The Error of Prediction for a Simultaneous Equation Model

Alexander Gorobets

ERASMUS RESEARCH INSTITUTE OF MANAGEMENT

REPORT SERIES
RESEARCH IN MANAGEMENT

BIBLIOGRAPHIC DATA AND CLASSIFICATIONS		
Abstract	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.	
Library of Congress Classification (LCC)	5001-6182	Business
	5546-5548.6	Office Organization and Management
	H 61.4	Forecasting
Journal of Economic Literature (JEL)	M	Business Administration and Business Economics
	M 10 L 2	Business Administration: general Firm Objectives, Organization and Behaviour
	C 53	Forecasting and other model applications
European Business Schools Library Group (EBSLG)	85 A	Business General
	100B 240 B	Organization Theory (general) Information Systems Management
	250 E	Forecasting (statistics)
Gemeenschappelijke Onderwerpsontsluiting (GOO)		
Classification GOO	85.00	Bedrijfskunde, Organisatiekunde: algemeen
	85.05 85.08	Management organisatie: algemeen Organisatiesociologie, organisatiepsychologie
	31.80	Toepassingen van de wiskunde
Keywords GOO	Bedrijfskunde / Bedrijfseconomie	
	Organisatieleer, informatietechnologie, prestatiebeoordeling	
	Simultane vergelijkingen, modellen, specificatie, voorspellingen	
Free keywords	simultaneous equations, misspecification, prediction	

THE ERROR OF PREDICTION FOR A SIMULTANEOUS EQUATION MODEL

By Alexander Gorobets

Erasmus Research Institute of Management, Erasmus University Rotterdam, Burg. Oudlaan 50 ,
P.O. Box 1738, 3000 DR Rotterdam, The Netherlands.; agorobets@fbk.eur.nl

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.

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

(2.1) $$Y\Gamma + XB + U = 0,$$

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 (2.1) is

(2.2) $$Y = X\Pi + V,$$

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

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

Suppose that the model (2.2) 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

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

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

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

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 (2.2) may be partitioned as

(3.1) $$Y = X\Pi + V = \begin{bmatrix} X_1 & X_2 \end{bmatrix}\begin{bmatrix} \Pi_1 \\ \Pi_2 \end{bmatrix} + V = X_1\Pi_1 + X_2\Pi_2 + V,$$

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

$$(3.2) \qquad\qquad \hat{y}_\tau = \hat{P}' x_{1\tau},$$

where $\hat{P}$ 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

$$(3.3) \qquad\qquad y_\tau = \Pi' x_\tau + v_\tau = \Pi_1' x_{1\tau} + \Pi_2' x_{2\tau} + v_\tau,$$

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

The error of prediction is

$$(3.4) \qquad\qquad \varepsilon_\tau = y_\tau - \hat{y}_\tau = (\Pi_1' - \hat{P}') x_{1\tau} + \Pi_2' x_{2\tau} + v_\tau.$$

The first result may now be stated.

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

$$(3.5) \qquad \Omega_b = (1 + x_{1\tau}'(X_1'X_1)^{-1} x_{1\tau})\Omega + (x_{1\tau}'C\Pi_2 - x_{2\tau}'\Pi_2)'(x_{1\tau}'C\Pi_2 - x_{2\tau}'\Pi_2),$$

where $C = (X_1'X_1)^{-1} X_1' X_2$.

Proof. See the appendix to this paper.

The first term of (3.5) 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 (2.4)). A consistent estimate of Ω_b is obtained by using these estimates in (3.5).

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 (2.3) we have

$$(3.6) \qquad \hat{\Omega}_t = (1 + x_{1\tau}'(X_1'X_1)^{-1} x_{1\tau})\hat{\Omega} + (x_{1\tau}'\tilde{L} - x_{2\tau}')\tilde{D}(x_{1\tau}'\tilde{L} - x_{2\tau}')'\hat{\Omega},$$

where $\tilde{L} = (X_1'X_1)^{-1} X_1' X_2 = C$, $\tilde{D} = (X_2'\tilde{R}X_2)^{-1}$, $\tilde{R} = I_N - X_1(X_1'X_1)^{-1} X_1'$.

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

(3.7) $$g\widehat{\Pi}_2\widehat{\Pi}'_2g' \leq \text{tr}(g\widetilde{D}g'\widehat{\Omega})$$

or

(3.8) $$\left|J + \widehat{\Pi}'_2g'g\widehat{\Pi}_2\right| \leq \left|J + g\widetilde{D}g'\widehat{\Omega}\right|.$$

The left-hand side of (3.7) 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

(4.1) $$Y = X\Pi + X_A\Pi_A + V = \begin{bmatrix} X & X_A \end{bmatrix}\begin{bmatrix} \Pi \\ \Pi_A \end{bmatrix} + V = WG + V,$$

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

(4.2) $$\widehat{y}_\tau = \widehat{G}'w_\tau,$$

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 (3.3). Then the error of prediction is

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

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

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

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 (4.4) 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 (4.4).

Comparison between (2.3) and (4.4) 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 (3.5)

We have

$$
\begin{aligned}
\Omega_b &= E(\varepsilon_\tau \varepsilon_\tau') = E((y_\tau - \widehat{y}_\tau)(y_\tau - \widehat{y}_\tau)') = \\
&= E(((\Pi_1' - \widehat{P}')x_{1\tau} + \Pi_2' x_{2\tau} + v_\tau)((\Pi_1' - \widehat{P}')x_{1\tau} + \Pi_2' x_{2\tau} + v_\tau)') = \\
&= E((\Pi_1 - \widehat{P})'x_{1\tau}x_{1\tau}'(\Pi_1 - \widehat{P}) + (\Pi_1 - \widehat{P})'x_{1\tau}x_{2\tau}'\Pi_2 + \\
&\quad + (\Pi_1 - \widehat{P})'x_{1\tau}v_\tau' + \Pi_2' x_{2\tau}x_{1\tau}'(\Pi_1 - \widehat{P}) + \Pi_2' x_{2\tau}x_{2\tau}'\Pi_2 + \\
&\quad + \Pi_2' x_{2\tau}v_\tau' + v_\tau x_{1\tau}'(\Pi_1 - \widehat{P}) + v_\tau x_{2\tau}'\Pi_2 + v_\tau v_\tau') = \\
&= E((\Pi_1 - \widehat{P})'x_{1\tau}x_{1\tau}'(\Pi_1 - \widehat{P})) + E(\Pi_1 - \widehat{P})'x_{1\tau}x_{2\tau}'\Pi_2 + \\
&\quad + \Pi_2' x_{2\tau}x_{1\tau}'E(\Pi_1 - \widehat{P}) + \Pi_2' x_{2\tau}x_{2\tau}'\Pi_2 + \Omega
\end{aligned}
\tag{A.1}
$$

The 1$^{\text{st}}$ term of (A.1) can be evaluated by using the following facts:

$$\widehat{P} = (X_1'X_1)^{-1}X_1'Y = (X_1'X_1)^{-1}X_1'(X_1\Pi_1 + X_2\Pi_2 + V) =$$

(A.2)
$$= \Pi_1 + (X_1'X_1)^{-1}X_1'X_2\Pi_2 + (X_1'X_1)^{-1}X_1'V =$$

$$= \Pi_1 + C\Pi_2 + (X_1'X_1)^{-1}X_1'V,$$

(A.3)
$$E(\widehat{P}) = E(\Pi_1 + C\Pi_2 + (X_1'X_1)^{-1}X_1'V) = \Pi_1 + C\Pi_2.$$

Then

$$E((\Pi_1 - \widehat{P})'x_{1\tau}x_{1\tau}'(\Pi_1 - \widehat{P})) =$$

$$= E((C\Pi_2 + (X_1'X_1)^{-1}X_1'V)'x_{1\tau}x_{1\tau}'(C\Pi_2 + (X_1'X_1)^{-1}X_1'V)) =$$

(A.4)
$$= E((C\Pi_2)'x_{1\tau}x_{1\tau}'C\Pi_2 + V'X_1(X_1'X_1)^{-1}x_{1\tau}x_{1\tau}'C\Pi_2 +$$

$$+ (C\Pi_2)'x_{1\tau}x_{1\tau}'(X_1'X_1)^{-1}X_1'V + V'X_1(X_1'X_1)^{-1}x_{1\tau}x_{1\tau}'(X_1'X_1)^{-1}X_1'V) =$$

$$= (C\Pi_2)'x_{1\tau}x_{1\tau}'C\Pi_2 + E(V'X_1(X_1'X_1)^{-1}x_{1\tau}x_{1\tau}'(X_1'X_1)^{-1}X_1'V).$$

The last term in (A.4) can be simplified as follows

(A.5)
$$E(V'X_1(X_1'X_1)^{-1}x_{1\tau}x_{1\tau}'(X_1'X_1)^{-1}X_1'V) = x_{1\tau}'(X_1'X_1)^{-1}x_{1\tau}\Omega.$$

The 2$^{\text{nd}}$ and 3$^{\text{rd}}$ terms of (A.1) are derived by using (A.3):

(A.6)
$$E(\Pi_1 - \widehat{P})'x_{1\tau}x_{2\tau}'\Pi_2 = -(C\Pi_2)'x_{1\tau}x_{2\tau}'\Pi_2.$$

Collecting terms, we obtain (3.6) in the text:

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

(A.7)
$$- (C\Pi_2)'x_{1\tau}x_{2\tau}'\Pi_2 - \Pi_2'x_{2\tau}x_{1\tau}'C\Pi_2 + \Pi_2'x_{2\tau}x_{2\tau}'\Pi_2 + \Omega =$$

$$= (1 + x_{1\tau}'(X_1'X_1)^{-1}x_{1\tau})\Omega + (x_{1\tau}'C\Pi_2 - x_{2\tau}'\Pi_2)'(x_{1\tau}'C\Pi_2 - x_{2\tau}'\Pi_2).$$

B. Derivation of equation (4.4)

We have

$$
\begin{aligned}
\Omega_o &= E(\varepsilon_\tau \varepsilon_\tau') = E((y_\tau - \hat{y}_\tau)(y_\tau - \hat{y}_\tau)') = \\
&= E(((\overline{G}' - \hat{G}')w_\tau + v_\tau)((\overline{G}' - \hat{G}')w_\tau + v_\tau)') = \\
&= E((\overline{G}' - \hat{G}')w_\tau w_\tau'(\overline{G} - \hat{G}) + (\overline{G}' - \hat{G}')w_\tau v_\tau' + \\
&\quad + v_\tau w_\tau'(\overline{G} - \hat{G}) + v_\tau v_\tau') = E((\overline{G} - \hat{G})'w_\tau w_\tau'(\overline{G} - \hat{G})) + \Omega.
\end{aligned}
$$

(B.1)

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

(B.2) $\quad \hat{G} = \begin{bmatrix} \hat{\Pi} \\ \hat{\Pi}_A \end{bmatrix} = (W'W)^{-1}W'Y = (W'W)^{-1}W'(W\overline{G} + V) = \overline{G} + (W'W)^{-1}W'V .$

Then

(B.3) $\qquad E((\overline{G} - \hat{G})'w_\tau w_\tau'(\overline{G} - \hat{G})) = E(V'W(W'W)^{-1}w_\tau w_\tau'(W'W)^{-1}W'V).$

The right side of (B.3) is similar to the last term in (A.4) and then by analogy we have

(B.4) $\qquad E(V'W(W'W)^{-1}w_\tau w_\tau'(W'W)^{-1}W'V) = w_\tau'(W'W)^{-1}w_\tau \Omega.$

Then

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

(B.5)

And, finally, collecting terms, we obtain (4.4) 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, and H.S. Odeh (1961): "The covariance matrices of reduced-form coefficients and of forecasts for a structural econometric model", *Econometrica*, 29, 556-573.

Gorobets, A.D. (2001): "Control objects structural identification method, which are described by the simultaneous equation systems", unpublished Ph.D. dissertation. Sevastopol National Technical University.

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

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

Kakwani, N.C., and 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., and 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.

Publications in the ERIM Report Series Research* in Management

ERIM Research Program: "Organizing for Performance"

2003

On The Future of Co-operatives: Talking Stock, Looking Ahead
George W.J. Hendrikse and Cees P. Veerman
ERS-2003-007-ORG
http://hdl.handle.net/1765/270

Governance of Chains and Networks: A Research Agenda
George W.J. Hendrikse
ERS-2003-018-ORG
http://hdl.handle.net/1765/298

Mystery Shopping: In-depth measurement of customer satisfaction
Martijn Hesselink, Ton van der Wiele
ERS-2003-020-ORG
http://hdl.handle.net/1765/281

Simultaneous Equation Systems Selection Method
Alexander Gorobets
ERS-2003-024-ORG
http://hdl.handle.net/1765/322

Stages Of Discovery And Entrepreneurship
Bart Nooteboom
ERS-2003-028-ORG
http://hdl.handle.net/1765/327

Change Of Routines: A Multi-Level Analysis
Bart Nooteboom and Irma Bogenrieder
ERS-2003-029-ORG
http://hdl.handle.net/1765/329

Generality, Specificity And Discovery
Bart Nooteboom
ERS-2003-030-ORG
http://hdl.handle.net/1765/330

Tracing Cold War in Post-Modern Management's Hot Issues
Slawomir J. Magala
ERS-2003-040-ORG
http://hdl.handle.net/1765/335

Networks in cultural, economic, and evolutionary perspective
Barbara Krug
ERS-2003-050-ORG

Creating competition & mastering markets
New entrants, monopolists, and regulators in transforming public utilities across the Atlantic
Willem Hulsink, Emiel Wubben
ERS-2003-051-ORG
http://hdl.handle.net/1765/434

Unhealthy Paradoxes of Healthy Identities
Slawomir J. Magala
ERS-2003-054-ORG
http://hdl.handle.net/1765/864

The Hidden Costs of Ubiquity: Globalisation and Terrorism
Barbara Krug and Patrick Reinmoeller
ERS-2003-062-ORG
http://hdl.handle.net/1765/993

Professional Elites in "Classless" Societies (from Marx to Debord)
Slawomir J. Magala
ERS-2003-069-ORG
http://hdl.handle.net/1765/974

Network effects on Entrepreneurial Processes: Start-ups in the Dutch ICT Industry 1990-2000
Willem Hulsink and Tom Elfring
ERS-2003-070-ORG
http://hdl.handle.net/1765/976