Asymptotic Properties of Full Information Estimators in Dynamic Autoregressive Simultaneous Equation Models

By

Phoebus J. Dhrymes and H. Erlat

Discussion Paper Number 30
December 1972

Asymptotic Properties of Full Information Estimators in Dynamic Autoregressive Simultaneous Equation Models

Phoebus J. Dhrymes and H. Erlat

1. Introduction

In a previous paper [3] we examined the problem of estimating, by maximum likelihood (ML) and three stage least squares-like methods, the parameters of the model

$$(1) \qquad y_{t\cdot} = y_{t\cdot}B + y_{t-1\cdot}C_0 + w_{t\cdot}C_1 + u_{t\cdot}, \qquad t = 1,2,\ldots,T$$

where $\{w_{t\cdot} : t = 0, \pm 1, \pm 2, \ldots\}$ is a sequence of s-element vectors of exogenous variables (which are uniformely bounded and nonstochastic), the error process obeys

$$(2) \qquad u_{t\cdot} = u_{t-1\cdot}R + \varepsilon_{t\cdot}$$

R is a stable matrix, $\{\varepsilon'_{t\cdot} : t = 0, \pm 1, \pm 2, \ldots\}$ is a sequence of independent identically distributed (i.i.d.) random variables such that

$$(3) \qquad E(\varepsilon'_{t\cdot}) = 0 \qquad Cov(\varepsilon'_{t\cdot}) = \Sigma$$

Σ being positive definite, and $y_{t\cdot}$ is the m-element vector of jointly dependent variables.

It was shown in [3] that the three stage least squares like procedure-termed there the full information dynamic autoregessive (FIDA) - satisfies, asymptotically, the same set of normal equations as the ML estimator, the difference being in the manner in which the jointly dependent variables are "purged" of their stochastic components. Both were shown to be consistent estimators; in addition, the asymptotic distribution of the estimators - though not explicitly obtained - was shown not to depend on the properties of the estimator of Σ beyond consistency but to depend on the asymptotic distribution properties of the estimator of R.

In this paper we derive explicity the joint asymptotic distribution of the FIDA estimator of B, C_0, C_1 and R.

2. Formulation of the Problem.

In connection with (1) and (2) we observe that we require that

(A.1) $(I - B)$ is nonsingular

(A.2) $C_0(I - B)^{-1}$, R are both stable matrices

Certain other assumptions will be invoked as the need for them arises.

We also observe that

$$(4) \qquad E(u'_{t\cdot}) = 0 , \qquad Cov(u'_{t\cdot}) = \sum_{i=0}^{\infty} R'^{i} \Sigma R^{i} = \Omega$$

and we assert that Ω is nonsingular.

The reduced form of (1) can be written as

$$(5) \qquad y_{t\cdot} = y_{t-1\cdot}\Pi_0 + w_{t\cdot}\Pi_{1\cdot} + v_{t\cdot}$$

where

$$(6) \qquad \Pi_0 = C_0(I - B)^{-1} , \quad \Pi_1 = C_1(I - B)^{-1} , \quad v_{t\cdot} = u_{t\cdot}(I - B)^{-1}$$

and the final form is

(7) $$y'_{t\cdot} = (I - \Pi'_0 L)^{-1} \Pi'_1 w'_{t\cdot} + (I - \Pi'_0 L)^{-1} v'_{t\cdot}$$

where L is the usual lag operator.

The FIDA estimator is obtained by minimizing

$$\mathrm{tr}\ \Sigma^{-1}(\tilde{Z}A^* - Z_{-1}A^*R)'\ (\tilde{Z}A^* - Z_{-1}A^*R)$$

with respect to A and R subject to a prior (consistent) estimator of Σ where

(8) $$A^* = (I - B', -C'_0, -C'_1)', \quad Z = (y_{t\cdot}, y_{t-1\cdot}, w_{t\cdot}), \quad t = 2,3,\ldots, T$$

and A^* is subject to the usual identifiability restrictions.

In the minimand above

(9) $$\tilde{Z} = (\tilde{y}_{t\cdot}, y_{t-1\cdot}, w_{t\cdot}) \quad \tilde{y}_{t\cdot} = (y_{t-1\cdot}, y_{t-2\cdot}, w_{t\cdot}, w_{t-1\cdot})\ (Q'Q)^{-1}\ Q'\ Y$$

$$Q = (y_{t-1\cdot}, y_{t-2\cdot}, w_{t\cdot}, w_{t-1\cdot}) \quad t = 2,3,\ldots,T \quad Y = (y_{ti}) \quad \begin{array}{l} t = 2,3,\ldots,T \\ i = 1,2,\ldots,m \end{array} .$$

i.e., the $y_{t\cdot}$ component of $\tilde{Z}$ is obtained from an ordinary least squares regression in the context of the reduced model

(10) $$y_{t\cdot} = y_{t-1\cdot}F_1 + y_{t-2\cdot}F_2 + WF_3 + W_{-1}F_4 + \varepsilon_{t\cdot}(I - B)^{-1} = q_{t\cdot}F + \varepsilon_{t\cdot}(I - B)^{-1}$$

where

(11) $$F_1 = R^* + \Pi_0, \ F_2 = -C_0R(I - B)^{-1}, \ F_3 = \Pi_1, \ F_4 = C_1R(I - B)^{-1}$$

$$R^* = (I - B)R(I - B)^{-1}, \ q_{t\cdot} = (y_{t-1\cdot} y_{t-2\cdot}, w_{t\cdot}, w_{t-1\cdot}) \ F = (F'_1, F'_2, F'_3, F'_4)'$$

REMARK 1. We observe that

$$F_2 = - C_0(I - B)^{-1}(I - B)R(I - B)^{-1} = - \Pi_0 R^* \tag{12}$$

Moreover R^* is stable if and only if R is. Consequently the second order difference equation in (10) is stable if and only if R^* and Π_0 are stable matrices, which is asserted by (A.2).

Generally, identification requirements will dictate that certain variables be absent from certain equations, i.e., that some elements in B, C_0, C_1 are known *a priori* to be null. Giving expression to these requirements will be greatly facilitated by introducing the selection matrices S_{ij}, $i = 1,2,\ldots,m$ $j = 1,2,3$ such that

$$YS_{i1} = Y_i, \quad Y_{-1}S_{i2} = {}_iY_{-1}, \quad WS_{i3} = W_i \qquad i = 1,2,\ldots,m \tag{13}$$

where Y_i, ${}_iY_{-1}$, W_i are respectively the matrices of observations on the jointly dependent, lagged endogenous and exogenous variables appearing in the right member of the i^{th} equation. Putting

$$S_i = \mathrm{diag}(S_{i1}, S_{i2}, S_{i3}), \quad S = \mathrm{diag}(S_1, S_2, \ldots, S_m) \tag{14}$$

we see that the i^{th} equation of (1) after having imposed the *a priori* restrictions may be written as

$$y_{\cdot i} = ZS_i\delta_{\cdot i} + u_{\cdot i} \qquad i = 1,2,\ldots,m \tag{15}$$

and the entire model may be written as

$$y = Z^*\delta + u \tag{16}$$

where

$$(17)\quad Z^* = (I_m \otimes Z)S,\ \delta = (\delta'_{\cdot 1},\delta'_{\cdot 2},\ldots,\delta'_{\cdot m})',\ u = (u'_{\cdot 1},u'_{\cdot 2},\ldots,u'_{\cdot m})'$$

$$\delta_{\cdot i} = (\beta'_{\cdot i},\gamma^{*\prime}_{\cdot i},\gamma'_{\cdot i})',\quad y = (y'_{\cdot 1},y'_{\cdot 2},\ldots,y'_{\cdot m})'$$

and $\beta_{\cdot i},\gamma^*_{\cdot i},\gamma_{\cdot i}$ are the i^{th} columns of B, C_0, C_1 respectively, after the elements known to be zero have been suppressed, while $y_{\cdot i}$, $u_{\cdot i}$ are, respectively, the i^{th} columns of $Y = (y_{ti})$ $u = (u_{ti})$, $t = 1,2,\ldots,T$ $i = 1,2,\ldots,m$.

The FIDA estimator of δ and R is given by the solution of the equations

$$(18)\quad [(\tilde{Z}^* - (\tilde{R}' \otimes I)Z^*_{-1})'(\tilde{\Sigma}^{-1} \otimes I)(\tilde{Z}^* - (\tilde{R}' \otimes I)Z^*_{-1})]\delta$$

$$= [\tilde{Z}^* - (\tilde{R}' \otimes I)Z^*_{-1}]'(\tilde{\Sigma}^{-1} \otimes I)\cdot[\tilde{y} - (\tilde{R}' \otimes I)y_{-1}]$$

$$\tilde{R} = (\tilde{U}'_{-1}\tilde{U}_{-1})^{-1}\tilde{U}'_{-1}\tilde{U},\quad \tilde{U} = Y - Z\tilde{A},\quad \tilde{A} = (\tilde{B},\tilde{C}_0,\tilde{C}_1)$$

where $\tilde{\Sigma}$ is a prior consistent estimate of Σ and

$$(19)\quad \tilde{Z}^* = (I_m \otimes \tilde{Z})$$

$\tilde{Z}$ being computed in accordance with (9).

Assuming no prior restrictions are imposed on R, writing $r_{\cdot i}$, $i = 1,2,\ldots,m$ for the i^{th} column of R and

$$(20)\quad r = (r'_{\cdot 1},r'_{\cdot 2},\ldots,r'_{\cdot m})'$$

we conclude[1] that, asymptotically,

---

[1] For details of the derivation see the Appendix

$$(21)\quad \begin{bmatrix} M & P_1 \\ P_2 & I \end{bmatrix} \sqrt{T} \begin{pmatrix} \hat{\delta} - \delta \\ \hat{r} - r \end{pmatrix} \sim \begin{bmatrix} I & 0 \\ 0 & I_m \otimes \Omega^{-1} \end{bmatrix} \frac{1}{\sqrt{T}} \begin{bmatrix} (\overline{Z}^* - (R' \otimes I)Z^*_{-1})' \; (\Sigma^{-1} \otimes I) \\ I_m \otimes U'_{-1} \end{bmatrix} \varepsilon$$

where $\overline{Z}^* = (I_m \otimes \overline{Z})$, $\overline{Z} = (\overline{Y}, Y_{-1}, W)$, $\overline{Y} = QF$ and

$$M = \operatorname*{plim}_{T\to\infty} \frac{1}{T} [\overline{Z}^* - (R' \otimes I)Z^*_{-1}]' \; (\Sigma^{-1} \otimes I) \; [\overline{Z}^* - (R' \otimes I)Z^*_{-1}]$$

$$(22)\quad P_1 = \operatorname*{plim}_{T\to\infty} \frac{1}{T} [\overline{Z}^* - (R' \otimes I)Z^*_{-1}]' \; (\Sigma^{-1} \otimes I)(I_m \otimes U_{-1})$$

$$P_2 = \operatorname*{plim}_{T\to\infty} \frac{1}{T} (I_m \otimes \Omega^{-1})(I_m \otimes U'_{-1})[Z^* - (R' \otimes I)Z^*_{-1}] = (\Sigma \otimes \Omega^{-1})P'_1$$

If we put

$$(23)\quad h'^{(i)}_{\cdot t} = (\delta_{i1}\bar{z}^{(1)}_{t\cdot} - r_{i1}z_{t-1}, \delta_{i2}\bar{z}^{(2)}_{t\cdot} - r_{i2}z^{(2)}_{t-1\cdot}, \ldots, \delta_{im}\bar{z}^{(m)}_{t\cdot} - r_{im}z^{(m)}_{t-1\cdot})$$

where δ_{ij} is the Kronecker delta, $\bar{z}^{(i)}_{t\cdot}$ is the row corresponding to the t^{th} observation (row) vector in $\overline{Z}S_i$ and $z^{(i)}_{t-1\cdot}$ is analogously defined for $Z_{-1} S_i$, then we can write compactly

$$(24)\quad \begin{bmatrix} M & P_1 \\ P_2 & I \end{bmatrix} \sqrt{T} \begin{pmatrix} \delta - \delta \\ \hat{r} - r \end{pmatrix} \sim \begin{bmatrix} I & 0 \\ 0 & I_m \otimes \Omega^{-1} \end{bmatrix} \frac{1}{\sqrt{T}} \sum_{t=2}^{T} \begin{pmatrix} H_t \\ I_m \otimes u'_{t-1\cdot} \end{pmatrix} \varepsilon'_{t\cdot}$$

where

$$(25)\quad H_t = (H'^{(1)}_t, \ldots, H'^{(m)}_t)', \qquad H^{(i)}_t = h^{(i)}_{\cdot t} \, \sigma^{i\cdot}$$

$\sigma^{i\cdot}$ is the i^{th} row of Σ^{-1} and $\varepsilon_{t\cdot} = (\varepsilon_{t1}, \varepsilon_{t2}, \ldots, \varepsilon_{tm})$ is the vector of structural errors at "time" t.

The vectors of the sum in the right member of (24) do not represent a sequence of independent random vectors. If in addition to (A.1), (A.2) we also assume

(A.3) The sequence $\{\varepsilon_{t\cdot}' : t = 0, \pm 1, \ldots\}$ has finite sixth order moments

(A.4) The exogenous variables are uniformly bounded nonstochastic

then it can be shown that the conditions of the Hoeffding-Robbins theorem [4] or [2] on m-dependent variables apply to a truncated vector sequence. The truncation may be determined by using the results in Mann and Wald [5] expecially[2] Lemma 1. We, thus, conclude that, asymptotically,

$$(26) \qquad \frac{1}{\sqrt{T}} \sum_{t=2}^{T} \begin{pmatrix} H_t \\ I_m \otimes u_{t-1\cdot}' \end{pmatrix} \varepsilon_{t\cdot}' \sim N(0, C^*)$$

where

$$(27) \qquad C^* = \begin{bmatrix} M & P_2'(I_m \otimes \Omega) \\ (I_m \otimes \Omega)P_2 & \Sigma \otimes \Omega \end{bmatrix}$$

[2] A somewhat more general theorem, i.e., one that utilizes assumptions less restrictive than (A.3) and (A.4) may be obtained by using the results in Billingsley [1]. Such results, however, are unfamiliar in the literature of econometrics and are not utilized here.

Consequently, in view of (22), we have

$$\sqrt{T}\begin{pmatrix}\hat{\delta}-\delta\\ \hat{r}-r\end{pmatrix} \sim N(0,\Phi_{FIDA}) \tag{28}$$

where

$$\Phi_{FIDA} = \begin{bmatrix} M & P_1 \\ P_1' & \Sigma^{-1}\otimes\Omega \end{bmatrix}^{-1} \tag{29}$$

We have therefore proved

THEOREM 1. Consider the model in (1), (2), (3) subject to the following conditions

(A.1) $(I-B)$ is nonsingular

(A.2) $C_0(I-B)^{-1}$, R are both stable matrices

(A.3) The sequence $\{\varepsilon_{t\cdot}' : t = 0, \pm 1, \pm 2, \ldots\}$ is one of i.i.d. random variables having finite sixth order moments

(A.4) The exogenous variable sequence $\{w_{t\cdot}' : t = 0, \pm 1, \pm 2, \ldots\}$ is nonstochastic uniformly bounded

(A.5) Σ as defined in (4) is an unrestricted positive definite matrix

(A.6) The matrix M, defined in (22), exists as a nonsingular nonstochastic probability limit of the right member of (22).

Then, the M.L. and FIDA estimators of the parameter vector $(\delta',r')'$ have the same asymptotic distribution which is given by

(30) $$\sqrt{T}\begin{pmatrix}\hat{\delta}-\delta\\ \hat{r}-r\end{pmatrix} \sim N(0,\Phi_{FIDA})$$

where Φ_{FIDA} is defined in (29) and (22).

COROLLARY 1. The marginal asymptotic distribution of the vector $\sqrt{T}\,(\hat{\delta}-\delta)$ is given by

(31) $$\sqrt{T}\,(\hat{\delta}-\delta) \sim N(0,C_{FIDA})$$

where

(32) $$C_{FIDA} = [M - P_1(\Sigma^{-1}\otimes\Omega)P_1']^{-1}$$

Proof. Obvious from the theorem.

COROLLARY 2. If $R = 0$ but this fact is not utilized in the estimation process, then there is asymptotic loss of efficiency in estimating δ.

Proof. If the information is utilized then the asymptotic distribution of $\sqrt{T}\,(\hat{\delta}-\delta)_{R=0}$ is normal with mean zero and covariance matrix M_0^{-1}, where M_0 is the matrix defined in (22) for the special case where $R = 0$. We observe that $P_1 \neq 0$ when $R = 0$. This immediately implies the corollary.

REMARK 2. Thus, here we incur a certain cost when autoregression in the errors is assumed, when in fact it is absent. This is to be contrasted to the case where the model contains no lagged endogenous variables. In such a case no loss in (asymptotic) efficiency results, when one assumes a higher order of autoregression than is, in fact, true.

REMARK 3. It is easily seen that the results of Theorem 1 specialize in the case where we deal with a single equation "system" containing a lagged endogenous variable and first order autoregressive error, to the result contained in Theorem 7.1 of [2, Ch. 7].

Indeed the development and results obtained herein are a direct generalizstion of the results contained in [2, Ch. 7], with respect to the dynamic demand model with first order autoregressive errors.

REMARK 4. It appears that no additional complications are entailed by the introduction of additional lags in the jointly dependent variables. No additional complications are introduced by considering higher order autoregressions in the errors except for the obvious computational burden of obtaining an expression for the covariance matrix of the structural errors.

REFERENCES

1. Billingsley, P., Convergence of Probability Measures, New York, Wiley, 1968.

2. Dhrymes, P. J., Distributed Lags: Problems of Estimation and Formulation, San Francisco, Holden-Day, 1971.

3. ____________, "Full Information Estimation of Dynamic Simultaneous Equations Models with Autoregressive Errors", Discussion Paper No. 203, University of Pennsylvania, March, 1971.

4. Hoeffding, W. and H. Robbins, "The Central Limit Theorem for Dependent Variables", Duke Mathematical Journal, vol. 15 (1943), pp. 773-780.

5. Mann, H. B. and A. Wald, "On the Statistical Treatment of Linear Stochastic Difference Equations", Econometrica, vol. 11 (1943), pp. 173-220.

APPENDIX

Here we made explicit the transition from (18) to (21) and hence to (24).

Substituting from (16) into (18) we obtain, for the $k+1^{st}$ iterate

$$(\tilde{\delta}-\delta)_{k+1} = [(\tilde{Z}^* - (\tilde{R}_k' \otimes I)Z^*_{-1})'(\tilde{\Sigma}^{-1} \otimes I)(\tilde{Z}^* - (\tilde{R}_k' \otimes I)Z^*_{-1})]^{-1} \quad \text{(A.1)}$$

$$\{[\tilde{Z} - (\tilde{R}_k' \otimes I)Z^*_{-1}]'(\tilde{\Sigma}^{-1} \otimes I) - [\varepsilon - [(\tilde{R}_k' - R') \otimes I]u_{-1}]\}$$

where in computing $\tilde{R}_k$ we have used the k^{th} iterate $\tilde{\delta}_k$.

Now we have

$$\tilde{R} - R = (\tilde{U}_{-1}'\tilde{U}_{-1})^{-1}\tilde{U}_{-1}[\tilde{U} - \tilde{U}_{-1}R] \quad \text{(A.2)}$$

But by definition

$$\tilde{U} = Y - Z\tilde{A} = U - Z(\tilde{A} - A) \quad \text{(A.3)}$$

Hence

$$\tilde{U} - \tilde{U}_{-1}R = E - Z(\tilde{A} - A) + Z_{-1}(\tilde{A} - A)R \quad \text{(A.4)}$$

Writing (A.2) in column form we find

$$(\tilde{r} - r)_k = [I_m \otimes (\tilde{U}_{-1}'\tilde{U}_{-1})^{-1}\tilde{U}_{-1}'] \, [\varepsilon - (Z^* - (R' \otimes I)Z^*_{-1}) \, (\tilde{\delta} - \delta)_K] \quad \text{(A.5)}$$

Moreover, we note that

$$[(\tilde{R} - R)_k' \otimes I]u_{-1} = (I_m \otimes U_{-1}) \, (\tilde{r} - r)_k \, . \quad \text{(A.6)}$$

Hence, the iteration scheme in (18) can be written more conveniently as

(A.7) $$\tilde{M}_k(\tilde{\delta}-\delta)_{k+1} = [\tilde{Z}^* - (\tilde{R}_k' \otimes I)Z^*_{-1}]'(\tilde{\Sigma}^{-1} \otimes I)\varepsilon$$

$$- [\tilde{Z}^* - (\tilde{R}_k' \otimes I)Z^*_{-1}]'(\tilde{\Sigma}^{-1} \otimes I)(I_m \otimes U_{-1})(\hat{r}-r)_k$$

(A.8) $$(\hat{r}-r)_k = [I_m \otimes (\tilde{U}_{-1}'\tilde{U}_{-1})^{-1}\tilde{U}_{-1}']\varepsilon$$

$$- [I_m \otimes (\tilde{U}_{-1}'\tilde{U}_{-1})^{-1}\tilde{U}_{-1}']\,(\tilde{Z}^* - (R' \otimes I)Z^*_{-1})\,(\tilde{\delta}-\delta)_k$$

For the converging iterate we can then write

(A.9) $$\sqrt{T}\begin{pmatrix}\hat{\delta}-\delta \\ \hat{r}-r\end{pmatrix} = \begin{bmatrix} \left(\frac{\tilde{M}}{T}\right)^{-1}\frac{1}{\sqrt{T}}[\tilde{Z}^* - (\tilde{R}' \otimes I)Z^*_{-1}]'(\tilde{\Sigma}^{-1} \otimes I) \\ I_m \otimes \left(\frac{\tilde{U}_{-1}'\tilde{U}_{-1}}{T}\right)^{-1}\frac{\tilde{U}_1'}{\sqrt{T}} \end{bmatrix}$$

$$\begin{bmatrix} 0 & \frac{1}{T}\left(\frac{\tilde{M}}{T}\right)^{-1}[\tilde{Z}^* - (\tilde{R}' \otimes I)Z^*_{-1}]'(\tilde{\Sigma}^{-1} \otimes I)(I_m \otimes U_{-1}) \\ \left[I_m \otimes \left(\frac{\tilde{U}_{-1}'\tilde{U}_{-1}}{T}\right)^{-1}\frac{\tilde{U}_{-1}'}{T}\right][\tilde{Z} - (R' \otimes I)Z^*_{-1}] & \end{bmatrix}\sqrt{T}\begin{pmatrix}\hat{\delta}-\delta \\ \hat{r}-r\end{pmatrix}$$

Multiplying through M the left by

$$\begin{bmatrix} \frac{\tilde{M}}{T} & 0 \\ 0 & T \end{bmatrix}$$

and rearranging terms we find

$$\text{(A.10)}\quad \begin{bmatrix} \frac{\tilde{M}}{T} & \frac{1}{T}(\tilde{Z}^* - (\tilde{R}' \otimes I)Z^*_{-1})'(\tilde{\Sigma}^{-1} \otimes I)(I_m \otimes U_{-1}) \\ \left(I_m \otimes \left(\frac{\tilde{U}'_{-1}\tilde{U}_{-1}}{T}\right)^{-1} \frac{\tilde{U}'_{-1}}{T}\right) (\tilde{Z}^* - (R' \otimes I)Z^*_{-1}) & \end{bmatrix} \sqrt{T}\begin{pmatrix} \hat{\delta} - \delta \\ \hat{r} - r \end{pmatrix}$$

$$= \frac{1}{\sqrt{T}} \begin{bmatrix} (\tilde{Z}^* - (\tilde{R}' \otimes I)Z^*_{-1})'(\tilde{\Sigma}^{-1} \otimes I) \\ I_m \otimes \left(\frac{\tilde{U}'_{-1}\tilde{U}_{-1}}{T}\right)^{-1} \tilde{U}'_{-1} \end{bmatrix} \varepsilon$$

It is, thus, easily verified that the solution vector in (A.10), i.e., $\sqrt{T}\begin{pmatrix} \hat{\delta} - \delta \\ \hat{r} - r \end{pmatrix}$ behaves asymptotically according to the relation given in (21).