The Journal of National Chengchi University Vol. 37, 38, 1978

Bayesian Analysis of Simultaneous Equation Models

by

Ing-tzer Wey

Professor, Graduate School of Business Administration

I. Introduction

The simultaneous equation models we shall consider are linear in the parameters and a generalization of the multivariate regression model. In an ordinary multivariate regression model, we have just one dependent or endogenous variable per epation. However, here in simultaneous equation models we may have more than one dependent variable appearing in each equation; that is, in the simultaneous equation model it is assumed that the data are generated by a model in the following form:

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

where $Y=(\underset{\sim}{y}_1, \underset{\sim}{y}_2, \cdots, \underset{\sim}{y}_m)$, an nxm matrix of observations on m dependent or endogenous variables whose variation is to be explained by the model, Γ is an mxm nonsingular matrix of coefficients for the endogenous variables, $X=(\underset{\sim}{x}_1, \underset{\sim}{x}_2, \cdots, \underset{\sim}{x}_k)$ is an nxk matrix of observations on k predetermined variables, B is a kxm matrix of coefficients for the predetermined variables, and $U=(\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_m)$ is an nxm matrix of random disturbance errors. The variables in X, the predetermined variables, may include lagged values of the endogenous and/or exogenous variables. In the latter category we include both nonstochastic and stochastic variables whose variation is determined outside the model. Stochastic exogenous variables, by definition, are assumed to be distributed independently of the random disturbance errors in U and to have distributions not involving any of the parameters of the model; that is, Γ, B, and the elements of the covariance matrix of U. Further, we assume that the rows of U are normally and independently distributed with zero means and, each with mxm positive definite symmetric covariariance matrix Σ.

We shall discuss first the Bayesian analysis of fully recursive model which is, from the estimation viewpoint, the simplest model among all simultaneous equation models. Next, the limited information Bayesian analysis of the simultaneous equation models will be studied. Finally, we shall discuss the full information Bayesian analysis of the simultaneous equation models. Since the Bayes estimator of a parameter θ with respect to the squared-error loss function is equal to the conditional expectation of θ with respect to the posterior distribution of θ given the sample, we shall derive the posterior distribuions of the parameters in the models under consideration.

II. Bayesian Analysis of Fully Recursive Model

We assume that our observations $Y=(\underset{\sim}{y}_1, \underset{\sim}{y}_2, \cdots, \underset{\sim}{y}_m)$ can be expressed by the following fully recursive model:

$$
\begin{aligned}
\underset{\sim}{y}_1 &= X_1\underset{\sim}{\beta}_1+\underset{\sim}{u}_1 \\
\underset{\sim}{y}_2 &= \gamma_{21}\underset{\sim}{y}_1+X_2\underset{\sim}{\beta}_2+\underset{\sim}{u}_2 \\
\underset{\sim}{y}_3 &= \gamma_{31}\underset{\sim}{y}_1+\gamma_{32}\underset{\sim}{y}_2+X_3\underset{\sim}{\beta}_3+\underset{\sim}{u}_3 \\
&\vdots \\
\underset{\sim}{y}_m &= \gamma_{m1}\underset{\sim}{y}_1+\gamma_{m2}\underset{\sim}{y}_2+\cdots+\gamma_{m,m-1}\underset{\sim}{y}_{m-1}+X_m\underset{\sim}{\beta}_m+\underset{\sim}{u}_m
\end{aligned}
\tag{2}
$$

where $\underset{\sim}{y}_i$ is an nxl vector of observations on the i-th endogenous variable, X_i is an nxk_i matrix with rank k_i of observations on k_i predetermined variables appearing in the i-th equation with coefficient vector $\underset{\sim}{\beta}_i$, a $k_i$xl column vector, $\underset{\sim}{u}_i$ is an nxl vector of distbance errors, and the γ_{ij} are scalar coefficients. Further, we assume that the $\underset{\sim}{u}_i$, i=1, 2, …, m, are normally distributed with zero means and covariance matrix

$$E(\underset{\sim}{u}\underset{\sim}{u}')=\Sigma \otimes I_n \tag{3}$$

where $\underset{\sim}{u}'=(\underset{\sim}{u}_1', \underset{\sim}{u}_2', \cdots, \underset{\sim}{u}_m')$, Σ is an mxm positive definite symmetric matrix and I_n is an nxn unit matrix.

Let $Z_i=(\underset{\sim}{u}_1, \underset{\sim}{u}_2, \cdots, \underset{\sim}{u}_{i-1}, X_i)$, $\underset{\sim}{\delta}_i'=(\gamma_i', \underset{\sim}{\beta}_i')$ with $\gamma_i'=(\gamma_{i1}, \gamma_{i2}, \cdots, \gamma_{i,i-1})$, and $\underset{\sim}{\delta}'=(\underset{\sim}{\delta}_1', \underset{\sim}{\delta}_2', \cdots, \underset{\sim}{\delta}_m')$. Then, we can write the system in (2) as

$$
\begin{pmatrix} \underset{\sim}{y}_1 \\ \underset{\sim}{y}_2 \\ \vdots \\ \underset{\sim}{y}_m \end{pmatrix}
=
\begin{pmatrix} Z_1 & & & O \\ & Z_2 & & \\ & & \ddots & \\ O & & & Z_m \end{pmatrix}
\begin{pmatrix} \underset{\sim}{\delta}_1 \\ \underset{\sim}{\delta}_2 \\ \vdots \\ \underset{\sim}{\delta}_m \end{pmatrix}
+
\begin{pmatrix} \underset{\sim}{u}_1 \\ \underset{\sim}{u}_2 \\ \vdots \\ \underset{\sim}{u}_m \end{pmatrix}
\tag{4}
$$

or

$$\underset{\sim}{y}=Z\underset{\sim}{\delta}+\underset{\sim}{u} \tag{5}$$

where Z denotes the block diagonal matrix on the right handside of (4) and $\underset{\sim}{y}$, the vector on the left hand side of (4). The likelihood functon for parameters $\underset{\sim}{\delta}$ and Σ is given by

$$L(\underset{\sim}{\delta}, \Sigma|\underset{\sim}{y}) \propto |\Sigma^{-1}|^{n/2} \exp\left[-\frac{1}{2}(\underset{\sim}{y}-Z\underset{\sim}{\delta})'\Sigma^{-1}\otimes I_n(\underset{\sim}{y}-Z\underset{\sim}{\delta})\right]$$

$$\propto |\Sigma^{-1}|^{n/2} \exp\left[-\frac{1}{2}\mathrm{tr}(A\Sigma^{-1})\right] \tag{6}$$

where

$$A=\begin{pmatrix} (\underset{\sim}{y}_1-Z_1\underset{\sim}{\delta}_1)'(\underset{\sim}{y}_1-Z_1\underset{\sim}{\delta}_1) \cdots (\underset{\sim}{y}_1-Z_1\underset{\sim}{\delta}_1)'(\underset{\sim}{y}_m-Z_m\underset{\sim}{\delta}_m) \\ \vdots \qquad\qquad\qquad \vdots \\ (\underset{\sim}{y}_m-Z_m\underset{\sim}{\delta}_m)'(\underset{\sim}{y}_1-Z_1\underset{\sim}{\delta}_1) \cdots (\underset{\sim}{y}_m-Z_m\underset{\sim}{\delta}_m)'(\underset{\sim}{y}_m-Z_m\underset{\sim}{\delta}_m) \end{pmatrix}$$

an mxm symmetric matrix.

We assume that little is known, a priori, about the parameters, the elements of $\underset{\sim}{\delta}$ and the $m(m+1)/2$ distinct elements of Σ. As our diffuse prior probability density function (pdf), we assume that the elements of $\underset{\sim}{\delta}$ and those of Σ are independently distributed; that is,

$$f(\underset{\sim}{\delta}, \Sigma)=f_1(\underset{\sim}{\delta})\, f_2(\Sigma) \tag{7}$$

In (7), using the invariance theory due to Jeffreys, H. (1961), we take

$$f_1(\underset{\sim}{\delta})=\text{constant} \tag{8}$$

and

$$f_2(\Sigma) \propto |\Sigma|^{-(m+1)/2} \tag{9}$$

If we denote σ^{ij} as the (i, j)th element of the inverse of Σ, the Jacobian of the transformation of the $m(m+1)/2$ variables,

$$(\sigma_{11}, \sigma_{12}, \ldots, \sigma_{mm}) \text{ to } (\sigma^{11}, \sigma^{12}, \ldots, \sigma^{mm})$$

is

$$J=\left|\frac{\partial(\sigma_{11}, \sigma_{12}, \ldots, \sigma_{mm})}{\partial(\sigma^{11}, \sigma^{12}, \ldots, \sigma^{mm})}\right|=|\Sigma|^{m+1} \tag{10}$$

Consequently, the prior pdf in (9) implies the following prior pdf on the $m(m+1)/2$ distinct elements of Σ^{-1}:

$$f_3(\Sigma^{-1}) \propto |\Sigma^{-1}|^{-(m+1)/2} \tag{11}$$

Geisser, S. (1962, 1965) comments on the diffuse prior pdf's in (8) and (9) and points out that (11) would result from taking an informative prior pdf on Σ^{-1} in the Wishart pdf form and allowing the degrees of freedom in the prior pdf to be zero. On combining (8) and (11) with (6), the joint posterior pdf for δ and Σ^{-1} is

$$p(\delta, \Sigma^{-1}|\underset{\sim}{y}) \propto |\Sigma^{-1}|^{n-(m+1)/2} \exp\left[-\frac{1}{2}(\underset{\sim}{y}-Z\delta)'\Sigma^{-1}\otimes I_n(\underset{\sim}{y}-Z\delta)\right]$$

$$\propto |\Sigma^{-1}|^{n-(m+1)/2} \exp\left[-\frac{1}{2}\mathrm{tr}(A\Sigma^{-1})\right] \qquad (12)$$

From (12) it is seen that the conditional posterior pdf for δ, given Σ^{-1}, is in the multivariate normal form with mean

$$E(\delta|\Sigma^{-1}, \underset{\sim}{y}) = [Z'(\Sigma^{-1}\otimes I_n)Z]^{-1}Z'(\Sigma^{-1}\otimes I_n)\underset{\sim}{y} \qquad (13)$$

and the conditional covariance matrix

$$\mathrm{Cov}(\delta|\Sigma^{-n}, \underset{\sim}{y}) = [Z'(\Sigma^{-1}\otimes I_n)Z]^{-1} \qquad (14)$$

In large sample, a Bayes estimator of δ can be obtained from using a consistent estimator of Σ in the conditional mean (13). In ssmall sample, however, it is better to obtain the marginal posterior pdf for δ and base inferences on it rather than rely on conditional results. To obtain the marginal posterior pdf for δ we can use properties of the Wishart pdf to integrate (12) with respect to the distinct elements of Σ^{-1}. This yields

$$p_1(\delta\,|\,\underset{\sim}{y}) \propto |A|^{-n/2} \qquad (15)$$

as the joint posterior pdf for the elements of δ. We have thus proved the following theorem.

[**Theorem 1**] Let $\underset{\sim}{y}_1, \underset{\sim}{y}_2, \cdots, \underset{\sim}{y}_m$ be nxl vectors of observations on m endogeneous variables and generated by the following fully recursive model:

$$\begin{pmatrix} \underset{\sim}{y}_1 \\ \underset{\sim}{y}_2 \\ \vdots \\ \underset{\sim}{y}_m \end{pmatrix} = \begin{pmatrix} Z_1 & & & O \\ & Z_2 & & \\ & & \ddots & \\ O & & & Z_m \end{pmatrix} \begin{pmatrix} \delta_1 \\ \delta \\ \vdots \\ \delta_m \end{pmatrix} + \begin{pmatrix} \underset{\sim}{u}_1 \\ \underset{\sim}{u}_2 \\ \vdots \\ \underset{\sim}{u}_m \end{pmatrix}$$

or $\qquad \underset{\sim}{y} = Z\delta + \underset{\sim}{u}$

where Z_i's, δ_i's, Z, and δ have been defined in connection with (4) and (5). Assume that $\underset{\sim}{u}$ is normally distributed with

$$E(\underset{\sim}{u}) = \underset{\sim}{0} \qquad E(\underset{\sim}{u}\underset{\sim}{u}') = \Sigma\otimes I_n$$

where Σ is an mxm positive definite symmetric matrix. Let the diffuse prior pdf of δ and Σ^{-1} be as follows:

$$f(\delta, \Sigma^{-1}) \propto |\Sigma^{-1}|^{-(m+1)/2}$$

Then, the joint posterior pdf of δ and Σ^{-1} is given by

$$p(\delta, \Sigma^{-1}|y) \propto |\Sigma^{-1}|^{n-(m+1)/2} \exp\left[-\frac{1}{2}\mathrm{tr}(A\Sigma^{-1})\right]$$

where the matrix A has been defined in connection with (6). The conditional posterior pdf of δ, given Σ^{-1} is in the multivariate normal form with mean

$$E(\delta|\Sigma^{-1}, y) = [Z'(\Sigma^{-1}\otimes I_n)Z]^{-1} Z'(\Sigma^{-1}\otimes I_n)y$$

and the conditional covariance matrix

$$\mathrm{cov}(\delta|\Sigma^{-1}, \delta) = [Z'(\Sigma^{-1}\otimes I_n)Z]^{-1}$$

Furthermore, the marginal posterior pdf of δ is given by

$$p_1(\delta|y) \propto |A|^{-n/2}.$$

We note that the marginal posterior pdf of δ in (15) is in exactly the same form as that for the coefficients of the seemingly unrelated regression model discussed by Zellner, A. (1962).

III. Bayesian Limited Information Analysis

In this section, we take acconnt only the prior identifiable information pertaining to the prarameters of one single equation in an m simultaneous equations model. Such analysis will be useful, for example, when we are uncertain about the formulation of some or all of the other structual equations of a model or when we wish to build into our analysis just that part of the identifiable prior information relating to the parameters of one equation.

Consider a particular structural equation of an m simultaneous equations model, say the first:

$$y_1 - Y_1\gamma_1 = X_1\beta_1 + u_1 \qquad (16)$$

where y_1 is an nxl vector of observations on the endogenous variable whose coefficient is set equal to one by our normalization, Y_1 is an nxm_1 matrix of observations on m_1 other endogenous variables appearing in the first equation with coefficients not assumed equal to zero, X_1 is an nxk_1 matrix of observations on k_1 predetermined variables appearing in thə first equation with coefficients not assumed equal to zero, γ_1 and β_1 are m_1xl and k_1xl coefficient vectors, respecsvely, and u_1 is an nxl vector of serially uncorrelated normal disturbance errors, each with mean zero and variance σ_{11}. We assume that the parameters of (16) are identified by virtue of the restrictions imposed, that is, we can write (16) as

$$y_1 - Y_1\gamma_1 - Y_0\gamma_0 = X_1\beta_1 + X_0\beta_0 + u_1$$

with $\gamma_0 = 0$ and $\beta_0 = 0$, where $Y = (y_1, Y_1, Y_0)$, an nxm matrix of observations on all endogenous variables, and $X = (X_1, X_0)$, an nxk matrix of observations on all predetermined variables. The identifiable restrictions are $\gamma_0 = 0$ and $\beta_0 = 0$.

Since the identification problem for simultaneous equation models is often considered in terms of the relations between reduced form parameters and structural parameters, we shall discuss this approach in Bayesian terms by making use of the valuable work of Dreze, J.H. (1976). Since the reduced form of $Y\Gamma = XB + U$ is given by $Y = X\Pi + V$, where $\Pi = B\Gamma^{-1}$ and $V = U\Gamma^{-1}$, the reduced-form equations for (y_1, Y_1) can be written as follows:

$$(y_1, Y_1) = (X_1, X_0)\begin{pmatrix} \pi_{11} & \Pi_{01} \\ \pi_{10} & \Pi_{00} \end{pmatrix} + (v_1, V_1) \tag{17}$$

where π_{11} is a $k_1$xl vector, π_{10} is a $k_0$xl vector, Π_{01}, a k_1xm_1 matrix, and Π_{00}, a k_0xm_1 matrix. The nx(m_1+1) matrix (v_1, V_1) contains reduced-form disturbance errors, where v_1 is an nxl vector and V_1 is an nxm_1 matrix.

On postmultiplying both sides of (17) by $(1, -\gamma_1')'$ and equating the resulting coefficients with those appearing in (16), we obtain

$$\pi_{11} - \Pi_{01}\gamma_1 = \beta_1 \tag{18}$$

and $$\pi_{10} - \Pi_{00}\gamma_1 = \beta_0 = 0 \tag{19}$$

Using (18) and (19), we can express the reduced-form equations, given in (17), as

$$\begin{aligned} y_1 &= (X_1\Pi_{01} + X_0\Pi_{00})\gamma_1 + X_1\beta_1 + v_1 \\ &= X\Pi_0\gamma_1 + X_1\beta_1 + v_1 \end{aligned} \tag{20}$$

and $$Y_1 = X\Pi_0 + V_1 \tag{21}$$

where $\Pi_0' = (\Pi_{01}', \Pi_{00}')$. Under the assumption that the rows of (v_1, V_1) are normally and independently distributed, each with zero mean vector and $(m_1+1)\times(m_1+1)$ positive definite symmetric covariance matrix Ω_1, (21) is in the form of a multivariate regression model. The maximum likelihood estimator of Π_0 is given by

$$\hat{\Pi}_0=(X'X)^{-1}X'Y_1$$

Then, for given Π_0, say $\Pi_0=\hat{\Pi}_0$, (20) is in the form of a multiple regression model, i. e.

$$y_1=W_1\delta_1+v_1$$

where $W_1=(X\Pi_0,\ X_1),\quad \delta_1=\begin{pmatrix}\gamma_1\\ \beta_1\end{pmatrix}$

If the diffuse prior pdf of δ_1 and ω_{11}, the common variance of the elements of v_1, is taken as

$$f(\delta_1,\ \omega_{11})\propto \omega_{11}^{-1/2}$$

then the joint conditional pdf of δ_1 and ω_{11}, given $\Pi_0=\hat{\Pi}_0$ is

$$p(\delta_1,\ \omega_{11}|y_1,\ W_1)\propto \omega_{11}^{-(n+1)/2}\exp\left\{-\frac{1}{2\omega_{11}}[\nu s^2+(\delta-\hat{\delta}_1)'W_1'W_1(\delta_1-\hat{\delta}_1)]\right\} \quad (22)$$

where $\hat{\delta}_1=\begin{pmatrix}\hat{\gamma}_1\\ \hat{\beta}_1\end{pmatrix}=(W_1'W_1)^{-1}W_1'y_1,\qquad \nu=n-m_1-k_1$

$$s^2=(y_1-W_1'\hat{\delta}_1)'(y_1-W\hat{\delta}_1)/\nu$$

Integrating (22) with respect to ω_{11}, we ottain the conditional posterion pdf of $\hat{\delta}_1$, given $\Pi_0=\hat{\Pi}_0$ as follows:

$$p_1(\hat{\delta}_1|y_1,\ W_1)\propto[\nu s^2+(\delta_1-\hat{\delta}_1)'W_1'W_1(\delta_1-\hat{\delta}_1)]^{-n/2} \quad (23)$$

which is in the form of a multivariate Student t pdf. The conditional posterior mean, given $\Pi_0=\hat{\Pi}_0$ is $\hat{\delta}_1$ which is just the two-stage least squares (2SLS) estimator. It must be appreciated, however, that $\hat{\delta}_1$ may be a poor approximation to the unconditional posterior mean of δ_1 in small samples. To make inferences with a small sample, it is better to base on unconditional posterior pdf of δ_1 than to rely on the conditional result obtained in (23).

To obtain an unconditional posterior pdf for the parameters in (16), we employ an approach similar in certain respects to that put by Drèze, L.H. (1976). Combining (16) and the part of (17) relating to Y_1, we have

$$(y_1-Y_1\gamma_1,\ Y_1)=(X_1,\ X_0)\begin{pmatrix}\beta_1 & \Pi_{01}\\ \beta_0 & \Pi_{00}\end{pmatrix}+(u_1,\ V_1) \quad (24)$$

where we have already incorporated the prior information that certain endogenous variables do not appear in (16) and will later itnroduce $\beta_0=0$. Since $u_1=v_1-V_1\gamma_1$, the error matrix in (24) can be written as

$$(\underset{\sim}{u}_1,\ V_1)=(\underset{\sim}{y}_1,\ V_1)\begin{pmatrix}1 & \underset{\sim}{0}' \\ -\underset{\sim}{\gamma}_1 & I\end{pmatrix}\equiv(\underset{\sim}{y}_1,\ V_1)C \tag{25}$$

Under the assumption that the rows of $(\underset{\sim}{y}_1,\ V_1)$ are normally and independently distributed, each with zero mean vector and positive definite symmetric $(m_1+1)\times(m_1+1)$ covariance matrix Ω_1, the covariance matrix for each row of $(\underset{\sim}{u}_1,\ V_1)$ is $\Omega=C'\Omega_1C$. Then the the likelihood function for the system in (24) is

$$L(\underset{\sim}{\gamma}_1, \underset{\sim}{\beta}, \Pi_0, \Omega \mid y_1, Y_1)\propto|\Omega|^{-n/2}\exp\left[-\frac{1}{2}\mathrm{tr}(W-X\Pi_*)'(W-X\Pi_*)\Omega^{-1}\right] \tag{26}$$

where $W=(\underset{\sim}{y}_1-Y_1\underset{\sim}{\gamma}_1,\ Y_1)$, $\Pi_*=(\underset{\sim}{\beta},\ \Pi_0)$, and $\underset{\sim}{\beta}'=(\underset{\sim}{\beta}_1',\ \underset{\sim}{\beta}_0')$ with the prior information $\underset{\sim}{\beta}_0=\underset{\sim}{0}$ to be incorporated in our prior pdf. If the diffuse prior pdf for the parameters is taken as

$$f(\underset{\sim}{\gamma}_1,\ \underset{\sim}{\beta}_1,\ \Pi_0, \Omega \mid \underset{\sim}{\beta}_0=\underset{\sim}{0})\propto|\Omega|^{-(m_1+1)/2} \tag{27}$$

then the posterior pdf is

$$p(\underset{\sim}{\gamma}_1, \underset{\sim}{\beta}, \Pi_0, \Omega \mid \underset{\sim}{y}_1, Y_1)\propto|\Omega|^{(n+m_1+2)/2}$$
$$\times\ \exp\left[-\frac{1}{2}\mathrm{tr}(W-X\Pi_*)'(W-X\Pi_*)\Omega^{-1}\right] \tag{28}$$

The posterior pdf given in (28) is in what Tiao, G. C. and Zellner, A. (1964) call the "inverted" Wishart form. On integrating (28) with respect to the elements of Ω, we obtain

$$p_1(\underset{\sim}{\gamma}_1, \underset{\sim}{\beta}, \Pi_0 \mid \underset{\sim}{y}_1, Y_1)\propto|(W-X\Pi_*)'(W-X\Pi_*)|^{-n/2}$$
$$\propto|G+(\Pi_*-\hat{\Pi}_*)'X'X(\Pi_*-\hat{\Pi}_*)|^{-n/2} \tag{29}$$

where $\hat{\Pi}_*=(X'X)^{-1}X'W$

$G=(W-X\hat{\Pi}_*)'(W-X\hat{\Pi}_*)$.

Since (29) has the form of a generalized Student t pdf (see Dickey, J. M., 1967), we can use this fact to integrate it with respect to the elements of Π_0 to obtain

$$p_2(\underset{\sim}{\gamma}_1, \underset{\sim}{\beta} \mid \underset{\sim}{y}_1, Y_1)\propto g_{11}^{(n-m_1-k)/2}[g_{11}+(\underset{\sim}{\beta}-\hat{\underset{\sim}{\beta}})'X'X(\underset{\sim}{\beta}-\hat{\underset{\sim}{\beta}})]^{-(n-m_1)/2}$$
$$\propto g_{11}^{-k/2}\left[1+(\underset{\sim}{\beta}-\hat{\underset{\sim}{\beta}})'\frac{X'X}{g_{11}}(\underset{\sim}{\beta}-\hat{\underset{\sim}{\beta}})\right]^{-(n-m_1)/2} \tag{30}$$

where g_{11} is the (1,1) element of G, i.e.,

$$g_{11}=(\underset{\sim}{y}_1-Y_1\underset{\sim}{\gamma}_1-X\hat{\underset{\sim}{\beta}})'(\underset{\sim}{y}_1-Y_1\underset{\sim}{\gamma}_1-X\hat{\underset{\sim}{\beta}})$$
$$=(\hat{\underset{\sim}{v}}_1-V_1\underset{\sim}{\gamma}_1)'(\hat{\underset{\sim}{v}}_1-V_1\underset{\sim}{\gamma}_1) \tag{31}$$

$$(\hat{v}_1, V_1)=(y_1, Y_1)-X(\hat{\pi}_1, \hat{\Pi}_0)$$

and $$\hat{\beta}=(X'X)^{-1}X_1(y_1-Y_1\gamma_1).$$

In order to express the quadratic form in β of (30) by the form in δ_1, we write

$$\begin{aligned} X(\beta-\hat{\beta}) &= X_1\beta_1+X_0\beta_0-X(X'X)^{-1}X'(y_1-Y_1\gamma_1) \\ &= X_1\beta_1-(\hat{y}_1-\hat{Y}_1\gamma_1) \\ &= -(\hat{y}_1-\hat{Z}_1\delta_1) \end{aligned}$$

$$\hat{y}_1=X(X'X)^{-1}X'y_1$$
$$\hat{Y}_1=X(X'X)^{-1}X'Y_1=X\hat{\Pi}_0$$
$$\hat{Z}_1=(\hat{Y}_1, X_1)$$

and the prior information $\beta_0=0$ has been used.

Thus $$(\beta-\hat{\beta})'X'X(\beta-\hat{\beta})=(\hat{y}_1-\hat{Z}_1\delta_1)'(\hat{y}_1-\hat{Z}_1\delta_1) \tag{32}$$

Using the 2SLS estimator of δ_1 given in (22), i.e.,

$$\begin{aligned} \tilde{\delta}_1=\begin{pmatrix} \tilde{\gamma}_1 \\ \tilde{\beta}_1 \end{pmatrix} &= (W_1'W_1)^{-1}W_1'y_1 \\ &= \begin{pmatrix} \hat{\Pi}_0'X'X\hat{\Pi}_0 & \hat{\Pi}_0'X'X_1 \\ X_1'X\hat{\Pi}_0 & X_1'X_1 \end{pmatrix}^{-1}\begin{pmatrix} \hat{\Pi}_0'X'y_1 \\ X_1'y_1 \end{pmatrix} \end{aligned} \tag{33}$$

we can write (32) in the following form:

$$\begin{aligned} (\hat{y}_1-\hat{Z}_1\delta_1)'(\hat{y}_1-\hat{Z}_1\delta_1) &= [(\hat{y}_1-\hat{Z}_1\tilde{\delta}_1-\hat{Z}_1(\delta_1-\tilde{\delta}_1)]'[\hat{y}_1-\hat{Z}_1\tilde{\delta}_1-\hat{Z}_1(\delta_1-\tilde{\delta}_1)] \\ &= (\delta_1-\tilde{\delta}_1)'\hat{Z}_1'\hat{Z}_1(\delta_1-\tilde{\delta}_1) \end{aligned}$$

since $\hat{Z}_1'(\hat{y}_1-\hat{Z}_1\tilde{\delta}_1)=0$ from (33),

$$\begin{aligned} \hat{y}_1-\hat{Z}_1\tilde{\delta}_1 &= \hat{y}_1-\hat{Y}_1\tilde{\gamma}_1-X_1\tilde{\beta}_1 \\ &= y_1-Y_1\tilde{\gamma}_1-X_1\tilde{\beta}_1-(\hat{v}_1-V_1\tilde{\gamma}_1) \\ &= \hat{u}_1-(\hat{v}_1-V_1\tilde{\gamma}_1)=0 \end{aligned}$$

Hence, the quadratic form in β of (30) can be expressed as

$$(\beta-\hat{\beta})'X'X(\beta-\hat{\beta})=(\delta_1-\tilde{\delta}_1)'\hat{Z}_1'\hat{Z}_1(\delta_1-\tilde{\delta}_1) \tag{34}$$

Therefore, (30) can be expressed as

$$p_2(\delta_1|y_1, Y_1)\propto g_{11}^{-k/2}[1+(\delta_1-\tilde{\delta}_1)'H(\delta_1-\tilde{\delta}_1)] \tag{35}$$

where $H=\hat{Z}_1'\hat{Z}_1/g_{11}$.

If g_{11} is independent of γ_1, then the posterior pdf of δ_1 given in (35) is in the multivariate Student t form centered at the 2SLS estimate $\hat{\delta}_1$. However, since the conditional posterior pdf of β_1, given γ_1, is in the multivariate Student t form, (35) can be integrated with respect to the k_1 elements of β_1 to yield the following marginal posterior pdf of γ_1:

$$p_3(\gamma_1|y_1, Y_1)\propto g_{11}^{-(k-k_1)/2}[1+(\gamma_1-\hat{\gamma}_1)'H_1(\gamma_1-\hat{\gamma}_1)]^{-(a+m_1)/2} \qquad (36)$$

where $a=n-2m_1-k_1$,

$$H_1=[\hat{Y}_1'\hat{Y}_1-\hat{Y}_1'X_1(X_1'X_1)^{-1}X_1'\hat{Y}_1]/g_{11}.$$

With regard to the posterior pdf of an element of β_1, say β_{11}, (35) can be integrated analytically with respect to the elements of β_1, other than β_{11}, to provide the joint posterior pdf of γ_1 and β_{11} which can be analyzed conveniently by using numerical integration techniques when the dimensionality of γ_1 is small.

Now we summarize the results obtained so far in the following theorem:
[**Theorem 2**] The reduced form of the simultaneous equation model $Y\Gamma=XB+U$ is given by $Y=X\Pi+V$, where $\Pi=B\Gamma^{-1}$ and $V=U\Gamma^{-1}$. Suppose that the first structural equation of $Y\Gamma=XB+U$ with a priori identifiable information $\gamma_0=0$ and $\beta_0=0$ imposed is given by $y_1-Y_1\gamma_1=X_1\beta_1+u_1$ and is expressed as follows:

$$(y_1-Y_1\gamma_1, Y_1)=(X_1, X_0)\begin{pmatrix}\beta_1 & \Pi_{01}\\ 0 & \Pi_{00}\end{pmatrix}+(v_1, V_1)\begin{pmatrix}1 & 0'\\ -\gamma_1 & I\end{pmatrix}$$

where y_1 is an nxl vector of observations on the endogenous variable whose coefficient is set equal to one, Y_1 is an nxm_1 matrix of observations on m_1 other endogenous variables appearing in the first equation with coefficients not assumed equal to zero, $X=(X_1, X_0)$ is an nxk matrix of observations with rank k on the k predetermined variables, with X_1, an nxk_1 matrix of observations on the k_1 predetermined variables appearing in the first equation with coefficients not assumed equal to zero, and X_0, an nxk_0 matrix of observations on k_0 predetermined variables excluded fron the first equation on a priori grounds. γ_1 and β_1 are $m_1$xl and $k_1$xl coefficient vectors, respectively, Π_{01} and Π_{00} are k_1xm_1 and k_0xm_1 matrices, respectively, the nx(m_1+1) matrix (v_1, V_1) contains reduced-form disturbance error terms, where v_1 is an nxl vector and V_1 is an nxm_1 matrix. Assume that the rows of (v_1, V_1) are normally and independently distributed, each with zero mean vector and $(m_1+1)\times(m_1+1)$ positive definite symmetric covariance matrix. If the diffuse prior pdf of the parameters is taken as

$$f(\gamma_1, \beta_1, \Pi_0, \Omega_*) \propto |\Omega_*|^{-(m_1+1)}$$

where $\Pi_0' = (\Pi_{01}', \Pi_{00}')$ and Ω_* is the covariance matrix for each row of (u_1, V_1). Then, the posterior pdf of $\delta_1' = (\gamma_1', \beta_1')$ is given by

$$p_2(\delta_1 | y_1, Y_1) \propto g_{11}^{-k/2} [1 + (\delta_1 - \hat{\delta}_1)'(\frac{\hat{Z}_1'\hat{Z}_1}{g_{11}})(\delta_1 - \hat{\delta}_1)]$$

where g_{11} is the (1,1) element of $G = (W - X\hat{\Pi}_*)'(W - X\hat{\Pi}_*)$,

$$\hat{\Pi}_* = (X'X)^{-1}X'W, \quad W = (y_1 - Y_1\hat{\gamma}_1, Y_1)$$

$$\hat{\delta}_1 = \begin{pmatrix} \hat{\gamma}_1 \\ \hat{\beta}_1 \end{pmatrix} = \begin{pmatrix} \hat{Y}'_1\hat{Y}_1 & X_1'\hat{Y}_1 \\ \hat{Y}_1'X_1 & X_1'X_1 \end{pmatrix}^{-1} \begin{pmatrix} \hat{Y}_1'y_1 \\ X_1'y_1 \end{pmatrix}$$

$$\hat{Y}_1 = X(X'X)^{-1}X'Y_1, \quad \hat{Z}_1 = (\hat{Y}_1, X_1).$$

The marginal posterior pdf of γ_1 is given by

$$p_3(\gamma_1 | y_1, \hat{Y}_1) \propto g_{11}^{-(k-k_1)/2} [1 + (\gamma_1 - \hat{\gamma}_1)'H_1(\gamma_1 - \hat{\gamma}_1)]^{-(a+m_1)/2}$$

where $a = n - 2m_1 - k_1$ and $H_1 = [\hat{Y}_1'\hat{Y}_1 - \hat{Y}_1'X_1(X_1'X_1)^{-1}X_1'\hat{Y}_1]/g_{11}$.

IV. BAYESIAN FULL INFORMATION ANALYSIS

In the preceeding section, we have drived the posterior pdf of the parameters in a single equation, using just the identifiable prior information for those parameters. We now derive the joint posterior pdf of the parameters appearing in all structual equations. Under the assumption that the rows of U are normally and independently distributed, each with zero mean vector and mxm positive definite symmetric covariance matrix Σ, the likelihood function is

$$L(\Gamma, B, \Sigma, Y) \propto |\Sigma|^{-n/2} |\Gamma|^n \exp[-\tfrac{1}{2}\mathrm{tr}(Y\Gamma - XB)'(Y\Gamma - XB)\Sigma^{-1}]$$

To incorporate the identifiable prior information that certain elements of (Γ, B) are equal to zero, let $\delta_i' = (\gamma_i', \beta_i')$ denote the i-th row of (Γ, B) by omitting $\gamma_{ii} \equiv 1$, $\beta_{ii} \equiv 1$ and the elements $\gamma_{ij} \equiv 0$, $\beta_{ij} \equiv 0$ corresponding to the identifiable prior information and let $Z_i = (Y_i, X_i)$ denote the submatrix of (Y, X) consisting of those columns whose indices correspond to the elements of δ_i. Then, the simultaneous equations model with the identifiable prior information incorporated can be written as follows:

$$\begin{pmatrix} \underset{\sim}{y}_1 \\ \underset{\sim}{y}_2 \\ \vdots \\ \underset{\sim}{y}_m \end{pmatrix} = \begin{pmatrix} Z_1 & & & O \\ & Z_2 & & \\ & & \ddots & \\ O & & & Z_m \end{pmatrix} \begin{pmatrix} \underset{\sim}{\delta}_1 \\ \underset{\sim}{\delta}_2 \\ \vdots \\ \underset{\sim}{\delta}_m \end{pmatrix} + \begin{pmatrix} \underset{\sim}{u}_1 \\ \underset{\sim}{u}_2 \\ \vdots \\ \underset{\sim}{u}_m \end{pmatrix} \tag{37}$$

or $\underset{\sim}{y} = Z\underset{\sim}{\delta} + \underset{\sim}{u}$

The likelihood function is then written as

$$L(\underset{\sim}{\delta}, \Sigma|Y) \propto |\Sigma|^{-n/2}|\Gamma|^n \exp[-\tfrac{1}{2}\mathrm{tr}(\underset{\sim}{y}-Z\underset{\sim}{\delta})'(\underset{\sim}{y}-Z\underset{\sim}{\delta})(\Sigma^{-1}\otimes I_n)]$$

$$\propto |\Sigma|^{-n/2}|\Gamma'\hat{\Omega}\Gamma|^{n/2}\exp[-\tfrac{1}{2}n\mathrm{tr}(\hat{\Omega}\Omega^{-1})]$$

$$.\exp[-\tfrac{1}{2}\{(\underset{\sim}{\delta}-\hat{\underset{\sim}{\delta}})'M(\underset{\sim}{\delta}-\hat{\underset{\sim}{\delta}})+\hat{\underset{\sim}{y}}'[I_{mn}-\hat{Z}M^{-1}\hat{Z}')\hat{\underset{\sim}{y}}\}] \tag{38}$$

where $\hat{\underset{\sim}{\delta}} = M^{-1}\hat{Z}'(\hat{\Sigma}^{-1}\otimes I_n)\hat{\underset{\sim}{y}}$, $M = \hat{Z}'(\hat{\Sigma}^{-1}\otimes I_n)\hat{Z}$

$\hat{\underset{\sim}{y}} = X(X'X)^{-1}X'\underset{\sim}{y}$, $\hat{Y} = X(X'X)^{-1}X'Y = X\hat{\Pi}$, $\hat{Z} = (\hat{Y}, X)$

$\hat{\Sigma} = \hat{U}'\hat{U}/n$, $\hat{U} = (\hat{\underset{\sim}{u}}_1, \hat{\underset{\sim}{u}}_2, \ldots . \hat{\underset{\sim}{u}}_m)$, $\hat{\underset{\sim}{u}}_i = \underset{\sim}{y}_i - Z_i\hat{\underset{\sim}{\delta}}_i$

$\hat{\Omega} = (Y - X\hat{\Pi})'(Y - X\hat{\Pi})/n$

$\Sigma = \Gamma'\Omega\Gamma.$

If we employ the extended natural conjugate prior pdf of $(\underset{\sim}{\delta}, \Sigma)$ introduced by Dreze, J,H. and Morales, J.A. (1976) as

$$f(\underset{\sim}{\delta}, \Sigma) \propto |\Sigma|^{-(m+1)/2}|\Gamma|^{m+1}\exp[-\tfrac{1}{2}\mathrm{tr}[\Sigma^{-1}\otimes I_n)]$$

$$.\exp[-\tfrac{1}{2}(\underset{\sim}{\delta}-\underset{\sim}{\delta}_0)'K(\underset{\sim}{\delta}-\underset{\sim}{\delta}_0)] \tag{39}$$

where $\underset{\sim}{\delta}_0$ is the prior mean vector and K^{-1} is the prior c ovariance matrix. Then, on combining the likelihood function (38) with the prior pdf (39), we obtain the posterior pdf of $(\underset{\sim}{\delta}, \Omega)$ as follows:

$$p(\underset{\sim}{\delta}, \Omega|Y) \propto |\Omega|^{-(n+m+1)/2}\exp[-\frac{n}{2}\mathrm{tr}(\hat{\Omega}\Omega^{-1})]$$

$$.\exp[-\tfrac{1}{2}\{(\underset{\sim}{\delta}-\underset{\sim}{\delta}^*)'(M+K)(\underset{\sim}{\delta}-\underset{\sim}{\delta}^*)+S^*\}] \tag{40}$$

where $\underset{\sim}{\delta}^* = (M+K)^{-1}(M\hat{\underset{\sim}{\delta}}+K\underset{\sim}{\delta}_0)$

$S^* = \hat{\underset{\sim}{y}}'\hat{\underset{\sim}{y}} + \underset{\sim}{\delta}_0'K\underset{\sim}{\delta}_0 - \underset{\sim}{\delta}^{*\prime}(M+K)\underset{\sim}{\delta}^*.$

We see that in (40) Ω and $\underset{\sim}{\delta}$ are independently distributed, with the elemənts of Ω having an inverted Wishart pdf and those of $\underset{\sim}{\delta}$ having a multivariate normal pdf, with mean vector $\underset{\sim}{\delta}^*$ and covariance matrix $(M+K)^{-1}$.

Zellner, A. (p. 272, 1971) has considered the prior pdf as

$$f(\underset{\sim}{\delta}, \Sigma) \propto f_1(\underset{\sim}{\delta}) |\Sigma|^{-(m+1)/2}$$

and developed an asymptotic expansion of the posterior pdf of (δ, Ω) in which the leading term is given by

$$p(\delta, \Omega Y) \propto |\Omega|^{-(n+m+1)/2} \exp[-\frac{n}{2} tr(\hat{\Omega}\Omega^{-1})]$$

$$\cdot \exp[-\tfrac{1}{2}(\underset{\sim}{\delta} - \hat{\underset{\sim}{\delta}})' M (\underset{\sim}{\delta} - \hat{\underset{\sim}{\delta}})] \quad (41)$$

Now we summarize the results obtained in this section as follows:

[**Theorem 3**] The simultaneous equations model $Y\Gamma = XB + U$ with the identifiable prior information incorporated is given by $\underset{\sim}{y} = Z\underset{\sim}{\delta} + \underset{\sim}{u}$, where $\underset{\sim}{u}$ is normally distributed with zero mean vector and covariance matrix $\Sigma \otimes I_n$. If (39) is used as the extended natural conjugate prior pdf of $(\underset{\sim}{\delta}, \Sigma)$, then the posterior pdf of $(\underset{\sim}{\delta}, \Omega)$, where $\Sigma = \Gamma'\Omega\Gamma$, is obtained in (40).

REFERENCES

1. Dickey, J. M. Matricvariate generalizations of the multivariate t distribution and the inverted multivariate t distribution. Annals of Mathematical Statistics, 38: 511–518, 1967.
2. Dreze, J. H. Bayesian limited information analysis of the simultaneous equations model. Econometrica 44: 1045–76, 1976.
3. Dreze, J. H. and Morales, J. A. Bayesian full information analysis of simultaneous equations. Journal of American Statistical Association 71: 919–923, 1976.
4. Geisser, S. A Bayes approach for combining correlated estimates. Journal of American Statistical Association 60: 602–607, 1962.
5. Geisser, S. Bayesian estimaion in multivariate analysis. Annals of Mathematical Statistics 36: 150–159, 1965.
6. Jeffreys, H. Theory of Probability. Oxford: Clarendon, 1961.
7. Tiao, G. C. and Zellner, A. Bayes' theorem and the use of prior knowledge in regression analysis. Biometrika 51: 219–230, 1964.
8. Zellner, A. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. Journal of American Statistical Association 57: 348–368, 1962.
9. Zellner, A. Introduction to Bayesian Inference in Econometrics. John Wiley & Sons, Inc., New York, 1971.