ECONOMETRIC INSTITUTE

PREDICTIVE MOMENTS OF SIMULTANEOUS ECONOMETRIC MODELS A BAYESIAN APPROACH

H.K. van DIJK and T. KLOEK

REPORT 7614/E

ERASMUS UNIVERSITY ROTTERDAM, P.O. BOX 1738, ROTTERDAM, THE NETHERLANDS

Erasmus University Rotterdam
ECONOMETRIC INSTITUTE

Report 7614/E

PREDICTIVE MOMENTS OF SIMULTANEOUS ECONOMETRIC MODELS
A Bayesian Approach

by

H.K. van Dijk and T. Kloek

August 1976

PREDICTIVE MOMENTS OF SIMULTANEOUS ECONOMETRIC MODELS*

A Bayesian Approach

by

H.K. van Dijk and T. Kloek

ABSTRACT

Bayesian full-information predictive moments are derived using both exact and stochastic prior information about the structural and reduced form of an econometric model. The prior information on the nuisance parameters (constant terms and covariance matrix) is such that analytical integration is possible. The prior information on the economically interesting parameters is allowed to be much more flexible, so that numerical methods are required. The numerical part of the integration problem is being solved by Monte Carlo methods; compare [12]. In that paper we concentrated ourselves on the economically interesting parameters. For prediction purposes, however, we also need the constant terms. In this area a number of analytical results can be obtained. These are derived in the present paper. Small sample results are produced, contrary to the classical approach and Rothenberg's large sample Bayesian analysis [16].

Contents

* A preliminary version of this paper was presented at the First European Conference on Foundations and Applications of Bayesian Methods, held at Fontainebleau, June 1-4, 1976.

1. INTRODUCTION

This paper is part of a larger research project, which deals with a Bayesian full-information analysis of the simultaneous linear equation system. Nowadays there are various Bayesian approaches to this problem, see the recent survey by Rothenberg [17] and the references cited there, especially Drèze [5], Harkema [9], Morales [13], and Richard [15].

A particular difficult problem is the conflict between two - apparently reasonable - requirements on prior information, i.e., analytical tractability and richness of prior information. Rothenberg [16, section 6.4] points out that the class of analytically tractable priors is not rich enough. The main advantage of these priors is that integration can be performed analytically. This is important in view of the problems rising when numerical integration has to be performed in spaces with high dimensionality.

Our objective is to develop methods which are computationally efficient so that ultimately medium-size econometric models may be handled and which are flexible enough to allow for a rich set of possible prior densities on economically meaningful parameters.

We opted for a mixed analytical numerical approach. To reduce the dimensionality of the numerical integration problem we used analytical integration on a subset of parameters with non-informative or conjugate priors. On another subset of parameters, with informative priors, which are not subjected to many (mathematical) restrictions, we use numerical integration methods. This enables the researcher to make use of a much wider class of prior densities than the analytically tractable priors. Examples of such priors are given in [12, section 3]. In that paper we concentrated ourselves on the economically interesting parameters. For prediction purposes, however, we also need the constant terms. In this area a number of analytical results can be obtained. These are derived in the present paper. We advocate the use of Monte Carlo integration methods for the computation of posterior moments and marginal posterior densities. Monte Carlo has the advantage above standard numerical integration methods, like product rules, that it is easier to work with in large dimensions.[1] The application of Monte Carlo to Bayesian estimation problems was introduced by Kloek and van Dijk [12].

The subject matter of this paper is organized as follows. In section 2 expressions for the posterior moments of structural and reduced form parameters are derived except for the covariance matrix of the disturbances.

[1] For a survey of numerical integration methods see Haber [8].

In section 3, multiperiod predictive moments are derived, following a suggestion from Chow [2]. In section 4 some remarks are made about Bayesian approaches to full-information posterior analysis and prediction.

The notation used is a slight variant of Theil's [18] notation of the simultaneous equation system.

2. FULL INFORMATION POSTERIOR MOMENTS OF STRUCTURAL AND REDUCED FORM PARAMETERS

2.1. The Statistical Model

We consider the simultaneous linear equation model

$$Y\Gamma + XB = U \tag{2.1}$$

where the matrix Y consists of n observations on G current endogenous variables and the matrix X consists of n observations on K predetermined variables; U is an n×G matrix of unobservable disturbances. The jointly dependent variables are linked to the predetermined variables and the disturbances through the G×G matrix Γ and the K×G matrix B. These matrices contain constants, some of which are known a priori (see Assumption (5) below).

Several assumptions are specified with respect to the system (2.1), partly in order to keep the information processing at a tractable level.

ASSUMPTION 1. $|\Gamma| \neq 0$.

The determinant of Γ should not equal zero. So, the reduced form $Y = -XB\Gamma^{-1} + U\Gamma^{-1}$ exists.

ASSUMPTION 2. _The_ n _rows of_ U _are independently and identically distributed as_ $N(0, H^{-1})$, _where_ H _is a positive definite symmetric_ (PDS) _matrix_.

We abstract from serial correlation in the structural disturbances. The matrix H has the interpretation of a precision matrix of a multinormal process. H^{-1} is the variance covariance matrix. The PDS requirement on the variance covariance matrix implies that any identities in the model

(2.1) are removed by a preliminary substitution procedure; compare Rothenberg [16, Ch. 4, Appendix B].

ASSUMPTION 3. *The vectors* x_s', u_t', u_{t+1}', ..., u_n' *are independently distributed for any* $s \leq t$ *and any* $t = 1, \ldots, n$.

Here x_s' represents the s-th row of X and u_t' the t-th row of U. The predetermined variables are independently distributed of current and future values of the disturbances.

ASSUMPTION 4. *Rank* $(X) = K < n$.

Linear dependence between the columns of X is excluded.

Assumptions (1)-(3) enable us to write the likelihood function of the system (2.1) as

$$(2.2) \qquad \ell(Y;\, X,\, \Gamma,\, B,\, H) \propto |H|^{\frac{1}{2}n}|\Gamma|^{n}\exp[-\tfrac{1}{2}\mathrm{tr}(QH)]$$

where $|\Gamma|$ denotes the absolute value of the determinant of Γ and Q is defined by

$$(2.2a) \qquad Q = (Y\Gamma + XB)'(Y\Gamma + XB)$$

Using Assumption (4) we define

$$(2.3) \qquad \hat{\Pi} = (X'X)^{-1}X'Y$$

$$(2.4) \qquad \hat{B} = -\hat{\Pi}\Gamma$$

$$(2.5) \qquad W = (Y - X\hat{\Pi})'(Y - X\hat{\Pi})$$

Note that $\hat{\Pi}$ and W are classical (ordinary) least squares estimates, but that $\hat{B}$ is a "hybrid" containing estimated elements, $\hat{\Pi}$, and unknown elements, Γ. We can make use of (2.3) to find

$$(2.6) \qquad X'(Y - X\hat{\Pi}) = 0$$

If a constant term is present in each equation of the model so that X contains a column of unit elements to be denoted by ι, (2.6) implies

(2.6a) $$\iota'(Y - X\hat{\Pi}) = 0$$

Furthermore we use

(2.7) $$\begin{aligned} Y\Gamma + XB &= (Y + X\hat{B}\Gamma^{-1})\Gamma + X(B - \hat{B}) \\ &= (Y - X\hat{\Pi})\Gamma + X(B - \hat{B}) \end{aligned}$$

Now we write

(2.8) $$Q = \Gamma'W\Gamma + (B - \hat{B})'X'X(B - \hat{B})$$

This factorization will prove useful in the following subsections.

Usually, more information is available then we stated sofar. Plausible intervals of parameters as the marginal propensity to consume and the short term multiplier are immediate examples. We will define a rather wide class of prior distributions on the elements of Γ, B and H. This is summarized in assumptions 5 and 6.

ASSUMPTION 5. *Some restrictions on the elements of* Γ *and* B *are known exactly* a priori.

Examples are the unit diagonal elements of Γ following from normalization and a number of zero elements of Γ and B implied by zero identifying restrictions.[1] If preliminary substitution of identities has been performed (compare the comments on Assumption 2), these unit and zero restrictions may be replaced by other restrictions. The known parameter values and other restrictions are substituted in the likelihood function (2.2) and the prior density to be specified below.

The remaining parameters of Γ and B are unrestricted. We shall distinguish between two types: the constant terms (the first row of B, to be denoted β_0') and the economically interesting parameters, to be arranged in a vector θ. So we can write $\Gamma = \Gamma(\theta)$ and $B = B(\beta_0, \theta)$.

ASSUMPTION 6. *The stochastic prior information available can be described by the prior density*

$$p(\theta, \beta_0, H) \propto p(\theta)p^*(H)$$

[1] We follow the classical approach to identification.

$p^*(H)$ is either Wishart or has the limiting form

$$p^*(H) \propto |H|^{-\frac{1}{2}(G+1)}$$

No specific restrictions on $p(\theta)$ are introduced.

It follows from Assumption 6 that our prior information on the constant terms is (locally) uniform. With respect to H we shall confine ourselves in the remainder of this paper to the limiting form which is described by Zellner [19, p. 225 and 226], but we want to emphasize that generalization of our results to the case of a Wishart prior is straightforward. With respect to the economically interesting parameters θ, we make no specific restrictions in order to retain a maximum amount of flexibility. Examples of such priors are given in [12, Section 3]. The reason for the more restrictive assumptions on the priors of the constant terms and the precision matrix is that they allow us to do part of the integration analytically and, hence, save a substantial amount on numerical work. Summarizing our discussion we shall make use of the prior density

$$p(\theta, \beta_0, H) \propto p(\theta)|H|^{-\frac{1}{2}(G+1)} \tag{2.9}$$

How restrictive this set of assumptions is, remains to be investigated; compare also Rothenberg [17, pp. 419 and 420].

Combining the prior density (2.9) and the likelihood (2.2) one obtains, according to Bayes theorem, the joint posterior density

$$p(\theta, \beta_0, H; Y, X) \propto p(\theta)|H|^{\frac{1}{2}(n-G-1)}|\Gamma|^n \exp[-\tfrac{1}{2}tr(QH)] \tag{2.10}$$

This density is our point of departure for a Bayesian inference of the simultaneous equation model.

2.2. A Scheme of Integration Steps

In subsection 2.1 we started with the formulation of a Bayesian statistical model for the structural form of a linear system of simultaneous equations. In subsection 2.3 marginal posterior distributions of subsets of the parameterset (θ, β_0, H) are derived. We will use a two step integration procedure for this purpose. It may be useful to explain the sequence of operations, since we will apply these integration steps extensively.

We will always use the following order in the integration procedure. Firstly, we integrate analytically with respect to the inverse of the covariance matrix of the disturbances, H. Here we make use of the properties of the Wishart density function; see Anderson [1, Ch. 7] and Zellner [19, Appendix B]. In the second step we shall meet a conditional multivariate Student-t distribution of the vector of constant terms. This distribution is conditional on the remaining set of structural parameters.[1] We can make use of well-known properties of this distribution; see Raiffa and Schlaiffer [14, Ch. 8.3]. As a point of reference for reading the following subsections a scheme of the integration steps is presented in figure 1.

The starting point is the joint posterior of (θ, β_0, H), specified in equation (2.10). Then the Wishart step and the Student-t step are shown (together with their equation numbers). Next it is indicated that Monte Carlo is used in order to compute posterior moments of the remaining (unknown) structural and reduced form parameters, see subsection 2.4.

Posterior moments of the constant terms can be evaluated using Monte Carlo results and results of the analytical integration steps.[2] This is explained in 2.5.

2.3. Marginal Posterior Densities on the Structural Parameter Space

The specification of the prior information on H and β_0 enables us to integrate these analytically out of the joint posterior (2.10). The regions of integration for H, β_0 and θ are denoted by

$$A_1 = \{H \mid H \text{ is a G-dimensional positive definite symmetric matrix}\}$$

$$A_2 = \mathbb{R}^G$$

[1] One may interchange the order of integration operations and firstly integrate with respect to the vector of constant terms, by making use of properties of the multivariate normal distribution. After this, one integrates with respect to H. The application of this order in subsection 2.5 requires the evaluation of moments which have been investigated by Kaufman [11].

[2] Posterior first order moments of the variance covariance matrix of disturbances, H^{-1}, can be evaluated using Monte Carlo results and results from Kaufman [11].

FIGURE 1. SCHEME OF INTEGRATION STEPS

Posterior $p(\theta, \beta_0, H; Y, X)$; eq.(2.10)

↓

| Wishart step on H |

↓

$p(\theta, \beta_0; Y, X)$; eq.(2.12)

↓

| Completing squares and Student-t step on β_0 |

↓

$p(\theta; Y, X)$; eq.(2.21)

↓

| Monte Carlo on θ |

↓

$E(\theta; Y, X)$; eq.(2.24)

and

$E[\Pi_1(\theta); Y, X]$; eq.(2.30)

and A_3, the set of all θ vectors. The restrictions on A_3 may vary in different situations and will be discussed in subsection 2.4.

We start with the Wishart step on H and rewrite (2.10) as

(2.11) $$p(\theta, \beta_0, H; Y, X) \propto p(\theta)|\Gamma|^{n}|Q|^{-\frac{1}{2}n} \times |H|^{\frac{1}{2}(n-G-1)}|Q|^{\frac{1}{2}n} \exp[-\tfrac{1}{2}\mathrm{tr}(QH)]$$

Now it is easily seen that the second line in (2.11) equals a kernel of the Wishart density $W(H; Q^{-1}, n, G)$ under the conditions $n \geq G$ and $|Q| > 0$; compare Zellner, loc.cit.. So we have

(2.12) $$p(\theta, \beta_0; Y, X) = \int_{A_1} p(\theta, \beta_0, H; Y, X)dH \propto p(\theta)|\Gamma|^{n}|Q|^{-\frac{1}{2}n}$$

The density in (2.12) is, except for $p(\theta)$, of the generalized Student-t type; see Dickey [4] and Zellner [19, Appendix B5].

As an introduction to the second step, notice that β_0 only appears in the expression Q, given in equation (2.8). Consider then the following partitioning

$$(2.13)\qquad Q = \Gamma'W\Gamma + \begin{bmatrix} \beta_0' - \hat{\beta}_0' \\ B_1 - \hat{B}_1 \end{bmatrix}' \begin{bmatrix} n & \iota'X_1 \\ X_1'\iota & X_1'X_1 \end{bmatrix} \begin{bmatrix} \beta_0' - \hat{\beta}_0' \\ B_1 - \hat{B}_1 \end{bmatrix}$$

where $\beta_0' = (\beta_{01}, \beta_{02}, \ldots, \beta_{0G})$ is the first row vector of the matrix B and ι is a column vector of unit elements. We complete the square on β_0 as follows

$$(2.14)\qquad Q = \Gamma'W\Gamma + (\beta_0 - \hat{\beta}_0)n(\beta_0 - \hat{\beta}_0)'$$

$$+ (\beta_0 - \hat{\beta}_0)\iota'X_1(B_1 - \hat{B}_1) + (B_1 - \hat{B}_1)'X_1'\iota(\beta_0 - \hat{\beta}_0)'$$

$$+ (B_1 - \hat{B}_1)'X_1'X_1(B_1 - \hat{B}_1)$$

$$= \Gamma'W\Gamma + (B_1 - \hat{B}_1)'X_1'NX_1(B_1 - \hat{B}_1)$$

$$+ n[\beta_0' - \hat{\beta}_0' + \frac{1}{n}\iota'X_1(B_1 - \hat{B}_1)]'[\beta_0' - \hat{\beta}_0' + \frac{1}{n}\iota'X_1(B_1 - \hat{B}_1)]$$

where $N = I - (1/n)\iota\iota'$. Define

$$(2.15)\qquad Q_1 = \Gamma'W\Gamma + (B_1 - \hat{B}_1)'X_1'NX_1(B_1 - \hat{B}_1)$$

$$(2.16)\qquad t' = \beta_0' - [\hat{\beta}_0' - \frac{1}{n}\iota'X_1(B_1 - \hat{B}_1)]$$

so that (2.14)-(2.16) imply

$$(2.17)\qquad |Q| = |Q_1 + ntt'|$$

Now we make use of

$$(2.18)\qquad |Q_1 + ntt'| = |Q_1|(1 + nt'Q_1^{-1}t)$$

For a proof, see Dhrymes [3, Appendix A6]. Using (2.17) and (2.18), we rewrite (2.12) as

$$(2.19)\qquad p(\theta, \beta_0; Y, X) \propto p(\theta)|\Gamma|^n|Q_1|^{-\frac{1}{2}n}(1 + nt'Q_1^{-1}t)^{-\frac{1}{2}n}$$

Then we integrate with respect to β_0 as follows

$$(2.20) \qquad p(\theta;\, Y,\, X) = \int_{A_2} p(\theta,\, \beta_0;\, Y,\, X)d\beta_0$$
$$\propto p(\theta)|\Gamma|^n |Q_1|^{-\frac{1}{2}(n-1)}$$
$$\times \int_{A_2} |Q_1^{-1}|^{\frac{1}{2}}(1 + nt'Q_1^{-1}t)^{-\frac{1}{2}n} d\beta_0$$

The integrand in the last line is a kernel of the multivariate Student-t function $p(\beta_0';\, \hat{\beta}_0' - \frac{1}{n}\,\iota'X_1(B_1 - \hat{B}_1),\, nQ_1^{-1},\, n - G)$ under the conditions that Q_1 is a positive definite symmetric matrix and $n \geq G$; see Raiffa and Schlaifer [14, p. 256-259]. It follows that the integral is a constant, independent of Q_1 and that

$$(2.21) \qquad p(\theta;\, Y,\, X) \propto p(\theta)|\Gamma|^n |Q_1|^{-\frac{1}{2}(n-1)}$$

So we obtained the posterior density of $(\theta,\, \beta_0)$, marginal with respect to H in equation (2.12) and the posterior of θ, marginal with respect to $(H,\, \beta_0)$ in (2.21). Notice that when $p(\theta) \propto$ const., equation (2.12) becomes equivalent to a concentrated likelihood function of $(\theta,\, \beta_0)$, as defined by Hood and Koopmans [10, p. 191]. However, equation (2.21) differs from the concentrated likelihood function of θ. This difference is the increase in the exponent of Q_1 by one half.

2.4. Posterior Moments of θ and $\Pi_1(\theta)$

We are interested in the posterior moments of the structural and reduced form parameters. For many economically interesting priors, $p(\theta)$, these have to be computed using numerical integration methods. It is convenient, for the computations based on Monte Carlo principles, to use (2.15) and rewrite (2.21) as

$$(2.22) \qquad p(\theta;\, Y,\, X) \propto p(\theta)\kappa(\theta;\, Y,\, X)$$

where

$$(2.23) \qquad \kappa(\theta;\, Y,\, X) = |\Gamma|^n |Q_1|^{-\frac{1}{2}(n-1)}$$

The first order posterior moments of the structural parameters θ are defined as

$$(2.24) \qquad E(\theta; Y, X) = \frac{\int_{A_3} \theta\kappa(\theta; Y, X)p(\theta)d\theta}{\int_{A_3} \kappa(\theta; Y, X)p(\theta)d\theta}$$

if these integrals exist. Second order moments of θ are defined analogously to (2.24).

In order to define posterior moments of the reduced form parameters, without the constant terms, in case a prior is specified on the structural parameter space, we proceed by considering the reduced form of (2.1),

$$(2.25) \qquad Y = X\Pi + V$$

with

$$(2.26) \qquad \Pi = -B\Gamma^{-1}$$

and

$$(2.27) \qquad V = U\Gamma^{-1}$$

and partition (2.25) as

$$(2.28) \qquad Y = [\iota \;\vdots\; X_1] \begin{bmatrix} \pi_0' \\ \cdot\cdot \\ \Pi_1 \end{bmatrix} + V$$

where π_0' is the first row vector of Π containing the constant terms. From (2.26) and (2.28) we can write

$$(2.29) \qquad \Pi_1(\theta) = -B_1(\theta)[\Gamma(\theta)]^{-1}$$

which shows that the elements of Π_1 are functions of θ. Hence, under the condition that the integrals mentioned below exist, the posterior first order moments of the reduced form parameters Π_1 read

$$(2.30) \qquad E[\Pi_1(\theta)] = \frac{\int_{A_3} \Pi_1(\theta)\kappa(\theta; Y, X)p(\theta)d\theta}{\int_{A_3} \kappa(\theta; Y, X)p(\theta)d\theta}$$

Second order moments are defined analogously. By considering the integration over θ, we avoid difficult transformation problems involved in going from a distribution on $(\Gamma(\theta), B(\beta_0, \theta), H)$ to a distribution on the reduced form parameter set $(\Pi, \Gamma H \Gamma')$.

The integrals (2.24) and (2.30) exist if the integrands are bounded and the region of integration is bounded. This imposes restrictions on $\kappa(\theta; Y, X)$ and $p(\theta)$. Some simple sufficient conditions for the existence of the integrals are discussed in Kloek and Van Dijk [12]. It is argued there, that κ is bounded if W is positive definite, since we assumed that $|\Gamma| \neq 0$; see Assumption 1. The bounded region condition can be fulfilled by choosing truncated prior distributions. It is a topic of further research to investigate the restrictiveness of these conditions.

2.5. Posterior Moments of π_0 and β_0

For prediction purposes, one may be interested in the expected future value of a row of Y and its covariance matrix. Then it is necessary to compute the posterior first order moments of the constant terms of the reduced form parameters and their covariance matrix. Consider therefore

$$(2.31) \qquad \pi_0'(\beta_0, \theta) = -\beta_0'[\Gamma^{-1}(\theta)]$$

and the definition of the posterior expected value of π_0' is given by

$$(2.32) \qquad E(\pi_0') = \frac{\int_{A_3}\int_{A_2}\int_{A_1} -\beta_0'\Gamma^{-1} p(\theta, \beta_0, H; Y, X) dH d\beta_1 d\theta}{\int_{A_3}\int_{A_2}\int_{A_1} p(\theta, \beta_0, H; Y, X) dH d\beta_0 d\theta}$$

where $p(\theta, \beta_0, H; Y, X)$ has been specified in (2.10).

We tackle the integration of (2.32) by analyzing the numerator. One proceeds in the same way as in subsection 2.3. Firstly, the integration with respect to H is performed. This is the Wishart step; compare figure 1. In the second step one has to evaluate the first order moments of the multivariate Student-t function $p(\beta_0'; \hat{\beta}_0' - \frac{1}{n}\iota' X_1(B_1 - \hat{B}_1), nQ_1^{-1}, n - G)$. That is to say, one can write after the Wishart step that the numerator of (2.32) is proportional to

$$(2.33) \qquad \int_{A_3}\int_{A_2} -\beta_0'\Gamma^{-1} p(\theta, \beta_0; Y, X) d\theta d\beta_0$$

$$\propto -\int_{A_3}\left[\int_{A_2} \beta_0' |Q_1^{-1}|^{\frac{1}{2}} (1 + nt'Q_1^{-1}t)^{-\frac{1}{2}n} d\beta_0\right]$$

$$\times\ \Gamma^{-1}p(\theta)|\Gamma|^{n}|Q_1|^{-\frac{1}{2}(n-1)}d\theta$$

$$\propto -\int_{A_3}[\hat{\beta}_0' - \frac{1}{n}\iota'X_1(B_1 - \hat{B}_1)]\Gamma^{-1}p(\theta)|\Gamma|^{n}|Q_1|^{-\frac{1}{2}(n-1)}d\theta$$

$$= \int_{A_3}[\hat{\pi}_0' - \frac{1}{n}\iota'X_1(\Pi_1 - \hat{\Pi}_1)]p(\theta)|\Gamma|^{n}|Q_1|^{-\frac{1}{2}(n-1)}d\theta$$

The second proportionality sign in (2.33) indicates that use is made of the first order moments of the multivariate Student-t function of β_0, under the condition $n - G > 1$.

Now it follows from (2.6a) that

(2.34)
$$\iota'\left[Y - (\iota \vdots X_1)\begin{pmatrix}\hat{\pi}_0' \\ \vdots \\ \hat{\Pi}_1\end{pmatrix}\right]$$

$$= \iota'Y - n\hat{\pi}_0' - \iota'X_1\hat{\Pi}_1 = 0$$

Substitution of this result in (2.33) yields the numerator of $E(\pi_0')$, which is proportional to,

(2.35)
$$\int_{A_3}(\frac{1}{n}\iota'Y - \frac{1}{n}\iota'X_1\Pi_1)p(\theta)|\Gamma|^{n}|Q_1|^{-\frac{1}{2}(n-1)}d\theta$$

The same operations, which were used in the numerator, are applied to the denominator of (2.32), where the omitted multiplicative constants are the same. Then one obtains

(2.36)
$$E[\pi_0'(\theta)] = \frac{1}{n}\iota'Y - \frac{1}{n}\iota'X_1E[\Pi_1(\theta)]$$

Similarly

(2.37)
$$E[\beta_0'(\theta)] = -\frac{1}{n}\iota'YE[\Gamma(\theta)] + \frac{1}{n}\iota'X_1E[B_1(\theta)]$$

So we conclude that, once $E(\theta)$ and $E[\Pi_1(\theta)]$ are computed by numerical integration methods, the posterior expected values of π_0 and β_0 can be computed in a simple manner.

The computation of the covariance matrix of π_0 can be analyzed as follows

(2.38) $$\operatorname{cov}(\pi_0) = E(\pi_0\pi_0') - E(\pi_0)E(\pi_0')$$

and the evaluation of the posterior second order moments of π_0 proceeds along the same lines as in the case of $E(\pi_0)$. We start with the definition

(2.39) $$E(\pi_0\pi_0') = \frac{\int_{A_3}\int_{A_2}\int_{A_1} (\Gamma^{-1})'\beta_0\beta_0'\Gamma^{-1}p(\theta,\, \beta_0,\, H;\, Y,\, X)dHd\beta_0 d\theta}{\int_{A_3}\int_{A_2}\int_{A_1} p(\theta,\, \beta_0,\, H;\, Y,\, X)dHd\beta_0 d\theta}$$

where $p(\theta, \beta_0, H; Y, X)$ is given in equation (2.10). Again we consider the numerator. The integration with respect to H is performed as before and in the second step one has to evaluate

(2.40) $$\int_{A_3} (\Gamma^{-1})'[\int_{A_2} \beta_0\beta_0'|Q_1^{-1}|^{\frac{1}{2}}(1 + nt'Q_1^{-1}t)^{-\frac{1}{2}n}d\beta_0]$$
$$\times \Gamma^{-1}p(\theta)|\Gamma|^n|Q_1|^{-\frac{1}{2}(n-1)}d\theta$$

The expression within brackets is proportional to the second order moments of the multivariate Student-t distribution of β_0, defined in subsection 2.3 (under the condition $n - G > 2$). Then we can obtain

(2.41) $$\int_{A_3} (\Gamma^{-1})'\{\frac{n-G}{n(n-G-2)} Q_1 + [\hat{\beta}_0' - \frac{1}{n} \iota'X_1(B_1 - \hat{B}_1)]'$$
$$\times [\hat{\beta}_0' - \frac{1}{n} \iota'X_1(B_1 - \hat{B}_1)]\}\Gamma^{-1}p(\theta)|\Gamma|^n|Q_1|^{-\frac{1}{2}(n-1)}d\theta$$

As an intermediate step, use equations (2.4), (2.15), and (2.29) so that

(2.42) $$(\Gamma^{-1})'Q_1\Gamma^{-1} = (\Gamma^{-1})'[\Gamma'W\Gamma + (B_1 - \hat{B}_1)'X_1'NX_1(B_1 - \hat{B}_1)]\Gamma^{-1}$$
$$= W + (\Pi_1 - \hat{\Pi}_1)'X_1'NX_1(\Pi_1 - \hat{\Pi}_1)$$

When (2.42) is substituted back into (2.41) and when the same operations are applied to the denominator of (2.39) as to the numerator[1], we have

(2.43) $$E(\pi_0\pi_0') = \frac{n - G}{n(n - G - 2)} W$$
$$+ \frac{n - G}{n(n - G - 2)} E[(\Pi_1 - \hat{\Pi}_1)'X_1'NX_1(\Pi_1 - \hat{\Pi}_1)]$$

---

[1] The omitted numerical constants for numerator and denominator are the same.

$$+ E[(\frac{1}{n} \iota'Y - \frac{1}{n} \iota'X_1\Pi_1)'(\frac{1}{n} \iota'Y - \frac{1}{n} \iota'X_1\Pi_1)]$$

where the last term is obtained by making use of (2.4) and (2.34) in a similar way as in the evaluation of $E(\pi_0)$. Notice that the expectations in (2.43) are taken with respect to θ.

The covariance matrix of the constant terms of the reduced form parameters reads then

$$(2.44) \quad cov(\pi_0) = \frac{n - G}{n(n - G - 2)} W + \frac{n - G}{n(n - G - 2)} E[(\Pi_1 - \hat{\Pi}_1)'X_1'NX_1(\Pi_1 - \hat{\Pi}_1)]$$

$$+ E(\frac{1}{n} \iota'Y - \frac{1}{n} \iota'X_1\Pi_1)'(\frac{1}{n} \iota'Y - \frac{1}{n} \iota'X_1\Pi_1)$$

$$- [\frac{1}{n} \iota'Y - \frac{1}{n} \iota'X_1E(\Pi_1)]'[\frac{1}{n} \iota'Y - \frac{1}{n} \iota'X_1E(\Pi_1)]$$

$$= \frac{n - G}{n(n - G - 2)} W + \frac{n - G}{n(n - G - 2)} E[(\Pi_1 - \hat{\Pi}_1)'X_1'NX_1(\Pi_1 - \hat{\Pi}_1)]$$

$$+ \frac{1}{n} [E(\Pi_1'X_1' \frac{\iota\iota'}{n} X_1\Pi_1) - E(\Pi_1')X_1' \frac{\iota\iota'}{n} X_1E(\Pi_1)]$$

where all the expectations are taken with respect to θ. The expression (2.44) can be written in terms of variances, covariances and first moments. We will not do so, but concentrate on the evaluation of expressions like

$$(2.45) \qquad R = E[P'(\theta)MQ(\theta)]$$

where the expectation is taken with respect to θ. Let the positive semi-definite matrix M be m×m. Let P be an m×r matrix and Q an m×s matrix of random elements, which are functions of a random vector θ. Let M be independent of P and Q. Notice that in our case M stands for $X_1'NX_1$ and $X_1'\iota\iota'X_1/n$ and $P = Q = \Pi_1$. A typical element of R is[1]

$$(2.46) \qquad r_{ij} = E(\sum_{k=1}^{m} \sum_{h=1}^{m} p_{ki}m_{kh}q_{hj}) \qquad \begin{matrix} i = 1, \ldots, r \\ j = 1, \ldots, s \end{matrix}$$

$$= \sum_{k=1}^{m} \sum_{h=1}^{m} m_{kh}[cov(p_{ki}, q_{hj}) + E(p_{ki})E(q_{hj})]$$

So the matrix R is computed by simple (albeit tedious) calculations once

[1] Or $r_{ij} = tr[M\ E(q_jp_i')]$, where p_i is the i-th column of P and q_j is the j-th column of Q.

the covariance matrix of P and Q and the expectations of P and Q are known. A similar statement holds for the covariance matrix of π_0.

For prediction purposes, we finally want to know the covariance matrix of the constant terms, π_0, and the other reduced form parameters Π_1. We will discuss the evaluation of these cross moments for a particular column of Π_1, say π_{i}. Obviously,

$$\text{(2.47)} \qquad \operatorname{cov}(\pi_i, \pi_0') = E(\pi_i \pi_0') - E(\pi_i)E(\pi_0')$$

The evaluation of $E(\pi_i \pi_0')$ is done in the same way as the evaluation of $E(\pi_0)$, except that we carry along now the vector of π_i. Then one obtains, proceeding along the lines indicated in equations (2.32)-(2.36),

$$\text{(2.48)} \qquad E(\pi_i \pi_0') = E(\pi_i)\frac{1}{n}\,\iota'Y - E(\pi_i \frac{1}{n}\,\iota'X_1\Pi_1)$$

Using (2.48) and (2.36) one obtains for the cross moments of π_i and π_0

$$\text{(2.49)} \qquad \begin{aligned} \operatorname{cov}(\pi_i, \pi_0') &= E(\pi_i)\frac{1}{n}\,\iota'Y - E(\pi_i \frac{1}{n}\,\iota'X_1\Pi_1) \\ &\quad - E(\pi_i)\frac{1}{n}\,\iota'Y + E(\pi_i)\frac{1}{n}\,\iota'X_1E(\Pi_1) \\ &= E(\pi_i)\frac{1}{n}\,\iota'X_1E(\Pi_1) - E(\pi_i \frac{1}{n}\,\iota'X_1\Pi_1) \end{aligned}$$

The evaluation of the last term of (2.49) is done analogously to (2.45), except that instead of a matrix M, one has a vector $m_1 = \frac{1}{n}\,\iota'X_1$.

3. MULTIPERIOD PREDICTIVE MOMENTS WITH STRUCTURAL INFORMATION

Assume that values of the jointly dependent variables will be generated for the period n+r by the same stochastic economic system as described in section 2. So we have

$$\text{(3.1)} \qquad y_{n+r}'\Gamma + x_{n+r}'B = u_{n+r}'$$

Extending Assumptions (2)-(3) and (5) from section 2, we can formulate the predictive density of y_{n+r} conditional upon (x_{n+r}, Γ, B, H) as[1]

[1] Remember that $|\Gamma|$ stands for mod $|\Gamma|$ and that $\Gamma = \Gamma(\theta)$ and $B = B(\beta_0, \theta\ )$.

$$(3.2)\qquad p(y_{n+r};\, x_{n+r},\, \Gamma,\, B,\, H) \propto |H|^{\frac{1}{2}}|\Gamma|$$

$$\times \exp\{-\tfrac{1}{2}\mathrm{tr}[(y'_{n+r}\Gamma + x'_{n+r}B)'(y'_{n+r}\Gamma + x'_{n+r}B)H]\}$$

We are interested in predictive moments of y_{n+r}, incorporating prior information on the structural parameters, i.e., we want to consider predictive moments, marginal with respect to (θ, β_0, H) but conditional on the sample information (Y, X) and the future values of the (truly) exogenous variables.

We start with considering the joint density of (y_{n+r}, θ, β_0, H) conditional upon (Y, X, x_{n+r}) as the product of the predictive density (3.2) and the posterior density of (θ, β_0, H); see equation (2.10). So one can write

$$(3.3)\qquad p(y_{n+r},\, \theta,\, \beta_0,\, H;\, Y,\, X,\, x_{n+r}) \propto$$

$$p(\theta)|\Gamma|^{n+1}|H|^{\frac{1}{2}(n-G)}\exp[-\tfrac{1}{2}\mathrm{tr}(Q^*H)]$$

where

$$(3.4)\qquad Q^* = (Y\Gamma + XB)'(Y\Gamma + XB) + (y'_{n+r}\Gamma + x'_{n+r}B)'(y'_{n+r}\Gamma + x'_{n+r}B)$$

In order to derive predictive moments we will again use a two step integration procedure. The first step is the Wishart step with respect to H, i.e., we make use of the properties of the Wishart density function $W(H;\, Q^{*-1},\, G,\, n)$. Note that the conditions $|Q^*| > 0$ and $n \geq G$ are fulfilled since the matrix Q given in equation (2.8) is required to be positive definite and we work already under the condition $n \geq G$. Then one obtains

$$(3.5)\qquad p(y_{n+r},\, \theta,\, \beta_0;\, Y,\, X,\, x_{n+r}) \propto p(\theta)|\Gamma|^{n+1}|Q^*|^{-\frac{1}{2}(n+1)}$$

which is written as

$$(3.6)\qquad p(y_{n+r},\, \theta,\, \beta_0;\, Y,\, X,\, x_{n+r}) \propto p(\theta)|A|^{-\frac{1}{2}(n+1)}$$

where

$$(3.7)\quad |A|^{-\frac{1}{2}(n+1)} = |\Gamma|^{n+1}|Q^*|^{-\frac{1}{2}(n+1)}$$

$$= |(Y - X\Pi)'(Y - X\Pi) + (y'_{n+r} - x'_{n+r}\Pi)'(y'_{n+r} - x'_{n+r}\Pi)|^{-\frac{1}{2}(n+1)}$$

using $\Pi = - B\Gamma^{-1}$ and (3.4).

Define now

$$(3.8)\quad A_1 = (Y - X\Pi)'(Y - X\Pi)$$

$$t^{*\prime} = y'_{n+r} - x'_{n+r}\Pi$$

and make use of[1]

$$(3.9)\quad |A_1 + t^*t^{*\prime}| = |A_1||1 + t^{*\prime}A_1^{-1}t^*|$$

Then we reformulate (3.6) as

$$(3.10)\quad p(y_{n+r}, \theta, \beta_0; Y, X, x_{n+r}) \propto p(\theta)$$

$$\times |A_1|^{-\frac{1}{2}(n+1)}|1 + t^{*\prime}A_1 t^*|^{-\frac{1}{2}(n+1)}$$

Conditional upon $(\theta, \beta_0, Y, X, x_{n+r})$ we have in (3.10) a kernel of the multivariate Student-t function of y_{n+r}, $p(y'_{n+r}; x'_{n+r}\Pi, A_1^{-1}, n - G + 1)$. Therefore we have

$$(3.11)\quad E_{y_{n+r}}(y_{n+r}; \theta, \beta_0, Y, X, x_{n+r}) = x'_{n+r}\Pi$$

under the condition $n > G$, where the subscript y_{n+r} means that the expectation is taken with respect to y_{n+r}. Furthermore, under the condition $n > G + 1$, the covariance matrix of y_{n+r}, conditional upon $(\theta, \beta_0, Y, X, x_{n+r})$ is

$$(3.12)\quad \text{cov}(y_{n+r}; \theta, \beta_0, Y, X, x_{n+r}) = \frac{n - G + 1}{n - G - 1} A_1$$

We are interested in moments of y_{n+r}, marginal with respect to (θ, β_0)

[1] Notice that $A_1 = \Gamma^{-1\prime}Q\Gamma^{-1}$, where Q is given in equation (2.8) and that equation (3.9) corresponds to equation (2.18).

In case of the expected value of y_{n+r}, marginal with respect to (θ, β_0), we have

$$(3.13) \qquad E(y_{n+r}; Y, X, x_{n+r}) = E_{\theta,\beta_0}[E_{y_{n+r}}; \theta, \beta_0, Y, X, x_{n+r})]$$

$$= x'_{n+r}E_{\theta,\beta_0}(\Pi)$$

In case of the covariance matrix of y_{n+r}, marginal with respect to (θ, β_0), we proceed as follows

$$(3.14) \qquad \text{cov}(y_{n+r}; Y, X, x_{n+r}) = E_{\theta,\beta_0}[E_{y_{n+r}}(y_{n+r}y'_{n+r}; \theta, \beta_0, Y, X, x_{n+r})]$$

$$- [x'_{n+r}E_{\theta,\beta_0}(\Pi)]'[x'_{n+r}E_{\theta,\beta_0}(\Pi)]$$

Now the first integral at the right hand side of (3.14) equals

$$(3.15) \qquad E_{\theta,\beta_0}\left[\frac{n - G + 1}{n - G - 1} A_1 + \Pi' x_{n+r}x'_{n+r}\Pi\right]$$

$$= \frac{n - G + 1}{n - G - 1} E_{\theta,\beta_0}[(Y - X\Pi)'(Y - X\Pi)] + E_{\theta,\beta_0}(\Pi' x_{n+r}x'_{n+r}\Pi)$$

In the last line above we meet expressions which are computed in a similar way as described at the end of subsection 2.5 equation (2.45).

If a simultaneous equation system does not contain any lagged endogenous variables, formulae (3.13) and (3.15) suffice for the expected value and covariance matrix of the jointly dependent variables for any period $r = 1, 2, \ldots$. Note that for any period beyond the period n+1, say n+s, the expected value and the covariance matrix of y_{n+s} differ only in the term x_{n+s}.

In case lagged endogenous variables are present, one may proceed as follows[1]; compare Chow [2]. Consider the partitioning

$$(3.16) \qquad x'_{n+r}\Pi = (y'_{n+r-1} \;\vdots\; z'_{n+r})\begin{bmatrix} C(\theta) \\ \hdashline D(\beta_0, \theta) \end{bmatrix}$$

where y'_{n+r-1} is a k_1-vector of lagged endogenous variables and z'_{n+r} a $(K - k_1)$-vector of exogenous variables. The matrices C and D are appropriate submatrices of Π. Applying a well-known substitution procedure

[1] We treat the case of one period lagged endogenous variables. The generalization to more periods is straightforward, see Chow [2].

for y_{n+r-1}, y_{n+r-2}, ..., y_{n+1} in (3.16), one can obtain

(3.17) $$E(y'_{n+r};\ Y, X, z_{n+1}, \ldots, z_{n+r}) = y'_n E(C^r) + z'_{n+1}E(DC^{r-1}) + z'_{n+2}E(DC^{r-2}) + \ldots + z'_{n+r}E(D)$$

where the expectation at the left hand side of the equality sign is taken with respect to y_{n+r}, y_{n+r-1}, ..., y_{n+1} (or, equivalently, u_{n+r}, ..., u_{n+1}), and the expectations at the right hand side are taken with respect to β_0 and θ.

So, for $r = 1$, the first order moments and the covariance matrix of the jointly dependent variables, marginal with respect to (θ, β_0), can be determined once the posterior first and second order moments of Π are known (equations (3.13) and (3.15) directly apply). Furthermore when $r = 2$, the expected value of y'_{n+r}, marginal with respect to (θ, β_0), can be determined once the second order posterior moments of Π are known. Notice that a covariance matrix of y_{n+2} requires the evaluation of fourth order posterior moments of Π. For $r = 3$, one would need to evaluate sixth order moments, and so on.

4. FINAL REMARKS

We obtained small sample results for first and second order multi period predictive moments, using prior information on the structural parameters. The wide class of prior densities allowed makes numerical integration on a subset of parameters necessary, for which we advocate Monte Carlo methods [12].

The classical approach to prediction in simultaneous equation models is limited to asymptotic results; compare Goldberger, Nagar and Odeh [7]. Rothenberg's [16] Bayesian approach is asymptotic as well.

It is noteworthy that we did not make use of a criterion function, such as a quadratic loss function, to derive multi period predictive moments, compare Chow's [2] approach in this respect.

The feasibility of our approach to predictions a long period ahead, say ten years or more, remains doubtful, especially the computation of estimates of second order moments. But medium term predictions, say two or three years ahead, appear worth the (computational) effort. In addition, it is possible to estimate total multipliers with our approach if one wants to obtain an idea about the long run implications of a dynamic econometric model; see [12].

It appears interesting to compare our approach of allowing richness of prior information with the non-informative approach of Zellner [19, Ch. 9]; with the natural conjugate approach of Harkema [9]; with the extended natural conjugate of Drèze [5], Drèze and Morales [6], Morales [13] and Richard [15].

Further research obviously consists of numerical experiments with the algebraic results presented here, for instance, with Klein's model I.

REFERENCES

[1] Anderson, T.W., An Introduction to Multivariate Statistical Analysis, New York: Wiley, 1958.

[2] Chow, G.C., "Multiperiod Predictions from Stochastic Difference Equations by Bayesian Methods", Ch. 8.1 of Fienberg S.E. and A. Zellner (eds.), Studies in Bayesian Econometrics and Statistics, Amsterdam: North Holland, 1975.

[3] Dhrymes, P.J., Econometrics: Statistical Foundations and Applications, New York: Harperand Row, 1970.

[4] Dickey, J.M., "Matricvariate Generalizations of the Multivariate t Distribution and the Inverted Multivariate t Distribution", Annals of Mathematical Statistics, 38 (1967), 511-519.

[5] Drèze, J.H., "Limited Information Estimation from a Bayesian Viewpoint", Econometrica, 44 (1976), 1045-1075.

[6] Drèze, J.H., and J.A. Morales, "Bayesian Full Information Analysis of the Simultaneous Equations Model", CORE Discussion Paper 7031, Louvain, 1970.

[7] Goldberger, A.S., Nagar, A.L., and H.S. Odeh, "The Covariance Matrices of Reduced-Form Coefficients and of Forecasts for a Structural Econometric Model", Econometrica, 29 (1961), 556-573.

[8] Haber, S., "Numerical Evaluation of Multiple Integrals", SIAM Review, 12 (1970), 481-526.

[9] Harkema, R., Simultaneous Equations, A Bayesian Approach, Rotterdam: University Press, 1971.

[10] Hood, W.C. and T.C. Koopmans, eds., Studies in Econometric Method, New York: Wiley, 1953.

[11] Kaufman, G.M., "Some Bayesian Moment Formulae", CORE Discussion Paper 6710, Louvain, 1967.

[12] Kloek, T., and H.K. van Dijk, "Bayesian Estimates of Equation System Parameters, An Application of Integration by Monte Carlo", Report 7511/E of the Econometric Institute, Rotterdam (revised).

[13] Morales, J.A., Bayesian Full Information Structural Analysis, Berlin: Springer Verlag, 1971.

[14] Raiffa, H., and R. Schlaiffer, Applied Statistical Decision Theory, Boston: Graduate School of Business Administration, Harvard University 1961.

[15] Richard, J.F., Posterior and Predictive Densities for Simultaneous Equation Models, Berlin: Springer Verlag, 1973.

[16] Rothenberg, T.J., Efficient Estimation with A Priori Information, New Haven and London: Yale University Press, 1973.

[17] Rothenberg, T.J., "Bayesian Analysis of Simultaneous Equations Models", Ch. 9.3 of Fienberg, S.E., and A. Zellner (eds.), Studies in Bayesian Econometrics and Statistics, Amsterdam: North Holland, 1975.

[18] Theil, H., Principles of Econometrics, New York: John Wiley, 1971.

[19] Zellner, A., An Introduction to Bayesian Inference in Econometrics, New York: John Wiley, 1971.