Bayesian Analysis of Cointegration in a Structural Error Correction Model*

Chew Lian Chua[†]
Melbourne Institute of Applied Economic and Social Research,
The University of Melbourne

Peter M. Summers[‡]
Melbourne Institute of Applied Economic and Social Research
and
Federal Reserve Bank of Kansas City

July 9, 2004

Abstract

This paper analyses cointegration in structural error correction models (SECM). We combine the Bayesian methods of Kleibergen and Paap (2002) for analysis of cointegration in the error correction model (ECM) with those of Waggoner and Zha (2003) for estimating the structural parameters in Bayesian structural vector autoregressions (BSVAR). When we apply these methods to simulated data, Bayes factors are able to select the appropriate cointegrating vectors, and the actual values of the structural parameters are well-estimated. We then analyse the data from King et. al. (1991), and find five cointegrating vectors among the six variables.

Key words: structural error correction model; cointegration; Bayesian; structural parameters; singular value decomposition.

JEL Classification: C11, C15, C32, C52.

*Work in progress, not to be cited yet. Both authors acknowledge the financial support from the Faculty of Economics and Commerce, The University of Melbourne. We retain sole responsibility for any errors.

†E-mail: mchua@unimelb.edu.au, Tel: +61 3 83442144.

‡E-mail: p.summers@unimelb.edu.au, Tel : +61 3 83442143.

1 Introduction

It is well known that when a vector autoregressive (VAR) model is represented as an error correction model, it can be employed for the analysis of cointegration. Pioneered by Granger (1981), Granger and Weiss (1983) and Engle and Granger (1987), cointegration is intuitively attractive for economic modelling, particularly for macroeconomic time series. Because it allows clear differentiation between the short-run variation and adjustment towards long-run equilibrium among the economic series, and because the estimated long-run relationships can often be given a theoretical interpretation, cointegrated VAR models have become one of the major workhorses in applied macroeconomics. Recent work in this area has extended and refined Bayesian methods of analysis for cointegrated VARs. Kleibergen and Paap (2002), Sugita (2001), and Amisano (2003) discuss different ways of detecting the presence of cointegration. Strachan (2003) provides a procedure for providing valid estimates for the cointegrating vectors. Kleibergen and van Dijk (1994) discuss the consequences of local non-identification. Bauwens and Giot (1997) use Gibbs sampling to estimate the cointegrating relations.

If, however, one is interested in the structural analysis, one can only analyse the structural from the reduced-form parameters (See Johansen and Juselius (1992), Johansen (1995), Fisher et. al. 2000). Just like any system of simultaneous equations, this approach is characterized by an identification problem; there is not enough information in the reduced model to recover all the structural parameters. If one is to follow Sims's (1980) type of identification, i.e. recovering the structural parameters from the reduced-form covariance matrix, at least $n(n-1)/2$ structural parameters have to be restricted before the rest of the structural parameters can be imputed, where n is the number of equations in the system. Recent papers on structural vector autoregression (SVAR) models by Leeper et. al. (1996), Sims and Zha (1998), Sims and Zha (1999) and Waggoner and Zha (2003) allow the structural parameters in SVARs to be estimated directly using Bayesian methods.

In this paper, we focus on the analysis of cointegration and the structural parameters for a SVAR model in a Bayesian framework. We represent the SVAR as a structural error correction model , and show that the analysis of cointegration for the SECM follows that of the ECM. We adopt the methods of Kleibergen and Paap (2002) in analysing cointegration in the ECM and apply it to the SECM. A parameter is constructed from a singular value decomposition that reflects the presence of cointegration. For the estimation of structural parameters, we employ the methods of Waggoner and Zha (2003) in which a Gibbs sampler is used to estimate the structural parameters in a BSVAR. Two by-products arise from this paper. First, we generalise a theorem of Waggoner and Zha (2003) to allow for non-zero prior means for the structural parameters. Second, we provide a more efficient way drawing from the posterior distributions in which we circumvent the use of the Metropolis-Hastings algorithm.

The organisation of this paper is as follows. In Section 2, we present the structural error correction model. We derive the SECM from the SVAR model. In Section 3, we apply the methods of Kleibergen and Paap (2002) to analyse the presence of cointegration. We specify priors for the parameters of the linear SECM, and translate these priors to those of the unrestricted and the restricted SECM. We then derive the posterior distributions for the parameters of the different models. In Section 4, we provide a Gibbs sampler for drawing from the posterior pdfs of the structural parameters. We consider selecting the possible number of cointegration relations using Bayes factors in Section 5. In Section 6, we provide two illustrations of the model. The first is a simulated example, while in the second example we apply the model to the data analyzed by King, Plosser, Stock and Watson (1991). Section 7 concludes the paper.

2 Structural Error Correction Model

This section shows how a SVAR can be represented as a SECM. One advantage of this specification is that both contemporaneous relationships among the endogenous variables and the analysis of cointegration can be presented in a single model. Consider a structural vector autoregressive model

$$A_0' Y_t = d + \sum_{i=1}^{k} A_i' Y_{t-i} + \varepsilon_t, \qquad t = 1, 2, ..., T, \tag{1}$$

where k is the number of lags, $Y_t, Y_{t-1}, ..., Y_{t-k}$ are $(n \times 1)$ vectors of observations, A_0 is a $(n \times n)$ structural coefficient matrix, $A_1, A_2, ..., A_k$ are $(n \times n)$ lag coefficient matrices, d is a $(n \times 1)$ vector of constant terms, and ε_t is a vector of i.i.d structural shocks that is assumed to be distributed as

$$\varepsilon_t | Y_{t-s} \sim N(0, I) \qquad for\ 0 < s < t.$$

By subtracting $A_0' Y_{t-1}$ from both sides of equation (1), and adding and subtracting $\sum_{i}^{k-1} \sum_{j=i}^{k} A_j' Y_{t-i}$ on the right hand side of equation (1), we arrive at a structural error correction model (SECM)[1].

[1] Alternatively, Johansen (1995) shows that the SECM can be derived from the Error Correction Model (ECM). The ECM takes the following form

$$\triangle Y_t = c + \varsigma' \beta' Y_{t-1} + \sum B_i^{*\prime} \triangle Y_{t-i} + \mu_t$$

where $\mu_t = (A_0')^{-1} \varepsilon_t$, ς' is the adjustment matrix. Multiply the ECM by A_0' to get the SECM

$$A_0'\triangle Y_t = d + \Pi' Y_{t-1} + \sum_{i=1}^{k-1} B_i' \triangle Y_{t-i} + \varepsilon_t, \tag{2}$$

where $\Pi' = \sum\limits_{j=1}^{k} A_j' - A_0'$, and $B_i' = -\sum\limits_{j=i+1}^{k} A_j'$, $i = 1, ..., k-1$. The characteristic equation for equation (2) is

$$Y_t = (A_0^{-1})' \sum_{i=1}^{k} A_i' Y_{t-i}. \tag{3}$$

Denoting L to be the lag operator, such that $L^k Y_t = Y_{t-k}$, then the above equation can be written as

$$((A_0^{-1})' \sum_{i=1}^{k} A_i' L^i - I_n) Y_t = 0.$$

To determine the existence of cointegration, we evaluate the rank of

$$(A_0^{-1})' \sum_{i=1}^{k} A_i' - I_n. \tag{4}$$

If equation (4) has zero rank, the series Y_t contains n unit roots. On the other hand, if it has full rank n, the univariate series in Y_t are all stationary. Cointegration is present only when the rank of equation (4) lies between 0 and n. Equation (4) can be re-arranged into $rank\left((A_0^{-1})'(\sum\limits_{i=1}^{k} A_i' - A_0')\right)$ and since $\sum\limits_{i=1}^{k} A_i' - A_0' = \Pi'$, it simplifies to

$$rank\left((\Pi A_0^{-1})'\right).$$

The rank is clearly dependent on

$$rank\left(\Pi A_0^{-1}\right) \leq \min\left\{rank(A_0^{-1}), rank(\Pi)\right\}.$$

Since by assumption A_0^{-1} is a nonsingular matrix and so has full rank, the determination of cointegration depends solely on Π. Henceforth, the analysis of cointegration follows that of the ECM. We can decompose Π into two full $(n \times r)$ matrices, β and α' :

$$\Pi = \beta\alpha,$$

$$A_0' \triangle Y_t = d + \alpha'\beta' Y_{t-1} + \sum B_i' \triangle Y_{t-i} + \varepsilon_t$$

where $d = A_0'c$, $\alpha' = A_0'\varsigma'$and $B_i' = A_0' B_i^{*\prime}$. Note that β is invariant to this transformation.

where β contains the cointegration vectors that reflects the long-run relations between the univariate series in Y_t, and α is the adjustment matrix that indicates the speed of adjustment to the equilibria $\beta' Y_t$.

Rewriting the SECM in matrix notation gives

$$\triangle Y A_0 = Y_{-1}\Pi + XB + \varepsilon, \tag{5}$$

where $\triangle Y = (\triangle Y_{k+1} \ ... \ \triangle Y_T)'$, $Y_{-1} = (Y_k \ ... \ Y_{T-1})'$, $\varepsilon = (\varepsilon_{k+1} \ ... \ \varepsilon_T)'$, $X = (X'_{k+1} \ ... \ X'_T)'$, $X_t = (1 \ \triangle Y'_{t-1} \ ... \ \triangle Y'_{t-k+1})$, $B = (d \ B_1 \ ... \ B_{k-1})'$, B is a $(q \times n)$ matrix, and $q = (k-1)n + 1$. As is well known, the individual parameters in $\beta\alpha$ (for a given rank r), are only identified up to a rotation. Normalisation is carried out so that α and β are estimable. One common way of normalising α and β is

$$\beta = \begin{pmatrix} I_r \\ -\beta_2 \end{pmatrix}. \tag{6}$$

3 Method of Kleibergen and Paap: Singular Value Decomposition Approach

Following the methods of Kleibergen and Paap, we decompose Π as follows:

$$\begin{aligned} \Pi &= \beta\alpha + \beta_\perp \lambda \alpha_\perp \qquad (7) \\ &= \begin{pmatrix} \beta & \beta_\perp \end{pmatrix} \begin{pmatrix} I_r & 0 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} \alpha \\ \alpha_\perp \end{pmatrix}. \end{aligned}$$

The attractiveness of this decomposition is that when λ is restricted to zero it reflects the presence of cointegration. The matrices $\alpha_\perp$and $\beta_\perp$ are perpendicular to α and β (i.e. $\alpha_\perp \alpha' \equiv \beta' \beta_\perp \equiv 0$), and $\alpha_\perp \alpha'_\perp \equiv \beta'_\perp \beta_\perp \equiv I_{n-r}$. The decomposition in equation (7) corresponds to a singular value decomposition of Π, which is

$$\Pi = USV' \tag{8}$$

where U and V are $(n \times n)$ orthonormal matrices ($U'U = V'V = I_n$), and S is an $(n \times n)$ diagonal matrix containing the non-negative singular values of Π (in decreasing order). Partition U, S and V, respectively, as

$$U = \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix}, S = \begin{pmatrix} S_1 & 0 \\ 0 & S_2 \end{pmatrix}, \text{ and } V = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}.$$

Kleibergen and Paap (2002, Appendix A) show that

$$\begin{aligned} \alpha &= U_{11} S_1 (V'_{11} V'_{21}), \qquad (9) \\ \beta_2 &= -U_{21} U_{11}^{-1}, \qquad (10) \\ \lambda &= (U'_{22} U_{22})^{-\frac{1}{2}} U_{22} S_2 V'_{22} (V_{22} V'_{22})^{-\frac{1}{2}}. \qquad (11) \end{aligned}$$

The number of non-zero singular values in S determines the rank of Π and hence the number of cointegrating vectors. When λ is restricted to zero, and because $(U_{22}U_{22}')^{-\frac{1}{2}}U_{22}$ and $V_{22}'(V_{22}V_{22}')^{-\frac{1}{2}}$ are orthonormal matrices, this means that S_2 is restricted to zero as well (and so implies rank reduction in Π.)

The model in equation (5) can be reparameterised as

$$\triangle Y A_0 = Y_{-1}\beta\alpha + Y_{-1}\beta_{\perp}\lambda\alpha_{\perp} + XB + \varepsilon. \tag{12}$$

Using the terminology of Kleibergen and Paap, from here on, when λ is not restricted equation (12) will be known as an unrestricted SECM. When λ is restricted to zero it will be denoted as a restricted SECM, and we will refer to equation (5) as a linear SECM.

3.1 Prior Specification

Denote a_i, b_i, and π_i to be the i^{th} column of A_0, B and Π respectively. We assume that the joint prior pdfs for the parameters of the models are

$$\begin{aligned} p(\Pi, A_0, B) &= \prod_{i=1}^{n} p(\pi_i)p(a_i)p(b_i|a_i) && \text{for linear SECM,} \\ p(\alpha, \lambda, \beta_2, A_0, B) &= p(\alpha, \lambda, \beta_2)\prod_{i=1}^{n} p(a_i)p(b_i|a_i) && \text{for unrestricted SECM,} \\ p(\alpha, \beta_2, A_0, B) &= p(\alpha, \beta_2)\prod_{i=1}^{n} p(a_i)p(b_i|a_i) && \text{for restricted SECM.} \end{aligned}$$

3.1.1 Prior for (a_i, b_i)

The priors for a_i and b_i are specified such that $p(a_i)$ and $p(b_i|a_i)$ are multivariate normal distributions.

$$\begin{aligned} a_i &\sim N(\overline{a}_i, \overline{O}_i), \\ b_i|a_i &\sim N(\overline{P}_i a_i, \overline{H}_i), \end{aligned}$$

where $\overline{a}_i$ is an $(n \times 1)$ vector of prior means of a_i, $\overline{O}_i$ is an $(n \times n)$ prior covariance matrix.of a_i, $\overline{H}_i$ is a $(q \times q)$ conditional prior covariance matrix of $b_i, \overline{P}_i$ is a $(q \times n)$ matrix that allows for different interactions of a_i. If $\overline{P}_i$ is null, then $p(b_i|a_i)$ is independent of a_i. One advantage of having this prior specification is that the random walk for Bayesian SVARs can be applied, as in Sims and Zha (1998). Essentially, we can nudge the SECM towards a random walk model in $\triangle Y_{t-i}$. In this paper, we also adopt the methods of Waggoner and Zha (2003) dealing with linear parameter restrictions in the SVAR model. Instead of assuming the prior means of the structural parameters to be zero as in their paper, we extend the prior mean of A_0 to take a general form (having non-zero or zero mean).

Following Waggoner and Zha (2003) assume that some elements in a_i and b_i are restricted,

$$\begin{aligned} Q_i a_i &= 0, \\ R_i b_i &= 0, \end{aligned}$$

where Q_i is an $(n \times n)$ matrix of rank p_i, and R_i is a $(q \times q)$ matrix of rank r_i that impose the restrictions. We make the assumption that the diagonal elements of A_0 are unrestricted which guarantees that A_0 is a nonsingular matrix. Suppose that there exists F_i (an $(n \times p_i)$ matrix) and M_i (a $(q \times r_i)$ matrix) such that the columns of F_i and M_i, respectively, are orthonormal for the null space of Q_i and R_i. a_i and b_i can then be expressed as

$$\begin{aligned} a_i &= F_i \gamma_i, \\ b_i &= M_i g_i. \end{aligned}$$

Waggoner and Zha (2003) consider priors on γ_i of the form $\gamma_i \sim N(0, \widetilde{O}_i)$. We prove in Appendix A that if the prior is more general, the prior of γ_i and the conditional prior of g_i given γ_i are, respectively,

$$\gamma_i \sim N(\widetilde{F}_i \overline{a}_i, \widetilde{O}_i),$$

and

$$g_i | \gamma_i \sim N(\widetilde{P}_i \gamma_i, \widetilde{H}_i),$$

where $\widetilde{F}_i = \widetilde{O}_i F_i' \overline{O}_i^{-1}$, $\widetilde{O}_i = \left[F_i' \overline{O}_i^{-1} F_i + F_i' \overline{P}_i' \overline{H}_i^{-1} \overline{P}_i F_i - \widetilde{P}_i' \widetilde{H}_i^{-1} \widetilde{P}_i \right]^{-1}$, $\widetilde{H}_i = \left(M_i' \overline{H}_i^{-1} M_i \right)^{-1}$ and $\widetilde{P}_i = \widetilde{H}_i M_i' \overline{H}_i^{-1} \overline{P}_i F_i$.

3.1.2 Prior for π_i

The prior for π_i is assumed to be a multivariate normal distribution with mean $\overline{\pi}_i$ and an $(n \times n)$ covariance matrix $\overline{\Phi}_i$.

$$\pi_i \sim N(\overline{\pi}_i, \overline{\Phi}_i).$$

For ease of derivation of the prior for (α, λ, β) and (α, β), we express the above equation in term of Π

$$\begin{aligned} p(\Pi) &= \prod_{i=1}^{n} p(\pi_i) \\ &= (2\pi)^{-\frac{1}{2}n^2} \left| \overline{\Sigma}_\Pi \right|^{-\frac{1}{2}} \exp \left[-\frac{1}{2} (vec(\Pi) - \overline{\pi})' \overline{\Sigma}_\Pi^{-1} (vec(\Pi) - \overline{\pi}) \right], \end{aligned}$$

where $\overline{\pi} = \begin{pmatrix} \overline{\pi}_1' & ... & \overline{\pi}_n' \end{pmatrix}'$ and $\overline{\Sigma}_\Pi = diag(\overline{\Phi}_1, ..., \overline{\Phi}_n)$.

3.1.3 Prior for (α, λ, β)

Instead of placing priors on α, λ and β_2, the joint prior for (α, λ, β) can be derived from $p(\Pi)$

$$\begin{aligned} p(\alpha, \lambda, \beta_2) &= p(\Pi) \left| \frac{\partial \Pi}{\partial(\alpha, \lambda, \beta_2)} \right| \\ &\propto \exp\left[-\frac{1}{2}(vec(\beta\alpha + \beta_\perp \lambda \alpha_\perp) - \overline{\pi})' \overline{\Sigma}_\Pi^{-1} (vec(\beta\alpha + \beta_\perp \lambda \alpha_\perp) - \overline{\pi}) \right] \left| \frac{\partial \Pi}{\partial(\alpha, \lambda, \beta_2)} \right|. \end{aligned}$$

3.2 Prior for (α, β)

In the restricted SECM, $p(\alpha, \beta_2)$ is a conditional prior of (α, λ, β) given $\lambda = 0$

$$\begin{aligned} p(\alpha, \beta_2) &= \frac{p(\alpha, \lambda, \beta_2)|_{\lambda=0}}{p(\lambda)|_{\lambda=0}} \left| \frac{\partial \Pi}{\partial(\alpha, \lambda, \beta_2)} \right| \Big|_{\lambda=0} \\ &\propto \exp\left[-\frac{1}{2}(vec(\beta\alpha) - \overline{\pi})' \overline{\Sigma}_\Pi^{-1} (vec(\beta\alpha) - \overline{\pi}) \right] |\beta'\beta|^{\frac{1}{2}(n-r)} |\alpha\alpha'|^{\frac{1}{2}(n-r)}. \end{aligned}$$

Here $p(\lambda)|_{\lambda=0} = \int\int p(\alpha, \lambda, \beta_2)|_{\lambda=0} \partial\alpha \partial\beta_2$ is a normalising constant, which plays a crucial role in the determination of cointegration as it is part of the marginal likelihoods. Since $p(\lambda)|_{\lambda=0}$ is analytically intractable, we estimate it using the simulation techniques of Chen (1994). See Appendix B for details. For a derivation of the Jacobian terms $\left| \frac{\partial \Pi}{\partial(\alpha,\lambda,\beta_2)} \right| \Big|_{\lambda=0}$ and $\left| \frac{\partial \Pi}{\partial(\alpha,\lambda,\beta_2)} \right|$, refer to Appendix B of Kleibergen and Paap (2002).

3.3 Likelihood functions

Linear SECM: The likelihood function for $\triangle Y_1$, $\triangle Y_2, \ldots \triangle Y_T$ conditional on the initial observations $\triangle Y_0$, $\triangle Y_{-1}, \ldots \triangle Y_{-k+2}$ is

$$\begin{aligned} p(\triangle Y | A_0, B, \Pi) &= (2\pi)^{-\frac{1}{2}Tn} |A_0|^T \times \\ &\quad \exp\left[-\frac{1}{2} tr\left((\triangle Y A_0 - Y_{-1}\Pi - XB)'(\triangle Y A_0 - Y_{-1}\Pi - XB) \right) \right], \end{aligned} \tag{13}$$

which can be written in terms of the free parameters,

$$\begin{aligned} p(\triangle Y | \gamma, g, \Pi) &\propto |[F_1\gamma_1| \ldots |F_n\gamma_n]|^T \times \\ &\quad \exp\left[-\frac{1}{2} \sum_{i=1}^{n} (\triangle Y F_i \gamma_i - Y_{-1}\pi_i - X M_i g_i)'(\triangle Y F_i \gamma_i - Y_{-1}\pi_i - X M_i g_i) \right], \end{aligned}$$

where $\gamma = (\begin{array}{ccc} \gamma_1' & \ldots & \gamma_n' \end{array})'$, and $g = (\begin{array}{ccc} g_1' & \ldots & g_n' \end{array})'$.

Unrestricted and Restricted SECM: The likelihood function for the unrestricted SECM is

$$p(\triangle Y|\gamma, g, \alpha, \lambda, \beta_2) = p(\triangle Y|\gamma, g, \Pi)|_{\Pi=\beta\alpha+\beta_\perp\lambda\alpha_\perp},$$

and the restricted model is

$$p(\triangle Y|\gamma, g, \alpha, \beta_2) = p(\triangle Y|\gamma, g, \Pi)|_{\Pi=\beta\alpha}.$$

3.4 Posterior Distributions and Sampling Schemes

We consider the respective posterior distributions for the parameters of the three models.

Linear SECM: It can be shown that the marginal posterior pdfs for γ is

$$p(\gamma|\triangle Y) \propto |[F_1\gamma_1|...|F_n\gamma_n]|^T \prod_{i=1}^{n} \exp\left(-\frac{T}{2}\left(\gamma_i - \widehat{F}_i\underline{\gamma}_i\right)' O_i^{-1}\left(\gamma_i - \widehat{F}_i\underline{\gamma}_i\right)\right),$$

the conditional posterior pdfs for Π given γ

$$p(\Pi|\gamma, \triangle Y) \propto \prod_{i=1}^{n} \exp\left(-\frac{1}{2}(\pi_i - Q_i\underline{\pi}_i)'\Phi_i^{-1}(\pi_i - Q_i\underline{\pi}_i)\right),$$

and the conditional posterior pdfs for g given Π and γ

$$p(g|\gamma, \Pi, \triangle Y) \propto \prod_{i=1}^{n} \exp\left[-\frac{1}{2}\left(g_i - P_i\underline{g}_i\right)' H_i^{-1}\left(g_i - P_i\underline{g}_i\right)\right].$$

In these expressions, we have used the following:

$$O_i^{-1} = \frac{1}{T}\widehat{O}_i^{-1},$$

$$\widehat{O}_i = \left(F_i'\triangle Y'\triangle Y F_i + \widetilde{O}_i^{-1} + \widetilde{P}_i'\widetilde{H}_i^{-1}\widetilde{P}_i - P_{1i}'H_i^{-1}P_{1i} - Q_{1i}'\Phi_i^{-1}Q_{1i}\right)^{-1},$$

$$\widehat{F}_i = \left[\begin{array}{cc}\widehat{F}_{1i} & \widehat{F}_{2i}\end{array}\right],$$

$$\widehat{F}_{1i} = \widehat{O}_i\widetilde{O}_i^{-1}\widetilde{F}_i,$$

$$\widehat{F}_{2i} = \widehat{O}_iQ_{1i}'\Phi_i^{-1}Q_{2i},$$

$$\underline{\gamma}_i = \left[\begin{array}{c}\overline{a}_i \\ \overline{\pi}_i\end{array}\right],$$

$$\Phi_i = \left(Y_{-1}'Y_{-1} + \overline{\Phi}_i^{-1} - P_{2i}'H_i^{-1}P_{2i}\right)^{-1},$$

$$Q_i = \left[\begin{array}{cc} Q_{1i} & Q_{2i} \end{array} \right],$$

$$Q_{1i} = \Phi_i Y_{-1}' \triangle Y F_i + P_{2i}' H_i^{-1} P_{1i},$$

$$Q_{2i} = \Phi_i \overline{\Phi}_i^{-1},$$

$$\underline{\pi}_i = \left[\begin{array}{c} \gamma_i \\ \overline{\pi}_i \end{array} \right],$$

$$H_i = \left(M_i' X' X M_i + \widetilde{H}_i^{-1} \right)^{-1},$$

$$P_i = \left[\begin{array}{cc} P_{1i} & P_{2i} \end{array} \right],$$

$$\underline{g}_i = \left[\begin{array}{c} \gamma_i \\ \pi_i \end{array} \right],$$

$$P_{1i} = H_i (M_i' X' \triangle Y F_i + \widetilde{H}_i^{-1} \widetilde{P}_i),$$

and

$$P_{2i} = -H_i M_i' X' Y_{-1}.$$

See Appendix C for derivation.

As $p(\Pi|\gamma, \triangle Y)$ and $p(g|\gamma, \Pi, \triangle Y)$ are multivariate normal, sampling of these distributions is straightforward. The only difficulty lies in sampling $p(\gamma| \triangle Y)$ because it is not of any known distribution. It turns out that when the prior mean of A_0 is equal to zero, the methods of Waggoner and Zha (2003) can be applied. They show that using a Gibbs sampler to draw from $p(\gamma| \triangle Y)$, drawing γ_i conditional on γ_j, $j \neq i$ is equivalent to drawing independently from a multivariate normal distribution with zero mean and variances $\frac{1}{T}$ and a univariate distribution which is equivalent to taking square roots of the draws from a gamma distribution.

We further show that when the prior mean of A_0 is not equal to zero results similar to those of Waggoner and Zha (2003) hold. Specifically, the draws of γ_i conditional on γ_j, $j \neq i$ are equivalent to independent draws from a multivariate normal distribution with a nonzero mean and variances $\frac{1}{T}$ and a univariate distribution. As far as we know the univariate distribution cannot be transformed into any recognisable form. In the next section, we provide a strategy for drawing from this univariate distribution. For the moment, let assume that it is possible to draw from the univariate distribution. A sampling scheme for the linear SECM is then

For $i = 1, ..., n$.

1. Specify starting values for γ_i.

2. Draw $\gamma_i^{(j+1)}$ from $p(\gamma_i| \triangle Y)$ using the methods described in section 4.

3. Draw $\pi_i^{(j+1)} \sim N(\underline{\pi}_i, \Phi_i|\gamma_i^{(j+1)})$ for $i = 1, ..., n$.

4. Draw $g_i^{(j+1)} \sim N(\underline{g}_i, H_i|\gamma_i^{(j+1)}, \pi_i^{(j+1)})$ for $i = 1, ..., n$.

5. Set $j = j + 1$. Return to step 2.

Unrestricted SECM: The posterior pdfs are similar to those found in the linear SECM, except that Π is expressed in terms of α, λ and β_2,

$$p(\alpha, \lambda, \beta_2|\gamma, \triangle Y) \propto p(\Pi|\gamma, \triangle Y))|_{\Pi=\beta\alpha+\beta_\perp\lambda\alpha_\perp} \left|\frac{\partial \Pi}{\partial(\alpha, \lambda, \beta_2)}\right|,$$

and

$$p(g|\gamma, \alpha, \beta_2, \triangle Y) \propto p(g|\gamma, \Pi, \triangle Y)|_{\Pi=\beta\alpha+\beta_\perp\lambda\alpha_\perp}$$

As shown by Kleibergen and Paap (2002),obtaining the draws for α, λ and β_2 are relatively straightforward, we simply decompose the draws of Π using equation (8) and compute α, λ and β_2 using (9), (10) and (11).

Restricted SECM: For the restricted SECM, the posterior pdfs are similar to those of the unrestricted SECM. In this case, λ is restricted to zero

$$p^*(\alpha, \beta_2|\gamma, \triangle Y) \propto p(\Pi|\gamma, \triangle Y))|_{\Pi=\beta\alpha} \left|\frac{\partial \Pi}{\partial(\alpha, \lambda, \beta_2)}\right|\Bigg|_{\lambda=0}, \tag{14}$$

$$p(g|\gamma, \alpha, \beta_2, \triangle Y) \propto p(g|\gamma, \Pi, \triangle Y)|_{\Pi=\beta\alpha},$$

where $\beta' = \begin{pmatrix} I_r & -\beta_2 \end{pmatrix}'$. As $p^*(\alpha, \beta_2|\gamma, \triangle Y)$ is not of any recognisable distribution, one approach to drawing this conditional distribution is to employ a Metropolis-Hastings (MH) algorithm for each draw of γ. This approach is rather inefficient due to the fact that for every draw the MH algorithm requires a burn-in period. A more efficient approach is to obtain the draws of α, β_2 from the singular value decomposition of Π. In Appendix D, we show that $p(\alpha, \lambda, \beta_2|\gamma, \triangle Y)$ can be expressed as $p(\lambda|\alpha, \beta_2, \gamma, \triangle Y) \times p^*(\alpha, \beta_2|\gamma, \triangle Y)$ which implies that α, β_2 can be obtained directly from the decomposition of Π.

4 Gibbs Sampler for $p(\gamma|\triangle Y)$

As mentioned in the previous section, when the prior mean of A_0 is zero, $p(\gamma|\triangle Y)$ is simulated using the Gibbs simulator of Waggoner and Zha. This section generalises Theorem 2 of Waggoner and Zha (2003) by allowing the prior mean of A_0 to be non-zero, and provides a sampler for the univariate distribution.

The generalised theorem is

Theorem 1 *The random vector γ_i conditional on $\gamma_1, ..., \gamma_{i-1}, \gamma_{i+1}, ..., \gamma_n$ with mean of $\widehat{F}_i\underline{\gamma}_i$ is a linear function of p_i independent random variables κ_j such that*

(a) the density function of κ_1 is proportional to

$$|\kappa_1|^T \exp\left(-\frac{T}{2}(\kappa_1 - \overline{\kappa}_1)^2\right), \tag{15}$$

(b) for $2 \leq j \leq p_i$, κ_j is normally distributed with mean $\overline{\kappa}_j$ and variance $\frac{1}{T}$.

(See appendix E for proof). Given that $\widehat{F_i \underline{\gamma}_i}$ is known, $\overline{\kappa}_i$ can be computed

$$\begin{bmatrix} \overline{\kappa}_1 \\ \vdots \\ \overline{\kappa}_{p_i} \end{bmatrix} = [w_1|\cdots|w_{p_i}]^{-1} T_i^{-1} \widehat{F_i \underline{\gamma}_i},$$

where T_i is the Choleski decomposition of O_i, w_j is constructed such that $F_i T_i w_j$ is perpendicular to the linear combination of $F_j a_j$, $j \neq i$. See Waggoner and Zha (2003, p357) for the construction of w_j. Note that, when $\widehat{F_i \underline{\gamma}_i}$ is zero, implying that $\overline{\kappa}_1, ..., \overline{\kappa}_{p_i}$ are zero, then the distribution (15) becomes that of Waggoner and Zha. We compute γ_i as

$$\gamma_i = T_i \sum_{j=1}^{p_i} \kappa_j w_j.$$

Drawing from $|\kappa_1|^T \exp\left(-\frac{T}{2}(\kappa_1 - \overline{\kappa}_1)^2\right)$ is not feasible. It cannot be translated to any recognisible distribution except when $\overline{\kappa}_1$ is zero, when it becomes the square root of a gamma distribution. To analyse the properties of the distribution of κ_1, we plot this distribution under various values $\overline{\kappa}_1 = 0, 0.1, 0.2, 0.3$ and 0.4 with $T = 10$ (see Figure 1). It is observed that

- Starting from $\overline{\kappa}_1 = 0$, the distribution is a symmetric distribution.
- When $\overline{\kappa}_1 \neq 0$, there exists a dominant mode. The region in which the dominant mode lies corresponds to the sign of $\overline{\kappa}_1$.
- As $|\overline{\kappa}_1|$ increases, the bimodal distribution tends toward a unimodal distribution When $|\overline{\kappa}_1| > 3T^{-1}$ the less dominant mode becomes insignificant.
- The distribution is discontinuous at $\kappa_1 = 0$. This means that the normalising constants (area under the curve) are approximately.

$$\begin{aligned} C &= \int_{-\infty}^{\approx 0} |\kappa_1|^T \exp\left(-\frac{T}{2}(\kappa_1 - \overline{\kappa}_1)^2\right) dx + \int_{\approx 0}^{\infty} |\kappa_1|^T \exp\left(-\frac{T}{2}(\kappa_1 - \overline{\kappa}_1)^2\right) dx \\ &= C_- + C_+. \end{aligned}$$

- $C_- > C_+$ if $\overline{\kappa}_1 < 0$, and vice versa, and $C_- = C_+$ if $\overline{\kappa}_1 = 0$.
- Both tails decay towards zero as κ_1 goes to $-\infty$ or $+\infty$.

Using the above observed information, we construct a sampler to draw from equation (15) that is based on the concepts of the Griddy-Gibbs sampler of Ritter and Tanner (1992). The sampler is as follows:

1. Define the range R_L.and R_H, where R_L is the lower limit and R_H is the upper limit.

2. Specify a set of gridpoints $(k_1 < k_2 < ..... < k_J)$ within R_L and R_H.

3. Using numerical integration methods (such as Simpson's rule), compute the area between k_j and k_{j+1} for $j = 1, 2, ...J - 1$:

$$Area_j = \int_{k_j}^{k_{j+1}} |\kappa_1|^T \exp\left(-\frac{T}{2}(\kappa_1 - \overline{\kappa}_1)^2\right) dx.$$

4. Compute the normalising constants.

$$C = \sum_{j}^{J-1} Area_j.$$

5. Randomly draw ϱ from a uniform distribution with range between 0 and C.

6. Obtain k_1^i by numerical interpolation of the inverted C.

5 Testing for Cointegration Rank

Unlike the classical approach of selecting an appropriate cointegration rank, the Bayesian approach compares the degree of evidence for each possible cointegration rank r. The selection of cointegration rank is then choosing the strongest evidence. Bayes factors, posterior odds and posterior probabilities are the common tools used for this purpose. These tools are computed from the marginal likelihoods. In the SECM, there are $n + 1$ possible values of the cointegration rank r. Denote the marginal likelihood of rank r as $p(\triangle Y|rank = r)$ and the prior probabilities that the cointegration vector r is correct as $\Pr(rank = r)$ (such that $\sum_{i=0}^{n} \Pr(rank = i) = 1.$). Then the Bayes factor that compares the model with rank r to a model with rank n is given by

$$BF(r|n) = \frac{p(\triangle Y|rank = r)}{p(\triangle Y|rank = n)} = \frac{\iiint p(\alpha, \beta_2, \gamma, g)p(\triangle Y|\alpha, \beta_2, \gamma, g)d\alpha d\beta_2 d\gamma dg}{\iiiint p(\alpha, \lambda, \beta_2, \gamma, g)p(\triangle Y|A\alpha, \lambda, \beta_2, \gamma, g)d\alpha d\lambda d\beta_2 d\gamma dg}, \tag{16}$$

for $r = 0, ..., n$. Multiplying the Bayes factors (16) by the ratio of prior probabilities (the prior odds) gives the posterior odds in favor of the rank r model,

$$PO(r|n) = BF(r|n) \times \frac{\Pr(rank = r)}{\Pr(rank = n)},$$

Clearly, when $\Pr(rank = r) = \Pr(rank = n)$, the posterior odds and the Bayes factors are equivalent. Selecting an appropriate cointegration vector amounts to choosing the highest $BF(r|n)$ (in the case of equal prior odds) or highest $PO(r|n)$. However, they do not offer a direct interpretation on whether model averaging or model selection is desirable, especially so for comparing more than two models. A better way of selecting an appropriate cointegration rank is through posterior probabilities:[2]

$$\Pr(rank = r|\triangle Y) = \frac{p(\triangle Y|rank = r)}{\sum_{i=0}^{n} p(\triangle Y|rank = i)} = \frac{PO(r|n)}{\sum_{i=0}^{n} PO(i|n))}.$$

When all the posterior probabilities are less then one, we would average measures of interest such as forecasts or impulse responses implied by the cointegrating vectors according to their probabilities. While, the posterior probability of one cointegration rank is equal to one, we will then select the model with that rank:

$$p(\theta|\triangle Y) = \sum_{i=0}^{n} p(\theta|\triangle Y, rank = i) \Pr(rank = i|\triangle Y),$$

where θ is the quantity of interest, $p(\theta|\triangle Y, rank = i)$ is the posterior pdf of θ implied by i cointegrating vectors, and $p(\theta|\triangle Y)$ is the averaged posterior pdf.

5.1 Computing Bayes factors

In this paper, we compute the Bayes factors by first computing the marginal likelihoods $p(\triangle Y|rank = r)$. However, the marginal likelihoods are analytically intractable. One common way of estimating them is through the methods of Gelfand and Dey (1994) that use the draws of posterior parameters. They show that if $f(\alpha, \beta_2, \gamma, g)$ is any pdf, then the expectation of $\frac{f(\alpha,\beta_2,\gamma,g)}{p(\alpha,\beta_2,\gamma,g)p(\triangle Y|\alpha,\beta_2,\gamma,g)}$ with respect to the joint posterior pdf of $\alpha, \beta_2, \gamma, g$ is equal to $p(\triangle Y|rank = r)^{-1}$ (this is the harmonic mean estimator of the marginal likelihood). For proof see Geweke (1999) or Koop (2003, p105).

[2] For example, suppose that the posterior odds for M_1, M_2 and M_3 with respect to M_3 are 10, 30 and 1 which implies that M_2 is the most likely model. However, expressing these odds as posterior model probabilities shows that the probability of M_2 being correct is about 73%. The overall evidence for M_2 is not as strong as its relative evidence (to M_1). In this case model averaging may be preferred.

The only requirement for the harmonic mean estimator is that the space of $\alpha, \beta_2, \gamma, g$ must be within the support of the posterior space of $\alpha, \beta_2, \gamma, g$. We have assumed $f(\alpha, \beta_2, \gamma, g)$ to be a truncated multivariate normal. The truncation is to ensure that $f(\alpha^{(i)}, \beta_2^{(i)}, \gamma^{(i)}, g^{(i)})$ is within the support of the posterior space for each i. In short, the expression for $f(\alpha, \beta_2, \gamma, g)$ is

$$f(\alpha, \beta_2, \gamma, g) = p^{-1}(2\pi)^{-w/2} \left|\widehat{\Sigma}_M\right|^{-1/2} \exp\left[-\frac{1}{2}\left((\psi - \widehat{\psi}_M)'\widehat{\Sigma}_M^{-1}(\psi - \widehat{\psi}_M)\right)\right] I(\Xi)$$

where $\psi = \begin{bmatrix} \alpha' & \beta_2' & \gamma' & g' \end{bmatrix}'$, $\widehat{\psi}_M = \frac{1}{M}\sum_{i=1}^{M}\psi^{(i)}$, $\widehat{\Sigma}_M = \frac{1}{M}\sum_{i=1}^{M}(\psi^{(i)} - \widehat{\psi}_M)(\psi^{(i)} - \widehat{\psi}_M)'$, $p \in (0,1)$, w is the number of parameters and $I(\Xi)$ is an indicator function, equal to 1 when $(\psi - \widehat{\psi}_M)'\widehat{\Sigma}_M^{-1}(\psi - \widehat{\psi}_M) \leq \chi^2_{1-p}(w)$, and zero otherwise. $\chi^2_{1-p}(w)$ is the $(1-p)$ per cent critical value from a chi-square distribution.

Thus,

$$p(\triangle Y|rank = r)^{-1} \approx \frac{1}{M}\sum_{i=1}^{M}\frac{f(\alpha^{(i)}, \beta_2^{(i)}, \gamma^{(i)}, g^{(i)})}{p(\alpha^{(i)}, \beta_2^{(i)}, \gamma^{(i)}, g^{(i)})p(\triangle Y|\alpha, \beta_2^{(i)}, \gamma^{(i)}, g^{(i)})}.$$

6 Two Illustrations

6.1 Simulated Example

To illustrate the methods, we simulated four sets of series, each with 150 observations, from four data generating processes. The four DGPs contain 0, 1, 2 and 3 cointegrating vectors, respectively.

DGP 1

$$\begin{bmatrix} 4.1 & 1.7 & -0.2 \\ 0 & 2.5 & 0.3 \\ 0 & 0 & 2.5 \end{bmatrix}' \triangle Y_t = \begin{bmatrix} 0.1 \\ 0.1 \\ 0.1 \end{bmatrix} + \begin{bmatrix} 0.1 & 0.3 & 2.3 \\ 0.2 & -1.2 & 2.4 \\ 1.3 & -0.1 & -0.4 \end{bmatrix}' \triangle Y_{t-1} + \varepsilon_t,$$

DGP 2

$$\begin{aligned} \begin{bmatrix} 4.1 & 1.7 & -0.2 \\ 0 & 2.5 & 0.3 \\ 0 & 0 & 2.5 \end{bmatrix}' \triangle Y_t &= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0.4 & 0.7 \end{bmatrix}' \begin{bmatrix} 1 \\ -0.01 \\ -0.8 \end{bmatrix}' Y_{t-1} \\ &+ \begin{bmatrix} 0.1 & 0.3 & 2.3 \\ 0.2 & -1.2 & 2.4 \\ 1.3 & -0.1 & -0.4 \end{bmatrix}' \triangle Y_{t-1} + \varepsilon_t, \end{aligned}$$

DGP 3

$$\begin{bmatrix} 4.1 & 1.7 & -0.2 \\ 0 & 2.5 & 0.3 \\ 0 & 0 & 2.5 \end{bmatrix}' \triangle Y_t = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0.3 & 0.1 & 0.1 \\ 0.1 & -0.2 & 0.1 \end{bmatrix}' \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1.5 & -0.1 \end{bmatrix}' Y_{t-1}$$
$$+ \begin{bmatrix} 0.1 & 0.3 & 2.3 \\ 0.2 & -1.2 & 2.4 \\ 1.3 & -0.1 & -0.4 \end{bmatrix}' \triangle Y_{t-1} + \varepsilon_t,$$

DGP 4

$$\begin{bmatrix} 4.1 & 1.7 & -0.2 \\ 0 & 2.5 & 0.3 \\ 0 & 0 & 2.5 \end{bmatrix}' \triangle Y_t = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0.23 & 0.19 & 0.17 \\ 0.43 & 0.33 & 0.32 \\ -0.2 & -0.23 & -0.23 \end{bmatrix}' \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}' Y_{t-1}$$
$$+ \begin{bmatrix} 0.1 & 0.3 & 2.3 \\ 0.2 & -1.2 & 2.4 \\ 1.3 & -0.1 & -0.4 \end{bmatrix}' \triangle Y_{t-1} + \varepsilon_t,$$

where $\varepsilon_t \sim N(0, I_3)$.

The prior means for A_0, B and Π are assumed to be a (3×3) matrix of ones, a zero matrix, and a zero matrix respectively. The specification of covariances for A_0 and B are similar to those of Sims and Zha, in that a set of hyper-parameters is used to control the standard deviations. The conditional standard deviation for the elements of B associated with lag l of variable j in equation i is assumed to be

$$\frac{\tau_1 \tau_2}{\sigma_j l^{\tau_3}},$$

the conditional standard deviation for the constants is

$$\tau_1 \tau_4,$$

and the standard deviation for the nonrestricted elements of A_0 is

$$\frac{\tau_1}{\sigma_j}.$$

In these expressions, τ_1 controls the tightness of belief on A_0, τ_2 controls overall tightness of beliefs around the random walk prior, τ_3 controls the rate at which prior variance shrinks with increasing lag length, τ_4 controls the tightness on the constant terms, σ_j is the sample standard deviation of residuals from a univariate regression of the j^{th} element of $\triangle Y_t$ on its first difference and lagged level. As for the prior covariance matrix of Π, we assume that $\overline{\Phi}_i$ is equal to $\tau_5 \left(\triangle Y_{-1}' \triangle Y_{-1}\right)^{-1}$. This assumption means that the elements of Π are independent across equations, but within the equations, the elements are reacting according to $\tau_5 \left(\triangle Y_{-1}' \triangle Y_{-1}\right)^{-1}$. The parameter τ_5 controls the overall information entering the prior. Since in this exercise we are concerned with choosing the right cointegrating vectors, firstly the hyper-parameters are assigned so that the overall priors are fairly uninformative. Specifically, the assigned values are $\tau_1 = 30$, $\tau_2 = 30, \tau_3 = 1, \tau_4 = 30$ and $\tau_5 = 20$; Secondly, we give equal prior probability to each possible value of, $r = 0, 1, ..., 3$.

For each of the models in each of the DGPs, we produce 12 000 draws and discard the first 2000 draws. All the programs are written in Matlab. Using a Pentium IV 3.02 GHz, the programs took about 30-40 minutes to complete the draws for all the models in each of the DGPs.

Table 1 indicates the marginal likelihoods, the Bayes factors and the posterior probabilities for the four DGPs. Looking at the third column, the log Bayes factors are able to select the correct rank for each of the DGPs. The probability of selecting the correct rank are about 0.73 for DGP1, 0.84 for DGP 2 0.53 for DGP 3 and one in DGP 4. For the first three cases, an averaging process may be performed.

Figure 2 shows the marginal posterior parameters' pdfs from the SECM having $r = 1$ for DGP.2. In the graphs, the vertical lines are the actual values. The results indicate that the sampling techniques are appropriate as it can be seen that the estimated marginal pdfs cover the actual values.

As the whole, the SECM together with the Bayesian techniques serve as a useful tool in handling cointegrating analysis and estimation of the structural parameters.

6.2 Empirical Example

For an empirical example, we use data from King et. al. (1991; KPSW hereafter). These data are quarterly U.S. data from 1954(1) to 1988(4) that consist of per capita real consumption expenditure c, per capita gross private domestic fixed investment i, per capita private gross national product y, real money balances $m - p$, the interest rate on three-month U.S. Treasury bills r, and the annualised rate of inflation using the GNP deflator $\triangle P$. All variables except the interest rate and inflation are in logarithms. In their paper, KPSW derived an econometric procedure that uses the cointegrating relationships to decompose the structural shocks into permanent effects and transitory effects. Three cointegrating relationships are suggested from economic theory; these are a balanced-growth trend (c and y), an inflation/money-growth trend (i and y), and a real-interest-rate stochastic trend ($(m - p)$, y and i). Fisher et. al. (2000) combine the KPSW procedure with contemporaneous restrictions to identify the structural shocks. They impose three contemporaneous zero restrictions in A_0: y variable in c equation, y variable in i equation, and $\triangle P$ variable in $(m - p)$ equation.

In this section, we test for the number of cointegrating relationships together with the three contemporaneous relations suggested by Fisher et.

al. (2000). We follow the same ordering of variables as KPSW, that is $Y = \left[\begin{array}{cccccc} y & c & i & m-p & r & \triangle P \end{array}\right]$. Furthermore, we assume fairly uninformative priors on the parameters so that the data will be the main determiner of the posterior outcomes. We assign equal prior probabilities for each of the possible cointegrating vectors and set the hyper-parameters as follows: $\tau_1 = \tau_2 = \tau_3 = \tau_4 = 3$ and $\tau_5 = 2$. We use a model with two lagged differences.

The log-marginal likelihoods, the log Bayes factors and the posterior probabilities are reported in Table 2. The results suggest that there is no evidence for three cointegrating vectors. The Bayes factors indicate a strong evidence for five cointegrating vectors, and the posterior model probability for this model is almost one. Indeed, the hypothesis test performed in KPSW rejected the three cointegrating vectors hypothesis at the 10-percent level but not at the 5-percent level. Nevertheless, they maintained that the existence of three cointegrating vectors because it provided a good qualitative description of the cointegrating system.

7 Conclusions

This paper presents a structural error correction model which provides concurrent analysis of cointegration and estimation of the structural parameters. Set in a Bayesian framework, we provide accounts on the specification of priors for the parameters, derivation of posterior pdfs and the sampling techniques used. The main methods used in this paper are partially based on the methods of Kleibergen and Paap (2002) for analysis of cointegration in the ECM, and the methods of Waggoner and Zha (2003) for estimating of the structural parameters in BSVAR. Through the simulated series, the estimated results show that the Bayes factors are able to select the appropriate ranks, and the posterior marginal pdfs cover the actual values. Through the KPSW empirical example, the estimated Bayes factors indicate existence of five cointegrating vectors instead of three cointegration vectors suggested by the theory.

A Prior of γ_i and g_i

Given that $p(a_i)$ and $p(b_i|a_i)$, respectively, are $N(\overline{a}_i, \overline{O}_i)$ and $N(\overline{P}_i a_i, \overline{H}_i)$, and $a_i = F_i\gamma_i$ and $b_i = M_i g_i$. Then, the joint prior pdf for γ_i and b_i is

$$p(\gamma_i, b_i) \propto \exp\left[-\frac{1}{2}\left((F_i\gamma_i - \overline{a}_i)'\overline{O}_i^{-1}(F_i\gamma_i - \overline{a}_i) + (M_i g_i - \overline{P}_i F_i\gamma_i)'\overline{H}_i^{-1}(M_i g_i - \overline{P}_i F_i\gamma_i)\right)\right].$$

Concentrate on the terms within the exponential and complete the squares

$$\begin{aligned}
& (F_i\gamma_i - \overline{a}_i)'\overline{O}_i^{-1}(F_i\gamma_i - \overline{a}_i) + (M_i g_i - \overline{P}_i F_i\gamma_i)'\overline{H}_i^{-1}(M_i g_i - \overline{P}_i F_i\gamma_i) \\
= \; & \gamma_i' F_i'\overline{O}_i^{-1}F_i\gamma_i - 2\gamma_i' F_i'\overline{O}_i^{-1}\overline{a}_i + \overline{a}_i'\overline{O}_i^{-1}\overline{a}_i + g_i' M_i'\overline{H}_i^{-1}M_i g_i - \\
& 2g_i' M_i'\overline{H}_i^{-1}\overline{P}_i F_i\gamma_i + \gamma_i' F_i'\overline{P}_i'\overline{H}_i^{-1}\overline{P}_i F_i\gamma_i
\end{aligned}$$

Let $\widetilde{H}_i = \left(M_i'\overline{H}_i^{-1}M_i\right)^{-1}$

$$\begin{aligned}
= \; & \gamma_i'(F_i'\overline{O}_i^{-1}F_i + F_i'\overline{P}_i'\overline{H}_i^{-1}\overline{P}_i F_i)\gamma_i - 2\gamma_i' F_i'\overline{O}_i^{-1}\overline{a}_i + g_i'\widetilde{H}_i^{-1}g_i - \\
& 2g_i'\widetilde{H}_i^{-1}\widetilde{H}_i M_i'\overline{H}_i^{-1}\overline{P}_i F_i\gamma_i + \overline{a}_i'\overline{O}_i^{-1}\overline{a}_i
\end{aligned}$$

Let $\widetilde{P}_i = \widetilde{H}_i M_i'\overline{H}_i^{-1}\overline{P}_i F_i$

$$\begin{aligned}
= \; & \gamma_i'(F_i'\overline{O}_i^{-1}F_i + F_i'\overline{P}_i'\overline{H}_i^{-1}\overline{P}_i F_i)\gamma_i - 2\gamma_i' F_i'\overline{O}_i^{-1}\overline{a}_i + g_i'\widetilde{H}_i^{-1}g_i - 2g_i'\widetilde{H}_i^{-1}\widetilde{P}_i\gamma_i \\
& +\gamma_i'\widetilde{P}_i'\widetilde{H}_i^{-1}\widetilde{P}_i\gamma_i - \gamma_i'\widetilde{P}_i'\widetilde{H}_i^{-1}\widetilde{P}_i\gamma_i + \overline{a}_i'\overline{O}_i^{-1}\overline{a}_i \\
= \; & \gamma_i'(F_i'\overline{O}_i^{-1}F_i + F_i'\overline{P}_i'\overline{H}_i^{-1}\overline{P}_i F_i - \widetilde{P}_i'\widetilde{H}_i^{-1}\widetilde{P}_i)\gamma_i - 2\gamma_i' F_i'\overline{O}_i^{-1}\overline{a}_i \\
& +(g_i - \widetilde{P}_i\gamma_i)'\widetilde{H}_i^{-1}(g_i - \widetilde{P}_i\gamma_i) + \overline{a}_i'\overline{O}_i^{-1}\overline{a}_i
\end{aligned}$$

Let $\widetilde{O}_i = \left[F_i'\overline{O}_i^{-1}F_i + F_i'\overline{P}_i'\overline{H}_i^{-1}\overline{P}_i F_i - \widetilde{P}_i'\widetilde{H}_i^{-1}\widetilde{P}_i\right]^{-1}$

$$= \gamma_i'\widetilde{O}_i^{-1}\gamma_i - 2\gamma_i'\widetilde{O}_i^{-1}\widetilde{O}_i F_i'\overline{O}_i^{-1}\overline{a}_i + (g_i - \widetilde{P}_i\gamma_i)'\widetilde{H}_i^{-1}(g_i - \widetilde{P}_i\gamma_i) + \overline{a}_i'\overline{O}_i^{-1}\overline{a}_i$$

Let $\widetilde{F}_i = \widetilde{O}_i F_i'\overline{O}_i^{-1}$

$$\begin{aligned}
= \; & \gamma_i'\widetilde{O}_i^{-1}\gamma_i - 2\gamma_i'\widetilde{O}_i^{-1}\widetilde{F}_i\overline{a}_i + \overline{a}_i'\widetilde{F}_i'\widetilde{O}_i^{-1}\widetilde{F}_i\overline{a}_i - \overline{a}_i'\widetilde{F}_i'\widetilde{O}_i^{-1}\widetilde{F}_i\overline{a}_i \\
& +(g_i - \widetilde{P}_i\gamma_i)'\widetilde{H}_i^{-1}(g_i - \widetilde{P}_i\gamma_i) + \overline{a}_i'\overline{O}_i^{-1}\overline{a}_i \\
= \; & (\gamma_i - \widetilde{F}_i\overline{a}_i)'\widetilde{O}_i^{-1}(\gamma_i - \widetilde{F}_i\overline{a}_i) + (g_i - \widetilde{P}_i\gamma_i)'\widetilde{H}_i^{-1}(g_i - \widetilde{P}_i\gamma_i) + \overline{a}_i'(\overline{O}_i^{-1} - \widetilde{F}_i'\widetilde{O}_i^{-1}\widetilde{F}_i)\overline{a}_i
\end{aligned}$$

Replace the above expression back into the exponential and absorb $\overline{a}_i'(\overline{O}_i^{-1} - \widetilde{F}_i'\widetilde{O}_i^{-1}\widetilde{F}_i)\overline{a}_i$ into proportionality gives

$$p(\gamma_i, b_i) \propto \exp\left[-\frac{1}{2}\left((\gamma_i - \widetilde{F}_i\overline{a}_i)'\widetilde{O}_i^{-1}(\gamma_i - \widetilde{F}_i\overline{a}_i) + (g_i - \widetilde{P}_i\gamma_i)'\widetilde{H}_i^{-1}(g_i - \widetilde{P}_i\gamma_i)\right)\right].$$

Then, it can be shown that the prior of γ_i and conditional prior of g_i given γ_i are

$$\gamma_i \sim N(\widetilde{F}_i \overline{a}_i, \widetilde{O}_i),$$

and

$$g_i|\gamma_i \sim N(\widetilde{P}_i \gamma_i, \widetilde{H}_i).$$

B Computing normalised $p(\lambda)|_{\lambda=0}$

For computation of normalised constant, we follow technique of Chen (1994). We know for the fact that

$$\begin{aligned}
\iint p^*(\alpha, \beta_2) d\alpha d\beta_2 &= \iint \frac{p(\alpha, \lambda, \beta_2)|_{\lambda=0}}{p(\lambda)|_{\lambda=0}} \left| \frac{\partial \Pi}{\partial(\alpha, \lambda, \beta_2)} \right|\Big|_{\lambda=0} d\alpha d\beta_2 = 1 \\
p(\lambda)|_{\lambda=0} &= \iint p(\alpha, \lambda, \beta_2)|_{\lambda=0} \left| \frac{\partial \Pi}{\partial(\alpha, \lambda, \beta_2)} \right|\Big|_{\lambda=0} d\alpha d\beta_2 \\
&= \frac{\iint p(\alpha, \lambda, \beta_2)|_{\lambda=0} \left| \frac{\partial \Pi}{\partial(\alpha, \lambda, \beta_2)} \right|\Big|_{\lambda=0} d\alpha d\beta_2}{\iiint p(\alpha, \lambda, \beta_2) d\alpha d\lambda d\beta_2} \\
&= \frac{\iint p(\alpha, \lambda, \beta_2)|_{\lambda=0} \left| \frac{\partial \Pi}{\partial(\alpha, \lambda, \beta_2)} \right|\Big|_{\lambda=0} \left(\int h(\lambda|\alpha, \beta_2) d\lambda \right) d\alpha d\beta_2}{\iiint p(\alpha, \lambda, \beta_2) d\alpha d\lambda d\beta_2} \\
&= \frac{\iiint p(\alpha, \lambda, \beta_2)|_{\lambda=0} \left| \frac{\partial \Pi}{\partial(\alpha, \lambda, \beta_2)} \right|\Big|_{\lambda=0} h(\lambda|\alpha, \beta_2) d\alpha d\lambda d\beta_2}{\iiint p(\alpha, \lambda, \beta_2) d\alpha d\lambda d\beta_2} \qquad (17)
\end{aligned}$$

where $\int h(\lambda|\alpha, \beta_2) d\lambda$ and $\iiint p(\alpha, \lambda, \beta_2) d\alpha d\lambda d\beta_2$ will integrate to 1. $h(\lambda|\alpha, \beta_2)$ is a proper conditional density which appropriate the conditional prior of λ. Henceforth, to estimate $p(\lambda)|_{\lambda=0}$, we

- draw $\Pi^{(i)}$ from $p(\Pi)$ for $i = 1, 2, ..., M$,
- svd $\Pi^{(i)}$ into $\alpha^{(i)}, \lambda^{(i)}$ and $\beta_2^{(i)}$ for $i = 1, 2, ..., M$
- and average the following to get an estimate for

$$p(\lambda)|_{\lambda=0} \approx \frac{1}{M} \sum_{i=1}^{M} \frac{p(\alpha^{(i)}, \lambda^{(i)}, \beta_2^{(i)})|_{\lambda=0} \left| \frac{\partial \Pi}{\partial(\alpha^{(i)}, \lambda^{(i)}, \beta_2^{(i)})} \right|\Big|_{\lambda^{(i)}=0} h(\lambda^{(i)}|\alpha^{(i)}, \beta_2^{(i)})}{p(\alpha^{(i)}, \lambda^{(i)}, \beta_2^{(i)})}.$$

An appropriate $h(\lambda|\alpha, \beta_2)$ is found to be

$$h(\lambda|\alpha, \beta_2) \propto (2\pi)^{-\frac{1}{2}(n-r)^2} \left|\overline{\Sigma}_\lambda\right|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\left(vec(\lambda) - \overline{\lambda}\right)' \overline{\Sigma}_\lambda^{-1} \left(vec(\lambda) - \overline{\lambda}\right)\right],$$

where $\overline{\lambda} = \overline{\Sigma}_\lambda(\alpha'_\perp \otimes \beta_\perp)'\overline{\Sigma}_\Pi^{-1}(\overline{\pi} - vec(\beta\alpha))$, and $\overline{\Sigma}_\lambda = \left((\alpha'_\perp \otimes \beta_\perp)'\overline{\Sigma}_\Pi^{-1}(\alpha'_\perp \otimes \beta_\perp)\right)^{-1}$ (See the subsection for detail).

It can then be shown that ratio of the integrands in equation (17) can be simplified to

$$= \frac{\left|\beta'\beta\right|^{\frac{1}{2}(n-r)} \left|\alpha\alpha'\right|^{\frac{1}{2}(n-r)} (2\pi)^{-\frac{1}{2}(n-r)^2} \left|\overline{\Sigma}_\lambda\right|^{-\frac{1}{2}}}{\exp\left[\frac{1}{2}\overline{\lambda}'\overline{\Sigma}_\lambda^{-1}\overline{\lambda}\right] \left|\frac{\partial\Pi}{\partial(\alpha,\lambda,\beta_2)}\right|}.$$

B.1 Approximating the conditional prior of λ given α, β_2

Concentrate on the terms within the exponential in $p(\alpha, \lambda, \beta_2)$

$$\begin{aligned}
& (vec(\beta\alpha + \beta_\perp\lambda\alpha_\perp) - \overline{\pi})'\overline{\Sigma}_\Pi^{-1}(vec(\beta\alpha + \beta_\perp\lambda\alpha_\perp) - \overline{\pi}) \\
= & (vec(\beta\alpha) + vec(\beta_\perp\lambda\alpha_\perp) - \overline{\pi})'\overline{\Sigma}_\Pi^{-1}(vec(\beta\alpha) + vec(\beta_\perp\lambda\alpha_\perp) - \overline{\pi}) \\
= & (vec(\beta_\perp\lambda\alpha_\perp) - (\overline{\pi} - vec(\beta\alpha)))'\,\overline{\Sigma}_\Pi^{-1}(vec(\beta\alpha) - (\overline{\pi} - vec(\beta\alpha)) \\
= & ((\alpha'_\perp \otimes \beta_\perp)vec(\lambda) - (\overline{\pi} - vec(\beta\alpha)))'\overline{\Sigma}_\Pi^{-1}((\alpha'_\perp \otimes \beta_\perp)vec(\lambda) - (\overline{\pi} - vec(\beta\alpha))) \\
= & vec(\lambda)'(\alpha'_\perp \otimes \beta_\perp)'\overline{\Sigma}_\Pi^{-1}(\alpha'_\perp \otimes \beta_\perp)vec(\lambda) - 2vec(\lambda)'(\alpha'_\perp \otimes \beta_\perp)'\overline{\Sigma}_\Pi^{-1}(\overline{\pi} - vec(\beta\alpha)) \\
& +(\overline{\pi} - vec(\beta\alpha))'\overline{\Sigma}_\Pi^{-1}(\overline{\pi} - vec(\beta\alpha))
\end{aligned}$$

Let $\overline{\Sigma}_\lambda = \left((\alpha'_\perp \otimes \beta_\perp)'\overline{\Sigma}_\Pi^{-1}(\alpha'_\perp \otimes \beta_\perp)\right)^{-1}$

$$\begin{aligned}
= & vec(\lambda)'\overline{\Sigma}_\lambda^{-1}vec(\lambda) - 2vec(\lambda)'\overline{\Sigma}_\lambda^{-1}\overline{\Sigma}_\lambda(\alpha'_\perp \otimes \beta_\perp)'\overline{\Sigma}_\Pi^{-1}(\overline{\pi} - vec(\beta\alpha)) \\
& +(\overline{\pi} - vec(\beta\alpha))'\overline{\Sigma}_\Pi^{-1}(\overline{\pi} - vec(\beta\alpha))
\end{aligned}$$

Let $\overline{\lambda} = \overline{\Sigma}_\lambda(\alpha'_\perp \otimes \beta_\perp)'\overline{\Sigma}_\Pi^{-1}(\overline{\pi} - vec(\beta\alpha))$

$$\begin{aligned}
= & vec(\lambda)'\overline{\Sigma}_\lambda^{-1}vec(\lambda) - 2vec(\lambda)'\overline{\Sigma}_\lambda^{-1}\overline{\lambda} + \overline{\lambda}'\overline{\Sigma}_\lambda^{-1}\overline{\lambda} - \overline{\lambda}'\overline{\Sigma}_\lambda^{-1}\overline{\lambda} \\
& +(\overline{\pi} - vec(\beta\alpha))'\overline{\Sigma}_\Pi^{-1}(\overline{\pi} - vec(\beta\alpha)) \\
= & \left(vec(\lambda) - \overline{\lambda}\right)'\overline{\Sigma}_\lambda^{-1}\left(vec(\lambda) - \overline{\lambda}\right) - \overline{\lambda}'\overline{\Sigma}_\lambda^{-1}\overline{\lambda} \\
& +(\overline{\pi} - vec(\beta\alpha))'\overline{\Sigma}_\Pi^{-1}(\overline{\pi} - vec(\beta\alpha))
\end{aligned}$$

Thus an appropriate conditional prior for λ would be

$$h(\lambda|\alpha, \beta_2) \propto (2\pi)^{-\frac{1}{2}(n-r)^2}\left|\overline{\Sigma}_\lambda\right|^{-\frac{1}{2}}\exp\left[-\frac{1}{2}\left(vec(\lambda) - \overline{\lambda}\right)'\overline{\Sigma}_\lambda^{-1}\left(vec(\lambda) - \overline{\lambda}\right)\right]$$

C Marginal and Conditional Posterior Pdfs of Linear SECM

The joint posterior distribution is given as

$$
\begin{aligned}
& p(\Pi, \gamma, g| \triangle Y) \\
\propto\ & p(\triangle Y|\gamma, g, \Pi) p(\gamma, g) p(\Pi) \propto p(\triangle Y|\gamma, g, \Pi) \prod_{i=1}^{n} p(\gamma_i) p(g_i|\gamma_i) \prod_{i=1}^{n} p(\pi_i) \\
\propto\ & |[F_1\gamma_1|...|F_n\gamma_n]|^T \exp\left[-\frac{1}{2}\sum_{i=1}^{n}(\triangle Y F_i\gamma_i - Y_{-1}\pi_i - XM_i g_i)'(\triangle Y F_i\gamma_i - Y_{-1}\pi_i - XM_i g_i)\right] \times \\
& \prod_{i=1}^{n}\left|\widetilde{O}_i\right|^{-\frac{1}{2}}\left|\widetilde{H}_i\right|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\left((\gamma_i - \widetilde{F}_i\overline{a}_i)'\widetilde{O}_i^{-1}(\gamma_i - \widetilde{F}_i\overline{a}_i) + (g_i - \widetilde{P}_i\gamma_i)'\widetilde{H}_i^{-1}(g_i - \widetilde{P}_i\gamma_i)\right)\right] \times \\
& \prod_{i=1}^{n}|\overline{\Phi}_i|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\pi_i - \overline{\pi}_i)'\overline{\Phi}_i^{-1}(\pi_i - \overline{\pi}_i)\right] \\
\propto\ & |[F_1\gamma_1|...|F_n\gamma_n]|^T \times \\
& \exp\left[-\frac{1}{2}\sum_{i=1}^{n}\left(\begin{array}{c} (\triangle Y F_i\gamma_i - Y_{-1}\pi_i - XM_i g_i)'(\triangle Y F_i\gamma_i - Y_{-1}\pi_i - XM_i g_i) + \\ (\gamma_i - \widetilde{F}_i\overline{a}_i)'\widetilde{O}_i^{-1}(\gamma_i - \widetilde{F}_i\overline{a}_i) + (g_i - \widetilde{P}_i\gamma_i)'\widetilde{H}_i^{-1}(g_i - \widetilde{P}_i\gamma_i) \\ +(\pi_i - \overline{\pi}_i)'\overline{\Phi}_i^{-1}(\pi_i - \overline{\pi}_i) \end{array}\right)\right].
\end{aligned}
$$

Concentrate on the terms of the i^{th} equation within the exponential

$$
\begin{aligned}
& (\triangle Y F_i\gamma_i - Y_{-1}\pi_i - XM_i g_i)'(\triangle Y F_i\gamma_i - Y_{-1}\pi_i - XM_i g_i) + (\gamma_i - \widetilde{F}_i\overline{a}_i)'\widetilde{O}_i^{-1}(\gamma_i - \widetilde{F}_i\overline{a}_i) \\
& (g_i - \widetilde{P}_i\gamma_i)'\widetilde{H}_i^{-1}(g_i - \widetilde{P}_i\gamma_i) + (\pi_i - \overline{\pi}_i)'\overline{\Phi}_i^{-1}(\pi_i - \overline{\pi}_i) \\
=\ & \gamma_i' F_i' \triangle Y' \triangle Y F_i\gamma_i - 2\pi_i' Y_{-1}' \triangle Y F_i\gamma_i - 2g_i' M_i' X' \triangle Y F_i\gamma_i + 2g_i' M_i' X' Y_{-1}\pi_i + \pi_i' Y_{-1}' Y_{-1}\pi_i \\
& + g_i' M_i' X' X M_i g_i + \gamma_i'\widetilde{O}_i^{-1}\gamma_i - 2\gamma_i'\widetilde{O}_i^{-1}\widetilde{F}_i\overline{a}_i + \overline{a}_i'\widetilde{F}_i'\widetilde{O}_i^{-1}\widetilde{F}_i\overline{a}_i + g_i'\widetilde{H}_i^{-1}g_i \\
& -2g_i'\widetilde{H}_i^{-1}\widetilde{P}_i\gamma_i + \gamma_i'\widetilde{P}_i'\widetilde{H}_i^{-1}\widetilde{P}_i\gamma_i + \pi_i'\overline{\Phi}_i^{-1}\pi_i - 2\pi_i'\overline{\Phi}_i^{-1}\overline{\pi}_i + \overline{\pi}_i'\overline{\Phi}_i^{-1}\overline{\pi}_i \\
=\ & g_i'\left(M_i' X' X M_i + \widetilde{H}_i^{-1}\right) g_i - 2g_i'\left(\left(M_i' X' \triangle Y F_i + \widetilde{H}_i^{-1}\widetilde{P}_i\right)\gamma_i - M_i' X' Y_{-1}\pi_i\right) \\
& +\pi_i'\left(Y_{-1}' Y_{-1} + \overline{\Phi}_i^{-1}\right)\pi_i - 2\pi_i'\left(Y_{-1}'\triangle Y F_i\gamma_i + \overline{\Phi}_i^{-1}\overline{\pi}_i\right) \\
& +\gamma_i'(F_i'\triangle Y'\triangle Y F_i + \widetilde{O}_i^{-1} + \widetilde{P}_i'\widetilde{H}_i^{-1}\widetilde{P}_i)\gamma_i - 2\gamma_i'\widetilde{O}_i^{-1}\widetilde{F}_i\overline{a}_i \\
& +\overline{a}_i'\widetilde{F}_i'\widetilde{O}_i^{-1}\widetilde{F}_i\overline{a}_i + \overline{\pi}_i'\overline{\Phi}_i^{-1}\overline{\pi}_i
\end{aligned}
$$

Let $H_i = \left(M_i' X' X M_i + \widetilde{H}_i^{-1}\right)^{-1}$, $P_i = \left[\begin{array}{cc} P_{1i} & P_{2i} \end{array}\right]$ and $\underline{g}_i = \left[\begin{array}{c} \gamma_i \\ \pi_i \end{array}\right]$.

Where $P_{1i} = H_i(M_i'X'\triangle Y F_i + \widetilde{H}_i^{-1}\widetilde{P}_i)$ and $P_{2i} = -H_iM_i'X'Y_{-1}$

$$\begin{aligned} &= g_i'H_i^{-1}g_i - 2g_i'H_i^{-1}P_i\underline{g}_i + \underline{g}_i'P_i'H_i^{-1}P_i\underline{g}_i - \underline{g}_i'P_i'H_i^{-1}P_i\underline{g}_i \\ &\quad +\pi_i'\left(Y_{-1}'Y_{-1} + \overline{\Phi}_i^{-1}\right)\pi_i - 2\pi_i'\left(Y_{-1}'\triangle Y F_i\gamma_i + \overline{\Phi}_i^{-1}\overline{\pi}_i\right) \\ &\quad +\gamma_i'(F_i'\triangle Y'\triangle Y F_i + \widetilde{O}_i^{-1} + \widetilde{P}_i'\widetilde{H}_i^{-1}\widetilde{P}_i)\gamma_i - 2\gamma_i'\widetilde{O}_i^{-1}\widetilde{F}_i\overline{a}_i \\ &\quad +\overline{a}_i'\widetilde{F}_i'\widetilde{O}_i^{-1}\widetilde{F}_i\overline{a}_i + \overline{\pi}_i'\overline{\Phi}_i^{-1}\overline{\pi}_i \end{aligned}$$

$$\begin{aligned} &= \left(g_i - P_i\underline{g}_i\right)' H_i^{-1}\left(g_i - P_i\underline{g}_i\right) \\ &\quad +\pi_i'\left(Y_{-1}'Y_{-1} + \overline{\Phi}_i^{-1} - P_{2i}'H_i^{-1}P_{2i}\right)\pi_i - 2\pi_i'\left(Y_{-1}'\triangle Y F_i\gamma_i + \overline{\Phi}_i^{-1}\overline{\pi}_i\right) \\ &\quad +\gamma_i'(F_i'\triangle Y'\triangle Y F_i + \widetilde{O}_i^{-1} + \widetilde{P}_i'\widetilde{H}_i^{-1}\widetilde{P}_i - P_{1i}'H_i^{-1}P_{1i})\gamma_i - 2\gamma_i'\widetilde{O}_i^{-1}\widetilde{F}_i\overline{a}_i \\ &\quad +\overline{a}_i'\widetilde{F}_i'\widetilde{O}_i^{-1}\widetilde{F}_i\overline{a}_i + \overline{\pi}_i'\overline{\Phi}_i^{-1}\overline{\pi}_i \end{aligned}$$

Let $\Phi_i = \left(Y_{-1}'Y_{-1} + \overline{\Phi}_i^{-1} - P_{2i}'H_i^{-1}P_{2i}\right)^{-1}$, $Q_i = \begin{bmatrix} Q_{1i} & Q_{2i} \end{bmatrix}$ and $\underline{\pi}_i = \begin{bmatrix} \gamma_i \\ \overline{\pi}_i \end{bmatrix}$. Where $Q_{1i} = \Phi_iY_{-1}'\triangle Y F_i + P_{2i}'H_i^{-1}P_{1i}$ and $Q_{2i} = \Phi_i\overline{\Phi}_i^{-1}$.

$$\begin{aligned} &= \left(g_i - P_i\underline{g}_i\right)' H_i^{-1}\left(g_i - P_i\underline{g}_i\right) \\ &\quad +\pi_i'\Phi_i^{-1}\pi_i - 2\pi_i'\Phi_i^{-1}Q_i\underline{\pi}_i + \underline{\pi}_i'Q_i'\Phi_i^{-1}Q_i\underline{\pi}_i - \underline{\pi}_i'Q_i'\Phi_i^{-1}Q_i\underline{\pi}_i \\ &\quad +\gamma_i'(F_i'\triangle Y'\triangle Y F_i + \widetilde{O}_i^{-1} + \widetilde{P}_i'\widetilde{H}_i^{-1}\widetilde{P}_i - P_{1i}'H_i^{-1}P_{1i})\gamma_i - 2\gamma_i'\widetilde{O}_i^{-1}\widetilde{F}_i\overline{a}_i \\ &\quad +\overline{a}_i'\widetilde{F}_i'\widetilde{O}_i^{-1}\widetilde{F}_i\overline{a}_i + \overline{\pi}_i'\overline{\Phi}_i^{-1}\overline{\pi}_i \end{aligned}$$

$$\begin{aligned} &= \left(g_i - P_i\underline{g}_i\right)' H_i^{-1}\left(g_i - P_i\underline{g}_i\right) + (\pi_i - Q_i\underline{\pi}_i)'\Phi_i^{-1}(\pi_i - Q_i\underline{\pi}_i) \\ &\quad +\gamma_i'(F_i'\triangle Y'\triangle Y F_i + \widetilde{O}_i^{-1} + \widetilde{P}_i'\widetilde{H}_i^{-1}\widetilde{P}_i - P_{1i}'H_i^{-1}P_{1i} - Q_{1i}'\Phi_i^{-1}Q_{1i})\gamma_i \\ &\quad -2\gamma_i'\widetilde{O}_i^{-1}\widetilde{F}_i\overline{a}_i + \overline{a}_i'\widetilde{F}_i'\widetilde{O}_i^{-1}\widetilde{F}_i\overline{a}_i + \overline{\pi}_i'\left(\overline{\Phi}_i^{-1} - Q_{2i}'\Phi_i^{-1}Q_{2i}\right)\overline{\pi}_i \end{aligned}$$

Let $\widehat{O}_i = \left(F_i'\triangle Y'\triangle Y F_i + \widetilde{O}_i^{-1} + \widetilde{P}_i'\widetilde{H}_i^{-1}\widetilde{P}_i - P_{1i}'H_i^{-1}P_{1i} - Q_{1i}'\Phi_i^{-1}Q_{1i}\right)^{-1}$, $\widehat{F}_i = \begin{bmatrix} \widehat{F}_{1i} & \widehat{F}_{2i} \end{bmatrix}$ and $\underline{\gamma}_i = \begin{bmatrix} \overline{a}_i \\ \overline{\pi}_i \end{bmatrix}$. Where $\widehat{F}_{1i} = \widehat{O}_i\widetilde{O}_i^{-1}\widetilde{F}_i$ and $\widehat{F}_{2i} =$

$\widehat{O}_i Q'_{1i}\Phi_i^{-1}Q_{2i}$

$$
\begin{aligned}
&= \left(g_i - P_i\underline{g}_i\right)' H_i^{-1}\left(g_i - P_i\underline{g}_i\right) + (\pi_i - Q_i\underline{\pi}_i)'\Phi_i^{-1}(\pi_i - Q_i\underline{\pi}_i) \\
&\quad +\gamma_i'\widehat{O}_i^{-1}\gamma_i - 2\gamma_i'\widehat{O}_i^{-1}\widehat{F}_i\underline{\gamma}_i + \underline{\gamma}_i'\widehat{F}_i'\widehat{O}_i^{-1}\widehat{F}_i\underline{\gamma}_i - \underline{\gamma}_i'\widehat{F}_i'\widehat{O}_i^{-1}\widehat{F}_i\underline{\gamma}_i \\
&\quad +\overline{a}_i'\widetilde{F}_i'\widetilde{O}_i^{-1}\widetilde{F}_i\overline{a}_i + \overline{\pi}_i'\left(\overline{\Phi}_i^{-1} - Q'_{2i}\Phi_i^{-1}Q_{2i}\right)\overline{\pi}_i \\
&= \left(g_i - P_i\underline{g}_i\right)' H_i^{-1}\left(g_i - P_i\underline{g}_i\right) + (\pi_i - Q_i\underline{\pi}_i)'\Phi_i^{-1}(\pi_i - Q_i\underline{\pi}_i) \\
&\quad + \left(\gamma_i - \widehat{F}_i\underline{\gamma}_i\right)'\widehat{O}_i^{-1}\left(\gamma_i - \widehat{F}_i\underline{\gamma}_i\right) - \underline{\gamma}_i'\widehat{F}_i'\widehat{O}_i^{-1}\widehat{F}_i\underline{\gamma}_i \\
&\quad +\overline{a}_i'\widetilde{F}_i'\widetilde{O}_i^{-1}\widetilde{F}_i\overline{a}_i + \overline{\pi}_i'\left(\overline{\Phi}_i^{-1} - Q'_{2i}\Phi_i^{-1}Q_{2i}\right)\overline{\pi}_i
\end{aligned}
$$

Replace the above expression back into the exponential term and absorb $\overline{a}_i'\widetilde{F}_i'\widetilde{O}_i^{-1}\widetilde{F}_i\overline{a}_i$, $\underline{\gamma}_i'\widehat{F}_i'\widehat{O}_i^{-1}\widehat{F}_i\underline{\gamma}_i$ and $\overline{\pi}_i'\left(\overline{\Phi}_i^{-1} - Q'_{2i}\Phi_i^{-1}Q_{2i}\right)\overline{\pi}_i$ into proportionality.

$$
\begin{aligned}
p(\Pi,\gamma,g|\triangle Y) \quad \propto \quad & |[F_1\gamma_1|...|F_n\gamma_n]|^T \times \\
& \exp\left[-\frac{1}{2}\sum_{i=1}^n \left(\begin{array}{c} \left(g_i - P_i\underline{g}_i\right)' H_i^{-1}\left(g_i - P_i\underline{g}_i\right) \\ +(\pi_i - Q_i\underline{\pi}_i)'\Phi_i^{-1}(\pi_i - Q_i\underline{\pi}_i) \\ +\left(\gamma_i - \widehat{F}_i\underline{\gamma}_i\right)'\widehat{O}_i^{-1}\left(\gamma_i - \widehat{F}_i\underline{\gamma}_i\right) \end{array}\right)\right]
\end{aligned}
$$

It can then be shown that

$$
p(\gamma|\triangle Y) \propto |[F_1\gamma_1|...|F_n\gamma_n]|^T \prod_{i=1}^n \exp\left(-\frac{T}{2}\left(\gamma_i - \widehat{F}_i\underline{\gamma}_i\right)' O_i^{-1}\left(\gamma_i - \widehat{F}_i\underline{\gamma}_i\right)\right),
$$

$$
p(\Pi|\gamma,\triangle Y) \propto \prod_{i=1}^n \exp\left(-\frac{1}{2}(\pi_i - Q_i\underline{\pi}_i)'\Phi_i^{-1}(\pi_i - Q_i\underline{\pi}_i)\right),
$$

and

$$
p(g|\Pi,\gamma,\triangle Y) \propto \prod_{i=1}^n \exp\left(-\frac{1}{2}\left(g_i - P_i\underline{g}_i\right)' H_i^{-1}\left(g_i - P_i\underline{g}_i\right)\right),
$$

where $O_i^{-1} = \frac{1}{T}\widehat{O}_i^{-1}$.

D Proof of $p(\alpha, \lambda, \beta_2|\gamma, \triangle Y) \propto p^*(\alpha, \beta_2|\gamma, \triangle Y)p(\lambda|\alpha, \beta_2, \gamma, \triangle Y)$

Given that the conditional posterior of α, β_2 given γ is

$$\begin{aligned}
p(\alpha,\lambda,\beta_2|\gamma,\triangle Y) &\propto p(\Pi|\gamma,\triangle Y))|_{\Pi=\beta\alpha}\left|\frac{\partial\Pi}{\partial(\alpha,\lambda,\beta_2)}\right|\Big|_{\lambda=0} \\
&\propto \prod_{i=1}^{n}\exp\left(-\frac{1}{2}(\pi_i-Q_i\underline{\pi}_i)'\Phi_i^{-1}(\pi_i-Q_i\underline{\pi}_i)\right)|_{\Pi=\beta\alpha}\left|\frac{\partial\Pi}{\partial(\alpha,\lambda,\beta_2)}\right|\Big|_{\lambda=0} \\
&\propto \exp\left(-\frac{1}{2}(vec(\Pi)-\underline{\pi})'\Sigma_\Pi^{-1}(vec(\Pi)-\underline{\pi})\right)|_{\Pi=\beta\alpha}\left|\frac{\partial\Pi}{\partial(\alpha,\lambda,\beta_2)}\right|\Big|_{\lambda=0} \\
&\propto \exp\left(-\frac{1}{2}(vec(\beta\alpha)-\underline{\pi})'\Sigma_\Pi^{-1}(vec(\beta\alpha)-\underline{\pi})\right)|\beta'\beta|^{\frac{1}{2}(n-r)}|\alpha\alpha'|^{\frac{1}{2}(n-r)},
\end{aligned}$$

where $\underline{\pi} = (\ (Q_i\underline{\pi}_1)'\ \ ...\ \ (Q_n\underline{\pi}_n)'\ )'$ and $\Sigma_\Pi = diag(\Phi_1, ..., \Phi_n)$.

Now, consider the conditional posterior of α, λ, β_2 given γ

$$\begin{aligned}
p(\alpha,\lambda,\beta_2|\gamma,\triangle Y) &\propto p(\Pi|\gamma,\triangle Y))|_{\Pi=\beta\alpha+\beta_\perp\lambda\alpha_\perp}\left|\frac{\partial\Pi}{\partial(\alpha,\lambda,\beta_2)}\right| \\
&\propto \exp\left(-\frac{1}{2}(vec(\Pi)-\underline{\pi})'\Sigma_\Pi^{-1}(vec(\Pi)-\underline{\pi})\right)|_{\Pi=\beta\alpha+\beta_\perp\lambda\alpha_\perp} \\
&\quad\times\left|\frac{\partial\Pi}{\partial(\alpha,\lambda,\beta_2)}\right|
\end{aligned}$$

$$\begin{aligned}
p(\alpha,\lambda,\beta_2|\gamma,\triangle Y) &\propto p(\Pi|\gamma,\triangle Y))|_{\Pi=\beta\alpha+\beta_\perp\lambda\alpha_\perp}\left|\frac{\partial\Pi}{\partial(\alpha,\lambda,\beta_2)}\right| \\
&\propto \exp\left(-\frac{1}{2}(vec(\Pi)-\underline{\pi})'\Sigma_\Pi^{-1}(vec(\Pi)-\underline{\pi})\right)|_{\Pi=\beta\alpha+\beta_\perp\lambda\alpha_\perp}\left|\frac{\partial\Pi}{\partial(\alpha,\lambda,\beta_2)}\right| \\
&\propto \exp\left(-\frac{1}{2}(vec(\beta\alpha+\beta_\perp\lambda\alpha_\perp)-\underline{\pi})'\Sigma_\Pi^{-1}(vec(\beta\alpha+\beta_\perp\lambda\alpha_\perp)-\underline{\pi})\right) \\
&\quad\times\left|\frac{\partial\Pi}{\partial(\alpha,\lambda,\beta_2)}\right| \\
&\propto \exp\left[-\frac{1}{2}(vec(\beta\alpha)+vec(\beta_\perp\lambda\alpha_\perp)-\underline{\pi})'\Sigma_\Pi^{-1}(vec(\beta\alpha)+vec(\beta_\perp\lambda\alpha_\perp)-\underline{\pi})\right] \\
&\quad\times\left|\frac{\partial\Pi}{\partial(\alpha,\lambda,\beta_2)}\right|
\end{aligned}$$

$$
\begin{aligned}
&\propto \exp\left[-\frac{1}{2}(vec(\beta\alpha)-\underline{\pi})'\Sigma_{\Pi}^{-1}(vec(\beta\alpha)-\underline{\pi})\right]\times \\
&\quad \exp\left[-\frac{1}{2}vec(\beta_{\perp}\lambda\alpha_{\perp})'\Sigma_{\Pi}^{-1}vec(\beta_{\perp}\lambda\alpha_{\perp})-vec(\beta_{\perp}\lambda\alpha_{\perp})'\Sigma_{\Pi}^{-1}(vec(\beta\alpha)-\underline{\pi})\right]\left|\frac{\partial\Pi}{\partial(\alpha,\lambda,\beta_2)}\right| \\
&\propto \exp\left[-\frac{1}{2}(vec(\beta\alpha)-\underline{\pi})'\Sigma_{\Pi}^{-1}(vec(\beta\alpha)-\underline{\pi})\right]\left|\beta'\beta\right|^{\frac{1}{2}(n-r)}\left|\alpha\alpha'\right|^{\frac{1}{2}(n-r)}\times \\
&\quad \exp\left[-\frac{1}{2}vec(\beta_{\perp}\lambda\alpha_{\perp})'\Sigma_{\Pi}^{-1}vec(\beta_{\perp}\lambda\alpha_{\perp})-vec(\beta_{\perp}\lambda\alpha_{\perp})'\Sigma_{\Pi}^{-1}(vec(\beta\alpha)-\underline{\pi})\right]\left|\frac{\partial\Pi}{\partial(\alpha,\lambda,\beta_2)}\right| \\
&\quad \times\left|\beta'\beta\right|^{-\frac{1}{2}(n-r)}\left|\alpha\alpha'\right|^{-\frac{1}{2}(n-r)}. \\
&= p(\lambda|\alpha,\beta_2,\gamma,\triangle Y)p^*(\alpha,\beta_2|\gamma,\triangle Y)
\end{aligned}
$$

where

$$
\begin{aligned}
p(\lambda|\alpha,\beta_2,\gamma,\triangle Y) \;\propto\; & \exp\left[-\frac{1}{2}vec(\beta_{\perp}\lambda\alpha_{\perp})'\Sigma_{\Pi}^{-1}vec(\beta_{\perp}\lambda\alpha_{\perp})-vec(\beta_{\perp}\lambda\alpha_{\perp})'\Sigma_{\Pi}^{-1}(vec(\beta\alpha)-\underline{\pi})\right] \\
& \times\left|\frac{\partial\Pi}{\partial(\alpha,\lambda,\beta_2)}\right|\left|\beta'\beta\right|^{-\frac{1}{2}(n-r)}\left|\alpha\alpha'\right|^{-\frac{1}{2}(n-r)}.
\end{aligned}
$$

This implies that α, β_2 can be computed from singular value decomposition of Π.

E Proof of Generalised Theorem of Waggoner and Zha

Given that we define

$$\gamma_i = T_i \sum_{i=1}^{p_i} \kappa_j w_j, \tag{18}$$

and

$$\begin{bmatrix} \overline{\kappa}_1 \\ \vdots \\ \overline{\kappa}_{p_i} \end{bmatrix} = [w_1|\cdots|w_{p_i}]^{-1} T_i^{-1} \widehat{F}_i \underline{\gamma}_i.$$

The immediate preceding equation implies that

$$\widehat{F}_i \underline{\gamma}_i = T_i \sum_{i=1}^{p_i} \overline{\kappa}_j w_j. \tag{19}$$

Then, the conditional posterior pdf of γ_i given $\gamma_1, ..., \gamma_{i-1}, \gamma_{i+1}, ..., \gamma_n$ and π_i is

$$p(\gamma_i|\gamma_1, ..., \gamma_{i-1}, \gamma_{i+1}, ..., \gamma_n, \Pi, \triangle Y) \propto |[F_1\gamma_1|...|F_n\gamma_n]|^T \exp\left[-\frac{T}{2}\left(\gamma_i - \widehat{F}_i\underline{\gamma}_i\right)' O_i^{-1}\left(\gamma_i - \widehat{F}_i\underline{\gamma}_i\right)\right]$$

Replace γ_i and $\widehat{F}_i\underline{\gamma}_i$ with the expressions found in equations (18) and (19) respectively into the above equation, then it can be proof that

$$\begin{aligned}
& \left|[F_1\gamma_1|...|T_i \sum_{i=1}^{p_i} \kappa_j w_j|...|F_n\gamma_n]\right|^T \times \\
& \exp\left[-\frac{T}{2}\left(T_i \sum_{i=1}^{p_i} \kappa_j w_j - T_i \sum_{i=1}^{p_i} \overline{\kappa}_j w_j\right)' O_i^{-1} \left(T_i \sum_{i=1}^{p_i} \kappa_j w_j - T_i \sum_{i=1}^{p_i} \overline{\kappa}_j w_j\right)\right] \\
\propto \quad & |\kappa_1|^T \exp\left(-\frac{T}{2}(\kappa_1 - \overline{\kappa}_1)^2\right) \prod_{j=2}^{p_i} \exp\left(-\frac{T}{2}(\kappa_j - \overline{\kappa}_j)^2\right).
\end{aligned}$$

References

[1] Amisano, G., 2003. Bayesian inference in cointegrated systems. Discussion paper 9903, Dipartimento di Scienze Economiche, Universita degli studi di Brescia.

[2] Bauwens, L., Giot, P., 1997. A Gibbs sampling approach to cointegration. Computational Statistics, 13 (3), 339-368.

[3] Chen, M.-H., 1994. Importance-weighted marginal Bayesian posterior density estimation. Journal of the American Statistical Association 89, 818–824.

[4] Engle, R.F., Granger, C.W.J., 1987. Co-integration and error correction: representation, estimation, and testing. Econometrica 55, 251–276.

[5] Fisher, L.A., Huh, H-S, Summers, P. 2000. Structural Identification of Permanent Shocks in Vector Error Correction Models: a Generalization. Journal of Macroeconomics, 22, 53-68.

[6] Gelfand, A.E., Dey, D.K., 1994. Bayesian model choice: asymptotics and exact calculations. Journal of the Royal Statistical Society Series B, 56, 501-514.

[7] Geweke, J. 1999. Using Simulation Methods For Bayesian Econometric Models: Inference, Development, and Communication. Econometric Review, 18, 1-73.

[8] Granger, C.W.J., 1981. Some properties of time series data and their use in econometric model specifcation. Journal of Econometrics 16, 121–130.

[9] Granger, C.W.J., Weiss, A.A., 1983. Time series analysis of error-correction models, Studies in Econometrics, Time Series, and Multivariate Statistics, Academic Press, New York, 255–278.

[10] Johansen, S. 1995. Identifying restrictions of linear equations with applications to simultaneous equations and cointegration, Journal of Econometrics, 69, 111-132.

[11] Johansen, S., Juselius, K. 1992. Testing structural hypotheses in a multivariate cointegration analysis of the PPP and UIP for UK, Journal of Econometrica, 53, 211-244.

[12] Kleibergen, F., Paap, R. 2002. Priors, Posterior Odds and Bayes Factors in Bayesian Analyses of Cointegration, Journal of Econometrics, 111, 223-249.

[13] King, R.G., Plosser, C.J., Stock, S.H., Watson, M.W. 1991. Stochastic Trends and Economic Fluctuations, American Economic Review, 81, 819-840.

[14] Kleibergen, F., van Dijk, H.K. 1994. On the shape of the likelihood posterior in cointegration models. Econometric Theory 10, 514–551.

[15] Koop, G. 2003. Bayesian Econometrics. Wiley.

[16] Leeper, E., Sims, C.A., Zha, T. 1996, What does monetary policy do?, Brookings Papers on Economic Activity 2, 1-63.

[17] Ritter. C., and Tanner, M.A. 1992. Facilitating the Gibbs Sampler: The Gibbs Stopper and the Griddy-Gibbs Sampler, Journal of the American Statistical Association, 87, 861-868.

[18] Sims, C.A., 1980. Macroeconomics and Reality. Econometrica 48, 1-48.

[19] Sims, C.A., Zha, T. 1998. Bayesian methods for dynamic multivariate models. International Economic Review 39, 949-968.

[20] Sims, C.A., Zha, T. 1999. Error bands for impulse responses. Econometrica 67, 1113 - 1155.

[21] Strachan, R.W., 2003. Valid Bayesian estimation of the cointegrating error correction model. Journal of Business and Economic Statistics 21, 185-195.

[22] Sugita, K. 2001. Bayesian cointegration analysis. Warwick Economic Research Papers, 591. University of Warwick.

[23] Waggoner, D.F., Zha, T. 2003. A Gibbs simulator for restricted VAR models. Journal of Economic Dynamics and Control 28, 349-366.

Figure 1.Distributions of κ_1

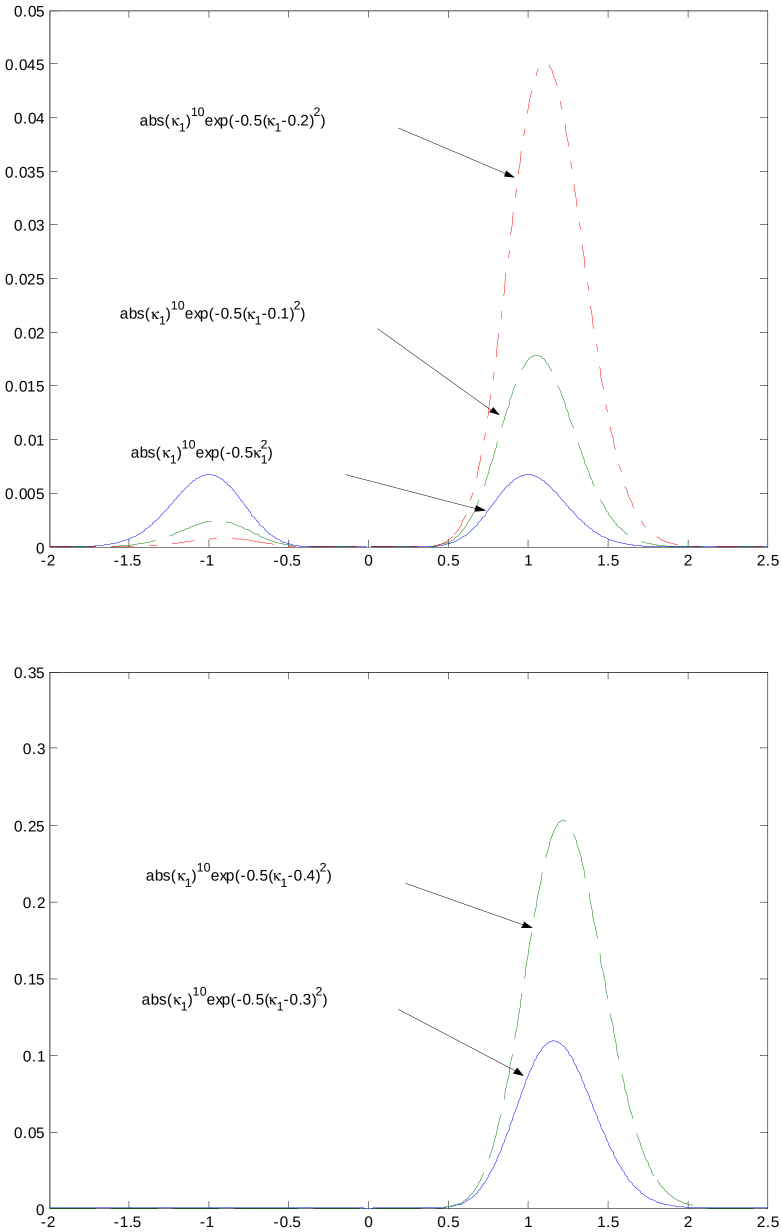

Figure 2 Marginal posterior parameters' pdfs and actual parameter values for $r = 1$ and DGP 2

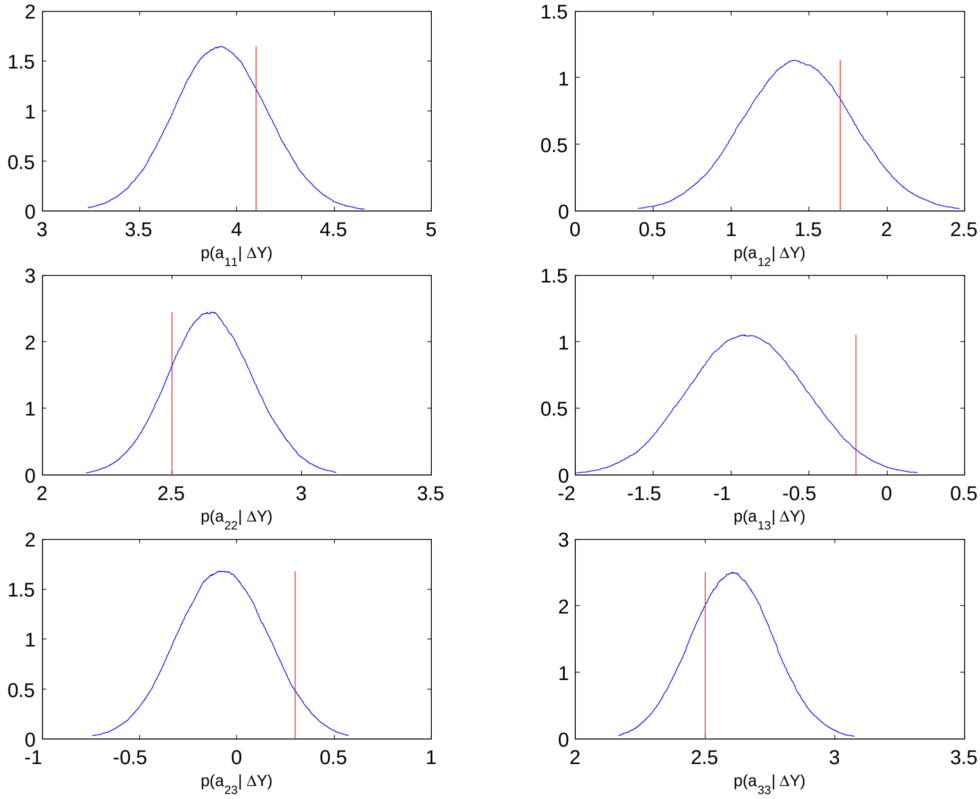

Figure 2 continues....

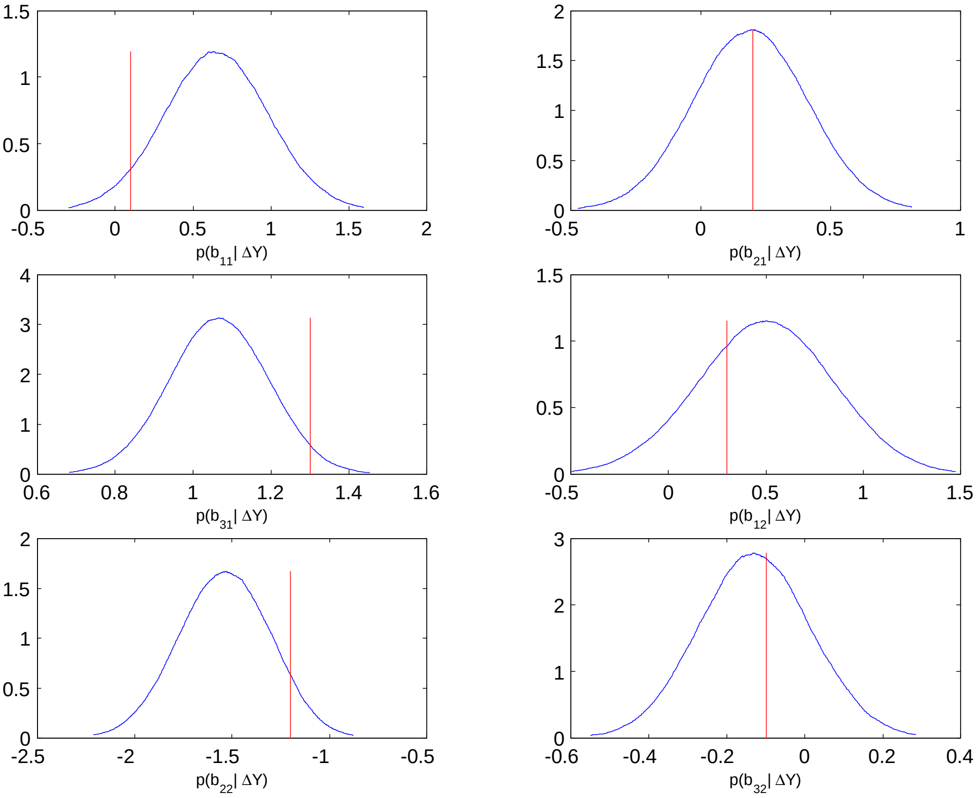

Figure 2 continues

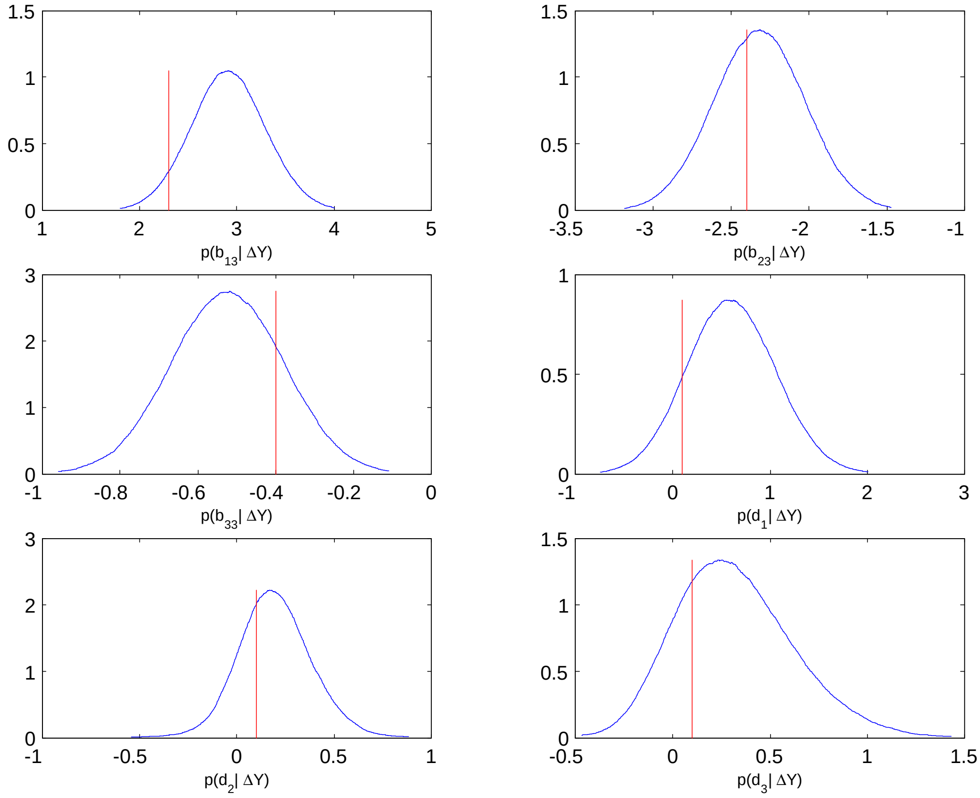

Figure 2 continues

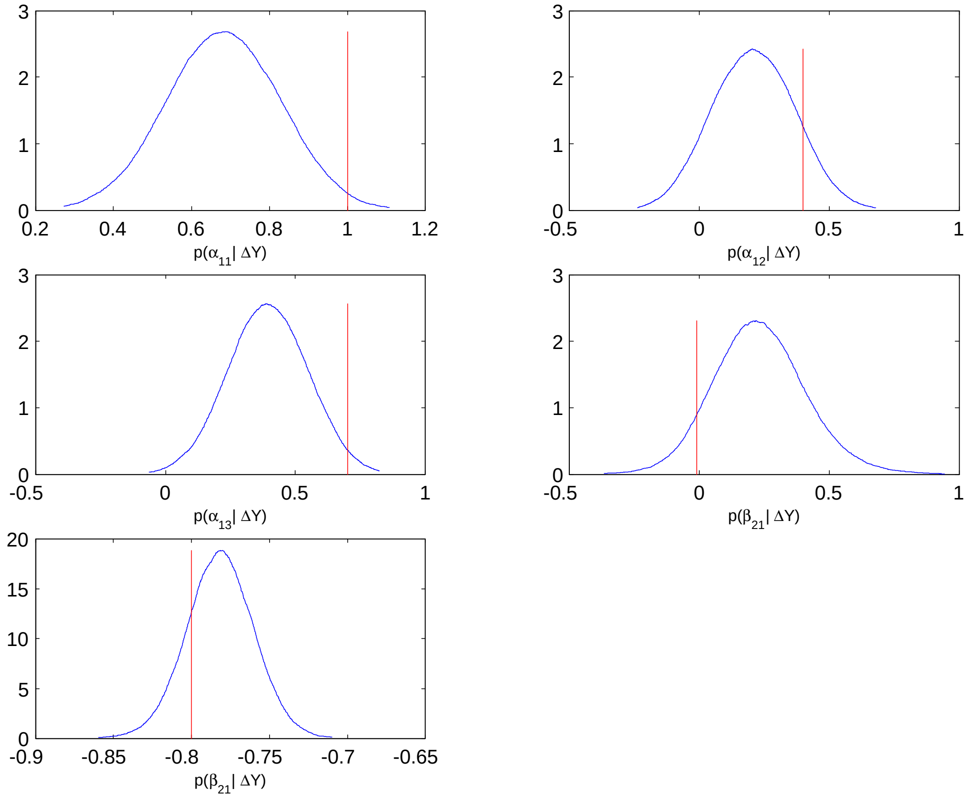

Table 1: Marginal likelihoods, Bayes factors and posterior probabilities

r	Log Marginal Likelihoods	Log Bayes Factors (r|3)	Posterior Probabilities
		DGP 1	
0	277.27	**8.06**	0.73
1	276.02	6.82	0.21
2	274.74	5.54	0.06
3	269.20	0	0
		DGP 2	
0	222.19	-3.34	0
1	231.31	**5.78**	0.84
2	229.66	4.12	0.16
3	225.53	0	0
		DGP 3	
0	237.52	-10.84	0
1	256.45	8.09	0.47
2	256.57	**8.21**	0.53
3	248.36	0	0
		DGP 4	
0	185.19	-46.01	0
1	202.87	-28.33	0
2	210.99	-20.21	0
3	231.20	**0**	1

Table 2: Marginal likelihoods, Bayes factors and posterior probabilities for the KPSW example

r	Log Marginal Likelihoods	Log Bayes Factors (r|6)	Posterior Probabilities
0	2415.7	13.699	0
1	2410.8	8.7578	0
2	2426.6	24.641	0.008
3	2422.5	20.457	0
4	2415.2	13.209	0
5	2431.5	29.514	0.992
6	2402	0	0