

SVERIGES RIKSBANK  
WORKING PAPER SERIES

156



# Monetary Policy Analysis in a Small Open Economy using Bayesian Cointegrated Structural VARs

*Mattias Villani and Anders Warne*

DECEMBER 2003

# MONETARY POLICY ANALYSIS IN A SMALL OPEN ECONOMY USING BAYESIAN COINTEGRATED STRUCTURAL VARs\*

MATTIAS VILLANI<sup>†</sup> AND ANDERS WARNE<sup>‡</sup>

SVERIGES RIKSBANK WORKING PAPER SERIES  
No. 156  
DECEMBER 2003

## ABSTRACT

Structural VARs have been extensively used in empirical macroeconomics during the last two decades, particularly in analyses of monetary policy. Existing Bayesian procedures for structural VARs are at best confined to a severely limited handling of cointegration restrictions. This paper extends the Bayesian analysis of structural VARs to cover cointegrated processes with an arbitrary number of cointegrating relations and general linear restrictions on the cointegration space. A reference prior distribution with an optional small open economy effect is proposed and a Gibbs sampler is derived for a straight-forward evaluation of the posterior distribution. The methods are used to analyze the effects of monetary policy in Sweden.

KEYWORDS: Structural, Vector autoregression, Monetary policy, Impulse responses, Counterfactual experiments.

JEL CLASSIFICATION: C11, C32, E52.

---

\*The authors thanks Eric M. Leeper for helpful comments on an earlier version of the paper. All non-Bayesian computations were performed with Anders Warne's software package *Structural VAR* which is freely available at <http://www.texlips.net/warne/code.html>. The opinions expressed herein are those of the author(s) and do not necessarily represent those of the European Central Bank or the Executive Board of Sveriges Riksbank.

<sup>†</sup>*Sveriges Riksbank, Research Department, SE-10337 Stockholm, Sweden, and Department of Statistics, Stockholm University. E-mail: mattias.villani@riksbank.se.*

<sup>‡</sup>*European Central Bank, Postfach 160319, 60311 Frankfurt am Main, Germany. E-mail: anders.warne@ecb.int.*

## 1. INTRODUCTION

Following the seminal paper by Sims (1980), structural vector autoregressions (SVAR) have become one of the most frequently used methods for empirical analysis of, e.g., the effects of monetary policy on output and prices. In these early days, SVARs were identified by assuming a recursive structure for the variables in the system and the parameters were estimated using least squares. The main tools for analysing the effects of structural shocks were impulse response functions and forecast error variance decompositions.

During the last two decades we have seen a number of important developments to the basic methodology. First, non-recursive identification procedures have been suggested (e.g., Bernanke (1986) and Sims (1986)), allowing variables to be simultaneously determined. When the vector of variables includes so called *fast* variables, such as interest rates and exchange rates, this is often a desirable feature. Second, Bayesian methods, often based on the so called Minnesota prior (e.g., Litterman, 1986), have emerged and proved useful mainly in forecasting exercises but also in policy analysis. Third, the link between VARs and cointegration in error correction models has been established (e.g., Engle and Granger, 1987) and is nowadays widely used to study long-run relations in data driven by unit-root (random walk) processes.

In contrast to the classical approach, with sampling distributions depending on whether the data generating process is stationary or not, the order of integration of the data has no bearing on Bayesian inference of time series data. Cointegration restrictions may nevertheless be useful for imposing long run structure on the otherwise loosely structured VARs, and they are especially important from a forecasting point of view as the imposed long run relations can be shown to be satisfied asymptotically at long forecast horizons (Engle and Yoo, 1987; Christoffersen and Diebold, 1998).

Sims and Zha (1998) propose a Bayesian approach to structural VAR modelling where the prior is approximated via the use of dummy observations similar to the mixed estimation approach of Theil and Goldberger (1961). Two of the dummy variables produce a shrinkage effect toward either none or a single cointegrating relation, without specifying the form of this relation.

In this paper we propose a more general framework for Bayesian analysis of cointegrated structural VARs (CSVAR) which allow for an arbitrary number of cointegrating relations and general linear over-identifying restrictions on the cointegration vectors. The approach extends some of Villani's (2000, 2003) work on reduced-form cointegrated VARs. In addition, we suggest extensions to the Minnesota prior both in the direction of dealing with potential exogeneity (as opposed to Cushman and Zha, 1997, who impose exact exogeneity) of certain variables and treatment of seasonality. Concerning the simultaneity of the model our approach is based on ideas presented in Sims and Zha (1999) and Waggoner and Zha (2003b).

As an illustration of our Bayesian CSVAR method we apply it to Swedish quarterly data covering 1975–2001. The identifying assumptions we consider are based on contemporaneous restrictions and, apart from imposing unit roots, are similar to those considered by Kim and Roubini (2000) for the non-US G7 countries. The vector of variables is also similar, but the foreign variables are TCW weighted rather than represented by the US.

We find that the identified monetary policy shock generally has the expected effects on the domestic variables. In particular, the price and exchange rate puzzles (rising prices and impact depreciation of the domestic currency following a monetary contraction) reported by Jacobson, Jansson, Vredin and Warne (2002) are no longer present. Moreover, the monetary policy shock is primarily important for short run fluctuations in the domestic interest rate (and to a lesser extent for the exchange rate), but has only negligible effects at business cycle frequencies. We also consider a few counterfactual policy experiments using conditional forecasts, as suggested

by Leeper and Zha (2003), and find that quite different interest rate paths are consistent with the idea of modest policy interventions.

The remainder of the paper is structured as follows. In the next section we briefly introduce the CSVAR model to establish notation. The prior distribution of its parameters is discussed in detail in Section 3, while the main results on conditional posterior distributions are presented in Section 4. In Section 5 we present the empirical application to the Swedish data, while Section 6 concludes. Finally, proofs of the theorems are given in the Appendices.

## 2. COINTEGRATED STRUCTURAL VARS

The cointegrated structural VAR (CSVAR) is of the form

$$(2.1) \quad \Gamma_0' \Delta x_t = \alpha \beta' x_{t-1} + \Gamma(L) \Delta x_t + \Phi' d_t + \varepsilon_t, \quad t = 1, \dots, T,$$

where  $x_t$  is  $p$ -dimensional vector containing the endogenous variables at time  $t$  and  $d_t$  is a  $q$ -dimensional vector of non-modelled exogenous variables (including lags), deterministic trends and seasonal components<sup>1</sup>.  $\Gamma_0$  is a  $p \times p$  non-singular matrix of contemporaneous coefficients,  $\Gamma(\lambda) = \sum_{i=1}^{k-1} \Gamma_i' \lambda^i$ , is a  $p \times p$  matrix polynomial with complex valued argument  $\lambda$ , and  $L$  is the usual lag operator such that  $L^j x_t = x_{t-j}$ . The columns of  $\beta$  ( $p \times r$ ) are the cointegration vectors which defines  $r$  stationary long run relations between the endogenous variables and  $\alpha$  ( $p \times r$ ) is the matrix of adjustment coefficients. The structural shocks,  $\varepsilon_t$ , are assumed to be i.i.d. Gaussian with zero mean and identity covariance matrix, following the tradition in structural VAR modelling, see e.g. Leeper, Sims and Zha (1996) for a motivation.

Following Cushman and Zha (1997), the vector of observations  $x_t$  is partitioned as  $x_t = (x_{1,t}', x_{2,t}')'$ , where  $x_2$  is assumed to be exogenous with respect to  $x_1$ . In the small open economy setting,  $x_2$  represents the foreign variables and  $x_1$  the domestic variables, whose effect on  $x_2$  is assumed to be close to zero *a priori*, see Section 3.

A convenient form of the model in (2.1) for the whole sample of  $T$  time periods is obtained by transposing (2.1) for each time period and stacking the resulting row vectors in matrices, yielding

$$(2.2) \quad X \Gamma_0 = W \beta \alpha' + Q \Gamma + E,$$

where the  $t$ th row of  $X$ ,  $W$ ,  $Q$  and  $E$  is given by  $\Delta x_t'$ ,  $x_{t-1}'$ ,  $(\Delta x_{t-1}', \dots, \Delta x_{t-k+1}', d_t')$  and  $\varepsilon_t'$ , respectively, and  $\Gamma = (\Gamma_1, \dots, \Gamma_{k-1}, \Phi)'$ .  $D = \{X, W, Q\}$  will be used as a short hand for the available data.

## 3. IDENTIFYING RESTRICTIONS AND PRIOR DISTRIBUTIONS

**3.1. Prior distribution on  $\Gamma_1, \dots, \Gamma_k$ .** Let  $\gamma_{ij}^{(k)}$  be the  $(i, j)$ th element of  $\Gamma_k'$ ,  $i, j = 1, \dots, p$ , i.e.  $\gamma_{ij}^{(k)}$  is the coefficient on lag  $k$  of the  $j$ th endogenous variable in equation  $i$ . We will assume the  $\gamma_{ij}^{(k)}$  to be independent (across  $i, j$  and  $k$ ) with prior distribution

$$\gamma_{ij}^{(k)} \sim N[0, s^2(\gamma_{ij}^{(k)})], \quad i, j = 1, \dots, p; k = 1, 2, \dots,$$

where  $s(\gamma_{ij}^{(k)})$  is the prior standard deviation of  $\gamma_{ij}^{(k)}$ .

The starting point for our prior is the so called Minnesota prior (Litterman, 1986; Robertson and Tallman, 1999; Sims and Zha, 1998) with

$$s(\gamma_{ij}^{(k)}) = \frac{\lambda_b}{k^{\lambda_l} \sigma_j},$$

---

<sup>1</sup>Constant and trends may be restricted to the cointegration space simply by moving the relevant deterministic component from  $d_t$  to  $x_{t-1}$  as explained in Johansen (1995, Ch. 5).

where  $\lambda_b > 0$  determines the overall tightness of the prior on  $\Gamma_1, \dots, \Gamma_k$  around zero,  $\lambda_l > 0$  controls the rate at which the coefficients are shrunk toward zero with increasing lag length and  $\sigma_1, \dots, \sigma_p$  are scale factors which takes into account the possibly differing measurement units of the endogenous variables. Following Sims and Zha (1998), no scale factor is used in the numerator of  $s(\gamma_{ij}^{(k)})$ , as the structural shocks in  $\varepsilon_t$ , are normalized to have unit variance.

We extend the Minnesota prior in two directions. First, rather than imposing the restriction that the rest of the world ( $x_2$ ) is exogenous with respect to the small economy ( $x_1$ ) with probability one, as in Cushman and Zha (1997), we add an additional hyperparameter  $\lambda_e$  to control the tightness *around* the exogeneity restriction. Secondly, instead of using a uniform shrinkage with respect to the lag length, we use different shrinking pattern on the coefficients on seasonal and non-seasonal lags. Note that if the time series have a season of  $s$  time periods then the seasonal coefficients in the SCVAR model appear in  $\Gamma_{s-1}, \Gamma_s, \Gamma_{2s-1}, \Gamma_{2s}, \dots$ . The extended Minnesota prior sets

$$s(\gamma_{ij}^{(k)}) = \frac{\lambda_b \lambda_e^{I_i(j)} \lambda_s^{I_s(k)}}{f_s(k) \lambda_l \sigma_j},$$

where  $I_i(j)$  is an exogeneity indicator which equals 1 if variable  $j$  is assumed exogenous in equation  $i$  and zero otherwise,  $0 < \lambda_e \leq 1$  determines the strength of the beliefs in the exogeneity assumption,  $I_s(k)$  is a seasonality indicator which takes the value 1 for  $k = s - 1, s, 2s - 1, 2s, \dots$  and zero otherwise, and

$$f_s(k) = \begin{cases} \frac{k}{s-1} & \text{if } (s-1) \mid k \\ \frac{k}{s} & \text{if } s \mid k \\ k & \text{otherwise,} \end{cases}$$

where  $m \mid n$  denotes that  $m$  divides  $n$ , i.e. that  $n = cm$  for some integer  $c$ . The hyperparameter  $\lambda_s$  ( $0 < \lambda_s \leq 1$ ) determines the importance of coefficients on seasonal lags relative coefficients on non-seasonal lags. Note how the  $f_s(k)$  function upgrades the importance of the seasonal lags compared to the non-seasonal ones.

**3.2. Prior Distribution on  $\Phi$ .** The following prior will be used for  $\Phi$

$$\text{vec } \Phi \sim N_{pq}(0, \lambda_d^2 I_{pq}),$$

where  $\lambda_d$  is a shrinkage factor.

In summary, the prior on  $\Gamma = (\Gamma'_1, \dots, \Gamma'_{k-1}, \Phi')'$  is

$$\text{vec } \Gamma \sim N_{p^2(k-1)+pq}(0, \Omega_\Gamma),$$

where

$$\Omega_\Gamma = P' \begin{pmatrix} \Psi & 0 \\ 0 & \lambda_d^2 I_{pq} \end{pmatrix} P$$

and  $\Psi$  is a diagonal matrix with typical element  $s^2(\gamma_{ij}^{(k)})$  and  $P$  is the commutation matrix (Harville, 1997, Sec. 16.3) which maps  $\text{vec}(\Gamma')$  into  $\text{vec } \Gamma$ .

**3.3. Prior Distribution on Contemporaneous Coefficients in  $\Gamma_0$ .** Let  $\gamma_i^{(0)}$  be the contemporaneous coefficients in the  $i$ th equation of the system, i.e.  $\gamma_i^{(0)}$  is the  $i$ th column of  $\Gamma_0$ . The assumed identity covariance matrix of the structural shocks fixes the scale of  $\Gamma_0$ , but leaves the model invariant to rotations of  $\Gamma_0$ . General linear restrictions on each column of  $\Gamma_0$  may be used to settle this indeterminacy

$$(3.1) \quad \gamma_i^{(0)} = G_i \psi_i, \quad i = 1, \dots, p,$$

where  $G_i$  ( $p \times s_i$ ) determines the restrictions and  $\psi_i$  is a  $s_i$ -vector of unrestricted coefficients.

Under the restrictions in (3.1), the model is still invariant to sign switches in the equations, as discussed in Sims and Zha (1999) and Waggoner and Zha (2003a). This may be settled by restricting one element in each  $\psi_i$  to be positive; see Section 4 for more details.

The fact that  $\Gamma_0$  contains both location and scale components makes it difficult to specify a prior distribution on  $\psi_1, \dots, \psi_p$ . A way out of this problem would be to use a model where the location and scale components of  $\Gamma_0$  are modelled as separate parameters. This leads to other difficulties, however, such as potentially improper posteriors (Sims and Zha, 1994). Here we settle with a uniform prior on  $\psi_1, \dots, \psi_p$ , leaving this aspect of the overall prior to future work. We note that sign restrictions on the contemporaneous coefficients advocated by Uhlig (1997) may be imposed simply by restricting the domain of  $\Gamma_0$ . The posterior computations remain exactly the same apart from a rejection step in the algorithm following a  $\Gamma_0$ -draw in conflict with the sign restrictions.

**3.4. Prior Distribution on the Adjustment Coefficients in  $\alpha$ .** The prior distribution on  $\alpha$  is modelled conditional on  $\beta$  as

$$\text{vec } \alpha | \beta \sim N_{pr}[0, (\beta' K_1 \beta)^{-1} \otimes K_2]$$

where  $K_1 = \text{diag}(\sigma_1^2, \dots, \sigma_p^2)$ , the  $\sigma$ 's are the scaling factors used in the  $\Gamma$ -prior, and

$$K_2 = \begin{pmatrix} \lambda_\alpha^2 I_{p_1} & 0 \\ 0 & \lambda_\alpha^2 \lambda_e^2 I_{p_2} \end{pmatrix}.$$

To understand the idea behind this prior, note that, conditional on  $\beta$ ,  $\alpha$  is the (partial) coefficient in the regression of  $\Gamma_0' \Delta x_t$  on the predictor vector  $\beta' x_{t-1}$ . A shrinkage prior on a regression coefficient should take into account the scale of both the response variables and the predictors.  $x_{t-1}$  is not known when the prior is formulated, however, and the best one can do is to include the scale factors  $\sigma_1, \dots, \sigma_p$ . This explains the appearance of  $(\beta' K_1 \beta)^{-1}$  in the prior conditional covariance matrix of  $\alpha$ . Since the response variables are assumed to have unit variances and zero covariances, identity matrices are used in  $K_2$ .

$K_2$  also contains the exogeneity parameter  $\lambda_e$ , which was also used in the  $\Gamma$ -prior, allowing the lower left block of  $\alpha \beta'$  (which contains the effects of the domestic variables on the foreign variables) to be more heavily shrunk toward zero compared to the other elements of  $\alpha \beta'$ . The adjustment coefficients corresponding to an entirely foreign relation, i.e. a cointegration vector with zero restrictions on the coefficients on all domestic variables, should of course not be given the additional  $\lambda_e$ -shrinkage.

**3.5. Prior Distribution on the Cointegration Vectors in  $\beta$ .** It is well known that the cointegration vectors are only determined up to arbitrary linear combinations. Here we will use general linear identifying restrictions on each cointegration vector to identify  $\beta$

$$(3.2) \quad \beta_i = h_i + H_i \phi_i, \quad i = 1, \dots, r,$$

where  $h_i$  is a vector containing the fixed part of  $\beta_i$ , including normalizations,  $H_i$  ( $p \times d_i$ ) determines the restrictions and  $\phi_i$  is a  $d_i$ -vector of free coefficients.

We suggest the following prior on  $\phi_1, \dots, \phi_r$  as a suitable reference prior

$$(3.3) \quad p(\phi_1, \dots, \phi_r) \propto |\beta' K_1 \beta|^{-p/2},$$

where  $K_1 = \text{diag}(\sigma_1^2, \dots, \sigma_p^2)$  and the  $\sigma$ 's are the scaling factors used in the  $\Gamma$ -prior. The properties of this prior has been investigated in detail under just-identifying restrictions with  $K_1 = I_p$  by Villani (2000, 2003), where it is shown that it assigns the same prior probability to each possible cointegration space of dimension  $r$  (more precisely, it implies a Haar invariant distribution over

the quotient space of all cointegration spaces). If  $\sigma_i^2 \neq \sigma_j^2$  for any  $i \neq j$ , the prior on  $\phi_1, \dots, \phi_r$  no longer has this property, but the prior in (3.3) should nevertheless constitute a reasonable reference prior, especially if the time series have widely differing scales, and has the additional benefit of yielding straightforward posterior computations.

In summary, the overall prior distribution depends on the scaling factors  $\sigma_1, \dots, \sigma_p$ , and the following six hyperparameters

$\lambda_b$	:	Baseline shrinkage
$\lambda_l$	:	Lag length shrinkage
$\lambda_s$	:	Seasonality shrinkage
$\lambda_\alpha$	:	Adjustment shrinkage
$\lambda_e$	:	Exogeneity shrinkage
$\lambda_d$	:	Deterministic shrinkage

The scaling factors  $\sigma_1, \dots, \sigma_p$  are not hard to elicit given a proper understanding of the analyzed time series, but are for convenience usually estimated from data as the residual standard deviation in univariate autoregressive models fitted to the each of the series, see e.g. Litterman (1986) and Sims and Zha (1998).

#### 4. POSTERIOR DISTRIBUTION

The joint posterior distribution of the SCVAR parameters is intractable. In this section we derive the necessary results for a numerical evaluation of the posterior distributions via the Gibbs sampler. The Gibbs sampler (e.g. Tierney, 1994) simulates from the joint posterior distribution of model parameters by iteratively generating draws from the full conditional posterior distributions, i.e. the posterior distribution of a group of model parameters conditional on all other parameters in the model. Denoting the full set of model parameters by  $\theta$ , the end result of the Gibbs sampling is a sequence of draws  $\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(n)}$  from the posterior distribution. The draws are not independent but the following can be shown to hold under certain conditions (Tierney, 1994) which are satisfied here

$$g(\theta^{(i)}) \xrightarrow{d} p[g(\theta)|\mathcal{D}]$$

where  $\xrightarrow{d}$  denotes convergence in distribution and  $g(\cdot)$  is any well-behaved real valued function. Thus, given the draws from the posterior of  $\theta$ , the posterior distribution of any function of the model parameters, e.g. the impulse response functions or forecast error variance decompositions, are immediately available simply by applying the  $g$ -function to each posterior draw and subsequently using some density estimator.

The next theorem gives the full conditional posteriors of  $\Gamma$ ,  $\alpha$  and  $\phi = (\phi'_1, \dots, \phi'_r)'$ .

*Theorem 1.*

- Full conditional posterior of  $\Gamma$

$$\text{vec } \Gamma | \Gamma_0, \alpha, \beta, \mathcal{D} \sim N_{p[p(k-1)+q]}(\bar{\mu}_\Gamma, \bar{\Omega}_\Gamma),$$

where  $\bar{\Omega}_\Gamma^{-1} = (I_p \otimes Q'Q) + \Omega_\Gamma^{-1}$  and  $\bar{\mu}_\Gamma = \bar{\Omega}_\Gamma \text{vec}[Q'(X\Gamma_0 - W\beta\alpha)']$ .

- Full conditional posterior of  $\alpha$

$$\text{vec } \alpha | \Gamma_0, \Gamma, \beta, \mathcal{D} \sim N_{pr}(\bar{\mu}_\alpha, \bar{\Omega}_\alpha),$$

where  $\bar{\Omega}_\alpha^{-1} = (\beta'W'W\beta \otimes I_p) + (\beta'K_1\beta \otimes K_2^{-1})$  and  $\bar{\mu}_\alpha = \bar{\Omega}_\alpha \text{vec}[(X\Gamma_0 - Q\Gamma)'W\beta]$ .  
 If  $\lambda_e = 1$ , then  $\bar{\Omega}_\alpha^{-1}$  simplifies to  $\lambda_\alpha^{-2}[\beta'(W'W + K_1)\beta] \otimes I_p$ .

- Full conditional posterior of  $\phi = (\phi'_1, \dots, \phi'_r)'$

$$\phi|\Gamma_0, \Gamma, \alpha, \mathcal{D} \sim N_d(\bar{\mu}_\phi, \bar{\Omega}_\phi),$$

where  $d = \sum_{i=1}^r d_i$ ,  $\bar{\Omega}_\phi^{-1} = H'AH$ ,  $A = \alpha'\alpha \otimes W'W + \alpha'K_2^{-1}\alpha \otimes K_1$ ,  
 $\bar{\mu}_\phi = \bar{\Omega}_\phi H' \{\text{vec}[W'(X\Gamma_0 - Q\Gamma)\alpha] - Ah\}$ ,  $H = \text{diag}(H_1, \dots, H_r)$  and  $h = (h'_1, \dots, h'_r)'$ . If  
 $\lambda_e = 1$ , then  $A$  simplifies to  $\alpha'\alpha \otimes (W'W + \lambda_\alpha^{-2}K_1)$ .

*Proof.* See the appendix.  $\square$

The full conditional posterior distributions of the simultaneous effects,  $\psi_i$ ,  $i = 1, \dots, p$ , do not belong to a known family of distributions and direct sampling is therefore not feasible. A solution to this problem for the structural VAR model without cointegration restrictions is given in Waggoner and Zha (2003b, Theorem 1). The Waggoner and Zha (2003b) solution cannot be applied directly when cointegration restrictions are allowed for, but a straightforward modification of their result is as follows. First we need the following definition.

*Definition 1.* A random variable  $X$  follows the *absolute normal distribution*  $AN(\mu, \rho)$  if it has density function

$$f_{AN}(x; \mu, \rho) = c|x|^{\frac{1}{\rho}} \exp\left[-\frac{1}{2\rho}(x - \mu)^2\right], \quad x \in \mathbb{R},$$

where  $c$  is a normalizing constant,  $\rho \in \mathbb{R}^+$  and  $\mu \in \mathbb{R}$ .

In the next theorem, let  $B_{-i}$  equal the matrix  $B$  with the  $i$ th column deleted,  $B_\perp$  is the orthogonal complement of  $B$ ,  $\text{Chol}(B)$  is the Choleski root of  $B$  such that  $B = \text{Chol}(B)\text{Chol}(B)'$ ,  $\|\cdot\|$  is the usual Euclidean norm and  $\stackrel{d}{=}$  denotes equality in distribution. With these preliminaries at hand we can prove the following extension of Theorem 1 in Waggoner and Zha (2003b).

*Theorem 2.* The full conditional posterior of the coefficients in the  $i$ th simultaneous relation is

$$\psi_i|\Gamma_{0,-i}, \alpha, \beta, \Gamma, \mathcal{D} \stackrel{d}{=} R_i \sum_{j=1}^{s_i} \xi_j v_j,$$

where  $R_i = \text{Chol}[T(G'_i X' X G_i)^{-1}]$ ,  $\xi_1 \sim AN(\hat{\xi}_1, T^{-1})$ ,  $\xi_j \sim N(\hat{\xi}_j, T^{-1})$ , for  $j = 2, \dots, s_i$ ,  $\hat{\xi}_j = \mu'_{\psi_i} R_i^{-1} v_j$ ,  $v_1 = R'_i G'_i \Gamma_{0-i\perp} / \|R'_i G'_i \Gamma_{0-i\perp}\|$ ,  $(v_2, \dots, v_{s_i}) = v_{1\perp}$  and  $\mu_{\psi_i} = (G'_i X' X G_i)^{-1} G'_i X' z_i$ , where  $z_i$  is the  $i$ th column of  $W\beta\alpha' + Q\Gamma$ .

*Proof.* See the appendix.  $\square$

The sign restriction mentioned in Section 2 is easily imposed by replacing a generated  $\psi_i$  which does not satisfy the sign restriction with the  $\psi_i$  generated in the previous iteration of the Gibbs sampler. Thus, if the sign restriction is not satisfied in a given iteration,  $\psi_i$  is not updated at this iteration.<sup>2</sup> A good choice of normalizing variables (see Waggoner and Zha, 2003a, for an automatic method) usually gives a small number of the rejected draws. In general, an equation should be normalized on a variable whose coefficient is likely to have negligible posterior probability in a neighborhood of zero.

<sup>2</sup>This updating step of the algorithm then becomes a Metropolis-Hastings step, rather than a pure Gibbs step (Gilks, Richardson and Spiegelhalter, 1996).



In order to use Theorem 2 in a Gibbs sampler we need to be able to efficiently generate variates from the absolute normal distribution. A simple, very accurate and readily sampled approximation can be derived from the fact that the  $AN(\mu, \rho)$ -distribution is bimodal with modes at  $\frac{\mu}{2} \pm \frac{1}{2}\sqrt{(\mu^2 + 4)}$ . Furthermore, the curvature at  $x = x_0$  is

$$-\left[\frac{d^2}{dx^2} \ln f(x)\right]^{-1}\bigg|_{x=x_0} = \frac{x_0^2}{(1+x_0^2)}\rho.$$

These two facts can be used to build the following mixture normal approximation to the  $AN(\mu, \rho)$  distribution

$$f_{AN}(x; \mu, \rho) \approx vN(x; \mu_1, \sigma_1^2) + (1-w)N(x; \mu_2, \sigma_2^2),$$

where  $N(x; \cdot, \cdot)$  is used as a shorthand for the density of a normal distribution,  $\mu_1 = \frac{\mu}{2} - \frac{1}{2}\sqrt{(\mu^2 + 4)}$ ,  $\mu_2 = \frac{\mu}{2} + \frac{1}{2}\sqrt{(\mu^2 + 4)}$ ,  $\sigma_i^2 = \frac{\mu_i^2}{(1+\mu_i^2)}\rho$ ,  $i = 1, 2$ , and  $w = [1 + \exp(2\mu\rho^{-1})]^{-1}$ . It is easily seen that, for a given  $\rho$ , the approximation error is maximal when  $\mu = 0$ . The accuracy of this approximation increases inversely with  $\rho$  and even in the worst scenario when  $\mu = 0$  it is already very accurate for  $\rho = 0.1$ . In our use of the absolute normal distribution,  $\rho = T^{-1}$ , where  $T$  is the length of the time series, so the approximation can, for all practical purposes, be taken as exact.

## 5. MONETARY POLICY IN A SMALL OPEN ECONOMY

**5.1. Data and Model.** Kim and Roubini (2000) studies the effects of monetary policy in the six non-US G7 countries using a structural VAR system with seven variables and non-recursive identifying restrictions on the contemporaneous coefficients. Their model avoids most of the puzzling results obtained in other studies and their conclusions seem to be rather robust across countries. Cushman and Zha (1997) use a similar scheme to identify Canadian monetary policy shocks in a larger model which also includes trade flows.

Following Kim and Roubini (2000), the data set analyzed here consists of quarterly observations from 1975:1 to 2001:4 on five Swedish variables: real GDP ( $y$ ), CPI ( $p$ ), M0 ( $m$ ), the three month bills rate in annual terms ( $i$ ), the nominal exchange rate (number of Swedish kronas needed to buy one unit of foreign currency), and two foreign variables: the oil price in US dollars ( $oil$ ) and the TCW-weighted interest rate in annual terms ( $if$ ). We shall also consider a model where CPI is replaced by a price index, where house mortgage interest expenditures, indirect taxes and subsidies have been excluded (UND1X). All variables except the two interest rates are in logs. The series are graphed in Figure 1.

Due to a break in the seasonal pattern in the late seventies, two different sets of seasonal dummies are used: one covering the period 1975:1 to 1979:4 and the other covering the period 1980:1 to 2001:4. To control for devaluations of the Swedish krona, we add five dummy variables at quarters 1976:4, 1977:2, 1977:3, 1981:3 and 1982:4. Two additional dummy variables are included in the analysis to account for the abandonment of the fixed exchange rate in 1992:4 followed by the introduction of an explicit inflation target by Sveriges Riksbank in 1993:1. The latter dummy takes the value zero during 1975:1-1992:4 and is equal to one from 1993:1 and onwards.

We use  $k = 4$  lags in the SCVAR model. Standard information criteria suggested that this is probably excessive (SBC (Schwarz, 1978) and Hannan and Quinn's HQ criteria (Quinn, 1980) both suggested  $k = 1$  whereas AIC (Akaike, 1974) preferred  $k = 4$ ), but we found the seasonal component to be better modelled with four lags. Because of the prior on  $\Gamma$ , which shrinks longer lags more heavily toward zero, increasing the lag length is not that costly in terms of lost precision.

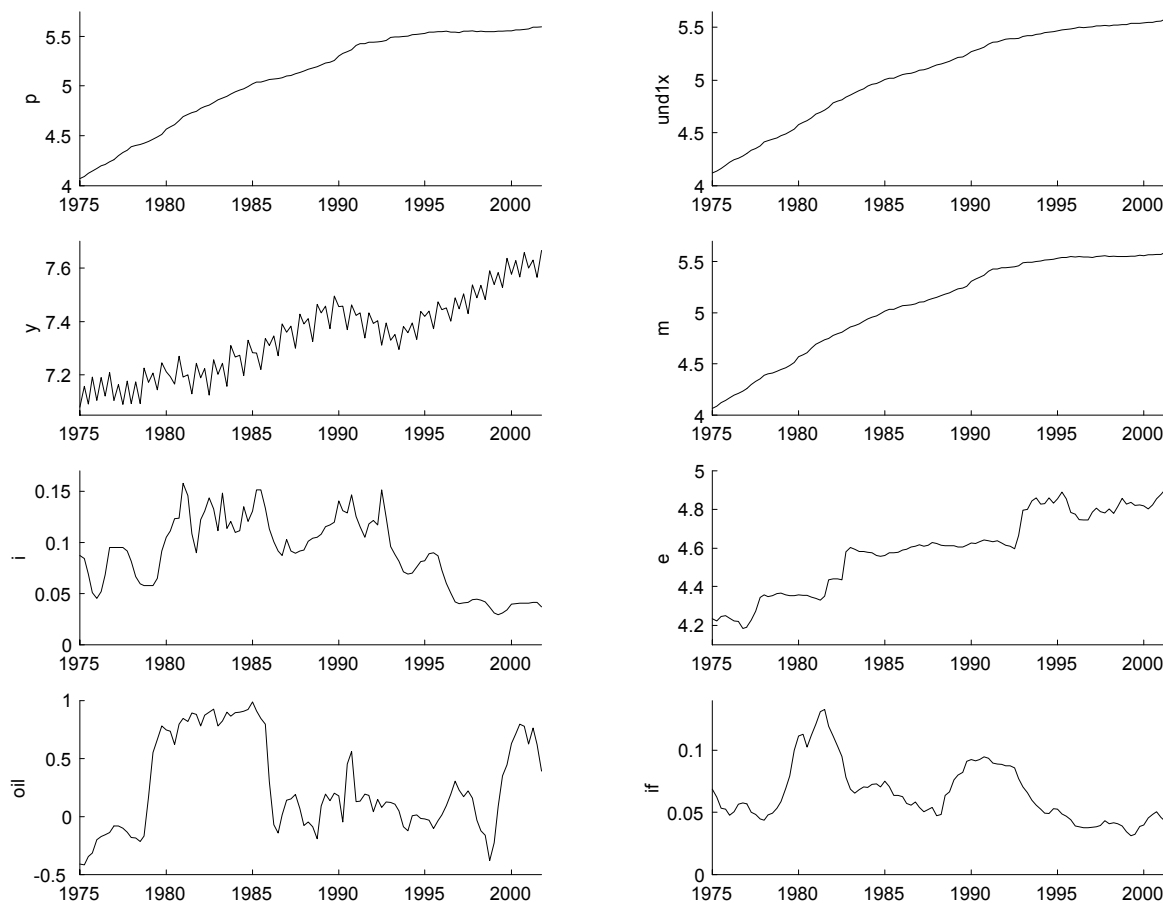


FIGURE 1. Swedish macro data 1975:1-2001:4.

To determine a value for the cointegration rank, the number of long run relations, we may use the Schwarz (1978) approximation of the posterior distribution of  $r$ :  $\Pr(r = 0|\mathcal{D}) = 0.116$ ,  $\Pr(r = 1|\mathcal{D}) = 0.863$ ,  $\Pr(r = 2|\mathcal{D}) = 0.02$ , and essentially zero for other  $r$ . The rank estimation procedure based on the trace test (Johansen, 1995) suggested that  $r = 2$  on 1% significance level and  $r = 3$  on the 5%-level using the asymptotic critical values. The corresponding test based on bootstrapped critical values suggested that  $r = 1$  on 1% significance level and  $r = 2$  on the 5%-level. The HQ criterion chose  $r = 2$  whereas AIC gave  $r = 5$  as optimal rank. In light of these results and the well known tendency of the Schwarz approximation to favor too small models, the ensuing analysis conditions on  $r = 2$ . Note, however, that all results conditional on a specific pair of lag length and cointegration rank (e.g. the impulse responses or forecasts) may be averaged over  $k$  and  $r$  with the posterior distribution as weight function (Draper, 1995).

One immediate candidate for a long run relation is the equality of the domestic and foreign interest rate, and we will take this as one of the cointegration vectors. The second cointegration vector is assumed to be a domestic relation normalized on  $m$ , with no other restrictions on the coefficients of the domestic variables. The specification of the two cointegration vectors is accomplished by six over-identifying restrictions. The Schwarz approximation of the posterior odds ratio of the restricted model to the unrestricted model yields the odds ratio 7.85, thus favoring the restricted cointegration space. The likelihood ratio test statistic is 19.27 which is

highly significant with respect to its asymptotic  $\chi_6^2$ -distribution, but the bootstrapped  $p$ -value of 0.3685 (Gredenhoff and Jacobson, 2001) confirms the finding of the Schwarz approximation.

**5.2. Identification of the Structural Shocks.** Successful application of structural VARs hinges on a proper identification of the structural shocks. Cushman and Zha (1997) and Kim and Roubini (2000) claim that non-recursive identifying restrictions are crucial for solving the price and exchange rate puzzles mentioned in the introduction. Their argument is that traditional recursive (Choleski) identification imposes restrictions which are incredible for a small open economy in that it requires either the exchange rate to be excluded from the monetary policy equation or the interest rate to be excluded from the exchange rate equation.

When applying the identifying scheme of Kim and Roubini (2000) to our data set, the interest and exchange rate equations turned out to be virtually inseparable, manifested in very uncertain impulse response functions. We therefore base our analysis on a different set of identifying restrictions, which still has many of the features of Kim and Roubini's specification, including the above mentioned simultaneity between the exchange and interest rate. The model used here is

$$(5.1) \quad \Gamma_0' \Delta x_t = \begin{pmatrix} \gamma_{11} & 0 & 0 & 0 & 0 & \gamma_{61} & 0 \\ \gamma_{12} & \gamma_{22} & 0 & 0 & 0 & \gamma_{62} & 0 \\ \gamma_{13} & \gamma_{23} & \gamma_{33} & \gamma_{43} & 0 & 0 & \gamma_{73} \\ 0 & \gamma_{24} & \gamma_{34} & \gamma_{44} & \gamma_{54} & \gamma_{64} & \gamma_{74} \\ 0 & \gamma_{25} & 0 & \gamma_{45} & \gamma_{55} & \gamma_{65} & \gamma_{75} \\ 0 & 0 & 0 & 0 & 0 & \gamma_{66} & 0 \\ 0 & 0 & 0 & 0 & 0 & \gamma_{67} & \gamma_{77} \end{pmatrix} \begin{pmatrix} \Delta y_t \\ \Delta p_t \\ \Delta m_t \\ \Delta i_t \\ \Delta e_t \\ \Delta oil_t \\ \Delta if_t \end{pmatrix} = \Upsilon_t + \begin{pmatrix} \varepsilon_{y,t} \\ \varepsilon_{p,t} \\ \varepsilon_{md,t} \\ \varepsilon_{mp,t} \\ \varepsilon_{e,t} \\ \varepsilon_{oil,t} \\ \varepsilon_{if,t} \end{pmatrix},$$

where  $\Upsilon_t = \alpha\beta'x_{t-1} + \Gamma(L)\Delta x_t + \Phi'd_t$ .

The differences between our identifying assumptions and those used in Kim and Roubini (2000) are:

- The foreign interest rate appears in both the money demand and monetary policy equation (third and fourth equation, respectively).
- The CPI appears in the monetary policy equation. Since we use quarterly data and not monthly as in Kim and Roubini, the argument that the current periods price level is unobserved by the monetary authority at the time of the policy decision is highly questionable.
- Output is excluded from the exchange rate equation, motivated by the substantial time lags in the publication of swedish GDP followed by significant revisions.
- Money does not appear in exchange rate equation. This restriction is imposed to reduce the high degree of simultaneity between money, the interest and exchange rate.

Our identifying scheme imposes four over-identifying restrictions on  $\Gamma_0$ . The posterior odds ratio comparing the model with these four over-identifying restrictions and the model with just-identifying restrictions is 0.359, indicating some preference for the just-identified model. Since all reasonable just-identifying schemes turns out to be unable to separate the money demand, monetary policy and exchange rate shocks, we nevertheless opt for the over-identified model.

Table 1 displays the presence or absence of contemporaneous effects of the structural shocks. Note that all shocks are allowed to influence money, the interest rate and the exchange rate contemporaneously.

**5.3. Results.** We use the following prior hyperparameters:  $\lambda_b = 0.3$ ,  $\lambda_l = 1$ ,  $\lambda_s = 0.5$ ,  $\lambda_\alpha = 1$ ,  $\lambda_e = 0.1$ ,  $\lambda_d = 10$ . The scale factors  $\sigma_1, \dots, \sigma_p$  are estimated from data as explained in Section 3. The results are insensitive to modest changes in the prior hyperparameters.

Variable	Shock						
	<i>y</i>	<i>p</i>	<i>md</i>	<i>mp</i>	<i>e</i>	<i>oil</i>	<i>if</i>
<i>y</i>	*	0	0	0	0	*	0
<i>p</i>	*	*	0	0	0	*	0
<i>m</i>	*	*	*	*	*	*	*
<i>i</i>	*	*	*	*	*	*	*
<i>e</i>	*	*	*	*	*	*	*
<i>oil</i>	0	0	0	0	0	*	0
<i>if</i>	0	0	0	0	0	*	*

TABLE 1. Contemporaneous effects of shocks. Stars indicate non-zero coefficients.

Variable						
<i>y</i>	<i>p</i>	<i>m</i>	<i>i</i>	<i>e</i>	<i>oil</i>	<i>if</i>
75.74 (66.6, 84.6)	—	—	—	—	0.11 (-1.3, 1.6)	—
-5.97 (-18.7, 6.2)	149.18 (132.5, 166.7)	—	—	—	-0.81 (-2.3, 0.7)	—
-21.80 (-35.5, -8.2)	-18.81 (-45.1, 7.7)	76.76 (42.5, 98.7)	-32.49 (-86.9, 17.9)	—	—	-46.15 (-98.8, 8.1)
—	-7.30 (-37.7, 23.1)	51.65 (15.5, 86.3)	90.03 (58.0, 114.7)	-21.42 (-41.5, -4.9)	-1.56 (-3.3, 0.2)	-67.26 (-116.8, -16.7)
—	-26.97 (-52.8, -2.3)	—	52.28 (20.5, 90.4)	51.02 (38.7, 59.8)	0.70 (-0.7, 2.2)	44.67 (-0.1, 90.9)
—	—	—	—	—	7.62 (6.8, 8.5)	—
—	—	—	—	—	-2.25 (-3.5, -1.0)	224.40 (199.1, 250.9)

TABLE 2. Posterior means and 90% probability intervals (in parantheses) for the unrestricted elements in  $\Gamma_0$ .

After 2500 burn-in draws, an additional 50.000 draws were simulated from the posterior distribution of the model parameters with the Gibbs sampler in Theorem 1 and 2 and every tenth draw was then saved for the inference. This subsampling reduces the autocorrelation of the draws and keeps storage demands manageable.

The first cointegration relation is, as mentioned above, the spread between the Swedish and the foreign interest rate. The second cointegrating relations is not as easily identified from data, although it can be concluded from tests of restrictions that it is probably a domestic relation. The posterior median estimate of the second cointegration relation, where the coefficients on the oil price and the foreign interest rate have been set to zero, is given by

$$m = \begin{matrix} 0.44 & y & +0.01 & p & -0.18 & i & +1.84 & e, \\ (-0.08, 0.79) & & (-0.55, 0.34) & & (-2.13, 1.36) & & (1.04, 3.40) & \end{matrix}$$

with symmetric 90% probability intervals in parantheses. Additional restrictions, e.g. that *m* and *p* have the same coefficient but with opposite signs, were not supported by data and gave rise to unreasonable estimates of the remaining unrestricted parameters.

Next we turn to the estimated simultaneous relations between the variables in Table 2. For example, the fourth relation may be interpreted as the central bank's reaction function. According to the estimated coefficients, if domestic prices increase then the Riksbank raises the interest rate. Similarly, if the nominal exchange rate weakens or the foreign interest rate increases, then the

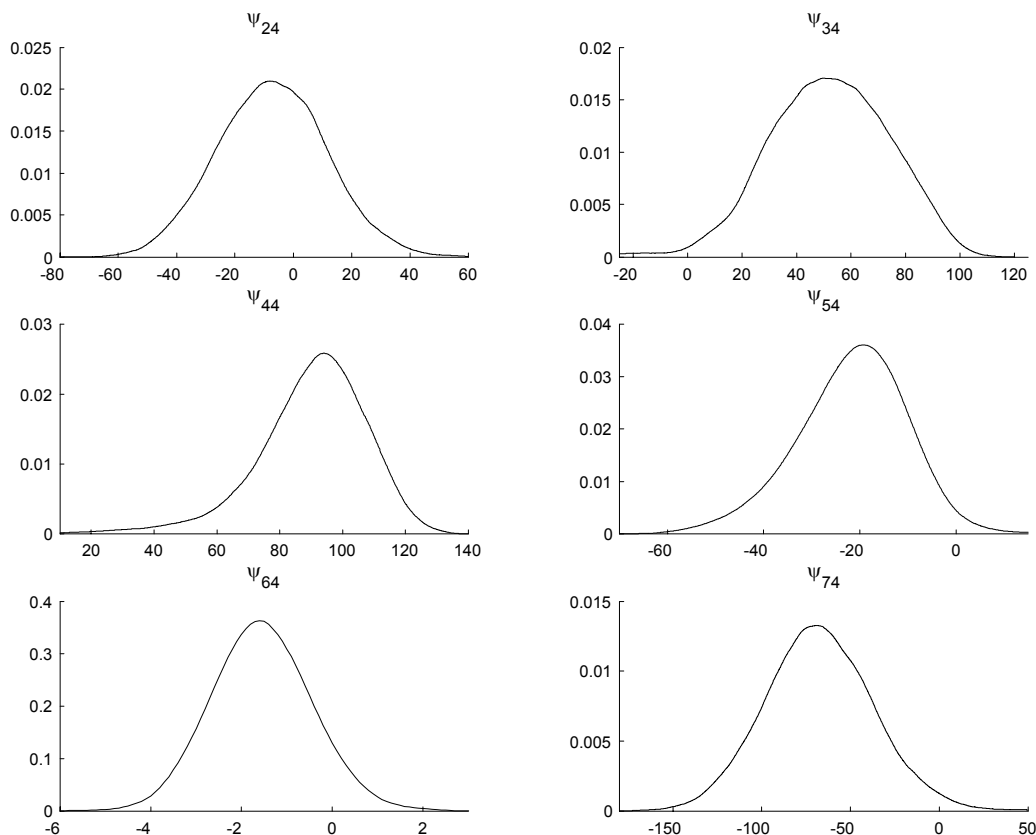


FIGURE 2. Posterior distributions of the simultaneous coefficients in the monetary policy equation:  $\psi_{24}p_t + \psi_{34}m_t + \psi_{44}i_t + \psi_{54}e_t + \psi_{64}oil_t + \psi_{74}if_t$ .

Riksbank also reacts by raising the domestic interest rate. Output changes, on the other hand, can only influence the domestic interest rate with a lag. The posterior distributions of the simultaneous coefficients, graphed in Figure 2, show that the money stock, domestic and foreign interest rate, the exchange rate and the oil price probably enter the reaction function whereas the coefficients on CPI may well be zero; a possible explanation for why the Riksbank does not react contemporaneously to CPI may be the inability of the Riksbank to see current periods CPI due to one month time lag in the publication in combination with the revisions following the preliminary figure.

To learn more about what this model suggest about the effects of monetary policy shocks we turn to Figure 3. The unconditional responses in the domestic variables along with 68 and 90 percent confidence intervals suggests that a positive monetary policy shock which raises the domestic interest rate during the first year affects the CPI after roughly 9 months when it begins to fall. Domestic output and the nominal exchange rate are affected more quickly, with a strengthening of the exchange rate and a weakening of output. There is a rather large posterior probability of a liquidity puzzle with an instantaneous increase of M0 from a contractionary monetary policy shock. This picture is reversed after a few months and after a approximately two years the money stock has fallen by roughly the same percentage as the CPI.

In the same figure we have also plotted the responses in the domestic variables from a monetary polic shock when the CPI measure has been replaced with the price variable used by the Riksbank for the inflation forecasts (UND1X). The main difference between this model and the former is

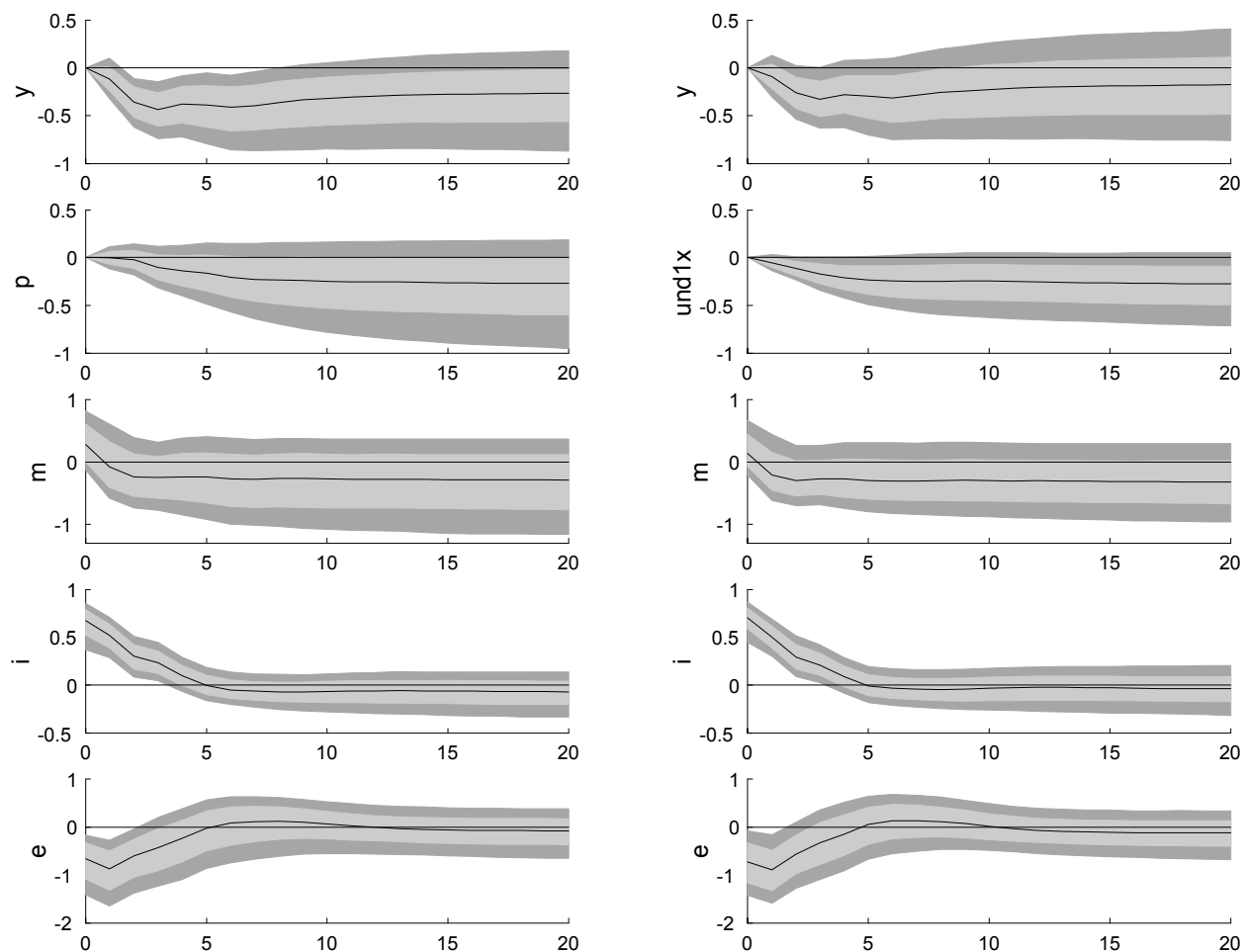


FIGURE 3. Impulse response functions over 20 quarters to a unit monetary policy shock.

that prices start to fall already after 3 months and that the probability of a liquidity puzzle is slightly smaller.

The forecast error variance decompositions for the monetary policy shock are given in Table 3. As expected, such shocks accounted for very little of the forecast uncertainty in the foreign interest rate and the oil price. Moreover, while they account for nearly half of the uncertainty in the domestic interest rate 1 quarter ahead, they seem to be important only for the nominal exchange rate. The uncertainty in the price level and output are only marginally affected by such shocks, suggesting that monetary policy surprises or mistakes are unimportant in the sense of being a sufficiently small part of the random behavior in these variables.

To further evaluate the empirical implications about monetary policy within our model, we turn to conditional forecasts. In particular, we focus on a period beginning in 1999:4 when the Riksbank first raised the repo rate from 2.90 to 3.25 on November 17, with a further increase by 50 basis points to 3.75 on February 9, 2000. Relative to the actual path of the repo rate we will consider three alternative interest rate paths generated exclusively by monetary policy shocks. That is, the interest rate path during 1999:4-2000:1 is fixed at the experiment's target by feeding a sequence of monetary policy shocks into the system. All other shocks are set to zero during these two quarters. In the subsequent quarters, the monetary policy shocks are set to zero (being

	$y$	$p$	$m$	$i$	$e$	$oil$	$i_f$
1	1.0 (0.0, 3.5)	0.4 (0.0, 1.6)	8.8 (1.4, 29.3)	45.2 (15.1, 71.6)	18.6 (1.6, 52.4)	0.1 (0.0, 0.4)	0.0 (0.0, 0.2)
4	8.0 (1.2, 18.3)	1.6 (0.1, 5.6)	9.6 (1.6, 26.5)	30.4 (9.6, 53.5)	15.6 (1.6, 45.8)	0.8 (0.0, 2.8)	0.4 (0.0, 1.5)
Horizon 8	9.3 (1.1, 23.0)	3.0 (0.1, 10.3)	9.9 (1.1, 30.9)	16.2 (4.9, 31.2)	14.2 (2.2, 39.5)	1.5 (0.0, 5.5)	0.9 (0.0, 3.5)
12	8.4 (0.8, 22.1)	3.5 (0.1, 12.4)	9.8 (0.8, 32.5)	10.8 (3.1, 20.6)	13.0 (2.3, 35.6)	1.8 (0.0, 6.7)	1.2 (0.0, 4.7)
16	7.5 (0.7, 20.8)	3.8 (0.1, 13.2)	9.6 (0.7, 33.2)	7.7 (2.2, 15.9)	11.8 (2.1, 32.1)	2.0 (0.0, 7.3)	1.4 (0.0, 5.5)

TABLE 3. Forecast error variance decomposition for the monetary policy shocks. Posterior means and 90% probability intervals (in parentheses)

controlled by the central bank) and all others shocks are drawn from their distributions. The first forecast path for the interest rate replicates its actual path in 1999:4-2000:1. The second forecast path assumes a constant interest rate, i.e., 2.90 for the two quarters. Thirdly, we study a scenario where the repo rate is lowered by 1 percent for 2 quarters. Since we do not use the repo rate in our model we impose the conditions on the closely related 3-month rate. Data up to 1999:3 are used in the estimation.

In Figure 4 we present the unconditional and the three conditional forecast paths for GDP and CPI inflation in annual growth terms, the domestic interest rate and the exchange rate. We find that the interest rate forecasts for the three conditional paths converge rather quickly and are roughly equal within a year and a half. All three conditional forecasts overshoot the realized interest rate after the first year and onwards, suggesting that the relatively low interest rate in 1999:4 (compared to the unconditional forecasts) will have to be compensated later on by an aggressive tightening of policy. The unconditional interest rate forecasts are fairly accurate over the whole forecasting horizon.

The inflation forecasts overshoot the realized path throughout the forecasting period, but are rather accurate in 2001. The tightening of policy does not exert a large downward pressure on prices. Regarding output growth, all forecasts indicates a general decline in growth, but the timing could have been better. The exchange rate forecasts suggest a weakening of the Krona, but not nearly by as much as was actually experienced in late 2000 and early 2001.

As a complement to Figure 4, we present probabilities of certain outcomes on GDP growth and CPI inflation for the three interventions and the unconditional forecast. Due to the relatively modest effects of the monetary policy shocks, there are no drastic differences between the scenarios. The probability of low GDP growth and the joint probability of *both* low growth and high inflation (3% is the upper bound of the Riksbanks inflation targeting interval) in 2000 is much larger for *Actual* scenario compared to the *Easing* scenario, however.

Counterfactual experiments of this type may, of course, be subject to the well known Lucas critique. Leeper and Zha (2002), who first introduced the concept of *modest* policy interventions, have suggested a simple metric for evaluating how unusual a conditional projection is relative to the typical size of the direct effects of policy. Their metric attempts to reflect how large the direct effects of a hypothetical intervention are and how likely the intervention is to trigger changes in agents' behavior. In Table 5 we report the Leeper-Zha modesty metric for the three policy scenarios and each domestic variable over the forecast period 1999:4–2001:4. When the metric, in absolute terms, is not greater than two, then the intervention is deemed modest and thus there is no reason for agents to change their beliefs about the policy regime and expectation formation effects may be assumed to be insignificant.

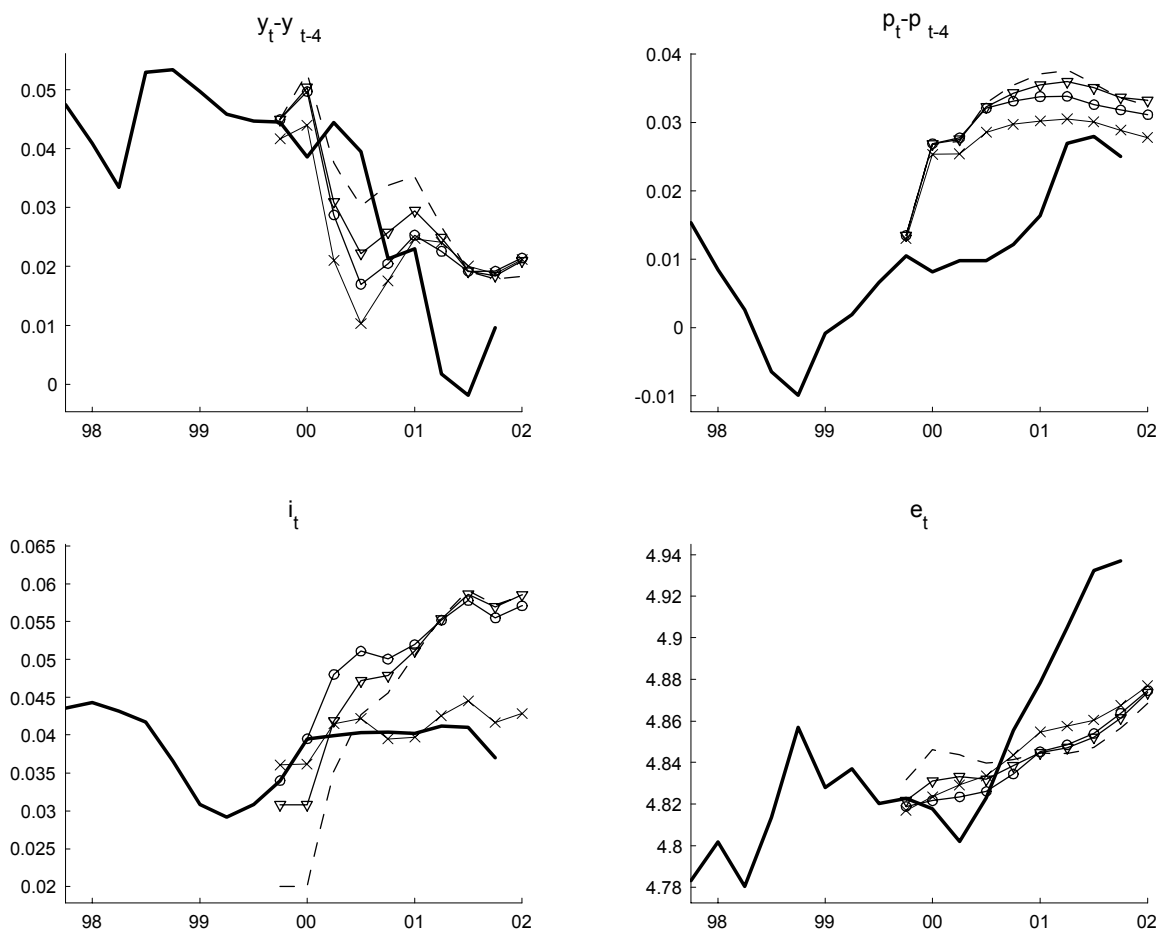


FIGURE 4. Forecasts conditional on different paths for the Swedish interest rate in the two quarters 1999:4 and 2000:1 generated exclusively by monetary policy shocks. Solid lines are the realized values. Unconditional forecast ( $- \times -$ ), Actual path:  $i_{99:4} = 0.0340$ ,  $i_{00:1} = 0.0395$  ( $-o-$ ). No change:  $i_{99:4} = i_{00:1} = 0.0308$ ,  $i_{00:1} = 0.0308$  ( $- \nabla -$ ). Easing:  $i_{99:4} = i_{00:1} = 0.02$  ( $- - -$ ).

Event	Scenario			
	Uncond.	Actual	Constant	Easing
<1% GDP growth 2000	0.391	0.292	0.196	0.098
<1% GDP growth 2001	0.388	0.347	0.352	0.366
<1% inflation 2000	0.204	0.051	0.044	0.039
<1% inflation 2001	0.266	0.164	0.137	0.141
>3% inflation 2000	0.496	0.584	0.623	0.654
>3% inflation 2001	0.483	0.531	0.567	0.565
<1% GDP growth and >3% inflation 2000	0.219	0.190	0.136	0.074
<1% GDP growth and >3% inflation 2001	0.211	0.209	0.223	0.236

TABLE 4. Posterior probability of certain events conditional on different monetary policy scenarios.



	$y$	$p$	$m$	$i$	$e$
Actual	0.6528	-0.6528	0.4956	-0.2478	0.4108
Constant	1.0393	-1.0393	0.0467	-1.0822	1.1724
Easing	2.3436	-2.3436	0.4548	-2.1165	2.3997

TABLE 5. Leeper-Zha modesty metric for the three monetary policy scenarios.

From Table 5 we find that the *Actual* and *Constant* scenarios result in modesty metrics below two for the five domestic variables. In the case of the one percent decrease, however, the metrics are above two for all variables but money. Hence, the actual path and the constant interest rate scenario may both be viewed as modest policy interventions and thus not likely to be subject to the Lucas critique. The third case, however, does not seem to be a modest intervention. Given that it implies an interest rate at 1.4 and 1.95 percent lower than the actual path over the two conditioning quarters, this does not seem unreasonable. However, the fact that the constant interest rate scenario means that the interest rate is 0.32 and 0.95 percent lower than the actual paths scenario suggests that a fairly broad range of hypothetical interest rate paths may be deemed as being consistent with modest policy interventions.

## 6. CONCLUDING REMARKS

This paper introduces a practicable Bayesian analysis of structural VARs with an arbitrary number of long run relations structured by general linear restrictions on the cointegration space. The structural part of the model is identified by restricting the contemporaneous relations among variables, following the tradition in the Bayesian structural VAR literature. There may be situations where long run identifying restrictions as in Blanchard and Quah (1989) are preferred, e.g. the restriction that monetary policy shocks have no long run impact on output. The related common trends approach introduced by King, Plosser, Stock and Watson (1991) and further developed by Warne (1993) connects the structural part of the model to its long run (cointegration) properties and provides a natural framework for the class of cointegrated processes studied here. A Bayesian analysis of models with long run restrictions is by no means straightforward and poses a real challenge for future methodological work.

## APPENDIX A. PROOF OF THEOREM 1

**A.1. The joint posterior.** The joint posterior distribution of  $\Gamma_0, \Gamma, \alpha$  and  $\beta$  conditional on the data  $D$  is

$$p(\Gamma_0, \alpha, \beta, \Gamma | D) \propto |\Gamma_0|^T \text{etr} [(X\Gamma_0 - W\beta\alpha' - Q\Gamma)'(X\Gamma_0 - W\beta\alpha' - Q\Gamma)] p(\Gamma)p(\alpha, \beta)$$

where  $\text{etr}(B) = \exp(-\frac{1}{2} \text{tr} B)$ , for any quadratic matrix  $B$ ,

$$p(\Gamma) \propto \exp[-\frac{1}{2}(\text{vec } \Gamma)' \Omega_\Gamma^{-1}(\text{vec } \Gamma)]$$

and

$$\begin{aligned} p(\alpha, \beta) &= p(\alpha|\beta)p(\beta) \\ &\propto |(\beta' K_1 \beta)^{-1} \otimes K_2|^{-1/2} |\beta' K_1 \beta|^{-p/2} \exp\left(-\frac{1}{2}(\text{vec } \alpha)'(\beta' K_1 \beta \otimes K_2^{-1})(\text{vec } \alpha)\right) \\ &\propto \text{etr}(\alpha' K_2^{-1} \alpha \beta' K_1 \beta). \end{aligned}$$

**A.2. Full conditional of  $\Gamma$ .** Follows easily from the treatment of the multivariate regression in Zellner (1971).

**A.3. Full Conditional of  $\alpha$ .** Using the same manipulations as for the multivariate regression model (Zellner, 1971) we obtain

$$p(\mathcal{D}|\Gamma_0, \Gamma, \alpha, \beta) \propto \exp\left[-\frac{1}{2}(\text{vec } \alpha - \text{vec } \hat{\alpha})'(\beta'W'W\beta \otimes I_p)(\text{vec } \alpha - \text{vec } \hat{\alpha})\right],$$

where  $\hat{\alpha} = (X\Gamma_0 - Q\Gamma)'W\beta(\beta'W'W\beta)^{-1}$ . The conditional prior of  $\alpha$  is

$$p(\alpha|\beta) \propto \exp\left[-\frac{1}{2}(\text{vec } \alpha)'(\beta'K_1\beta \otimes K_2^{-1})(\text{vec } \alpha)\right],$$

Thus,

$$\text{vec } \alpha|\Gamma_0, \beta, \Gamma, \mathcal{D} \sim N_{pr}(\bar{\mu}_\alpha, \bar{\Omega}_\alpha),$$

where  $\bar{\Omega}_\alpha^{-1} = (\beta'W'W\beta \otimes I_p) + (\beta'K_1\beta \otimes K_2^{-1})$  and

$$\bar{\mu}_\alpha = \bar{\Omega}_\alpha[(\beta'W'W\beta \otimes I_p) \text{vec } \hat{\alpha}] = \bar{\Omega}_\alpha \text{vec}(\hat{\alpha}\beta'W'W\beta) = \bar{\Omega}_\alpha \text{vec}[(X\Gamma_0 - Q\Gamma)'W\beta].$$

**A.4. Full conditional of  $\beta$ .** Let  $U = X\Gamma_0 - Q\Gamma$ . Straight-forward calculations (Zellner, 1971) give

$$\begin{aligned} p(\beta|\Gamma_0, \Gamma, \alpha, \mathcal{D}) &\propto \text{etr}[(U - W\beta\alpha')'(U - W\beta\alpha')] \text{etr}(\alpha'K_2^{-1}\alpha\beta'K_1\beta) \\ &= \exp\left\{-\frac{1}{2}[\text{vec } U - (\alpha \otimes W) \text{vec } \beta]'[\text{vec } U - (\alpha \otimes W) \text{vec } \beta]\right\} \\ &\quad \times \exp\left(-\frac{1}{2}(\text{vec } \beta)'(\alpha'K_2^{-1}\alpha \otimes K_1)(\text{vec } \beta)\right) \\ &\propto \exp\left\{-\frac{1}{2}(\text{vec } \beta - \text{vec } \tilde{\beta})'A(\text{vec } \beta - \text{vec } \tilde{\beta})\right\}, \end{aligned}$$

where  $A = \alpha'\alpha \otimes W'W + \alpha'K_2^{-1}\alpha \otimes K_1$ ,  $\text{vec } \tilde{\beta} = A^{-1}(\alpha'\alpha \otimes W'W) \text{vec } \hat{\beta}$  and

$$\text{vec } \hat{\beta} = [(\alpha'\alpha)^{-1} \otimes (W'W)^{-1}](\alpha' \otimes W') \text{vec } U = \text{vec}[(W'W)^{-1}W'(X\Gamma_0 - Q\Gamma)\alpha(\alpha'\alpha)^{-1}].$$

Inserting  $\text{vec } \hat{\beta}$  into the formula for  $\text{vec } \tilde{\beta}$  gives

$$\begin{aligned} \text{vec } \tilde{\beta} &= A^{-1}(\alpha'\alpha \otimes W'W) \text{vec}[(W'W)^{-1}W'(X\Gamma_0 - Q\Gamma)\alpha(\alpha'\alpha)^{-1}] \\ &= A^{-1} \text{vec}[W'(X\Gamma_0 - Q\Gamma)\alpha]. \end{aligned}$$

Using that  $\text{vec } \beta = h + H\phi$ , yields

$$\begin{aligned} p(\beta|\Gamma_0, \Gamma, \alpha, \mathcal{D}) &\propto \exp\left\{-\frac{1}{2}(h + H\phi - \text{vec } \tilde{\beta})'A(h + H\phi - \text{vec } \tilde{\beta})\right\} \\ &\propto \exp\left\{-\frac{1}{2}[A^{1/2}(\text{vec } \tilde{\beta} - h) - A^{1/2}H\phi]'[A^{1/2}(\text{vec } \tilde{\beta} - h) - A^{1/2}H\phi]\right\} \\ &= \exp\left\{-\frac{1}{2}(\phi - \bar{\mu}_\phi)'H'AH(\phi - \bar{\mu}_\phi)\right\}, \end{aligned}$$

where  $\bar{\Omega}_\phi^{-1} = H'AH$ ,  $\bar{\mu}_\phi = \bar{\Omega}_\phi H'A(\text{vec } \tilde{\beta} - h) = \bar{\Omega}_\phi H' \{\text{vec}[W'(X\Gamma_0 - Q\Gamma)\alpha] - Ah\}$ .

## APPENDIX B. PROOF OF THEOREM 2

B.1. **Full Conditional of  $\psi_i$ .** Let  $Z = W\beta\alpha' + Q\Gamma$ .

$$\begin{aligned} p(\psi_i|\Gamma_{0,-i}, \Gamma, \alpha, \beta, \mathcal{D}) &\propto |\Gamma_0|^T \text{etr}[(X\Gamma_0 - Z)'(X\Gamma_0 - Z)] \\ &\propto |\Gamma_0|^T \text{etr}[(\Gamma_0 - \hat{\Gamma}_0)'X'X(\Gamma_0 - \hat{\Gamma}_0)] \\ &= |\Gamma_0|^T \exp\left(-\frac{1}{2} \sum_{k=1}^p (G_k\psi_k - \hat{\gamma}_k)'X'X(G_k\psi_k - \hat{\gamma}_k)\right) \\ &\propto |\Gamma_0|^T \exp\left(-\frac{T}{2}(\psi_i - \mu_{\psi_i})'\Omega_{\psi_i}^{-1}(\psi_i - \mu_{\psi_i})\right) \end{aligned}$$

where  $\Omega_{\psi_i}^{-1} = T^{-1}G_i'Y'YG_i$ ,  $\hat{\Gamma}_0 = (\hat{\gamma}_1, \dots, \hat{\gamma}_p) = (X'X)^{-1}X'Z$  and  $\mu_{\psi_i} = (G_i'X'XG_i)^{-1}G_i'X'X\hat{\gamma}_i$ . Note that  $\hat{\gamma}_i = (X'X)^{-1}X'z_i$ , where  $z_i$  is the  $i$ th column of  $W\beta\alpha' + Q\Gamma$ , so that  $\mu_{\psi_i} = (G_i'X'XG_i)^{-1}G_i'X'z_i$ .

Let  $R_i = \text{Chol}(\Omega_{\psi_i})$  such that  $R_iR_i' = \Omega_{\psi_i}$  and, following Waggoner and Zha (2003b), decompose  $\psi_i$  as  $\psi_i = R_i \sum_{j=1}^{s_i} \beta_j w_j$ . Then

$$\begin{aligned} p(\psi_i|\Gamma_{0,-i}, \Gamma, \alpha, \beta, \mathcal{D}) &\propto |\Gamma_0|^T \exp\left(-\frac{T}{2}(\psi_i'\Omega_{\psi_i}^{-1}\psi_i - 2\mu_{\psi_i}'\Omega_{\psi_i}^{-1}\psi_i + \mu_{\psi_i}'\Omega_{\psi_i}^{-1}\mu_{\psi_i})\right) \\ &\propto |\Gamma_0|^T \exp\left(-\frac{T}{2} \left[ \left(\sum_{j=1}^{s_i} \beta_j w_j\right)' \left(\sum_{j=1}^{s_i} \beta_j w_j\right) - 2\mu_{\psi_i}'R_i'^{-1} \left(\sum_{j=1}^{s_i} \beta_j w_j\right) \right]\right) \\ &= |\Gamma_0|^T \exp\left(-\frac{T}{2} \left[ \sum_{j=1}^{s_i} (\beta_j^2 - 2\beta_j \mu_{\psi_i}'R_i'^{-1}w_j) \right]\right) \\ &\propto |\Gamma_0|^T \exp\left(-\frac{T}{2} \left[ \sum_{j=1}^{s_i} [(\beta_j - \hat{\beta}_j)^2] \right]\right), \end{aligned}$$

where  $\hat{\beta}_j = \mu_{\psi_i}'R_i'^{-1}w_j$ . As  $|\Gamma_0| = |(\gamma_1, \dots, G_iR_i \sum_{j=1}^{s_i} \beta_j w_j, \dots, \gamma_p)| \propto |\beta_1|$  (Waggoner and Zha, 2003b), we have

$$p(\psi_i|\Gamma_{0,-i}, \Gamma, \alpha, \beta, \mathcal{D}) \propto |\beta_1|^T \exp\left(-\frac{T}{2}(\beta_1 - \hat{\beta}_1)^2\right) \prod_{j=2}^{s_i} \exp\left(-\frac{T}{2}(\beta_j - \hat{\beta}_j)^2\right).$$

## REFERENCES

- [1] Akaike, H. (1974). A new look at the statistical model identification, *IEEE Trans. Automatic Control*, **19**, 716-727.
- [2] Bernanke, B. S. (1986), "Alternative Explanations of the Money-Income Correlation", *Carnegie-Rochester Conference Series on Public Policy*, 25, 49-100.
- [3] Blanchard, O. and Quah, D. (1989). The dynamic effects of aggregate demand and supply disturbances, *American Economic Review*, **79**, 655-673.
- [4] Cushman, D. O. and Zha, T. (1997). Identifying monetary policy in a small open economy under flexible exchange rates. *Journal of Monetary Economics*, **39** (3), 433-448.
- [5] Christoffersen, P. F. and Diebold, F. X. (1998). Cointegration and long-horizon forecasting, *Journal of Business and Economic Statistics*, **16**, 450-458.
- [6] Draper, D. (1995). Assessment and propagation of model uncertainty (with discussion), *Journal of the Royal Statistical Society B*, **57**, 45-97.

- [7] Engle, R. F. and Granger, C. W. J. (1987), Co-Integration and Error Correction: Representation, Estimation and Testing, *Econometrica*, 55, 251–276.
- [8] Engle, R. F. and Yoo, B. S. (1987). Forecasting and testing in cointegrated systems, *Journal of Econometrics*, **35**, 143-159.
- [9] Gilks W. R., Richardson, S. and Spiegelhalter, D. J. (1996). *Markov Chain Monte Carlo in Practice*. Chapman and Hall, London.
- [10] Gredenhoff, M. and Jacobson, T. (2001). Bootstrap testing linear restrictions on cointegrating vectors, *Journal of Business and Economic Statistics*, **19**, 63-72.
- [11] Harville, D. A. (1997). *Matrix Algebra From a Statistician's Perspective*, Springer-Verlag, New York.
- [12] Jacobson, T., Jansson, P., Vredin, A., and Warne, A. (2002), “Identifying the Effects of Monetary Policy Shocks in an Open Economy”, Sveriges Riksbank Working Paper Series No. 134.
- [13] Johansen, S. (1995). *Likelihood Based Inference on Cointegration in the Vector Autoregressive Model*. Oxford University Press, Oxford
- [14] Kim, S. and Roubini, N. (2000). Exchange rate anomalies in the industrial countries: a solution with a structural VAR approach. *Journal of Monetary Economics*, **45**, 561-586.
- [15] King, R. G., Plosser, C. I., Stock, J. H., and Watson, M. W. (1991), Stochastic Trends and Economic Fluctuations, *American Economic Review*, **81**, 819–840.
- [16] Leeper, E. M., Sims, C. A. and Zha, T. (1996). What does monetary policy do? *Brookings Papers on Economic Activity*, **2**, 1-78.
- [17] Leeper, E. M. and Zha, T. (2003). Modest policy interventions, *Journal of Monetary Economics*, **50**, 1673-1700.
- [18] Litterman, R. B. (1986). Forecasting with Bayesian vector autoregressions - five years of experience. *Journal of Business and Economic Statistics*, **4**, 25-37.
- [19] Lütkepohl, H. (1990), “Asymptotic Distributions of Impulse Response Functions and Forecast Error Variance Decompositions of Vector Autoregressive Models”, *Review of Economics and Statistics*, **72**, 116-125.
- [20] Quinn, B. G. (1980). Order determination for a multivariate autoregression, *Journal of the Royal Statistical Society*, **B42**, 182-185.
- [21] Robertson, J. C. and Tallman, E. W. (1999). Vector autoregressions: forecasting and reality, Federal Reserve Bank of Atlanta *Economic Review*, **84**, 4-18.
- [22] Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, **6**, 461-464.
- [23] Sims, C. A. (1980), Macroeconomics and Reality, *Econometrica*, **48**, 1–48.
- [24] Sims, C. A. and Zha, T. (1994). Error bands for impulse responses. Cowles Foundation Discussion Paper no. 1085.
- [25] Sims, C. A. and Zha, T. (1998). Bayesian methods for dynamic multivariate models, *International Economic Review*, **39**, 949-968.
- [26] Sims, C. A. and Zha, T. (1999). Error bands for impulse responses. *Econometrica*, **67** (5), 1113-1155.
- [27] Theil, H. and Goldberger, A. S. (1961). On pure and mixed estimation in econometrics, *International Economic Review*, **2**, 65-78.
- [28] Tierney, L. (1994). Markov chains for exploring posterior distributions (with discussion). *Annals of Statistics*, **22**, 1701-1762.
- [29] Uhlig, H. (2001). What are the effects of monetary policy on output? Results from an agnostic identification procedure, Discussion paper 28, Center for Economic Research, Tilburg University.
- [30] Villani, M. (2000). *Aspects of Bayesian Cointegration*. PhD Thesis, Stockholm University.
- [31] Villani, M. (2003). Bayesian reference analysis of cointegration, manuscript.
- [32] Waggoner, D. F. and Zha, T. (2003a). Likelihood-preserving normalization in multiple equation models, forthcoming in *Journal of Econometrics*.
- [33] Waggoner, D. F. and Zha, T. (2003b). A Gibbs sampler for restricted vector autoregressions, forthcoming in *Journal of Economic Dynamics and Control*.
- [34] Warne A. (1993), A common trends model: Identification, estimation and inference, Seminar paper no. 555, Institute for International Economic Studies, Stockholm.
- [35] Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics*, Wiley, New York.

# Earlier Working Papers:

Inflation Forecast Targeting: the Swedish Experience by <i>Claes Berg</i> .....	2000:100
Wage Effects of Mobility, Unemployment Benefits and Benefit Financing by <i>Hans Lindblad</i> .....	2000:101
A Bivariate Distribution for Inflation and Output Forecasts by <i>Mårten Blix</i> and <i>Peter Sellin</i> .....	2000:102
Optimal Horizons for Inflation Targeting by <i>Nicoletta Batini</i> and <i>Edward Nelson</i> .....	2000:103
Empirical Estimation and the Quarterly Projection Model: An Example Focusing on the External Sector by <i>Robert Amano</i> , <i>Don Coletti</i> and <i>Stephen Murchison</i> .....	2000:104
Conduction Monetary Policy with a Collegial Bord: The New Swedish Legislation One Year On by <i>Claes Berg</i> and <i>Hans Lindberg</i> .....	2000:105
Price-level targeting versus inflation targeting in a forward-looking model by <i>David Vestin</i> .....	2000:106
Unemployment and Inflation Regimes by <i>Anders Vredin</i> and <i>Anders Warne</i> .....	2000:107
An Expectations-Augmented Phillips Curve in an Open Economy by <i>Kerstin Hallsten</i> .....	2000:108
An alternative interpretation of the recent U.S. inflation performance by <i>Mikael Apel</i> and <i>Per Jansson</i> .....	2000:109
Core inflation and monetary policy by <i>Marianne Nessén</i> and <i>Ulf Söderström</i> .....	2000:110
Estimating the Implied Distribution of the Future Short-Term Interest Rate Using the Longstaff-Schwartz Model by <i>Peter Hördahl</i> .....	2000:111
Financial Variables and the Conduct of Monetary Policy by <i>Charles Goodhart</i> and <i>Boris Hofmann</i> .....	2000:112
Testing for the Lucas Critique: A Quantitative Investigation by <i>Jesper Lindé</i> .....	2000:113
Monetary Policy Analysis in Backward-Looking Models by <i>Jesper Lindé</i> .....	2000:114
UIP for short investments in long-term bonds by <i>Annika Alexius</i> .....	2000:115
Qualitative Survey Responses and Production over the Business Cycle by <i>Tomas Lindström</i> .....	2000:116
Supply stocks and real exchange rates by <i>Annika Alexius</i> .....	2000:117
Casuality and Regime Inference in a Markov Switching VAR by <i>Anders Warne</i> .....	2000:118
Average Inflation Targeting by <i>Marianne Nessén</i> and <i>David Vestin</i> .....	2000:119
Forecast-based monetary policy in Sweden 1992-1998: A view from within by <i>Per Jansson</i> and <i>Anders Vredin</i> .....	2000:120
What have we learned from empirical tests of the monetary transmission effect? by <i>Stefan Norrbin</i> .....	2000:121
Simple monetary policy rules and exchange rate uncertainty by <i>Kai Leitemo</i> and <i>Ulf Söderström</i> .....	2001:122
Targeting inflation with a prominent role for money by <i>Ulf Söderström</i> .....	2001:123
Is the Short-run Phillips Curve Nonlinear? Empirical Evidence for Australia, Sweden and the United States by <i>Ann-Charlotte Eliasson</i> .....	2001:124
An Alternative Explanation of the Price Puzzle by <i>Paolo Giordani</i> .....	2001:125
Interoperability and Network Externalities in Electronic Payments by <i>Gabriela Guibourg</i> .....	2001:126
Monetary Policy with Incomplete Exchange Rate Pass-Through by <i>Malin Adolfson</i> .....	2001:127
Micro Foundations of Macroeconomic Price Adjustment: Survey Evidence from Swedish Firms by <i>Mikael Apel</i> , <i>Richard Friberg</i> and <i>Kerstin Hallsten</i> .....	2001:128
Estimating New-Keynesian Phillips Curves on Data with Measurement Errors: A Full Information Maximum Likelihood Approach by <i>Jesper Lindé</i> .....	2001:129
The Empirical Relevance of Simple Forward- and Backward-looking Models: A View from a Dynamic General Equilibrium Model by <i>Jesper Lindé</i> .....	2001:130
Diversification and Delegation in Firms by <i>Vittoria Cerasi</i> and <i>Sonja Daltung</i> .....	2001:131
Monetary Policy Signaling and Movements in the Swedish Term Structure of Interest Rates by <i>Malin Andersson</i> , <i>Hans Dillén</i> and <i>Peter Sellin</i> .....	2001:132
Evaluation of exchange rate forecasts for the krona's nominal effective exchange rate by <i>Henrik Degré</i> , <i>Jan Hansen</i> and <i>Peter Sellin</i> .....	2001:133
Identifying the Effects of Monetary Policy Shocks in an Open Economy by <i>Tor Jacobsson</i> , <i>Per Jansson</i> , <i>Anders Vredin</i> and <i>Anders Warne</i> .....	2002:134

Implications of Exchange Rate Objectives under Incomplete Exchange Rate Pass-Through by <i>Malin Adolfson</i> .....	2002:135
Incomplete Exchange Pass-Through and Simple Monetary Policy Rules by <i>Malin Adolfson</i> .....	2002:136
Financial Instability and Monetary Policy: The Swedish Evidence by <i>U. Michael Bergman</i> and <i>Jan Hansen</i> .....	2002:137
Finding Good Predictors for Inflation: A Bayesian Model Averaging Approach by <i>Tor Jacobson</i> and <i>Sune Karlsson</i> .....	2002:138
How Important Is Precommitment for Monetary Policy? by <i>Richard Dennis</i> and <i>Ulf Söderström</i> .....	2002:139
Can a Calibrated New-Keynesian Model of Monetary Policy Fit the Facts? by <i>Ulf Söderström</i> , <i>Paul Söderlind</i> and <i>Anders Vredin</i> .....	2002:140
Inflation Targeting and the Dynamics of the Transmission Mechanism by <i>Hans Dillén</i> .....	2002:141
Capital Charges under Basel II: Corporate Credit Risk Modelling and the Macro Economy by <i>Kenneth Carling</i> , <i>Tor Jacobson</i> , <i>Jesper Lindé</i> and <i>Kasper Roszbach</i> .....	2002:142
Capital Adjustment Patterns in Swedish Manufacturing Firms: What Model Do They Suggest? by <i>Mikael Carlsson</i> and <i>Stefan Laséen</i> .....	2002:143
Bank Lending, Geographical Distance, and Credit risk: An Empirical Assessment of the Church Tower Principle by <i>Kenneth Carling</i> and <i>Sofia Lundberg</i> .....	2002:144
Inflation, Exchange Rates and PPP in a Multivariate Panel Cointegration Model by <i>Tor Jacobson</i> , <i>Johan Lyhagen</i> , <i>Rolf Larsson</i> and <i>Marianne Nessén</i> .....	2002:145
Evaluating Implied RNDs by some New Confidence Interval Estimation Techniques by <i>Magnus Andersson</i> and <i>Magnus Lomakka</i> .....	2003:146
Taylor Rules and the Predictability of Interest Rates by <i>Paul Söderlind</i> , <i>Ulf Söderström</i> and <i>Anders Vredin</i> .....	2003:147
Inflation, Markups and Monetary Policy by <i>Magnus Jonsson</i> and <i>Stefan Palmqvist</i> .....	2003:148
Financial Cycles and Bankruptcies in the Nordic Countries by <i>Jan Hansen</i> .....	2003:149
Bayes Estimators of the Cointegration Space by <i>Mattias Villani</i> .....	2003:150
Business Survey Data: Do They Help in Forecasting the Macro Economy? by <i>Jesper Hansson</i> , <i>Per Jansson</i> and <i>Mårten Löf</i> .....	2003:151
The Equilibrium Rate of Unemployment and the Real Exchange Rate: An Unobserved Components System Approach by <i>Hans Lindblad</i> and <i>Peter Sellin</i> .....	2003:152
Monetary Policy Shocks and Business Cycle Fluctuations in a Small Open Economy: Sweden 1986-2002 by <i>Jesper Lindé</i> .....	2003:153
Bank Lending Policy, Credit Scoring and the Survival of Loans by <i>Kasper Roszbach</i> .....	2003:154
Internal Ratings Systems, Implied Credit Risk and the Consistency of Banks' Risk Classification Policies by <i>Tor Jacobson</i> , <i>Jesper Lindé</i> and <i>Kasper Roszbach</i> .....	2003:155



Sveriges Riksbank

Visiting address: Brunkebergs torg 11

Mail address: se-103 37 Stockholm

Website: [www.riksbank.se](http://www.riksbank.se)

Telephone: +46 8 787 00 00, Fax: +46 8 21 05 31

E-mail: [registratorn@riksbank.se](mailto:registratorn@riksbank.se)