Tilburg University

A Bayesian analysis of simultaneous equation models by combining recursive analytical and numerical approaches

Steel, M.F.J.

Publication date:
1989

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):

Steel, M. F. J. (1989). *A Bayesian analysis of simultaneous equation models by combining recursive analytical and numerical approaches*. (CentER Discussion Paper; Vol. 1989-8). Unknown Publisher.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Center
for
Economic Research

Discussion paper

500
CBM
R

8414
1989

8

No. 8908

**A BAYESIAN ANALYSIS OF SIMULTANEOUS
EQUATION MODELS BY COMBINING RECURSIVE
ANALYTICAL AND NUMERICAL APPROACHES**

by Mark F.J. Steel

February, 1989

A BAYESIAN ANALYSIS OF SIMULTANEOUS EQUATION MODELS BY
COMBINING RECURSIVE ANALYTICAL AND NUMERICAL APPROACHES

Mark F.J. Steel^{*)}
Tilburg University

JEL code: 211

Keywords: Bayesian econometrics; SEM; hybrid Monte Carlo methods;
recursive analytical techniques; poly-t densities

Abstract:

Inference on simultaneous equation models from a Bayesian approach is still far from a routine procedure: serious numerical problems may occur in most practical situations. In this paper a unifying framework is proposed using recent analytical [see Steel (1988)] and well-documented poly-t based methods in order to reduce the dimension of the final Monte Carlo integration. Different integration sequences are examined and the implications for various model and prior assumptions are tabulated. Some preliminary empirical work suggests that considerable gains can be obtained.

Acknowledgements:

This work has benefited from discussions with Luc Bauwens. The usual disclaimer applies. The author gratefully acknowledges the support of the Royal Netherlands Academy of Arts and Sciences.

^{*)} Department of Econometrics

Tilburg University

P.O. Box 90153

5000 LE Tilburg, The Netherlands

1. Introduction

In spite of the numerous theoretical attractions of Bayesian methods, a stimulating discussion of which can be found in Poirier (1988), their implementation and wide-spread use in applied econometrics has developed at a much slower rate than could initially be expected from the pioneering work of Drèze (1962) and Zellner (1965).

One of the reasons for this lack of empirical applications undoubtedly lies in the rather heavy computational requirements of Bayesian methods in most "serious" applied work. This could be interpreted as the price one has to pay for conducting marginal inferences, rather than conditional ones as in the sampling-theoretical approach.

Straightforward exact analytical results have only been applicable to simple seemingly unrelated regression equation (SURE) models compatible with the extremely restrictive (matrix-form) structure induced by natural conjugate (NC) analyses. Recently, somewhat less straightforward, but, nevertheless, analytical results were obtained for SURE models under a more general type of restrictions, called the recursive limited information (RLI) structure, in Steel (1988). This RLI assumption essentially imposes a block recursive structure on the model, in the sense that, broadly speaking, all variables appearing in previous equations should also enter in the present equation. Analytical results can then be obtained, under suitable stochastic prior assumptions, by using recursive techniques. Heuristically, this may be understood if we consider only the first i equations and realize that equation i will contain all regressors appearing in that block of equations, and can thus be treated as unrestricted with respect to this part of the model.

Although the RLI approach extends the realm of analytical Bayesian inference considerably, it is still not sufficient to cover the full range of so-called simultaneous equation models, which are used frequently in time-series econometrics and which are described in Subsection 2.1.

In order to tackle such models two main lines of research have been developed in the literature.

- (i) Numerical integration by Monte Carlo importance sampling methods was introduced in econometrics by Kloek and van Dijk (1978) and further developed in e.g. van Dijk (1984), Bauwens (1984), Geweke (1987), and Zellner et al. (1988). The dimensionality of the integrals is not as crucial a problem as in the more common methods like Gaussian or Simpson's product rules, but it is still a factor to be reckoned with, certainly if we wish to use a Personal Computer for the calculations.
- (ii) Using the specific properties of so-called poly-t densities as described in Drèze (1977), we can combine analytical and relatively simple numerical techniques in order to obtain posterior results for one equation, which are, however, conditional upon all remaining equations.

It seems only natural to combine different methods, as each method has its own specific merits (and difficulties). Van Dijk (1984) and Naylor and Smith (1988) explore the mixing of product rules and importance sampling, whereas Geweke (1987, 1988) focuses on antithetic acceleration of Monte Carlo procedures. Kloek and van Dijk (1978) mention the combination of analytical and Monte Carlo methods, following the precept found in Hammersley and Handscomb (1979) that one should replace numerical techniques by analytical methods wherever possible.

Much in the same vein, it will be attempted here to find a hybrid strategy that combines

- (i) exact analytical methods based on the RLI approach and exploiting the recursive properties of a subset of the parameters,
- (ii) existing techniques for analysing [see Richard and Tompa (1980)] and drawing [see Bauwens and Richard (1985)] poly-t distributed random variables,
- (iii) Monte Carlo integration by importance sampling of either regression coefficients or the covariance matrix, the relative merits of which have been discussed in Richard and Steel (1988) for models of the SURE type.

Moreover, these methods will be applied in the above order of preference.

Section 2 of the paper sets the stage by introducing the model, a recursive transformation of the latter as this is felt to clarify the analysis, and the prior structure, consisting of both exact prior restrictions and stochastic assumptions on the unrestricted parameters.

The ensuing posterior analysis is considered in Section 3, where the first three subsections imply a Monte Carlo analysis on a subset of the regression coefficients, and Subsection 3.4 is devoted to Monte Carlo integration of the covariance matrix. Finally, Subsection 3.5 groups some results on computational complexity and existence of posterior moments. A fourth and final section ventures some concluding remarks. Although the numerical procedures used in this combined approach have all been subjected to empirical examination in several applications (see the references quoted above), one might be interested in empirical results obtained with the hybrid method advocated here. Section 4 gives some indication of the gain in dimensionality for two well-known models, but, in view of the length of the present paper, actual empirical applications have been deferred to a following one.

2. The Statistical Framework

2.1. The Simultaneous Equation Model

In this paper we shall concentrate upon a widely used class of models in (time-series) econometrics, namely the simultaneous equation model (SEM), as introduced in Haavelmo (1943) and, since then, treated in a classical fashion in most textbooks [see, e.g., Theil (1971) or Spanos (1986)], whereas a Bayesian analysis is provided in the seminal papers by Drèze (1962), Rothenberg (1963), and Zellner (1965), and later in e.g. Drèze and Richard (1983).

A sequential model is assumed where the observations at time t on n endogenous variables in $y_{(t)}$ will be described by a set of n "structural" equations. These equations serve to restrict the conditional expectation of $y_{(t)}$, given the information set at time t , to some linear manifold:

$$B'E(y_{(t)}|x_{(t)}) + \Gamma'x_{(t)} = 0, \quad (2.1)$$

where it is assumed that the information set at time t , containing weakly exogenous variables and lagged values of both endogenous and exogenous variables, counts a finite number m of variables that are relevant for the description of $y_{(t)}$ and are grouped in $x_{(t)}$.

Remark that the parameter matrices in (2.1) B and Γ are respectively of dimensions $n \times n$ and $m \times n$, and will typically be subject to all kinds of exact prior restrictions. This issue will be dealt with later in this subsection. The fact that B is square implies that (2.1) has a unique solution provided $|B| \neq 0$. We shall make this full rank assumption, implying that we are in the context of a "complete" SEM, whereas Richard (1984) rightly argues that, in general, we shall have fewer structural equations than endogenous variables, and shall, therefore, usually be in the realm of "incomplete" SEM. We fully agree with this point of view and reinterpret the n equations in (2.1) as the "completed" set of equations, containing both structural and "reduced" (e.g. instrumental variables) types of equations. However, we shall still allow for general linear restrictions in these completing equations, and we shall, therefore, avoid distinguishing them from the truly structural equations in our notation.

Stochastics are introduced by putting a Gaussian structure on $y_{(t)}|x_{(t)}$ with a constant covariance matrix V , implying that

$$YB + X\Gamma = U \quad (2.2)$$

$$\text{with } D(U) = f_{MN}^{T \times n}(U|0, \Sigma \otimes I_T), \quad (2.3)$$

where

$$\Sigma = B'VB,$$

$$Y = (y_{(1)} \dots y_{(T)})',$$

$$X = (x_{(1)} \dots x_{(T)})'.$$

and U is a $T \times n$ matrix of innovations with respect to the information set at each time period $t : 1 \rightarrow T$.

For the notation and definitions of density functions used here we refer to Appendix A of Drèze and Richard (1983).

In order to avoid conditionalization paradoxes of the type discussed in Kolmogoroff (1946), we wish to treat the exact prior restrictions that pertain to B and Γ in (2.2) explicitly. We assume, in line with most applications of SEM, that the covariance matrix Σ is not subject to restrictions, other than being in the class of positive definite symmetric (PDS) matrices. Let us, for notational convenience, rewrite (2.2) as

$$WA = U, \quad (2.4)$$

where we have implicitly defined

$$W = (Y : X), \quad (2.5)$$

and where

$$A = \begin{bmatrix} B \\ \Gamma \end{bmatrix} \quad (2.6)$$

groups all coefficients in the model. Let us now restrict the $(m+n) \times n$ elements of A in a linear way by considering its column expansion

$$\text{vec } A = \alpha = R\xi, \quad (2.7)$$

where R is a $(m+n)n \times \ell$ selection matrix of known constants and of full column rank, whereas the ℓ dimensional vector ξ groups all unrestricted parameters.

This characterization of the restrictions allows for linear restrictions both within and across equations and seems to imply enough flexibility to be empirically useful. Its general nature and, in particular, its advantages over restrictions expressed in matrix form (i.e. by directly restricting A instead of its column expansion) are discussed in more detail in Richard and Steel (1988) and Steel (1988). This matrix form can, of

course, be written as a special case of (2.7), where R takes a Kronecker product structure, say,

$$R = I_n \otimes F, \quad (2.8)$$

which, indeed, leads to

$$A = FZ, \quad (2.9)$$

provided $\zeta = \text{vec } Z$.

As already explained in the introduction, the recursive analytical method of Steel (1988) will be used to deal with part of the parameters in the model. It was found that inference can be conducted analytically for the coefficients of a SURE model provided it has a recursive limited information (RLI) structure and we possess diffuse prior information on the free parameters. Such an RLI assumption, in fact, says that the space spanned by the variables corresponding to the free parameters in equation i should include all variables used in previous equations.

In practice, this will boil down to an assumption of block recursivity, where all variables in the previous $(i-1)$ equations also appear (with unrestricted coefficients) in the i -th equation.

In order to use these analytical techniques to reduce the dimension of the parameter space left for numerical treatment, we single out those parameters for which these assumptions hold, i.e. they should not intervene in the Jacobian of the system (as in the SURE case) and they should, after possibly rearranging the equations, possess an RLI structure. In addition, their prior density has to belong to a special class as will be discussed later in Subsection 2.3.

This means we shall rewrite the restrictions in (2.7) as

$$\alpha = \alpha_1 + \alpha_2 = R^1 \zeta^1 + R^2 \zeta^2, \quad (2.10)$$

where ζ^2 should be amenable to analytical treatment and both R^1 and R^2 are of full column rank, respectively ℓ_1 and ℓ_2 , the sum of which will equal ℓ . The full column rank of the selection matrices implies, through basic

theorems on triangularization [see, e.g., Strang (1980) p. 97], that they can, without loss of generality, be rewritten in a block lower-triangular form, which induces the following recursive definition:

$$R_{(i)}^j = \begin{bmatrix} R_{(i-1)}^j & 0 \\ R_{i(i-1)}^j & R_{ii}^j \end{bmatrix}, \quad (2.11)$$

with R_{ii}^j of full column rank k_i^j , and its analogue on the vector ξ

$$\xi_{(i)}^j = \begin{bmatrix} \xi_{(i-1)}^j \\ \xi_i^j \end{bmatrix}, \quad (2.12)$$

where $j = 1, 2$ and $i : 1 \rightarrow n$ is the index¹⁾ referring to the equation in question; all this means that subscript (i-1) in parentheses refers to the previous (i-1) equations and subscript i pertains to equation i alone. If we now define

$$W_i^j = WR_{ii}^j, \quad (2.13)$$

and denote by $W_{(i)}^2$ the matrix of full column rank containing all the variables that have coefficients appearing in $\xi_{(i)}^2$, we can formally write down the requirements for ξ^2 in order to ensure its analytical tractability; we need, in particular, that

$$P_{w_i^2} W_{(i)}^2 = W_{(i)}^2, \quad (2.14)$$

i.e. the RLI assumption, where

$$P_{w_i^2} = W_{(i)}^2 W_{(i)}^{2+} = W_{(i)}^2 (W_{(i)}^{2'} W_{(i)}^2)^{-1} W_{(i)}^{2'}, \quad (2.15)$$

provided $W_{(i)}^2$ is of full column rank, which is assumed throughout the paper, and we require

$$|B| = f(\zeta^1) , \quad (2.16)$$

i.e. that $|B|$ is a function of ζ^1 only, and does not depend on ζ^2 . In Subsection 2.3 a third defining characteristic of ζ^2 is seen to be its prior dependence on Γ within a particular class of prior structures. For details and a proof of the sufficiency of (2.14), we refer to Steel (1988).

2.2. A Recursive Transformation

As argued in Steel (1988) a recursive transformation of our original model (2.2), (2.3), and (2.10) is not necessary for the argument, as all derivations can be done in both the original and the recursive parameterizations. It is, however, felt to somewhat clarify the analysis, and will, therefore, be introduced here.

Let Λ be the $n \times n$ lower triangular matrix that diagonalizes Γ , which was assumed to be PDS, and let us partition Λ as follows

$$\Lambda = \begin{bmatrix} 1 & & & 0 \\ -\lambda_{21} & 1 & & \\ -\lambda'_3 & & 1 & \\ -\lambda'_n & & & 1 \end{bmatrix} , \quad (2.17)$$

where also $\lambda'_i = (\lambda_{i1} \dots \lambda_{i,i-1})$; now define

$$\Lambda \Sigma \Lambda' = D_\omega = \begin{bmatrix} \omega_1^2 & & & 0 \\ & \ddots & & \\ 0 & & & \omega_n^2 \end{bmatrix} . \quad (2.18)$$

The transformation will, of course, also affect the coefficients A of the model, which will be mapped into, say, a matrix Δ

$$\Delta = A \Lambda' , \quad (2.19)$$

again subject to restrictions of the form

$$\delta = \text{vec } \Delta = (\Lambda \Theta I_{m+n}) R \zeta, \quad (2.20)$$

and, defining δ^j consistently with ζ^j , we write

$$\delta^j = (\Lambda \Theta I_{m+n}) R^j \zeta^j, \quad j = 1, 2. \quad (2.21)$$

Provided that the restrictions in (2.10) allow for a consistent solution in ζ^2 to exist, i.e. if

$$R^2 R^{2+} \alpha_2 = \alpha_2, \quad (2.22)$$

we can solve uniquely for ζ_i^2 in terms of ξ_i^2 , λ_i , and $\zeta_{(i-1)}^2$ from (2.21)

$$\zeta_i^2 = R_{ii}^{2+} \{ \xi_i^2 + [(\lambda_i' \Theta I) R_{(i-1)}^2 - R_{i(i-1)}^2] \zeta_{(i-1)}^2 \}, \quad (2.23)$$

where ξ^2 is partitioned as in (2.12).

Conformably with (2.10) we define, next,

$$A = A^1 + A^2,$$

or $\alpha_j = \text{vec } A^j = R^j \zeta^j$ ($j=1,2$), and we partition, as in (2.11) and (2.12),

$$A_{(i)}^j = (A_{(i-1)}^j \ a_i^j), \quad (2.24)$$

which allows us to write the likelihood of the model in the recursive parameterization as

$$\mathcal{L}(\zeta, \{\lambda_i, \omega_i^2\}; Y, X) \propto [f(\zeta^1)]^T$$

$$\prod_{i=1}^n (\omega_i^2)^{-\frac{1}{2}T} \exp -\frac{1}{2} \omega_i^{-2} (w a_i^1 - w A_{(i-1)}^1 \lambda_i + w \delta_i^2)' (w a_i^1 - w A_{(i-1)}^1 \lambda_i + w \delta_i^2), \quad (2.25)$$

where $f(\zeta^1)$ is $|B|$ from (2.16), and the coefficients are still restricted by

$$A_{(i)}^1 = R_{(i)}^1 \zeta_{(i)}^1, \quad (2.26)$$

and

$$\zeta_i^2 = R_{ii}^2 \zeta_i^2 - [(\lambda_i' \Theta I) R_{(i-1)}^2 - R_{i(i-1)}^2] \zeta_{(i-1)}^2. \quad (2.27)$$

Given the full column rank of R_{ii}^2 , it is obvious from (2.27) that a vector defined as

$$\zeta_i^2 = R_{ii}^{2*} \zeta_i^2 \quad (i:1 \rightarrow n) \quad (2.28)$$

of dimension k_i^2 , would not be subject to restrictions, given $\zeta_{(i-1)}^2$. From (2.23) or (2.27) it is then clear that ζ_i^2 is fully determined by knowledge of ζ_i^2 and λ_i , given $\zeta_{(i-1)}^2$, i.e. ζ^2 can be found recursively from all pairs $\{\zeta_i^2, \lambda_i\}$. As argued in Steel (1988), it is precisely the RLI assumption (2.14) that ensures that the likelihood function can be written in terms of ζ_i^2 , as it implies

$$W \zeta_i^2 = W_i^2 \zeta_i^2, \quad (2.29)$$

giving us the following representation of (2.25)

$$\mathcal{L}(\zeta, \{\lambda_i, \omega_i^2\}; Y, X) \propto [f(\zeta^1)]^T \prod_{i=1}^n (\omega_i^2)^{-\frac{1}{2}T} \exp -\frac{1}{2} \omega_i^{-2} (W a_i^1 - Z_i \zeta_i)^T (W a_i^1 - Z_i \zeta_i), \quad (2.30)$$

where

$$Z_i = (-W_i^2 : W A_{(i-1)}^1),$$

and

$$\zeta_i = \begin{bmatrix} \zeta_i^2 \\ \lambda_i \end{bmatrix}$$

is an unrestricted vector, allowing us to conduct analytical inference on the elements of $\mathbf{z}^2$ in a recursive way, given, of course, the vector $\mathbf{z}^1$, which is still implicitly present in $\mathbf{a}_1^1$ and $\mathbf{Z}_1$, both subject to the restrictions in (2.26).

2.3. The Stochastic Prior Structure

As already mentioned in Subsection 2.1 the recursive analytical approach in Steel (1988) is presented under diffuse prior information on the unrestricted coefficients of the model.

It is, however, not the only prior structure that allows for analytical inference, as we can use a specific type in the class of recursive extended natural-conjugate (RENC) prior densities introduced in Richard and Steel (1988), where we induce a linear dependence of one of the hyperparameters ($F_{i(i-1)}^0$ in their notation) on λ_i in order to preserve prior independence across equations in terms of the recursive parameters γ_i .

In particular, we choose for $(\mathbf{z}^2, \Sigma)$ the product of the following densities over all equations $i : 1 \rightarrow n$:

$$D(\omega_i^2) = f_{i\gamma}(\omega_i^2 | \omega_{i0}^2, \nu_{0-n+i}), \quad (2.31)$$

$$D(\lambda_i | \omega_i^2) = f_N^{i-1}(\lambda_i | \lambda_i^0, \omega_{i1}^2 \omega_{i(i-1)}^{0-1}), \quad (2.32)$$

and

$$D(\mathbf{z}_i^2 | \lambda_i, \omega_i^2, \mathbf{z}_{(i-1)}^2) = f_N^{i-1}(\mathbf{z}_i^2 | \mathbf{z}_{i0}^2 + \mathbf{R}_{ii}^{2+} [(\lambda_i' \boldsymbol{\theta} \mathbf{I}) \mathbf{R}_{(i-1)}^2 - \mathbf{R}_{i(i-1)}^2] | (\mathbf{z}_{(i-1)}^2 - \mathbf{z}_{(i-1)0}^2), \omega_{i1}^2 \mathbf{N}_i^{0-1}), \quad (2.33)$$

assuming $\mathbf{N}_1^0$ to be PDS and partitioning $\mathbf{z}_0^2$ conformably with $\mathbf{z}^2$. To facilitate the analysis, we shall assume prior independence between $(\mathbf{z}^2, \Sigma)$ and $\mathbf{z}^1$. Prior densities of $\mathbf{z}^1$ will be discussed at the end of this subsection.

Remark that the prior densities on $\{\lambda_i, \omega_i^2\}$ are compatible with an Inverted-Wishart prior structure on Σ of the form

$$D(\Sigma) = f_{IW}^n(\Sigma | \Sigma_0, \nu_0), \quad (2.34)$$

with $\nu_0 > n-1$ and where both Σ and Σ_0 are PDS matrices that decompose recursively into

$$\Sigma_{(i)} = \begin{bmatrix} \Sigma_{(i-1)} & \Sigma_{(i-1)} \lambda_i \\ \lambda_i' \Sigma_{(i-1)} & \omega_i^2 + \lambda_i' \Sigma_{(i-1)} \lambda_i \end{bmatrix},$$

if we define $\Sigma \equiv \Sigma_{(n)}$. In view of (2.23) and (2.28), we can now write the prior assumptions in terms of the recursive parameterization γ_i as (2.31) combined with

$$D(\gamma_i | \omega_i^2, \Sigma_{(i-1)}^2) = D(\gamma_i | \omega_i^2) = f_N^{k_i^2 + i - 1}(\gamma_i | \gamma_i^0, \omega_i^2 H_i^0)^{-1} \quad (2.35)$$

where

$$\gamma_i^0 = \begin{bmatrix} \Sigma_{i0}^2 - R_{ii}^{2+} [(\lambda_i^0)' \Theta I R_{(i-1)}^2 - R_{i(i-1)}^2] \Sigma_{(i-1)0}^2 \\ \lambda_i^0 \end{bmatrix}$$

and

$$H_i^0 = \begin{bmatrix} N_i^0 & N_i^0 R_{ii}^{2+} A_{(i-1)0}^2 \\ A_{(i-1)0}^{2'} R_{ii}^{2+} N_i^0 & \Sigma_{(i-1)}^0 + A_{(i-1)0}^{2'} R_{ii}^{2+} N_i^0 R_{ii}^{2+} A_{(i-1)0}^2 \end{bmatrix},$$

defining $\text{vec } A_{(i-1)0}^2 = R_{(i-1)}^2 \Sigma_{(i-1)0}^2$. Note that prior independence across equations is imposed in terms of $\{\gamma_i, \omega_i^2\}$.

Prior elicitation can be performed using the same kind of recursive formulae as found in Steel (1988), in particular $(\forall i: 1 \rightarrow n)$

$$E(\zeta_i^2) = \zeta_{i0}^2, \quad (2.36)$$

$$\text{Cov}(\zeta_i^2, \zeta_j^{2'}) = R_{ii}^{2+} [(\lambda_i^{0'} \Theta I) R_{(i-1)}^2 - R_{i(i-1)}^2] \text{Cov}(\zeta_{(i-1)}^2, \zeta_j^{2'}), \quad i > j, \quad (2.37)$$

and

$$\begin{aligned} \text{Var}(\zeta_i^2) &= \frac{\omega_{i0}^2}{\nu_{0-n+i-2}} \{N_i^{0-1} + R_{ii}^{2+} [\sum_{k,q=1}^{i-1} \Sigma_{(i-1)}^{0kq} R_k^2 \text{Cov}(\zeta_{(k)}^2, \zeta_{(q)}^{2'}) R_q^{2'}] R_{ii}^{2+} \} + \\ &R_{ii}^{2+} [(\lambda_i^{0'} \Theta I) R_{(i-1)}^2 - R_{i(i-1)}^2] \text{Var}(\zeta_{(i-1)}^2) [(\lambda_i^{0'} \Theta I) R_{(i-1)}^2 - R_{i(i-1)}^2]' R_{ii}^{2+}, \end{aligned} \quad (2.38)$$

defining $R_k^2 = (R_{k(k-1)}^2; R_{kk}^2)$ and denoting by superscripts the relevant elements of the inverse of $\Sigma_{(i-1)}^0$. The restrictive features of natural-conjugate (NC) prior structures, i.e. proportionality of the prior covariance matrices across equations [see e.g. Rothenberg (1963), Drèze and Richard (1983), and Steel (1988)], are avoided here by choosing a different precision matrix N_i^0 for every equation, breaking the proportionality between $\text{Var}(\zeta_i^2)$ and $\text{Var}(\zeta_j^2)$ ($i \neq j$) from (2.38), and by reasoning in terms of ζ^2 instead of A^2 while imposing linear restrictions in vec form, destroying the proportionality between $\text{Cov}(\zeta_i^2, \zeta_j^{2'})$ and $\text{Var}(\zeta_{(k)}^2)$ for $k = \min(i, j)$. Indeed, under the matrix form restrictions in (2.9) for A^2 we notice that a structure as in (2.8) leads to ($i > j$)

$$\text{Cov}(\zeta_i^2, \zeta_j^{2'}) = (\lambda_i^{0'} \Theta I) \text{Cov}(\zeta_{(i-1)}^2, \zeta_j^{2'}),$$

implying the latter type of proportionality, while (2.38) then simplifies to

$$\text{Var}(\zeta_i^2) = \frac{\omega_{i0}^2}{\nu_{0-n+i-2}} N_i^{0-1} + \sum_{k,q=1}^{i-1} \left[\frac{\omega_{i0}^2}{\nu_{0-n+i-2}} \Sigma_{(i-1)}^{0kq} + \lambda_{ik}^0 \lambda_{iq}^0 \right] \text{Cov}(\zeta_k^2, \zeta_q^{2'}),$$

still avoiding block-diagonal proportionality, which is imposed if we also give up the equation-specific precision matrices, i.e. we take $N_i^0 = N_0$, a PDS matrix, $\forall i : 1 \rightarrow n$, leading to

$$\text{Var}(\zeta^2) = \frac{1}{\nu_0 - n - 1} \Sigma_0 \Theta N_0^{-1},$$

the well-known NC result.

Apart from circumventing this NC problem, the specific class of prior densities in (2.31) - (2.33) has the advantage that its marginal prior moments are known analytically [which is generally not the case for other classes of extended natural-conjugate (ENC) prior densities, appearing e.g. in Drze and Richard (1983)], and thus its hyperparameters are easily elicited from assumptions on the prior moments. Remark, however, that there is a price to pay for analytical treatment here, which requires prior independence between γ_i and γ_j ($i \neq j$), as this restricts the way in which ζ_i^2 depends on $\zeta_{(i-1)}^2$ in prior mean [see (2.33)]. In fact, this means that the off-diagonal covariance structure in (2.37) is not as flexible as in the broader class of RENC densities in Richard and Steel (1988). On the other hand, the present structure can cope with general linear restrictions in vec form and might, in practice, prove sufficiently flexible. Of course, if we wish to incorporate prior information on ζ_1^2 , we should avoid being too diffuse on ω_1^2 , to prevent the well-known NC pitfall of inflating the relative prior precision [see Richard (1973)].

Finally, we can always choose to remain vague about the unrestricted parameters of the model by letting N_1^0 and Σ_0 approach zero matrices, giving as a limit case

$$D(\zeta^2, \Sigma) \propto |\Sigma|^{-\frac{1}{2}(\nu_0 + n + 1)} \prod_{i=1}^n (\omega_i^2)^{-\frac{1}{2}} \lambda_i^2,$$

whereas, to simplify the comparison with other studies, the more usual class of diffuse prior densities

$$D(\zeta^2, \Sigma) \propto |\Sigma|^{-\frac{1}{2}\mu_0} \quad (2.39)$$

is also considered, as in Steel (1988).

In fact, only in the special case where $\lambda_1^2 = \lambda_2/n$, $\forall i$, will the limit case exactly coincide with (2.39), implying $\mu_0 = \nu_0 + n + (\lambda_2/n) + 1$. When

performing the Monte Carlo analysis on the coefficients (Subsections 3.1-3.3) the only consequence of choosing between the limiting density and (2.39) will be a change in the existence of posterior moments, discussed in Subsection 3.5. Alternatively, a Monte Carlo analysis on the covariance matrix will be subject to slight changes in the formulae as indicated in Subsection 3.4.

All that remains now is to specify a prior density for those coefficients that are not amenable to analytical treatment, grouped in ζ^1 . Since we have to use numerical procedures for integrating ζ^1 anyway, as we shall discuss in the next section, we have a fully flexible choice in principle. It turns out, however, that matters are facilitated by choosing a prior structure such that $D(\zeta_n^1 | \zeta_{(n-1)}^1)$ takes on a Normal or a Student t form (in the informative case). An easily elicited prior is e.g. a product of independent Student densities:

$$D(\zeta^1) = \prod_{i=1}^n f_t^1(\zeta_i^1 | \zeta_{i0}^1, p_i^0, \nu_i^0) . \quad (2.40)$$

which has the advantage of being invariant with respect to normalization if $\nu_i^0 = 1$, i.e. the densities are of the Cauchy type; on the other hand, the latter choice for ν_i^0 prevents the existence of prior moments, which might complicate the elicitation process.

For the sake of generality, we shall, in the sequel, allow for any prior density of ζ^1 that is independent of (ζ^2, Σ) and that can be factorised as follows

$$D(\zeta^1) = D(\zeta_n^1 | \zeta_{(n-1)}^1) D(\zeta_{(n-1)}^1) , \quad (2.41)$$

where both factors can, of course, be diffuse densities of the uniform type, and where the first factor is either a Student or a Normal density in informative cases.

In summary, our prior assumptions on the entire parameter space are expressed by

$$D(\xi, \Sigma) = D(\xi^2, \Sigma) D(\xi^1), \quad (2.42)$$

where the first factor is the product of (2.33) and (2.34) [or, equivalently, of (2.31) - (2.33)] and the second factor is in the general class of densities (2.41).

3. Posterior Analysis

3.1. Recursive Analytical Results Given ξ^1

The defining characteristics of ξ^2 were given in Subsection 2.1 as the RLI structure (2.14), ensuring that the unrestricted y_{i2} appears in the likelihood function (2.30), and the condition that ξ^2 should not affect the Jacobian $|B|^T$ in (2.16). These requirements will lead to a separability of the likelihood function across equations in terms of the recursive parameters $\{y_i\}$, from which $(\xi^2, \{y_i\})$ can be found uniquely using (2.27) and (2.28), defining a one-to-one transformation.

A similar independence across equations was induced in the prior structure by the choices made in Subsection 2.3.

This implies that the analysis of the posterior densities of (ξ^2, Σ) , conditional upon ξ^1 , will become very simple, as in the more specific case discussed in Steel (1988).

In particular, reasoning in terms of the recursive parameterization $\{y_i, \omega_i^2\}$ we obtain $\forall i : 1 \rightarrow n$

$$D(\omega_i^2 | y_i, \xi^1, Y, X) = f_{iy}(\omega_i^2 | \omega_{i*}^2, \nu_* + k_i^2 + 2i - n - 1) \quad (3.1)$$

with

$$\nu_* = \nu_0 + T$$

$$\omega_{i*}^2 = \omega_{i0}^2 + (y_i - y_i^0)' H_i^0 (y_i - y_i^0) + (w_{a1}^1 - z_i y_i)' (w_{a1}^1 - z_i y_i),$$

and

$$D(\gamma_i | \zeta^1, Y, X) = f_t^{\lambda_i^2 + i - 1} (\gamma_i | \gamma_i^*, q_i^{*-1} H_i^*, \nu_{*+i-n}) , \quad (3.2)$$

where the hyperparameters are given by well-known NC formulae as

$$H_i^* = H_i^0 + Z_i' Z_i, \text{ assumed PDS,}$$

$$\gamma_i^* = H_i^{*-1} (H_i^0 \gamma_i^0 + Z_i' W a_i^1) ,$$

and

$$q_i^* = \omega_{i0}^2 + \gamma_i^{0'} H_i^0 \gamma_i^0 + a_i^{1'} W' W a_i^1 - \gamma_i^{*'} H_i^* \gamma_i^* , \text{ assumed positive,}$$

obviously depending on ζ^1 , the posterior density of which then becomes

$$D(\zeta^1 | Y, X) \propto [f(\zeta^1)]^T D(\zeta^1) \prod_{i=1}^n q_i^*^{-\frac{1}{2}(\nu_{*+i-n})} |H_i^*|^{-\frac{1}{2}} . \quad (3.3)$$

The factors in the latter kernel originate from, respectively, the Jacobian of the likelihood function [see (2.16)], the prior density of ζ^1 in (2.41), and the integrating constants from the posterior densities in (3.2). Its numerical treatment will be the subject of the next subsection. If, instead of the RENC prior structure for (ζ^2, Σ) (or its limit case), we use the popular class of diffuse prior densities in (2.39), the number of degrees of freedom in (3.2) and, consequently, the exponent of q_i^* in (3.3) are both changed to $\mu_0 + T - 2n - \lambda_i^2 + i - 1$.

It is now straightforward to use the densities in (3.2) in order to derive the moments of ζ^2 given ζ^1 by the same type of recursive formulae as found in Steel (1988). For example, the posterior mean can be obtained from

$$E(\zeta_i^2 | \zeta^1, Y, X) = \zeta_{i*}^2 + R_{ii}^{2*} [(\lambda_i^{*'} \theta I) R_{(i-1)}^2 - R_{i(i-1)}^2] E(\zeta_{(i-1)}^2 | \zeta^1, Y, X) , \quad (3.4)$$

where γ_{i*}^2 and λ_i^* are the appropriate subvectors of γ_i^* . The posterior covariance structure, given ζ^1 , can be found from exactly the same formulae as its prior equivalent in (2.37) and (2.38), replacing γ_i^0 , ω_{i0}^2 , H_i^0 , and ν_0 by the posterior hyperparameters γ_i^* , q_i^* , H_i^* , and ν_* , respectively.

Moments of λ_i or certain elements of Λ^{-1} , of interest for evaluating weak exogeneity, can be found through similar calculations as in Steel (1988), taking the appropriate hyperparameters, where it has to be stressed that all these posterior results for ζ^2 and Λ (or its transformations) are still conditional upon ζ^1 . If our main interests do not involve (ζ^2, Λ) this lack of marginal analytical moments is of minor concern; if (ζ^2, Λ) do appear as parameters of interest (possibly after some transformation) we can obtain the marginal moments as a "by-product" of the Monte Carlo analysis on ζ^1 , by evaluating the conditional moments at each drawing and computing the sample average with the usual weighting function [see e.g. Kloek and van Dijk (1978) or Bauwens (1984)]²⁾. Thanks to the known analytical properties of the conditional posterior densities (3.2), which can be interpreted as $D(\zeta_i^2, \lambda_i | \zeta_{(i-1)}^2, \zeta^1, Y, X)$, this will not increase the dimension of the numerical Monte Carlo integration. Details of the actual strategy proposed here will be given in Subsection 3.3.

3.2. Posterior Inference on ζ^1

The numerical part of the analysis will focus upon the λ_1 coefficients grouped in the vector ζ^1 , the posterior density of which was described in (3.3).

In order to provide a link with results stated elsewhere in the literature, consider the case where A^2 is simply matrix restricted as in (2.9), and the corresponding NC prior structure is assumed for its unrestricted coefficients. This translates into

$$R^2 = I \otimes F_2, \quad (3.5)$$

and

$$N_i^0 = N_0, \quad i: 1 \rightarrow n, \quad (3.6)$$

which we assume to be PDS, just like its posterior counterpart

$$N_* = N_0 + F_2' W' W F_2. \quad (3.7)$$

Applying some algebra to (3.3) then gives

$$D(\zeta^1 | Y, X) \propto [f(\zeta^1)]^T D(\zeta^1) \left| \Sigma_0 + Z_0^2 N_0 Z_0^2 + A^1' W' W A^1 - Z_*^2 N_* Z_*^2 \right|^{-\frac{1}{2} \nu_*}, \quad (3.8)$$

$$\text{with } N_* Z_*^2 = N_0 Z_0^2 - F_2' W' W A^1, \quad (3.9)$$

and where we have defined $A_0^2 = F_2 Z_0^2$ and $\text{vec } Z_0^2 = \zeta_0^2$. Notice that this result is a direct generalization of formula (6.5) in Drèze and Richard (1983)³⁾. In this specific case, the functional form of the posterior density on ζ is not affected by integrating out the matrix-restricted coefficients ζ^2 . Whenever ζ^2 has a more flexible RLI structure, the posterior density of ζ^1 can not be written as (3.8), but for the numerical analysis this will not make any difference, as the determinant in (3.8) is typically not the kernel of a matricvariate Student t density of A^1 , due to the general restrictions we can impose on these coefficients.

Let us, therefore, concentrate upon the posterior density as given in (3.3), which immediately allows for a factorization into $\zeta_n^1 | \zeta_{(n-1)}^1$ and $\zeta_{(n-1)}^1$. Indeed, the second factor, the prior density of ζ^1 , accommodates such a factorization by definition [see (2.41)], and the first factor, the Jacobian, can be decomposed as in Appendix B.4 of Drèze and Richard (1983) into a factor that no longer contains ζ_n^1 and a Student t kernel for $\zeta_n^1 | \zeta_{(n-1)}^1$, which is degenerate here if ζ_n^1 also includes some coefficients from Γ . As H_n^* does not depend on ζ_n^1 from the definition of Z_i following (2.30), the only other factor in the conditional posterior density for ζ_n^1 will be $(q_n^*)^{-\frac{1}{2} \nu_*}$, again a Student kernel. We have, therefore, the following factorization:

$$D(\xi^1 | Y, X) = D(\xi_n^1 | \xi_{(n-1)}^1, Y, X) D(\xi_{(n-1)}^1 | Y, X) , \quad (3.10)$$

If we take the Student option for $D(\xi_n^1 | \xi_{(n-1)}^1)$ in (2.41), the first factor is a 2-1 poly-t density⁴⁾, exactly as in Lemma 6.2 of Drèze and Richard (1983). The crucial difference with their analysis, however, is that we are now only reasoning in the space of ξ^1 , which is of dimension ℓ_1 , instead of the full ℓ -dimensional space of all unrestricted coefficients in ξ . As we shall argue later, this might be a major advantage in performing our numerical integrations.

If our prior assumptions on ξ^1 are changed to a diffuse (uniform) prior, the conditional posterior density for ξ_n^1 simplifies to a 1-1 poly-t density, pseudo-random drawings from which are easily obtained following Bauwens and Richard (1985). If, on the other hand, we wish to express our conditional prior opinions on ξ_n^1 by a Normal instead of a Student t density, its posterior counterpart will be the product of a 1-1 poly-t kernel and a Normal kernel, which can be analysed with the same techniques as 2-1 poly-t densities, using the results from Section 3 in Richard and Tompa (1980). For the next subsection we shall, therefore, focus on the Student option, whereas the Normal prior density will be used in Subsection 3.4.

3.3. The Monte Carlo Strategy

In order to perform the necessary numerical integrations, we shall use the method of Monte Carlo analysis with importance sampling, where we should like to perform our pseudo-random drawings according to a so-called importance function that closely approximates the density under scrutiny (i.e. the posterior density of ξ^1 in this case), yet allows for efficient algorithms to generate these drawings. As such techniques are treated in great detail in e.g. Hammersley and Handscomb (1979), van Dijk (1984), Bauwens (1984), and Hendry (1984), we shall refrain from a more extensive account of the underlying theory here.

In an effort to minimize the dimension of the actual Monte Carlo problem, we shall make full use of the poly-t structure of the conditional posterior for ζ_n^1 , as in the PTST-1 (poly-t Student, conditional posterior for equation 1) method described in Bauwens (1984). If we are interested in some integrable function $g(\mathbb{L}, \zeta)$ of the unrestricted parameters of our model, we first write its expectation as

$$E g(\mathbb{L}, \zeta) = E_{\zeta_{(n-1)}^1} E_{\zeta_n^1 | \zeta_{(n-1)}^1} E_{\wedge, \zeta^2 | \zeta^1} E_{\{\omega_i^2\} | \wedge, \zeta} [g(\mathbb{L}, \zeta) \cdot J], \quad (3.11)$$

with the Jacobian $J = \prod_{i=1}^{n-1} (\omega_i^2)^{n-i}$ as derived in Appendix C of Steel (1988), and where the subscripts indicate that we take the expected value with respect to the posterior density of those parameters. The last expectation is known analytically from (3.1), the third one can be evaluated analytically using recursive formulae and a one-to-one transformation on (3.2), and for the second one, corresponding to the first factor in (3.10), we can take advantage of the combined analytical and numerical techniques developed for poly-t densities in e.g. Richard and Tompa (1980) and Bauwens and Richard (1985). Finally, the only expectation that requires purely numerical Monte Carlo procedures for its evaluation is the first one, which is taken with respect to the second factor in (3.10), a density that defies analytical treatment altogether.

The evaluation of the second expectation in (3.11) can, in principle, be based on the algorithms for integrating poly-t densities as found in the computer software PTD, documented in Bauwens et al. (1981). The latter software can, however, not cope with moments of higher order than two or moments of nonlinear functions, and requires that we write the result of the last two, analytical, expectations as an explicit function of ζ_n^1 . Although the latter is not impossible, given the hyperparameters in (3.2) and the recursive formulae in (3.4) and as in (2.37) and (2.38), it might quickly lead to problems PTD cannot handle and might require considerable programming effort.

Therefore, we propose to draw values for ζ_n^1 , given $\zeta_{(n-1)}^1$, which can be done directly in the case of a 1-1 poly-t density, and requires some importance sampling when we are faced with a 2-1 poly-t density, using a 1-1

poly-t density as importance function. The latter case would, of course, increase the overall dimension of the Monte Carlo integration using importance sampling, but, from the results in Bauwens and Richard (1985), it seems that this type of numerical integration can be performed rather efficiently.

In summary, we suggest the following strategy for conducting posterior inference on $g(\Sigma, \zeta)$:

1. Draw a value for the $(\ell_1 - \ell_n^1)$ dimensional vector $\zeta_{(n-1)}^1$ according to some importance function $D_{IF}(\zeta_{(n-1)}^1 | Y, X)$. Denote the k -th value drawn by $\zeta_{(n-1)}^{1(k)}$.
2. Conditionally on $\zeta_{(n-1)}^{1(k)}$, draw a vector ζ_n^1 from a 1-1 poly-t density function in ℓ_n^1 dimensions, say $D_{1-1}(\zeta_n^1 | \zeta_{(n-1)}^{1(k)}, Y, X)$, which is either the actual conditional posterior density, or an approximation to it constructed according to the principles in Bauwens and Richard (1985). We have then obtained a drawing for the entire vector ζ^1 , which we call $\zeta^{1(k)}$.
3. Given $\zeta^{1(k)}$, evaluate $g(\Sigma, \zeta)$ analytically, using the conditional posterior densities in (3.1) and (3.2), and the recursive approach.

After N replications of these three steps, we approximate the actual expected value of $g(\Sigma, \zeta)$ by the weighted sample mean where now, for the sake of brevity, the analytical expectation is taken with respect to Σ

$$\hat{g} = \frac{\sum_{k=1}^N \int \int \int \Sigma, \zeta^2 | \zeta_{(n-1)}^{1(k)} g(\Sigma, \zeta^1(k), \zeta^2) w^{(k)} d\zeta^2 d\zeta^1 d\Sigma}{\sum_{k=1}^N w^{(k)}}, \quad (3.12)$$

with the weight function

$$w^{(k)} = \frac{D(\zeta_n^1(k) | \zeta_{(n-1)}^{1(k)}, Y, X) D(\zeta_{(n-1)}^1(k) | Y, X)}{D_{1-1}(\zeta_n^1(k) | \zeta_{(n-1)}^{1(k)}, Y, X) D_{IF}(\zeta_{(n-1)}^1(k) | Y, X)}, \quad (3.13)$$

the ratio of the actual posterior densities in (3.10) and the importance functions drawn from, both evaluated at the k -th value drawn. Note that, whenever the conditional posterior density of ζ_n^1 is of the 1-1 poly-t form, the weights will simplify to

$$w^{(k)} = \frac{D(\zeta_{(n-1)}^1 | Y, X)}{D_{IF}(\zeta_{(n-1)}^1 | Y, X)}, \quad (3.14)$$

and only the numerical integration over $\zeta_{(n-1)}^1$ will contribute to the estimated relative error bounds as defined in Kloek and van Dijk (1978) and Bauwens (1984). Even in the 2-1 poly-t case for $\zeta_n^1 | \zeta_{(n-1)}^1$, it is to be expected, given the encouraging results of Bauwens and Richard (1985), that the main difficulties in the Monte Carlo procedure will arise from integrating over $\zeta_{(n-1)}^1$, so that the choice of its importance function seems critical.

Bauwens (1984) suggests an overall Student t density

$$D_{IF}(\zeta_{(n-1)}^1 | Y, X) = f_t^{\lambda_1 - \lambda_n^1}(\zeta_{(n-1)}^1 | \zeta_{(n-1)}^1, P_*, \mu_*), \quad (3.15)$$

where the choice of hyperparameters can be based on, respectively, the posterior mode, the Hessian matrix and the order of existence of moments. Provided our prior $D(\zeta_{(n-1)}^1)$ in (2.41) falls in the class of ENC prior densities, the mode and the Hessian can be evaluated by FIML techniques, as explained in Morales (1971), and the order up to which posterior moments are sure to exist is given in Lemma 6.6 in Drèze and Richard (1983). A possible alternative method for choosing hyperparameters in (3.15) is to iteratively update them, based on the posterior moments resulting from small Monte Carlo runs, used only for "calibrating" the importance function.

Of course, (3.15) is still very "well-behaved", as it imposes both symmetry and unimodality. If such assumptions are strongly violated by the actual posterior density, we should consider using other classes of importance functions, possibly based on poly-t densities, along the lines of Bauwens' (1984) further suggestions.

As the posterior analysis of the last equation's coefficients in ζ^1 is based on the rather well-known (given some numerical effort) properties of its conditional posterior density, whereas this extra information is lacking in the case of $\zeta_{(n-1)}^1$, one might expect that the results are not fully invariant with respect to the ordering of the equations. Remember, however, that in practice this ordering will, to a large extent, be directed by the effort to group as many coefficients as possible in ζ^2 , which allows for (conditional) analytical treatment. Should this criterion still leave us with a choice, then one could think of putting an "equation of special interest" last, or one could be guided by considerations of dimension or covariance structure as in Bauwens (1984).

3.4. Monte Carlo Analysis of the Covariance Matrix

In the previous subsections we have adopted the usual practice of first integrating out the covariance matrix, using a recursive decomposition as in the prior specification (2.31) and (2.32). Indeed, combining the likelihood function (2.25), in terms of ζ , with the prior structure (2.31) - (2.33) we can deduce a similar factorization of the posterior density, given ζ , as

$$D(\omega_i^2 | \zeta, Y, X) = f_{iY}(\omega_i^2 | \omega_{i0}^2 + \lambda_i^0 \Gamma_{(i-1)}^0 \lambda_i^0 + (a_i^2 - a_{i0}^2)' R_{ii}^{2+} N_{ii}^{02+} (a_i^2 - a_{i0}^2) + a_i' W' W a_i - \lambda_i^{**} G_i^* \lambda_i^{**}, \nu_{*} - n + l_{i+1}^2) \quad (3.16)$$

$$D(\lambda_i | \omega_i^2, \zeta, Y, X) = f_N^{i-1}(\lambda_i | \lambda_i^{**}, \omega_i^2 G_i^{*-1}), \quad (3.17)$$

where a_i is the i -th column of A , just like a_i^2 in (2.24), the prior equivalent of which also holds in prior mean with subscript 0, and we define

$$G_i^* = \Gamma_{(i-1)}^0 + (A_{(i-1)}^2 - A_{(i-1)0}^2)' R_{ii}^{2+} N_{ii}^{02+} (A_{(i-1)}^2 - A_{(i-1)0}^2) + A_{(i-1)}' W' W A_{(i-1)}, \quad (3.18)$$

a relative precision matrix which we assume to be nonsingular, and

$$G_{ii}^* \lambda_{ii}^{**} = \Sigma_{(i-1)}^0 \lambda_i^0 + (A_{(i-1)}^2 - A_{(i-1)0}^2)' R_{ii}^{2*} N_{ii}^0 R_{ii}^{2*} (a_i^2 - a_{i0}^2) + A_{(i-1)}' W' W a_i \quad (3.19)$$

If A^2 happens to be matrix-restricted as in (2.9), leading to (3.5) and (3.6), the densities (3.16) and (3.17) are compatible with an Inverted-Wishart density on the full covariance matrix, which then takes the form

$$D(\Sigma | \mathcal{Y}, X) = f_{IW}^n(\Sigma | \Sigma_0 + (Z^2 - Z_0^2)' N_0 (Z^2 - Z_0^2) + A' W' W A, \nu_* + \frac{\lambda_2}{n}) \quad (3.20)$$

a straightforward generalization of formula (6.4) in Drèze and Richard (1983). It is, therefore, clear that only if the restrictions on A^2 are of a more general nature (i.e. vec form under the RLI assumption) we gain by treating the model equation by equation in integrating out the covariance structure.

Note that the integrating constant from (3.20) will lead to the following posterior density for the matrix of unrestricted coefficients Z^2

$$D(Z^2 | \mathcal{Y}^1, X) = f_{Mt}^{\frac{\lambda_2}{n} \times n} (Z^2 | Z_*, \Sigma_0 + Z_0^2' N_0 Z_0^2 - Z_*^2' N_* Z_*^2 + A^1' W' W A^1, N_*, \nu_*) \quad (3.21)$$

using the definitions in (3.7) and (3.9), a simple NC outcome, as expected, under matrix restrictions on A^2 . Finally, this gives us the same marginal posterior density as found in (3.8) for $\mathcal{Y}^1$.

Of course, the order of the integrations could be reversed. In particular, Richard and Steel (1988) find that in SURE models it might be preferable to integrate out the coefficients analytically (under an RENC prior structure) given Σ , and, subsequently, perform the Monte Carlo analysis on Σ , using an Inverted-Wishart importance function. The latter allows for a simple iterative strategy to calibrate its hyperparameters, and, heuristically, it is felt that the vec form restrictions in (2.7) might induce a larger departure from the NC posterior results for the coefficients on which the restrictions are imposed, than for the covariance matrix, which

is not subject to restrictions in our analysis. Let us, now, consider the possibility of a similar treatment of the SEM discussed here.

The RENC prior density of ζ_i^2 , given $(\lambda_i, \omega_i^2, \zeta_{(i-1)}^2)$ in (2.33) implies the following prior structure for the entire vector ζ^2 , given Σ :

$$D(\zeta^2 | \Sigma) = f_N^2(\zeta^2 | \zeta_0^2, (M_0^2)^{-1}) \quad (3.22)$$

where

$$M_0^2 = (I - L_0)^{-1} D_\omega^0 (I - L_0)^{-1} \quad (3.23)$$

with

$$D_\omega^0 = \begin{bmatrix} \omega_{11}^{2N_0-1} & & & 0 \\ & \ddots & & \\ 0 & & & \omega_{nn}^{2N_0-1} \end{bmatrix}, \quad (3.24)$$

and L_0 is a lower block-triangular matrix with zeros in the diagonal blocks and where the rest of the i -th block of rows is given by

$$L_{i(i-1)}^0 = R_{ii}^{2+} [(\lambda_i' \otimes I_m) R_{(i-1)}^2 - R_{i(i-1)}^2], \text{ for } i : 2 \rightarrow n. \quad (3.25)$$

Remark that under matrix restrictions on A^2 the structure of R^2 in (3.5) will imply

$$L_{i(i-1)}^0 = \lambda_i' \otimes I_m, \quad (3.26)$$

and, together with (3.6), this leads to

$$M_0^2 = \Sigma^{-1} \otimes N_0, \quad (3.27)$$

the usual NC result.

In our more general framework, we can combine the likelihood function with the prior structure in (2.34), (2.41), and (3.22) to first integrate out ζ^2 analytically, given the other parameters:

$$D(\zeta^2 | \zeta^1, \Sigma, Y, X) = f_N^{\mathcal{L}}(\zeta^2 | \zeta_*^2, (M_*^2)^{-1}) \quad (3.28)$$

with

$$M_*^2 = M_0^2 + R^{2'} (\Sigma^{-1} \Theta W' W) R^2 \quad (3.29)$$

$$M_{**}^{2,2} = M_0^{2,2} - R^{2'} (\Sigma^{-1} \Theta W' W) R^1 \zeta^1, \quad (3.30)$$

and where a PDS structure is assumed for M_*^2 .

We are then left with

$$D(\zeta^1, \Sigma | Y, X) \propto [f(\zeta^1)]^T D(\zeta^1) |M_0^2|^{\frac{1}{2}} |M_*^2|^{-\frac{1}{2}} |\Sigma|^{-\frac{1}{2}(\nu_* + n + 1)} \\ \exp -\frac{1}{2} [\text{tr } \Sigma^{-1} \Sigma_0 + \zeta_0^{2'} M_0^{2,2} \zeta_0^2 - \zeta_*^{2'} M_{**}^{2,2} \zeta_*^2 + \zeta^1{}' R^{1'} (\Sigma^{-1} \Theta W' W) R^1 \zeta^1], \quad (3.31)$$

from which we shall attempt to integrate out ζ^1 analytically, so that the subsequent Monte Carlo analysis can be performed on Σ . It now becomes obvious from (3.31) that we can only do this if $|B| = f(\zeta^1)$ is constant with respect to ζ^1 , i.e. if we are in a triangular⁵⁾ or a SURE system, and if $D(\zeta^1)$ is given a diffuse or Normal structure, say⁶⁾

$$D(\zeta^1) = f_N^{\mathcal{L}}(\zeta^1 | \zeta_0^1, (M_0^1)^+) . \quad (3.32)$$

Under these two simplifying assumptions, we obtain

$$D(\zeta^1 | \Sigma, Y, X) = f_N^{\mathcal{L}}(\zeta^1 | \zeta_*^1, (M_*^1)^{-1}) , \quad (3.33)$$

where we have defined

$$M_*^1 = M_0^1 + R^{1'} (\Sigma^{-1} \Theta W' W) R^1 - R^{1'} (\Sigma^{-1} \Theta W' W) R^2 (M_*^2)^{-1} R^{2'} (\Sigma^{-1} \Theta W' W) R^1 , \quad (3.34)$$

assumed nonsingular, and

$$M_{**}^1 \Sigma^1 = M_0^1 \Sigma_0^1 - R^{1'} (\Sigma^{-1} \Theta W' W) R^2 (M_{**}^2)^{-1} M_0^2 \Sigma_0^2, \quad (3.35)$$

leaving us with the following, generally intractable, posterior density for Σ :

$$D(\Sigma|Y, X) \propto |M_0^2|^{\frac{1}{2}} |M_{**}^2|^{-\frac{1}{2}} |M_{**}^1|^{-\frac{1}{2}} \exp -\frac{1}{2} \{ \Sigma_0^{2'} [M_0^2 - M_0^2 (M_{**}^2)^{-1} M_0^2] \Sigma_0^2 - \Sigma_{**}^{1'} M_{**}^1 \Sigma_{**}^1 \} \times \\ f_{IW}^n(\Sigma | \Sigma_0, \nu_{**}), \quad (3.36)$$

which simplifies somewhat if A^2 is matrix restricted, to become

$$D(\Sigma|Y, X) \propto |M_{**}^1|^{-\frac{1}{2}} \exp \frac{1}{2} \{ \Sigma_{**}^{1'} M_{**}^1 \Sigma_{**}^1 \} \times f_{IW}^n(\Sigma | \Sigma_0 + Z_0^{2'} F_2' W' M_F W F_2 Z_0^2, \nu_{**}), \quad (3.37)$$

where

$$M_F = I_T - W F_2 N_{**}^{-1} F_2' W'. \quad (3.38)$$

If we opt for a vague prior density of the simple type found in (2.39), we have to carry out slight changes to these formulae. In particular, M_0^2 becomes a zero matrix in (3.29), (3.30), and (3.35), whereas (3.36) is replaced by

$$D(\Sigma|Y, X) \propto |M_{**}^2|^{-\frac{1}{2}} |M_{**}^1|^{-\frac{1}{2}} \\ \exp \frac{1}{2} \{ \Sigma_{**}^{1'} M_{**}^1 \Sigma_{**}^1 \} |\Sigma|^{-\frac{1}{2}(\mu_0 + T)}, \quad (3.39)$$

which simplifies to

$$D(\Sigma|Y, X) \propto |M_{**}^1|^{-\frac{1}{2}} \\ \exp \frac{1}{2} \{ \Sigma_{**}^{1'} M_{**}^1 \Sigma_{**}^1 \} |\Sigma|^{-\frac{1}{2}(\mu_0 + T + (\ell_2/n))} \quad (3.40)$$

under matrix-form restrictions on A^2 .

From both (3.37) and (3.40) it becomes apparent that now a fully analytical treatment of all parameters requires that all coefficients be grouped in A^2 on which only matrix form restrictions are imposed, as this order of integration does not allow for exploiting recursive properties in any obvious way. In general, however, (3.36) will be used for a Monte Carlo analysis on Σ , using an importance function of the Inverted Wishart type, as in Richard and Steel (1988). The point to note here is that, in order to integrate out ξ^1 analytically by (3.33) we had to impose a Normal or diffuse prior density of ξ^1 and, much more importantly, we required that |B| did not involve any coefficients of the model, i.e., in practice, we limit ourselves to triangular systems or models of the SURE type. We consider this to be a major drawback of the approach described in this subsection.

Within the, rather narrow, bounds of its domain of application, however, the numerical performance of this approach may prove to be superior for the reasons outlined earlier. In addition, the dimension of the Monte Carlo integration may also be much smaller when drawing on Σ , as discussed in Steel (1987). Some further research is required to investigate the possibility of using the efficient algorithms in Richard and Tompa (1980) for analysing quadratic forms in Normal variables, in an effort to render this, numerically promising, approach feasible for SEM-type systems.

3.5. A Survey of Posterior Results

Having presented a collection of posterior results under two distinct integration sequences, we shall now attempt to systematically indicate the implications of both strategies under different model and prior assumptions.

Table 1 groups these results, reporting the type of densities appearing as posterior densities for ξ_n^1 given $\xi_{(n-1)}^1$, whenever we draw on the coefficients, along with the method we would suggest for conducting inference, as well as the dimension of the integrations by Monte Carlo (in square brackets) if the latter proves necessary. The term "direct simulation" is taken to imply that we can perform pseudo-random drawings from the actual

posterior density, and "PTD" indicates that we can use this computer software directly, provided our interest in ζ_n^1 is limited to integrating constants, first- and second-order moments, fractiles, or complete univariate marginal density functions, and all elements of ζ^1 are grouped in the last equation, following the discussion in Subsection 3.3.

As we know from the last subsection, the option of performing the Monte Carlo drawings on Σ is not presently available for the general SEM, indicated in the table by "N.A.". Other abbreviations used in the table are "NI" for noninformative, "N" for a Normal, and "St" for a Student t density.

As we implicitly define SEM in the table as a system leading to a nonconstant Jacobian $|B|^T$, and since ζ^1 is defined as the vector of coefficients that $|B|$ can possibly depend on, we feel that SEM is incompatible with the case where $\lambda_1 = 0$, as this would essentially imply we are in the case where $|B| = 1$ (or any other constant). The latter case is, computationally, the most simple one, and was analysed in Steel (1988), where attention is focused on the SURE model, i.e. B is taken to be the identity matrix. From our more general framework here it becomes obvious that these simple analytical results carry over to the case where B is triangular with constant diagonal elements, provided, of course, the RLI assumption continues to hold (i.e. $\lambda_1 = 0$). The case where $\lambda_n^1 > 0$ and we are noninformative on ζ_n^1 , given $\zeta_{(n-1)}^1$, is, in addition, equivalent to including ζ_n^1 in ζ^2 for triangular or SURE systems, which means the two lower left entries in the table coincide.

Remark that the Monte Carlo strategy as given in Subsection 3.3 strictly applies only to SEM systems with $\lambda_1 > \lambda_n^1 > 0$, the most complicated case, but that it can easily be adapted to cover the other, less demanding, cases appearing in Table 1 (barring the last column, of course).

The different entries in the table will also entail different consequences for the existence of posterior moments. The treatment of ζ^2 , however, is similar for all cases where we draw on the coefficients, and from (3.2) we obtain that the conditional posterior moments of ζ^2 , given ζ^1 , will exist at least up to order

Table 1. Overview of Posterior Inference

Monte Carlo (MC) on the coefficients (Subsections 3.1-3.3)					MC on Σ (3.4)
$\lambda_1 = 0$		$\lambda_1 = \lambda_n^1 > 0$		$\lambda_1 > \lambda_n^1 > 0$	
$D(\mathcal{Z}_n^1 \mathcal{Z}_{(n-1)}^1)$ $D(\mathcal{Z}_{(n-1)}^1)$	--	NI	N or St	NI	NI or N
	--	--	--	free	NI or N
SEM	--	1-1 poly-t	2-1 poly-t	1-1 poly-t	2-1 poly-t
	--	PTD or direct simulation	PTD or MC $[\lambda_n^1]$	direct simulation + MC $[\lambda_1 - \lambda_n^1]$	MC $[\lambda_1]$
$ B = 1$	analytically [see Steel (1988)]	St	2-0 poly-t	St	2-0 poly-t
		analytically	PTD or direct simulation	MC $[\lambda_1 - \lambda_n^1]$	direct simulation + MC $[\lambda_1 - \lambda_n^1]$
					MC $[\ln(n+1)]$

$$\rho < \nu_* - n + 1$$

under the RENC prior specification or its limiting noninformative case, and up to order

$$\rho < \mu_0 + T - 2n - \max_{i:1 \rightarrow n} \{\lambda_{i-1}^2\} - 1$$

if we choose the usual noninformative density in (2.39), which ensures the existence of posterior moments of order

$$\rho < \mu_0 + T - 2n - \lambda_n^2,$$

in view of the RLI assumption (i.e. $\max_i \{\lambda_i^2\} = \lambda_n^2$).

Of course, (3.2) is in terms of γ_i instead of ζ_i^2 , but it can easily be shown that moments for γ_i and for ζ_i^2 should exist up to the same order. Following Drèze (1977), we can also give sufficient conditions for the existence of posterior moments for ζ_n^1 , conditional on $\zeta_{(n-1)}^1$, provided its prior density in (2.41) is of the Student t form with τ_n^0 degrees of freedom and a PDS precision matrix. Assuming that q_n^* can be written as a non-degenerate kernel,⁷⁾ we obtain the existence results grouped in Table 2, where both the RENC prior structure as in (2.31)-(2.33), as the class of diffuse prior densities in (2.39) are considered, and we draw in the space of the coefficients.

Table 2. Existence of Posterior Moments for $\zeta_n^1 | \zeta_{(n-1)}^1$

$D(\zeta^2, \Sigma)$	RENC (2.31)-(2.33)		NI (2.39)	
	NI	St	NI	St
$D(\zeta_n^1 \zeta_{(n-1)}^1)$				
SEM	$\nu_0 - \lambda_n^1$	$\tau_n^0 + \nu_0$	$\mu_0 - n$ $- \lambda_n^2 - 1$	$\tau_n^0 + \mu_0 - n$ $- \lambda_n^2 - 1$
$ B = 1$	$\nu_0 - \lambda_n^1 + T$	$\tau_n^0 + \nu_0 + T$	$\mu_0 - n$ $- \lambda_n^2 + T - 1$	$\tau_n^0 + \mu_0 - n$ $- \lambda_n^2 + T - 1$

Conditional posterior moments will exist up to any order smaller than the expressions recorded in the table, where we have used the same abbreviations as in Table 1, and we have defined $\lambda_n = \lambda_n^1 + \lambda_n^2$.

4. Concluding Remarks

The aim of the present paper was to reduce the numerical complexity of the multivariate integrations required for Bayesian inference outside the restrictive natural conjugate framework.

This seems an essential requirement if we wish to extend the domain of application of Bayesian methods to models of some empirical relevance, like the popular simultaneous equation model discussed here.

Provided our prior assumptions are formulated within a Normal and Inverted-Wishart framework, we can treat some of the parameters analytically and we typically have a choice as to which parameters we integrate out first in this analytical fashion.

- (i) Given prior independence between ζ and Σ , and an Inverted-Wishart prior structure on Σ , we know its posterior density will be of the same functional form and we can treat it analytically. In the RENC prior density assumed here, however, ζ^2 does depend on Σ , leading to

an Inverted-Wishart result for Σ only under matrix-form restrictions on A^2 (i.e. the natural conjugate case). Nevertheless, an RLI structure on A^2 proves to be sufficient for analytical integration of Σ , exploiting its recursive decomposition and treating the system equation by equation.

The specific RENC prior density for ζ^2 in (2.33) then allows for its analytical treatment, using the recursive techniques explained in Steel (1988), given that ζ^2 does not intervene in the Jacobian of the model. Of course, we shall try to handle as many regression coefficients as possible by this analytical method.

The next step exploits the poly-t structure of the remaining parameters in one equation, given those in the rest of the model.

Finally, the latter are treated by Monte Carlo procedures using importance sampling.

- (ii) An alternative integration sequence proceeds from first using the Normal prior structure on ζ^2 to integrate it out analytically, given all other parameters, and then attempt to handle ζ^1 in an analytical way, given Σ . The snag of this strategy is that the latter integration can, at least presently, only be performed for triangular or SURE systems, seriously restraining its applicability.

An advantage, on the other hand, might be that the resulting Monte Carlo analysis of Σ could prove rather reliable and efficient, as suggested by Richard and Steel (1988).

Note that, in this approach, the RLI assumption becomes irrelevant as recursivity is not exploited here. Only in the limit case of matrix-form restrictions (natural conjugate⁸) will simplifications occur (to the well-known analytically tractable natural conjugate densities).

Although we have not yet undertaken any serious empirical work based on the strategies advocated in this paper, we can give some indication of the computational gain through approach (i) by considering the dimension of the Monte Carlo integrations, which, typically, bear the brunt of the hardware requirements.

Firstly, we consider a very small macroeconomic model from Johnston (1963), which was used in the seminal paper of Klok and van Dijk (1978).

Written in terms of deviations around the mean, as in Bauwens (1984), the model contains two behavioural equations (for consumption and investment) with three regression coefficients, and one identity. Of these three parameters one can be handled analytically in our framework, and one by poly-t based methods, which leaves just a simple univariate numerical integration.

A second model is Klein's (1950) Model I, which has been examined in reduced form by means of an extended natural conjugate approach in Harkema (1971) and a Monte Carlo analysis of which can be found in e.g. van Dijk and Kloek (1980) and Bauwens (1984). This small macroeconomic model comprises seven equations, three of which are stochastic and contain nine free coefficients. Solving out for a system explaining consumption, investment and private wage income, we can integrate out four coefficients analytically and one with the help of poly-t algorithms, which implies only four dimensions for the Monte Carlo integration.

The dimensional gain is thus considerable for both models examined here, and, in general, we expect our hybrid method in (i) to be much more efficient than Monte Carlo methods on the space of all regression coefficients.

This argument will, typically, also hold for the method in (ii) [see Steel (1987)], but additional work is needed to apply this approach to SEM systems.

An additional reason for preferring a partly analytical approach to a fully numerical one is, of course, the reliability of the inference, which, in a numerical Monte Carlo approach can always be subject to unexpected changes due to undominated⁹⁾ tails of the posterior density or an insufficient number of drawings. The latter argument is valid irrespective of computational considerations, and may, in smaller models, even be the main attraction of the hybrid approach proposed here.

Appendix. Rank assumptions made throughout the paper

model	prior	posterior	assumption
B [2.1]			nonsingular
$W_{(i)}^2$ [2.1]			full column rank
W_i^2 [2.1;3.2]			full column rank
R, R^j [2.1]			full column rank
$(\Rightarrow R_{ii}^j$ [2.1])			full column rank
$(\Rightarrow F, F_j$ [2.1;3.2])			full column rank
Σ [2.1;2.2]	Σ_0 [2.3]	G_i^* [3.4]	PDS
	N_i^0 [2.3]		PDS
	$(\Rightarrow H_i^0$ [2.3])	H_i^* [3.1]	PDS
	$(\Rightarrow M_0^2$ [3.4])	M_*^2 [3.4]	PDS
	N_0 [2.3;3.2]	N_* [3.2]	PDS
		M_*^1 [3.4]	PDS

Note: numbers in square brackets denote the relevant subsection(s). Implications of assumptions are given in parentheses. Indices j can be either 1 or 2 and $i : 1 \rightarrow n$.

Footnotes

1. It proves notationally convenient to define these recursions for $i : 1 \rightarrow n$ and thus allow $R_{(0)}^j$, $R_{1(0)}^j$, $\zeta_{(0)}^j$, and $A_{(0)}^j$ ($j=1,2$) to be of zero dimension.
2. One might think that evaluating the conditional moments for several values of ζ^1 (e.g. the posterior mean) could give a rough idea of marginal results.
This can, however, be extremely deceptive, especially for moments of order greater than one. Only if the prior information on (ζ^2, Σ) would strongly dominate the sample information, this might be a useful procedure, in view of the prior independence between (ζ^2, Σ) and ζ^1 in (2.42).
3. This becomes even more obvious for $N_0 = 0$, i.e. a diffuse prior structure on ζ^2 , when we obtain $|\Sigma_0 + A^{1'} W' M_F W A^1|$ for the determinant in (3.8), with $M_F = I_T - W F_2 (F_2' W' W F_2)^{-1} F_2' W'$, where the inverse exists due to the full column rank of $W_1^2 = W F_2$, $\forall i : 1 \rightarrow n$. Of course, (3.8) only becomes exactly equal to the result of Drèze and Richard (1983) if $F_2 = 0$, i.e. if we group all the free coefficients in ζ^1 , as they do.
4. For definitions of poly-t densities and an overview of their applications in econometrics, we refer the reader to the seminal paper of Drèze (1977). Suffice it to say here that the kernel of a general m-n poly-t density is defined as the ratio of the product of m Student t kernels over the product of n such kernels.
5. A triangular system is here taken to imply a triangular structure on the endogenous variables, i.e. B is upper- or lower-triangular, usually normalized to have ones on the diagonal, but the error covariance structure is still of a general form with $\Sigma = B'VB$ as in (2.3), instead of the independence inherent in a recursive system as in Subsection 2.2.

6. The use of the Moore-Penrose inverse in (3.32) is designed to allow for partially noninformative prior beliefs, where the corresponding rows and columns of M_0^1 can be fixed at zero.

7. This amounts to assuming that the relative precision matrix

$$W_n^{1'} [I - Z_n (H_n^0 + Z_n' Z_n)^{-1} Z_n'] W_n^1$$

is nonsingular. From Drèze (1977) we know that only the nondegenerate kernels from the numerator will contribute to these sufficient conditions for the existence of moments.

8. In order to find the exact natural conjugate analytical results we also need to impose a prior dependence of ξ^1 on Σ , if we wish to be informative on ξ^1 .

9. This refers to directions in which the far tail of the posterior density is located above the importance function. This situation may lead to tail ratios of posterior density over importance function values [i.e. the weights as defined in (3.13) or (3.14) here] that are many orders of magnitude larger than the ratios in regions of higher posterior density, attributing an overwhelming weight to a small number of drawings (sometimes only one) in the far tails. Of course, if the number of drawings goes to infinity, even such low probability drawings will cancel out, but in practice the size of the Monte Carlo runs is finite and serious biasing of the results may occur.

References

- Bauwens, L., 1984, Bayesian full information analysis of simultaneous equation models using integration by Monte Carlo (Springer Verlag, Berlin).
- Bauwens, L., J.P. Bulteau, P. Gille, L. Longree, M. Lubrano and H. Tompa, 1981, Bayesian regression program (BRP) user's manual, CORE Computing Report 81-A-01 (Université Catholique de Louvain, Louvain-la-Neuve).
- Bauwens, L. and J.F. Richard, 1985, A 1-1 poly-t random variable generator with applications to Monte Carlo integration, Journal of Econometrics 29, 19-46.
- Drèze, J.H., 1962, The Bayesian approach to simultaneous equations estimation, ONR research memorandum no. 67 (The Technological Institute, Northwestern University, Evanston, IL).
- Drèze, J.H., 1977, Bayesian regression analysis using poly-t densities, Journal of Econometrics 6, 329-354.
- Drèze, J.H. and J.F. Richard, 1983, Bayesian analysis of simultaneous equation systems, Ch. 9 in: Z. Griliches and M.D. Intriligator, eds., Handbook of Econometrics (North-Holland, Amsterdam).
- Geweke, J., 1987, Bayesian inference in econometric models using Monte Carlo integration, ISDS discussion paper no. 87-02 (Duke University, Durham, NC).
- Geweke, J., 1988, Antithetic acceleration of Monte Carlo integration in Bayesian inference, Journal of Econometrics 38, 73-89.
- Haavelmo, T., 1943, The statistical implications of a system of simultaneous equations, Econometrica 11, 1-12.

Hammersley, J.M. and D.C. Handscomb, 1979, Monte Carlo methods (Chapman and Hall, London).

Harkema, R., 1971, Simultaneous equations: A Bayesian approach (Universitaire Pers, Rotterdam).

Hendry, D.F., 1984, Monte Carlo experimentation in econometrics, Ch. 16 in: Z. Griliches and M.D. Intriligator, eds., Handbook of Econometrics (North-Holland, Amsterdam).

Johnston, J.J., 1963, Econometric methods (McGraw-Hill, New York).

Klein, L.R., 1950, Economic fluctuations in the United States, 1921-1941 (Wiley, New York).

Kloek, T. and H.K. van Dijk, 1978, Bayesian estimates of equation system parameters: An application of integration by Monte Carlo, Econometrica 46, 1-19.

Kolmogoroff, A., 1946, Grundbegriffe der Wahrscheinlichkeitsrechnung (Chelsea, New York).

Morales, J.A., 1971, Bayesian full information structural analysis (Springer Verlag, Berlin).

Naylor, J.C. and A.F.M. Smith, 1988, Econometric illustrations of novel numerical integration strategies for Bayesian inference, Journal of Econometrics 38, 103-125.

Poirier, D.J., 1988, Frequentist and subjectivist perspectives on the problems of model building in economics, Economic Perspectives 2, 121-144.

Richard, J.F., 1973, Posterior and predictive densities for simultaneous equation models (Springer-Verlag, Berlin).

Richard, J.F., 1984, Classical and Bayesian inference in incomplete simultaneous equation models, Ch. 4 in: D.F. Hendry and K.F. Wallis, eds., Econometrics and quantitative economics (Basil Blackwell, Oxford).

Richard, J.F. and M.F.J. Steel, 1988, Bayesian analysis of systems of seemingly unrelated regression equations under a recursive extended natural conjugate prior density, Journal of Econometrics 38, 7-37.

Richard, J.F. and H. Tompa, 1980, On the evaluation of poly-t density functions, Journal of Econometrics 12, 335-351.

Rothenberg, T., 1963, A Bayesian analysis of simultaneous equation systems, Econometric Institute report no. 6315 (Erasmus Universiteit, Rotterdam).

Spanos, A., 1986, Statistical foundations of econometric modelling (Cambridge University Press, Cambridge).

Steel, M.F.J., 1987, A Bayesian analysis of multivariate exogeneity: A Monte Carlo approach. Nouvelle série no. 168 (Université Catholique de Louvain, Louvain-la-Neuve).

Steel, M.F.J., 1988, Seemingly unrelated regression equation systems under diffuse stochastic prior information: A recursive analytical approach, CentER Discussion Paper 8805 (Katholieke Universiteit Brabant, Tilburg).

Strang, G., 1980, Linear algebra and its applications (Academic Press, New York).

Theil, H., 1971, Principles of econometrics (Wiley, New York).

van Dijk, H.K., 1984, Posterior analysis of econometric models using Monte Carlo integration (Erasmus Universiteit, Rotterdam).

Zellner, A., 1965, Bayesian and non-Bayesian analysis of simultaneous equation models, Paper presented at the First World Congress of the Econometric Society, Rome.

Zellner, A., L. Bauwens and H.K. van Dijk, 1988, Bayesian specification analysis and estimation of simultaneous equation models using Monte Carlo methods, Journal of Econometrics 38, 39-72.

Discussion Paper Series, CentER, Tilburg University, The Netherlands:

No.	Author(s)	Title
8801	Th. van de Klundert and F. van der Ploeg	Fiscal Policy and Finite Lives in Interdependent Economies with Real and Nominal Wage Rigidity
8802	J.R. Magnus and B. Pesaran	The Bias of Forecasts from a First-order Autoregression
8803	A.A. Weber	The Credibility of Monetary Policies, Policy-makers' Reputation and the EMS-Hypothesis: Empirical Evidence from 13 Countries
8804	F. van der Ploeg and A.J. de Zeeuw	Perfect Equilibrium in a Model of Competitive Arms Accumulation
8805	M.F.J. Steel	Seemingly Unrelated Regression Equation Systems under Diffuse Stochastic Prior Information: A Recursive Analytical Approach
8806	Th. Ten Raai and E.N. Wolff	Secondary Products and the Measurement of Productivity Growth
8807	F. van der Ploeg	Monetary and Fiscal Policy in Interdependent Economies with Capital Accumulation, Death and Population Growth
8901	Th. Ten Raai and P. Kop Jansen	The Choice of Model in the Construction of Input-Output Coefficient Matrices
8902	Th. Nijman and F. Palm	Generalized Least Squares Estimation of Linear Models Containing Rational Future Expectations
8903	A. van Soest, I. Woittiez, A. Kapteyn	Labour Supply, Income taxes and Hours Restrictions in The Netherlands
8904	F. van der Ploeg	Capital Accumulation, Inflation and Long-Run Conflict in International Objectives
8905	Th. van de Klundert and A. van Schaik	Unemployment Persistence and Loss of Productive Capacity: a Keynesian approach
8906	A.J. Markink and F. van der Ploeg	Dynamic Policy of Linear Models with Rational Expectations of Future Events: A Computer Package
8907	J. Osiewalski	Posterior Densities for Nonlinear Regression with Equicorrelated Errors
8908	M.F.J. Steel	A Bayesian Analysis of Simultaneous Equation Models by Combining Recursive Analytical and Numerical Approaches

P.O. BOX 90153, 5000 LE TILBURG, THE NETHERLANDS

Bibliotheek K. U. Brabant

17 000 01117381 3