COWLES COMMISSION DISCUSSION PAPER: STATISTICS: NO. 366

Estimation of a Single Equation in a Complete System of Stochastic Equations with Cross-Section, Time Series Data

S. G. Allen

October 8, 1951

1. Introduction.

The problem considered in this paper is the estimation of a linear equation containing variables for which both cross-section and time series observations are available. However the model specifies that all other equations in the linear system, of which the above equation is a member, contain only variables aggregated over the cross-sections.

Reasons for considering such a model are numerous. It might provide realistic description of a simple market of N buyers and a single seller. In any case, the economic units involved in an equation relating to their behavior as an aggregate are likely to be dissimilar for each equation of a macro-economic system; thus, if cross-section data are available to estimate any one equation in its "disaggregated" form, it is probably expedient to assume that the remaining equations of the system relate only to aggregate behavior.

It may be helpful to consider equation (2.1) below as a demand equation, where the variable $y_1(n, t)$ is the quantity demanded by the n^{th} economic unit at time t, and the other endogenous variables are variables common in value to all economic units at time t. As is usual in cross-section studies, it is assumed that the equation is sufficiently well specified that the random disturbance in the equation is independent of n.

The reader's familiarity with the relevant portions of [1] and [2] is assumed.

2. The Model.

Suppose there are observations for N economic units and for T time periods on the equation

$$(2.1)\quad y_1(n,t) + \frac{1}{N}[\beta_{12}\, y_2(t) + \ldots + \beta_{1G}\, y_G(t)] + \gamma_{11}\, z_1(n,t) + \ldots + \gamma_{1J}\, z_J(n,t) + \frac{1}{N}[\gamma_{1,J+1}\, z_{J+1}(t) + \ldots + \gamma_{1K}\, z_K(t)] = u_1(n,t), \qquad (n = 1, \ldots, N;\ t = 1, \ldots, T).$$

The variables $y_1(n,t)$ and $z^{(1)}(n,t) = [z_1(n,t) \ldots z_J(n,t)]$ are observable variables with values dependent on n, the variables $y^{(2)}(t) = [y_2(t) \ldots y_G(t)]$ and $z^{(2)}(t) = [z_{J+1}(t) \ldots z_K(t)]$ are observable variables with values independent of n, and the variable $u_1(n,t)$ is the unobservable random disturbance in the equation.

Let

$$(2.2)\quad \beta_{g1} \sum_{n=1}^{N} y_1(n,t) + \beta_{g2}\, y_2(t) + \ldots + \beta_{gG}\, y_G(t) + \gamma_{g1} \sum_{n=1}^{N} z_1(n,t) + \ldots + \gamma_{gJ} \sum_{n=1}^{N} z_J(n,t) + \gamma_{g,J+1}\, z_{J+1}(t) + \ldots + \gamma_{gK}\, z_K(t) = u_g(t), \qquad (g = 2, \ldots, G;\ t = 1, \ldots, T)$$

be a subsystem of G-1 equations which, together with the N equations (2.1) for n = 1, ..., N, constitute a complete system of stochastic difference equations in the N + G - 1 variables

$$(2.3)\quad y(t) = [y_1(1,t) \ldots y_1(N,t)\ y_2(t) \ldots y_G(t)].$$

The model further specifies that

$$(2.4) \qquad Eu_1(n,t)\, u_1(\nu, \tau) = \begin{cases} 0 & , \text{ for } n \neq \nu \text{ or } t \neq \tau \\ \omega_{11}, & \text{for } \begin{matrix} n = \nu = 1, \ldots, N \\ t = \tau = 1, \ldots, T \end{matrix} \end{cases}$$

$$(2.5) \qquad Eu_1(n,t)\, u_g(\tau) = \begin{cases} 0 & , \text{ for } t \neq \tau \\ \omega_{1g}, & \text{for } \begin{matrix} t = \tau = 1, \ldots, T \\ n = 1, \ldots, N \\ g = 2, \ldots, G \end{matrix} \end{cases}$$

and

$$(2.6) \qquad Eu_g(t)\, u_h(\tau) = \begin{cases} 0 & , \text{ for } t \neq \tau \\ \omega_{gh}, & \text{for } \begin{matrix} t = \tau = 1, \ldots, T \\ g, h = 2, \ldots, G \text{ .} \end{matrix} \end{cases}$$

Finally it is assumed that the random variables $u_I(t) = [u_1(1,t) \ldots u_1(N,t)]$ and $u_{II}(t) = [u_2(t) \ldots u_G(t)]$ are independent of the z-variables and have a joint normal distribution with zero means and covariance matrix

$$\Omega = \begin{bmatrix} \Omega_{I\,I} & \Omega_{I\,II} \\ \Omega_{II\,I} & \Omega_{II\,II} \end{bmatrix}$$

where

$$\Omega_{I\,I} \equiv Eu_I'(t)\, u_I(t) = \omega_{11} \cdot I_{NN}$$

$$\Omega_{I\,II} \equiv Eu_I'(t)\, u_{II}(t) = \iota_N' \cdot \Omega_{1,II}$$

$$\Omega_{1,II} \equiv Eu_1(n,t)\, u_{II}(t)$$

$$\Omega_{II\,II} \equiv Eu_{II}'(t)\, u_{II}(t),$$

where I_{NN} is the identity matrix of order N and ι_N is an n-dimensional vector with each element equal to one.

The problem is to derive maximum likelihood estimates of the elements of

$$B_1^{(2)} = (\beta_{12} \ldots \beta_{1G})$$

$$\Gamma_1^{(1)} = (\gamma_{11} \ldots \gamma_{1J})$$

$$\Gamma_1^{(2)} = (\gamma_{1,J+1} \ldots \gamma_{1K}),$$

ignoring all restrictions on values of the elements of

$$B_{II}^{(1)} = \begin{pmatrix} \beta_{21} \\ \vdots \\ \beta_{G1} \end{pmatrix} \qquad B_{II}^{(2)} = \begin{pmatrix} \beta_{22} & \cdots & \beta_{2G} \\ \vdots & & \vdots \\ \beta_{G2} & \cdots & \beta_{GG} \end{pmatrix}$$

$$\Gamma_{II}^{(1)} = \begin{pmatrix} \gamma_{21} & \cdots & \gamma_{2J} \\ \vdots & & \vdots \\ \gamma_{G1} & \cdots & \gamma_{GJ} \end{pmatrix} \qquad \Gamma_{II}^{(2)} = \begin{pmatrix} \gamma_{2,J+1} & \cdots & \gamma_{2K} \\ \vdots & & \vdots \\ \gamma_{G,J+1} & \cdots & \gamma_{GK} \end{pmatrix}.$$

Formally (2.1) for n = 1, ..., N is a subsystem of N equations. Thus the estimation problem is to derive the <u>simultaneous</u> maximum likelihood estimates of coefficients appearing in the subsystem (2.1) subject to linear restrictions on the coefficients. The latter are made explicit in the notation

(2.7) $\qquad B_I = (I_{NN} \quad \frac{1}{N} \cdot \iota_N' \cdot B_1^{(2)})$

(2.8) $\qquad \Gamma_I = (\Gamma_1^{(1)} \otimes I_{NN} \quad \frac{1}{N} \cdot \iota_N' \cdot \Gamma_1^{(2)}),$

where B_I and Γ_I are the matrices of coefficients of y(t) and

(2.9) $\qquad z(t) = [z_1(1,t) \ldots z_1(N,t) \ldots z_J(1,t) \ldots z_J(N,t) \; z_{J+1}(t) \ldots z_K(t)],$

respectively, in subsystem (2.1). Restrictions on coefficients in the subsystem (2.2) will be ignored except that the matrices of coefficients of y(t) and z(t) <u>are of the form</u>

(2.10) $\qquad B_{II} = (B_{II}^{(1)} \otimes \iota_N \quad B_{II}^{(2)})$

and

(2.11) $\qquad \Gamma_{II} = (\Gamma_{II}^{(1)} \otimes \iota_N \quad \Gamma_{II}^{(2)}),$

respectively.[1]

---

1. A $\otimes$ B denotes the Kronecker product of the matrices A and B. More specifically, if $A = (a_{ij})$, then $A \otimes B = (a_{ij} \cdot B)$.

3. The Aggregate Model.

Before the derivation of the estimates proposed in Section 2, it will be instructive to consider estimates of parameters in a model derivable from the previous one. Suppose the N equations (2.1) are summed to obtain

$$(3.1)\quad \sum_{n=1}^{N} y_1(n,t) + \beta_{12}\, y_2(t) + \dots + \beta_{1G}\, y_G(t) + \gamma_{11} \sum_{n=1}^{N} z_1(n,t) + \dots + \gamma_{1J} \sum_{n=1}^{N} z_J(n,t) + \gamma_{1,J+1}\, z_{J+1}(t) + \dots + \gamma_{1K}\, z_K(t) = \sum_{n=1}^{N} u_1(n,t).$$

Equation (3.1) and the G-1 equations (2.2) also constitute a complete system. This system may be written in the form

$$(3.2)\qquad B_{I_0}\, \Sigma y'(t) + \Gamma_{I_0}\, \Sigma z'(t) = \Sigma u_1(t)$$

$$(3.3)\qquad B_{II_0}\, \Sigma y'(t) + \Gamma_{II_0}\, \Sigma z'(t) = u'_{II}(t),$$

where

$$(3.4)\qquad y(n,t) = [y_1(n,t) \quad \tfrac{1}{N}\cdot y^{(2)}(t)]$$

$$(3.5)\qquad z(n,t) = [z^{(1)}(n,t) \quad \tfrac{1}{N}\cdot z^{(2)}(t)]$$

$$(3.6)\qquad \Sigma y(t) = \sum_{n=1}^{N} y(n,t), \quad \Sigma z(t) = \sum_{n=1}^{N} z(n,t),$$

$$(3.7)\qquad B_{I_0} = (1 \quad B_1^{(2)}) \quad \Gamma_{I_0} = (\Gamma_I^{(1)}\, \Gamma_I^{(2)})$$

$$(3.8)\qquad B_{II_0} = (B_{II}^{(1)} \quad B_{II}^{(2)}) \quad \Gamma_{II_0} = (\Gamma_{II}^{(1)}\, \Gamma_{II}^{(2)}).$$

The joint distribution of $\Sigma u_1(t)$ and $u_{II}(t)$ is then normal with zero means and covariance matrix

$$\Omega_0 = \begin{pmatrix} N\cdot\omega_{11} & N\,\Omega_{1,II} \\ N\,\Omega_{II,1} & \Omega_{II\,II} \end{pmatrix} .$$

Observable variables appear in this model as functions of only one index, namely t, as in the model considered in [1]. Derivation of estimates of (3.7) proceeds along the steps (S-60)-(S-115) and (S-124)-(S-135) of [1].

Corresponding to (S-60) of [1], the likelihood function of the parameters α_o and Ω_o is

$$(3.9) \qquad L_o(\alpha_o, \Omega_o) = \text{const} + \log|\det B_o| - \frac{1}{2}\log\det\Omega_o - \frac{1}{2}\,\text{tr}\,\alpha_o\,\Omega_o^{-1}\,\alpha_o\,M_{\Sigma x \Sigma x}$$

where

$$B_o = \begin{pmatrix} B_{I_o} \\ B_{II_o} \end{pmatrix}$$

$$\alpha_o = \begin{pmatrix} \alpha_{I_o} \\ \alpha_{II_o} \end{pmatrix}$$

$$\alpha_{I_o} = (B_{I_o} \ \Gamma_{I_o})$$

$$\alpha_{II_o} = (B_{II_o} \ \Gamma_{II_o})$$

$$\Sigma x(t) = [\Sigma y(t) \ \Sigma z(t)]$$

$$M_{ab} = \frac{1}{T}\sum_{t=1}^{T} a'(t)\,b(t). \ ^{2/}$$

If restrictions on (3.8) are ignored in deriving maximum likelihood estimates of (3.7), the task may be simplified by transforming the variables

---

2. Cf. definitions (3.4)-(3.6).

$\Sigma u_1(t)$ and $u_{II}(t)$ such that the covariance matrix of the transformed variables is

$$\Omega_o^* = \begin{pmatrix} N\omega_{11} & 0_{1,G-1} \\ 0_{G-1,1} & I_{G-1,G-1} \end{pmatrix} ,$$

where 0_{AB} is an A by B matrix with zero elements. The transformation accomplishing this is

$$(3.10) \qquad \Upsilon_o = \begin{pmatrix} 1 & 0_{1,G-1} \\ -\frac{1}{\omega_{11}} \Upsilon_{II\,II} \Omega_{II,1} & \Upsilon_{II\,II} \end{pmatrix}$$

where $\Upsilon_{II\,II}$ satisfies

$$(3.11) \qquad \Upsilon_{II\,II}(\Omega_{II\,II} - \frac{N}{\omega_{11}} \Omega_{II,1} \Omega_{1,II}) \Upsilon'_{II\,II} = I_{G-1,G-1}$$

The likelihood function of the transformed parameters $\alpha_o^* = \Upsilon_o \alpha_o$ and $\Omega_o^* = \Upsilon_o \Omega_o \Upsilon_o'$ is (corresponding to (S-102) of [1]).

$$(3.12) \qquad L_o^* (\alpha_{I_o}, \alpha^*_{II_o}, \omega_{11})$$

$$= \text{const} + \log\left|\det B_o^*\right| - \frac{1}{2} \log N\omega_{11}$$

$$- \frac{1}{2N\omega_{11}} \alpha_{I_o} M_{\Sigma x \Sigma x} \alpha'_{I_o} - \frac{1}{2} \operatorname{tr} \alpha^*_{II_o} M_{\Sigma x \Sigma x} \alpha'^*_{II_o}$$

where

$$B_o^* = \Upsilon_o B_o$$

$$\alpha^*_{II_o} = (-\frac{1}{\omega_{11}} \Upsilon_{II\,II} \Omega_{II,I} \quad \Upsilon_{II\,II}) \alpha_o .$$

The function (3.12) may be reduced to a function of α_{I_o} alone by "maximizing out" the parameters $\alpha^*_{II_o}$ and ω_{11}. Then the likelihood function may be maximized with respect to elements of α_{I_o} to obtain a system of equations in elements of α_{I_o}. The system of equations has a unique solution if necessary conditions for the identification of α_{I_o} are met.[3]

4. The likelihood function of α_I.

Let

$$\alpha_I = (B_I \quad \Gamma_I)$$

$$\alpha_{II} = (B_{II} \quad \Gamma_{II})$$

$$\alpha = \begin{pmatrix} \alpha_I \\ \alpha_{II} \end{pmatrix} .$$

Then the likelihood function of the parameters of the model defined in Section 2 may be written

$$(4.1) \qquad L(\alpha, \Omega) = \text{const} + \log |\det B| - \frac{1}{2} \log \det \Omega - \frac{1}{2} \operatorname{tr} \alpha \Omega^{-1} \alpha^1 M_{xx}$$

where

$$x(t) = [y(t) \; z(t)]$$

$$B = \begin{pmatrix} B_I \\ B_{II} \end{pmatrix} . \quad [4]$$

Analogous to the procedure outlined in Section 3, the variables $u_I(t)$, $u_{II}(t)$ may be transformed so that the covariance matrix of the transformed variables is

$$\Omega^* = \begin{pmatrix} \Omega_{I.I} & O_{N,G-1} \\ O_{G-1,N} & I_{G-1,G-1} \end{pmatrix} .$$

---

3. No systematic investigation of the identifiability of equation (2.1) or (3.1) will be attempted in this paper. It will be assumed that their identifiability is a property of both the models introduced. However it is not clear that necessary conditions are the same in both models.

4. Cf. (2.3), (2.7) - (2.11).

The required transformation is

$$\Upsilon = \begin{bmatrix} I_{NN} & 0_{N,G-1} \\ -\frac{1}{\omega_{11}} \Upsilon_{II\,II} \Omega_{II\,I} & \Upsilon_{II\,II} \end{bmatrix}$$

where $\Upsilon_{II\,II}$ satisfies

$$(4.2) \qquad \Upsilon_{II\,II}(\Omega_{II\,II} - \frac{1}{\omega_{11}} \Omega_{II\,I}\, \Omega_{I\,II})\Upsilon'_{II\,II} = I_{G-1,G-1} \,.$$

It will be noted that Υ preserves restrictions (2.7) and (2.8). Also, the transformed matrices

$$B^*_{II} = (-\frac{1}{\omega_{11}} \Upsilon_{II\,II}\Omega_{II\,I} \qquad \Upsilon_{II\,II})B$$

$$\Gamma^*_{II} = (-\frac{1}{\omega_{11}}\Upsilon_{II\,II}\Omega_{II\,I} \qquad \Upsilon_{II\,II}) \begin{bmatrix} \Gamma_I \\ \Gamma_{II} \end{bmatrix}$$

are in the forms of (2.10) and (2.11), respectively. Furthermore $\Upsilon_{II\,II}$ appearing in Υ is the same $\Upsilon_{II\,II}$ appearing in Υ_o, since (3.11) and (4.2) are equivalent. Therefore

$$(4.3) \qquad \det \Upsilon = \det \Upsilon_o \,.$$

Under the transformation the function (4.1) becomes

$$(4.4) \qquad L^*(\alpha_I, \alpha^*_{II}, \omega_{11}) = \text{const} + \log\left| \det B^* \right| - \frac{N}{2}\, \omega_{11} - \frac{1}{2\,\omega_{11}} \operatorname{tr} \alpha_I M_{xx} \alpha'_I - \frac{1}{2} \operatorname{tr} \alpha^*_{II} M_{xx} \alpha^{*\prime}_{II},$$

where $B^* = \Upsilon B$ and $\alpha^*_{II} = (B^*_{II} \quad \Gamma^*_{II})$. Due to the restrictions (2.7), (2.8), (2.10), and (2.11), the quadratic forms in (4.4) simplify to

$$(4.5) \qquad \operatorname{tr} \alpha_I M_{xx} \alpha'_I = \sum_{n=1}^{N} \alpha_{I_o} M_{x(n)x(n)} \alpha'_{I_o}$$

$$(4.6) \qquad \operatorname{tr} \alpha^*_{II} M_{xx} \alpha^{*\prime}_{II} = \operatorname{tr} \alpha^*_{II_o} M_{\Sigma x \Sigma x} \alpha^{*\prime}_{II_o} \,,$$

where

$$x(n,t) = [y(n,t)\ z(n,t)]$$

$$M_{a(n)b(n)} = \frac{1}{T}\sum_{t=1}^{T} a'(n,t)\ b(n,t).$$ 5/

Therefore in simplified form L^* of (4.4) is

$$(4.7)\qquad L^*(\alpha_{I_o}, \alpha^*_{II_o}, \omega_{11}) = \text{const} + \log|\det B^*|$$
$$- \frac{N}{2}\log \omega_{11} - \frac{1}{2\omega_{11}}\alpha_{I_o}\left(\sum_{n=1}^{N} M_{x(n)x(n)}\right)\alpha'_{I_o}$$
$$- \frac{1}{2}\operatorname{tr}\alpha^*_{II_o} M_{\Sigma x \Sigma x}\alpha^{*\prime}_{II_o}\ .$$

Due to the restrictions (2.7), (2.8), (2.10), and (2.11), two different types of moment matrices appear in L^*. Elements in one matrix are sums over n of cross products summed over t and in the other are sums over t of cross products of sums over n. In fact (4.6) is the quadratic form appearing in L^*_o of (3.12). Finally the Jacobian det B^* is the same as the Jacobian det B^*_o appearing in L^*_o. The equality (4.3) has already been established, and it can be easily verified that det B_o = det B.6/

The above are useful observations in connection with the next step in the derivation of estimates of α_I, namely, the reduction of L^* of (4.7) to a function of α_I and ω_{11}. It is now clear that the same system of (differential) equations result whether L^* is maximized with respect to α^*_{II} or L^*_o of (3.12) is maximized with respect to $\alpha^*_{II_o}$. The analogy to the derivation in [1], i.e., steps (S-104) to (S-113), is clearer if L^*_o is viewed as the

5. Cf. definitions (3.4) - (3.6).

6. The Jacobians would not be equal if the coefficient of $y_1(n,t)$ had not been chosen for normalization. The particular results given in this paper thus depend on this choice for normalization. However since the model introduced in Section 2 allows only one equation to appear in its "disaggregated" form (and hence allows the N component parts of only one aggregate variable to appear in the system), it seemed desirable to normalize so that the analogy between derivation of estimates in the "aggregate" and "disaggregate" models would be clearest.

function maximized.

The resulting system of equations is

$$(4.8) \qquad \left(\frac{\partial L_o^*}{\partial \alpha^*_{II_o}}\right) = \left(\frac{\partial \log |\det B_o^*|}{\partial \alpha^*_{II_o}}\right) - \alpha^*_{II_o} M_{\Sigma x \Sigma x} = 0.$$

And it can be shown that elements of $\alpha^*_{II_o}$ satisfying (4.8) also satisfy

$$(4.9) \qquad \alpha^*_{II_o} M_{\Sigma x \Sigma x} \alpha'_{II_o} = I_{G-1,G-1}$$

$$(4.10) \qquad \det B_o^* = \frac{\sqrt{B_{I_o} W_{\Sigma y \Sigma y} B'_{I_o}}}{\sqrt{\det W_{\Sigma y \Sigma y}}} ,$$

where

$$W_{\Sigma y \Sigma y} = M_{\Sigma y \Sigma y} - M_{\Sigma y \Sigma z} M^{-1}_{\Sigma z \Sigma z} M_{\Sigma z \Sigma y} .$$

Maximization of L^* of (4.7) with respect to ω_{11} yields the expression

$$(4.11) \qquad \omega_{11} = \frac{1}{N} \sum_{n=1}^{N} \alpha_{I_o} M_{x(n)x(n)} \alpha'_{I_o} .$$

With substitutions of (4.9) - (4.11), L^* may be rewritten as

$$(4.12) \qquad L^* = \text{const} + \frac{1}{2} \log B_{I_o} W_{\Sigma y \Sigma y} B'_{I_o} - \frac{N}{2} \log \alpha_{I_o} \left[\frac{1}{N} \sum_{n=1}^{N} M_{x(n)x(n)} \right] \alpha'_{I_o} .$$

Maximization of L_o^* of (3.12) with respect to $N \cdot \omega_{11}$ yields an expression

$$(4.13) \qquad N \cdot \omega_{11} = \alpha_{I_o} M_{\Sigma x \Sigma x} \alpha'_{I_o} .$$

With the aid (4.9), (4.10), and (4.13) L_o^* may be rewritten as

$$(4.14) \qquad L_o^* = \text{const} + \frac{1}{2} \log B_{I_o} W_{\Sigma y \Sigma y} B'_{I_o} - \frac{1}{2} \log \alpha_{I_o} M_{\Sigma x \Sigma x} \alpha'_{I_o} .$$

The chief difficulty in computing estimates maximizing L^* of (4.12) arises from the fact that the quadratic form (4.5) is raised to the N^{th} power in (4.12). (Recall that $\det \Omega_{II} = (\omega_{11})^N$.) Otherwise estimates maximizing L^* and those maximizing L_o^* would satisfy systems of equations identical <u>in form</u>. In particular conditions(353) and (355) of [2] do not hold for estimates maximizing L^*.

BIBLIOGRAPHY

[1] Cowles Commission Discussion Paper: Statistics, No. 310A.

[2] Cowles Commission Discussion Paper: Statistics, No. 310.