}x_{1\tau}\varOmega - \\ &- (C\varPi_2)'x_{1\tau}x_{2\tau}'\varPi_2 - \varPi_2'x_{2\tau}x_{1\tau}'C\varPi_2 + \varPi_2'x_{2\tau}x_{2\tau}'\varPi_2 + \varOmega = \\ &= (1 + x_{1\tau}'(X_1'X_1)^{-1}x_{1\tau})\varOmega + (x_{1\tau}'C\varPi_2 - x_{2\tau}'\varPi_2)'(x_{1\tau}'C\varPi_2 - x_{2\tau}'\varPi_2).\end{aligned} \tag{23}$$

B. Derivation of equation (16)

We have

$$\begin{aligned}\varOmega_o &= E(\varepsilon_\tau\varepsilon_\tau') = E((y_\tau - \widehat{y}_\tau)(y_\tau - \widehat{y}_\tau)') = \\ &= E(((\overline{G}' - \widehat{G}')w_\tau + v_\tau)((\overline{G}' - \widehat{G}')w_\tau + v_\tau)') = \\ &= E((\overline{G}' - \widehat{G}')w_\tau w_\tau'(\overline{G} - \widehat{G}) + (\overline{G}' - \widehat{G}')w_\tau v_\tau' + \\ &+ v_\tau w_\tau'(\overline{G} - \widehat{G}) + v_\tau v_\tau') = E((\overline{G} - \widehat{G})'w_\tau w_\tau'(\overline{G} - \widehat{G})) + \varOmega.\end{aligned} \tag{24}$$

The 1st term of (24) can be evaluated by using the following fact:

$$\widehat{G} = \begin{bmatrix}\widehat{\varPi} \\ \widehat{\varPi}_A\end{bmatrix} = (W'W)^{-1}W'Y = (W'W)^{-1}W'(W\overline{G} + V) = \overline{G} + (W'W)^{-1}W'V\,. \tag{25}$$

Then

$$E((\overline{G}-\widehat{G})'w_\tau w_\tau'(\overline{G}-\widehat{G})) = E(V'W(W'W)^{-1}w_\tau w_\tau'(W'W)^{-1}W'V). \tag{26}$$

The right side of (26) is similar to the last term in (20) and then by analogy we have

$$E(V'W(W'W)^{-1}w_\tau w_\tau'(W'W)^{-1}W'V) = w_\tau'(W'W)^{-1}w_\tau \Omega. \tag{27}$$

Then

$$\begin{aligned} w_\tau'(W'W)^{-1}w_\tau \Omega &= \begin{bmatrix} x_\tau' & x_{A\tau}' \end{bmatrix} \left(\begin{bmatrix} X' \\ X_A' \end{bmatrix} \begin{bmatrix} X & X_A \end{bmatrix} \right)^{-1} \begin{bmatrix} x_\tau' & x_{A\tau}' \end{bmatrix}' \Omega = \\ &= \begin{bmatrix} x_\tau' & x_{A\tau}' \end{bmatrix} \begin{bmatrix} (X'X)^{-1} + LDL' & -LD \\ -DL' & D \end{bmatrix} \begin{bmatrix} x_\tau' & x_{A\tau}' \end{bmatrix}' \Omega = \\ &= x_\tau'(X'X)^{-1}x_\tau \Omega + (x_\tau' L - x_{A\tau}')D(x_\tau' L - x_{A\tau}')'\Omega. \end{aligned} \tag{28}$$

And, finally, collecting terms, we obtain (16) in the text.

References

Amemiya, T. (1966) "On the use of principal components of independent variables in two-stage least-squares estimation" *International Economic Review* **7**, 283-303.

Dhrymes, P.J. (1973) "Restricted and unrestricted reduced forms: Asymptotic distribution and relative efficiency" *Econometrica* **41**, 119-134.

Goldberger, A.S., A.L. Nagar & H.S. Odeh (1961) "The covariance matrices of reduced-form coefficients and of forecasts for a structural econometric model" *Econometrica* **29**, 556-573.

Hocking, R.R. (1976) A Biometrics Invited Paper. "The Analysis and Selection of Variables in Linear Regression" *Biometrics* **32**, 1-49.

Hooper, J.W. & A. Zellner (1961) "The error of forecast for multivariate regression models" *Econometrica* **29**, 544-555.

Kakwani, N.C. & R.H. Court (1972) "Reduced form coefficient estimation and forecasting from a simultaneous equation model" *Australian Journal of Statistics* **14**, 143-160.

Nagar, A.L. & S.N. Sahay (1978) "The bias and mean squared error of forecasts from partially restricted reduced form" *Journal of Econometrics* **7**, 227-243.

Park, S.-B. (1982) "A forecasting property of the unrestricted, restricted, and partially restricted reduced-form coefficients" *Journal of Econometrics* **19**, 385-390.

Seber, G.A.F. (1977) *Linear Regression Analysis*, New York: Wiley.