# Generalized Two-Step Maximum Likelihood Estimation of Structural Vector Autoregressive Models partially identified with Short-Run Restrictions

Kyungho Jang*
University of Alabama at Birmingham

March 23, 2004

**Abstract**

This paper presents a generalized two-step maximum likelihood estimation method for partially identified vector autoregressive models. We suggest a likelihood ratio test for over-identification in a sub-system and derive the asymptotics for impulse responses and forecast-error variance decomposition for partially identified models. As an application, we consider an open economy model to investigate the effects of monetary policy on exchange rates and term structures. We find that exchange rates tend to overshoot and term structures have hump-shaped responses to monetary policy shocks.

*Keywords:* ML estimation, VAR model, Identification, Likelihood ratio test, Asymptotic distribution, Impulse response, Forecast-error variance decomposition, Monetary policy, Exchange rate
*JEL code:* C32, C51, E32, E52, F31

---

* Department of Finance, Economics and Quantitative Methods, University of Alabama at Birmingham, Birmingham, AL 35294. Tel. (205) 934-8833, Fax: (205) 975-4427, E-mail: kjang@uab.edu.
I thank Masao Ogaki for the guidance and S.D. Lee for comments.

# 1 Introduction

This paper develops a generalized two-step maximum likelihood (ML) estimation method and derives the asymptotics for impulse responses and forecast-error variance decomposition in partially identified vector autoregressive models. Various empirical studies have used impulse response analysis within the framework of vector autoregressive models since Sims (1980). For such analyses, Blanchard and Watson (1986), Bernanke (1986), and Blanchard (1989) imposed contemporaneous short-run restrictions, while Blanchard and Quah (1989) used long-run restrictions for identification. From the first-step ordinary least squares estimates, they estimated the structural parameters using Cholesky decomposition, generalized method of moments (GMM), or ML estimation in the second step. Gali (1992) is an exception, as an instrumental variables method was adopted to estimate IS-LM models with short- and long-run restrictions.

When partially identified models are considered, Cholesky decomposition can be used to estimate just-identified block recursive models, while the GMM can be used for over-identified models (see Eichenbaum and Evans, 1995; Bernanke and Mihov, 1998, among others). By contrast, two-step ML estimation is limited to fully identified models, since the second step of ML estimation involves the technically difficult process of constructing the likelihood function. We resolve this difficulty, transforming the reduced-form model in the second step. We provide a new ML estimation method for partially identified models and derive its asymptotic properties. We also suggest a likelihood ratio test for over-identifying restrictions in partially identified models.

Our results can be applied to many empirical studies of impulse responses to a subset of structural shocks. For illustration, we provide a two-step ML estimation of the block recursive VAR model that was used by Bernanke and Mihov (1998). In addition, we extend their model to an open economy and investigate the effects of monetary policy on exchange rates and term structures. We find that exchange rates tend to overshoot and term structures have hump-shaped responses to monetary policy shocks.

This paper is organized as follows. Section 2 develops ML estimation in VAR models that are partially identified with short-run restrictions. Section 3 derives the asymptotic properties of ML estimators, and Section 4 discusses the asymptotics of impulse responses and forecast-error variance. Section 5 extends Bernanke and Mihov (1998) to an open economy, and Section 6 contains our conclusions. Definitions and properties of the matrices and operators used in the text are summarized in Appendix A, while the proof of the lemma is provided in Appendix B. The data used for the application are described in Appendix C.

## 2 Generalized ML Estimation in VAR Models partially identified with Short-Run Restrictions

Suppose that an economy is described by an n-dimensional structural vector autoregressive (VAR) model[1]:

$$\mathbf{B}(L)\mathbf{y}_t = \mathbf{F}\mathbf{e}_t, \tag{2.1}$$

where $\mathbf{B}(L) = \mathbf{B}_0 - \sum_{i=1}^{p} \mathbf{B}_i L^i$, $\mathbf{e}_t$ is a vector of structural shocks with a mean of zero and variance $\mathbf{I}_n$, $L$ is the lag operator, and $\mathbf{I}_n$ is the $n$-dimensional identity matrix. We estimate the model using the ordinary least squares method in the first step with

[1] For simplicity, we assume that $\mathbf{y}_t$ is demeaned. This does not change our main results.

the corresponding reduced-form model:

$$\mathbf{A}(L)\mathbf{y}_t = \boldsymbol{\epsilon}_t, \tag{2.2}$$

where $\mathbf{A}(L) = \mathbf{I}_n - \sum_{i=1}^{p} \mathbf{A}_i L^i$, and $\boldsymbol{\epsilon}_t$ is *white noise* with a mean of zero and variance $\boldsymbol{\Sigma}$. From the first-step estimates, we estimate the structural parameters, $\mathbf{B}_0$ and $\mathbf{F}$ using ML estimation in the second step, and construct the remaining structural parameters and shocks using $\mathbf{B}_i = \mathbf{B}_0\mathbf{A}_i$ for $i = 1, \cdots, p$ and $\mathbf{e}_t = \mathbf{F}^{-1}\mathbf{B}_0\boldsymbol{\epsilon}_t$. Refer to Giannini (1992) for details of the ML estimation and its asymptotics for VAR models that are fully identified with short-run restrictions.

The VAR model in (2.1) is not econometrically identified in general. The block recursive assumption starts by partitioning $\mathbf{y}_t$ into three blocks as

$$\mathbf{y}_t = \begin{bmatrix} \mathbf{y}_{1t} \\ \mathbf{y}_{2t} \\ \mathbf{y}_{3t} \end{bmatrix}.$$

For example, $\mathbf{y}_{2t}$ is a set of $n_2$-dimensional policy indicators, and $\mathbf{y}_{1t}$ includes $n_1$ variables that are included in the information set when the Federal Reserve Bank (or Fed) implements a monetary policy, while $\mathbf{y}_{3t}$ contains $n_3$ variables that are excluded from the information set ($n = n_1 + n_2 + n_3$). Alternatively, $\mathbf{y}_{1t}$ does not respond to a monetary policy shock contemporaneously, while $\mathbf{y}_{3t}$ does. See Christiano, Eichenbaum, and Evans (1999) and Keating (1999) for an extended theoretical background. Throughout this paper, we assume that the second block of parameters and shocks is of interest given the following block recursive assumption:[2]

[2] As Christiano, Eichenbaum, and Evans (1999) noted, this block recursive system includes general classes of VAR models. For example, we may consider fully identified models with the choice of $n_1 = 0$ and $n_3 = 0$.

**Assumption 2.1.** $\mathbf{B}_0$ *is block lower triangular and* $\mathbf{F}$ *is block diagonal:*

$$\mathbf{B}_0 = \begin{bmatrix} \underset{(n_1\times n_1)}{\mathbf{B}_{11}} & \underset{(n_1\times n_2)}{\mathbf{0}} & \underset{(n_1\times n_3)}{\mathbf{0}} \\ \underset{(n_2\times n_1)}{\mathbf{B}_{21}} & \underset{(n_2\times n_2)}{\mathbf{B}_{22}} & \underset{(n_2\times n_3)}{\mathbf{0}} \\ \underset{(n_3\times n_1)}{\mathbf{B}_{31}} & \underset{(n_3\times n_2)}{\mathbf{B}_{32}} & \underset{(n_3\times n_3)}{\mathbf{B}_{33}} \end{bmatrix} \quad \textit{and} \quad \mathbf{F} = \begin{bmatrix} \underset{(n_1\times n_1)}{\mathbf{F}_{11}} & \underset{(n_1\times n_2)}{\mathbf{0}} & \underset{(n_1\times n_3)}{\mathbf{0}} \\ \underset{(n_2\times n_1)}{\mathbf{0}} & \underset{(n_2\times n_2)}{\mathbf{F}_{22}} & \underset{(n_2\times n_3)}{\mathbf{0}} \\ \underset{(n_3\times n_1)}{\mathbf{0}} & \underset{(n_3\times n_2)}{\mathbf{0}} & \underset{(n_3\times n_3)}{\mathbf{F}_{33}} \end{bmatrix}.$$

Denote the second block of $\mathbf{B}_0$ using $\bar{\mathbf{B}}_0 \equiv [\mathbf{B}_{21} \vdots \mathbf{B}_{22} \vdots \mathbf{0}]$. If $\mathbf{F}_{22}$ is the identity matrix and $\mathbf{B}_{22}$ is just-identified and lower triangular in addition to Assumption 2.1, then $\bar{\mathbf{B}}_0$ is identifiable and estimable using Cholesky decomposition. Otherwise, we may use ML estimation concentrating on the $n_2$-dimensional second block of the model:

$$\bar{\mathbf{B}}_0\mathbf{A}(L)\mathbf{y}_t = \mathbf{F}_{22}\mathbf{e}_{2t}.$$

It is impossible to estimate $\bar{\mathbf{B}}_0$ directly because its information matrix is singular. The lemma below suggests that the sub-model requires a certain transformation (or diagonalization) for ML estimation. In particular, we multiply a transformation matrix $\mathbf{M}$ by the reduced-form VAR model (2.2), which makes each block of transformed innovations mutually orthogonal.

**Lemma 2.1.** *Under Assumption 2.1, a block recursive VAR model with short-run restrictions is partially identifiable and two-step ML estimable by choosing a block lower triangular transformation matrix* $\mathbf{M}$ *such that*

$$\mathbf{B}_0 = \mathbf{B}_d\mathbf{M},$$

*where*

$$\mathbf{B}_d = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{22} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{B}_{33} \end{bmatrix} \text{ and } \mathbf{M} = \begin{bmatrix} \mathbf{I}_{n_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{M}_{21} & \mathbf{I}_{n_2} & \mathbf{0} \\ \mathbf{M}_{31} & \mathbf{M}_{32} & \mathbf{I}_{n_3} \end{bmatrix},$$

*where* $\mathbf{M}_{21} = -\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}$, $\mathbf{M}_{31} = -\boldsymbol{\Sigma}_{31}\boldsymbol{\Sigma}_{11}^{-1} + \mathbf{M}_{32}\mathbf{M}_{21}$, *and* $\mathbf{M}_{32} = -(\boldsymbol{\Sigma}_{32} - \boldsymbol{\Sigma}_{31}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12})(\boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12})^{-1}$.

*Proof.* See Appendix B. ∎

*Remark* 2.1. Lemma 2.1 is also valid under the alternative assumption that $\mathbf{B}_0$ is block diagonal and $\mathbf{F}$ is lower triangular. Consider the lower triangular matrix $\mathbf{Q}$ such that $\mathbf{F} = \mathbf{Q}\mathbf{F}_d$, where $\mathbf{F}_d$ is the block diagonal matrix of $\mathbf{F}$. From $\mathbf{B}_0\boldsymbol{\epsilon}_t = \mathbf{F}\mathbf{e}_t$ it follows that $\mathbf{B}_0^*\boldsymbol{\epsilon}_t = \mathbf{F}_d\mathbf{e}_t$, where $\mathbf{B}_0^*$ is a lower triangular matrix such that $\mathbf{B}_0^* = \mathbf{Q}^{-1}\mathbf{B}_0$. Note that the transformed matrices $\mathbf{B}_0^*$ and $\mathbf{F}_d$ satisfy Assumption 2.1. It is straightforward to show $\mathbf{B}_0^* = \mathbf{B}_0\mathbf{M}$ with the choice of $\mathbf{Q} = \mathbf{B}_0\mathbf{M}^{-1}\mathbf{B}_0^{-1}$. In particular, $\mathbf{Q} = \mathbf{M}^{-1}$ when $\mathbf{B}_0 = \mathbf{I}_n$. Therefore, we can use the same transformation matrix $\mathbf{M}$ for the ML estimation under the alternative assumptions.

Once the structural VAR model is diagonalized after transformation, we can concentrate on the ML estimation of $\mathbf{B}_{22}$ and $\mathbf{F}_{22}$ in the sub-system

$$\mathbf{B}_{22}\bar{\mathbf{M}}\mathbf{A}(L)\mathbf{y}_t = \mathbf{F}_{22}\mathbf{e}_{2t}, \tag{2.3}$$

where $\bar{\mathbf{M}} = [-\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} \vdots \mathbf{I}_{n_2} \vdots \mathbf{0}]$, which is the second block of $\mathbf{M}$. In particular, the transformation matrix $\bar{\mathbf{M}}$ makes the second block innovations, $\bar{\boldsymbol{\epsilon}}_{2t} = \bar{\mathbf{M}}\boldsymbol{\epsilon}_t$, mutually orthogonal to the first block innovations, $\boldsymbol{\epsilon}_{1t}$. In what follows, we assume that the second block of the model is partially identified by the following short-run restrictions:

**Assumption 2.2.** $\text{vec}(\mathbf{B}_{22}) = \mathbf{S}_b\mathbf{b}_s + \mathbf{s}_b$ *and* $\text{vec}(\mathbf{F}_{22}) = \mathbf{S}_f\mathbf{f}_s + \mathbf{s}_f$.

For computational purposes, we define $\mathbf{K}_{22} = \mathbf{F}_{22}^{-1}\mathbf{B}_{22}$, $\boldsymbol{\Phi}_{22} = \mathbf{B}_{22}^{-1}\mathbf{F}_{22}$, $\boldsymbol{\Lambda}_{22} = [\mathbf{B}_{22} \vdots \mathbf{F}_{22}]$, $\boldsymbol{\lambda}_s = (\mathbf{b}_s', \mathbf{f}_s')'$, $\mathbf{s}_\lambda = (\mathbf{s}_b', \mathbf{s}_f')'$, and denote the vectorization of the corresponding matrix with a lower case letter and the corresponding estimator with a caret. For example, $\mathbf{b}_{22} = \text{vec}(\mathbf{B}_{22})$ and $\hat{\hat{\mathbf{M}}} = [-\hat{\mathbf{\Sigma}}_{21}\hat{\mathbf{\Sigma}}_{11}^{-1} \vdots \mathbf{I}_{n_2} \vdots \mathbf{0}]$, where vec denotes the column stacking operator. We write the short-run restrictions in a compact form:

$$\boldsymbol{\lambda}_{22} = \mathbf{S}_{\lambda}\boldsymbol{\lambda}_s + \mathbf{s}_{\lambda}, \quad \text{where} \quad \mathbf{S}_{\lambda} = \begin{bmatrix} \mathbf{S}_b & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_f \end{bmatrix},$$

and designate the sample size, $T$, commutation matrix, $\mathbf{K}_{mn}$, duplication matrix, $\mathbf{D}_n$, elimination matrix, $\mathbf{L}_n$, and matrix $\mathbf{N}_n = \frac{1}{2}(\mathbf{I}_{n^2} + \mathbf{K}_{nn})$ as defined by Magnus (1988). See Appendix A for the definitions and properties of these matrices.

Finally, we assume that the second set of structural shocks, $\mathbf{e}_{2t}$, follows the identical independent multivariate standard normal distribution:

**Assumption 2.3.** $\mathbf{e}_{2t} \sim IIN(\mathbf{0}, \mathbf{I}_{n_2})$

The generalized two-step ML estimation (GMLE) of partially identified models is proposed in the following theorem:

**Theorem 2.1.** *(The GMLE of partially identified models with short-run restrictions) Under Assumptions 2.1–2.3, a VAR model partially identified with short-run restrictions can be estimated by generalized maximum likelihood estimation using the*

*(a) Likelihood function:*

$$L(\mathbf{B}_{22}, \mathbf{F}_{22}) = T \log|\mathbf{B}_{22}| - T \log|\mathbf{F}_{22}| - \frac{T}{2}\text{trace}(\mathbf{B}_{22}'\mathbf{F}_{22}'^{-1}\mathbf{F}_{22}^{-1}\mathbf{B}_{22}\hat{\hat{\mathbf{M}}}\hat{\mathbf{\Sigma}}\hat{\hat{\mathbf{M}}}') \qquad (2.4)$$

*(b) Gradient:*

$$\mathbf{g}(\mathbf{B}_{22}, \mathbf{F}_{22}) = T \begin{bmatrix} \mathbf{I}_{n_2} \otimes \mathbf{F}_{22}'^{-1} \\ -\mathbf{F}_{22}^{-1}\mathbf{B}_{22} \otimes \mathbf{F}_{22}'^{-1} \end{bmatrix} \left[ (\mathbf{I}_{n_2} \otimes \mathbf{F}_{22}')\text{vec}(\mathbf{B}_{22}'^{-1}) - (\hat{\hat{\mathbf{M}}}\hat{\mathbf{\Sigma}}\hat{\hat{\mathbf{M}}}' \otimes \mathbf{F}_{22}^{-1})\text{vec}(\mathbf{B}_{22}) \right],$$

*where $\otimes$ is the Kronecker product operator*

*(c) Information matrix:*

$$\mathbf{I}_T(\mathbf{B}_{22}, \mathbf{F}_{22}) = 2T \begin{bmatrix} \mathbf{B}_{22}^{-1}\mathbf{F}_{22} \otimes \mathbf{F}_{22}'^{-1} \\ -\mathbf{I}_{n_2} \otimes \mathbf{F}_{22}'^{-1} \end{bmatrix} \mathbf{N}_{n_2} \left[ \mathbf{F}_{22}'\mathbf{B}_{22}'^{-1} \otimes \mathbf{F}_{22}^{-1} \vdots -\mathbf{I}_{n_2} \otimes \mathbf{F}_{22}^{-1} \right]$$

*(d) Score algorithm:*

$$\boldsymbol{\lambda}_{s,i+1} = \boldsymbol{\lambda}_{s,i} + [\mathbf{I}_T(\boldsymbol{\lambda}_{s,i})]^{-1}\mathbf{g}(\boldsymbol{\lambda}_{s,i}),$$

*where* $\mathbf{g}(\boldsymbol{\lambda}_s) = \mathbf{S}'_\lambda \mathbf{g}(\mathbf{B}_{22}, \mathbf{F}_{22})$, $\mathbf{I}_T(\boldsymbol{\lambda}_s) = \mathbf{S}'_\lambda \mathbf{I}_T(\mathbf{B}_{22}, \mathbf{F}_{22})\mathbf{S}_\lambda$, *and* $i$ *denotes the iteration step, and*

*(e) if* $\mathbf{B}_{22}$ *or* $\mathbf{F}_{22}$ *is over-identified, the over-identifying restrictions are testable using a likelihood ratio test:*

$$LRT = 2(L_u - L_r),$$

*where* $L_u = -\frac{T}{2}\log|\hat{\bar{\mathbf{M}}}\hat{\boldsymbol{\Sigma}}\hat{\bar{\mathbf{M}}}'| - \frac{n_2 T}{2}$, $L_r = L(\hat{\mathbf{B}}_{22}, \hat{\mathbf{F}}_{22})$, *and* $LRT$ *is asymptotically* $\chi^2_{(q)}$*-distributed, where* $q$ *is the number of over-identifying restrictions.*

*Proof.* (a) Provided that the model is partially identifiable, we can concentrate on the second block of the structural model

$$\bar{\mathbf{B}}_0\mathbf{A}(L)\mathbf{y}_t = \mathbf{F}_{22}\mathbf{e}_{2t},$$

or

$$\bar{\mathbf{K}}_0\mathbf{A}(L)\mathbf{y}_t = \mathbf{e}_{2t},$$

where $\bar{\mathbf{K}}_0 = \mathbf{F}_{22}^{-1}\bar{\mathbf{B}}_{22}\bar{\mathbf{M}}$. As $\mathbf{e}_2$ is multivariate standard normally distributed, its p.d.f. follows

$$f(\mathbf{e}_2) = (2\pi)^{-\frac{1}{2}n_2 T}\exp(-\frac{1}{2}\sum_{t=1}^{T}\mathbf{e}'_{2t}\mathbf{e}_{2t}).$$

Therefore, from $\mathbf{e}_{2t} = \bar{\mathbf{K}}_0\boldsymbol{\epsilon}_t$ we get the p.d.f. of $\mathbf{y}$

$$f_y(\mathbf{y}) = (2\pi)^{-\frac{1}{2}n_2 T}\exp(-\frac{1}{2}\sum_{t=1}^{T}\boldsymbol{\epsilon}'_t\bar{\mathbf{K}}'_0\bar{\mathbf{K}}_0\boldsymbol{\epsilon}_t)|\bar{\mathbf{K}}_0\bar{\mathbf{K}}'_0|^{\frac{T}{2}}$$

and the log-likelihood function

$$L(\bar{\mathbf{K}}_0, \mathbf{A}(L)) = constant + \frac{T}{2}\log|\bar{\mathbf{K}}_0\bar{\mathbf{K}}_0'| - \frac{1}{2}\sum_{t=1}^{T}\boldsymbol{\epsilon}_t'\bar{\mathbf{K}}_0'\bar{\mathbf{K}}_0\boldsymbol{\epsilon}_t.$$

Consequently, the second-step log-likelihood function becomes

$$\begin{aligned} L(\mathbf{B}_{22}, \mathbf{F}_{22}, \hat{\mathbf{A}}(L)) &= constant + \frac{T}{2}\log|\bar{\mathbf{K}}_0\bar{\mathbf{K}}_0'| - \frac{1}{2}\sum_{t=1}^{T}\hat{\boldsymbol{\epsilon}}_t'\bar{\mathbf{K}}_0'\bar{\mathbf{K}}_0\hat{\boldsymbol{\epsilon}}_t \\ &= constant + T\log|\mathbf{K}_{22}| - \frac{T}{2}\text{trace}(\mathbf{K}_{22}'\mathbf{K}_{22}\hat{\hat{\mathbf{M}}}\hat{\boldsymbol{\Sigma}}\hat{\hat{\mathbf{M}}}') \\ &= constant + T\log|\mathbf{B}_{22}| - T\log|\mathbf{F}_{22}| - \frac{T}{2}\text{trace}(\mathbf{B}_{22}'\mathbf{F}_{22}'^{-1}\mathbf{F}_{22}^{-1}\mathbf{B}_{22}\hat{\hat{\mathbf{M}}}\hat{\boldsymbol{\Sigma}}\hat{\hat{\mathbf{M}}}'). \end{aligned}$$

(b) Taking the derivative of (2.4) with respect to $\boldsymbol{\Lambda}_{22}$ yields

$$\frac{\partial L}{\partial \text{vec}(\boldsymbol{\Lambda}_{22})'} = \frac{\partial L}{\partial \text{vec}(\mathbf{K}_{22})'}\frac{\partial \text{vec}(\mathbf{K}_{22})}{\partial \text{vec}(\boldsymbol{\Lambda}_{22})'},$$

where

$$\begin{aligned} \frac{\partial L}{\partial \text{vec}(\mathbf{K}_{22})'} &= T\left[\text{vec}(\mathbf{K}_{22}'^{-1}) - \text{vec}(\mathbf{K}_{22}\hat{\hat{\mathbf{M}}}\hat{\boldsymbol{\Sigma}}\hat{\hat{\mathbf{M}}}')\right]' \\ &= T\left[\text{vec}(\mathbf{K}_{22}'^{-1}) - (\hat{\hat{\mathbf{M}}}\hat{\boldsymbol{\Sigma}}\hat{\hat{\mathbf{M}}}' \otimes \mathbf{I}_{n_2})\text{vec}(\mathbf{K}_{22})\right]' \end{aligned} \quad (2.5)$$

and

$$\begin{aligned} \frac{\partial \text{vec}(\mathbf{K}_{22})}{\partial \text{vec}(\boldsymbol{\Lambda}_{22})'} &= \left[\frac{\partial \text{vec}(\mathbf{K}_{22})}{\partial \text{vec}(\mathbf{B}_{22})'} \vdots \frac{\partial \text{vec}(\mathbf{K}_{22})}{\partial \text{vec}(\mathbf{F}_{22})'}\right] \\ &= \left[\mathbf{I}_{n_2} \otimes \mathbf{F}_{22}^{-1} \vdots -\mathbf{B}_{22}'\mathbf{F}_{22}'^{-1} \otimes \mathbf{F}_{22}^{-1}\right]. \end{aligned}$$

Therefore, the gradient becomes

$$\begin{aligned} \mathbf{g}(\mathbf{B}_{22}, \mathbf{F}_{22}) &= \frac{\partial L}{\partial \text{vec}(\boldsymbol{\Lambda}_{22})} \\ &= T\begin{bmatrix} \mathbf{I}_{n_2} \otimes \mathbf{F}_{22}'^{-1} \\ -\mathbf{F}_{22}^{-1}\mathbf{B}_{22} \otimes \mathbf{F}_{22}'^{-1} \end{bmatrix}[(\mathbf{I}_{n_2} \otimes \mathbf{F}_{22}')\text{vec}(\mathbf{B}_{22}'^{-1}) - (\hat{\hat{\mathbf{M}}}\hat{\boldsymbol{\Sigma}}\hat{\hat{\mathbf{M}}}' \otimes \mathbf{F}_{22}^{-1})\text{vec}(\mathbf{B}_{22})]. \end{aligned}$$

(c) From the second derivative of (2.5) with respect to $\mathbf{K}_{22}$

$$\begin{aligned} \mathbf{H}_T(\mathbf{K}_{22}) = \frac{\partial L^2}{\partial \text{vec}(\mathbf{K}_{22})\partial \text{vec}(\mathbf{K}_{22})'} &= T\left[\frac{\partial \text{vec}(\mathbf{K}_{22}'^{-1})}{\partial \text{vec}(\mathbf{K}_{22})'}\right]' - T(\hat{\hat{\mathbf{M}}}\hat{\boldsymbol{\Sigma}}\hat{\hat{\mathbf{M}}}' \otimes \mathbf{I}_{n_2}) \\ &= -T\left[(\mathbf{K}_{22}^{-1} \otimes \mathbf{K}_{22}'^{-1})\mathbf{K}_{n_2n_2} + (\hat{\hat{\mathbf{M}}}\hat{\boldsymbol{\Sigma}}\hat{\hat{\mathbf{M}}}' \otimes \mathbf{I}_{n_2})\right] \end{aligned}$$

we get the information matrix

$$\begin{aligned}
\mathbf{I}_T(\mathbf{K}_{22}) = -E(\mathbf{H}_T(\mathbf{K}_{22})) &= T\left[(\mathbf{K}_{22}^{-1} \otimes \mathbf{K}_{22}'^{-1})\mathbf{K}_{n_2 n_2} + E(\hat{\hat{\mathbf{M}}}\hat{\mathbf{\Sigma}}\hat{\hat{\mathbf{M}}}') \otimes \mathbf{I}_{n_2}\right] \\
&= T\left[(\mathbf{K}_{22}^{-1} \otimes \mathbf{K}_{22}'^{-1})\mathbf{K}_{n_2 n_2} + \mathbf{K}_{22}^{-1}\mathbf{K}_{22}'^{-1} \otimes \mathbf{I}_{n_2}\right] \\
&= 2T(\mathbf{K}_{22}^{-1} \otimes \mathbf{I}_{n_2})\mathbf{N}_{n_2}(\mathbf{K}_{22}'^{-1} \otimes \mathbf{I}_{n_2}).
\end{aligned}$$

Therefore, the information matrix with respect to $\mathbf{\Lambda}_{22}$ becomes

$$\begin{aligned}
\mathbf{I}_T(\mathbf{B}_{22}, \mathbf{F}_{22}) &= \frac{1}{T}E(\mathbf{g}(\mathbf{B}_{22}, \mathbf{F}_{22})\mathbf{g}(\mathbf{B}_{22}, \mathbf{F}_{22})') \\
&= \begin{bmatrix} \mathbf{I}_{n_2} \otimes \mathbf{F}_{22}'^{-1} \\ -\mathbf{F}_{22}^{-1}\mathbf{B}_{22} \otimes \mathbf{F}_{22}'^{-1} \end{bmatrix} \frac{1}{T}E(\mathbf{g}(\mathbf{K}_{22})\mathbf{g}(\mathbf{K}_{22})') \left[\mathbf{I}_{n_2} \otimes \mathbf{F}_{22}^{-1} \vdots -\mathbf{B}_{22}'\mathbf{F}_{22}'^{-1} \otimes \mathbf{F}_{22}^{-1}\right] \\
&= \begin{bmatrix} \mathbf{I}_{n_2} \otimes \mathbf{F}_{22}'^{-1} \\ -\mathbf{F}_{22}^{-1}\mathbf{B}_{22} \otimes \mathbf{F}_{22}'^{-1} \end{bmatrix} \mathbf{I}_T(\mathbf{K}_{22}) \left[\mathbf{I}_{n_2} \otimes \mathbf{F}_{22}^{-1} \vdots -\mathbf{B}_{22}'\mathbf{F}_{22}'^{-1} \otimes \mathbf{F}_{22}^{-1}\right] \\
&= 2T \begin{bmatrix} \mathbf{B}_{22}^{-1}\mathbf{F}_{22} \otimes \mathbf{F}_{22}'^{-1} \\ -\mathbf{I}_{n_2} \otimes \mathbf{F}_{22}'^{-1} \end{bmatrix} \mathbf{N}_{n_2} \left[\mathbf{F}_{22}'\mathbf{B}_{22}'^{-1} \otimes \mathbf{F}_{22}^{-1} \vdots -\mathbf{I}_{n_2} \otimes \mathbf{F}_{22}^{-1}\right].
\end{aligned}$$

(d) A usual score algorithm applies, but the following transformation is necessary because $\mathbf{I}_T(\mathbf{B}_{22}, \mathbf{F}_{22})$ is singular. From $\boldsymbol{\lambda}_{22} = \mathbf{S}_\lambda \boldsymbol{\lambda}_s + \mathbf{s}_\lambda$ and

$$\begin{aligned}
\frac{\partial L}{\partial \boldsymbol{\lambda}_s'} &= \frac{\partial L}{\partial \boldsymbol{\lambda}_{22}'} \frac{\partial \boldsymbol{\lambda}_{22}}{\partial \boldsymbol{\lambda}_s'} \\
&= \mathbf{g}(\mathbf{B}_{22}, \mathbf{F}_{22})'\mathbf{S}_\lambda,
\end{aligned}$$

we get the gradient

$$\mathbf{g}(\boldsymbol{\lambda}_s) = \mathbf{S}_\lambda'\mathbf{g}(\mathbf{B}_{22}, \mathbf{F}_{22})$$

and the information matrix

$$\begin{aligned}
\mathbf{I}_T(\boldsymbol{\lambda}_s) &= \frac{1}{T}E(\mathbf{g}(\boldsymbol{\lambda}_s)\mathbf{g}(\boldsymbol{\lambda}_s)') \\
&= \mathbf{S}_\lambda' \frac{1}{T}E(\mathbf{g}(\mathbf{B}_{22}, \mathbf{F}_{22})\mathbf{g}(\mathbf{B}_{22}, \mathbf{F}_{22})')\mathbf{S}_\lambda \\
&= \mathbf{S}_\lambda'\mathbf{I}_T(\mathbf{B}_{22}, \mathbf{F}_{22})\mathbf{S}_\lambda.
\end{aligned}$$

(e) We write the log likelihood function for the sub-system following Giannini (1992), with some modifications

$$L(\mathbf{\Sigma}) \;=\; -\frac{T}{2}\log|\hat{\bar{\mathbf{M}}}\mathbf{\Sigma}\hat{\bar{\mathbf{M}}}'| - \frac{T}{2}\mathrm{trace}((\hat{\bar{\mathbf{M}}}\mathbf{\Sigma}\hat{\bar{\mathbf{M}}}')^{-1}\hat{\bar{\mathbf{M}}}\hat{\mathbf{\Sigma}}\hat{\bar{\mathbf{M}}}').$$

Therefore, the log likelihood ratio becomes

$$LRT \;=\; 2(L_u - L_r),$$

where

$$\begin{aligned} L_u &= L(\hat{\mathbf{\Sigma}}) = -\frac{T}{2}\log|\hat{\bar{\mathbf{M}}}\hat{\mathbf{\Sigma}}\hat{\bar{\mathbf{M}}}'| - \frac{n_2 T}{2} \quad \text{and} \\ L_r &= L(\mathbf{\Sigma}_{ML}) = T\log|\hat{\mathbf{B}}_{22}| - T\log|\hat{\mathbf{F}}_{22}| - \frac{T}{2}\mathrm{trace}(\hat{\mathbf{B}}_{22}'\hat{\mathbf{F}}_{22}'^{-1}\hat{\mathbf{F}}_{22}^{-1}\hat{\mathbf{B}}_{22}\hat{\bar{\mathbf{M}}}\hat{\mathbf{\Sigma}}\hat{\bar{\mathbf{M}}}') \end{aligned}$$

from $\hat{\bar{\mathbf{M}}}\mathbf{\Sigma}_{ML}\hat{\bar{\mathbf{M}}}' = (\hat{\mathbf{B}}_{22}'\hat{\mathbf{F}}_{22}'^{-1}\hat{\mathbf{F}}_{22}^{-1}\hat{\mathbf{B}}_{22})^{-1}$. ∎

*Example* 2.1. (Bernanke-Mihov Model, $n_3 = 0$)

It is straightforward to apply Theorem 2.1 to Bernanke and Mihov's (1998) model with the choice of $n_3 = 0$. Suppose that $\mathbf{B}_0$ is block lower triangular and $\mathbf{F}$ is block diagonal

$$\mathbf{B}_0 = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{0} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} \mathbf{F}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_{22} \end{bmatrix},$$

where $\mathbf{B}_{21}, \mathbf{B}_{22}$ and $\mathbf{F}_{22}$ are the structural parameters of interest. Write $\bar{\mathbf{B}}_0 = \mathbf{B}_{22}\bar{\mathbf{M}}$, where $\bar{\mathbf{M}} = [-\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1} \vdots \mathbf{I}_{n_2}]$. Provided that $\mathbf{B}_{22}$ and $\mathbf{F}_{22}$ are identifiable using short-run restrictions, the sub-system is two-step ML-estimable, as shown in Theorem 2.1. Alternatively, we may estimate $\mathbf{B}_{22}$ and $\mathbf{F}_{22}$ by solving

$$\mathbf{B}_{22}\hat{\bar{\mathbf{M}}}\hat{\mathbf{\Sigma}}\hat{\bar{\mathbf{M}}}'\mathbf{B}_{22}' = \mathbf{F}_{22}\mathbf{F}_{22}'$$

numerically for just-identified models or using the two-step GMM for over-identified models, as in Bernanke and Mihov (1998).

## 3 Asymptotics on ML estimators

It is often of interest to trace the dynamic responses and to decompose the forecast-error variance of economic variables. In order to provide confidence intervals for impulse responses, this section derives the asymptotic distribution of $\hat{\mathbf{\Phi}}_0 = \hat{\mathbf{B}}_0^{-1}\hat{\mathbf{F}}$. Once its asymptotic distribution is provided, it is straightforward to derive asymptotics for impulse responses and forecast-error variance decomposition, as shown in the next section.

We begin with the first-step OLS estimates. Let $\boldsymbol{\theta} = (\mathbf{a}', \boldsymbol{\sigma}')'$, where $\mathbf{a} = \text{vec}(\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_p)$ and $\boldsymbol{\sigma} = \text{vech}(\mathbf{\Sigma})$, where vech is the column stacking operator that stacks only the elements on and below the diagonal. It is well known that $\boldsymbol{\theta}$ is asymptotically normally distributed

$$\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} N(\mathbf{0}, \mathbf{\Sigma}_\theta),$$

where

$$\mathbf{\Sigma}_\theta = \begin{bmatrix} \mathbf{\Sigma}_a & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_\sigma \end{bmatrix} = \begin{bmatrix} [E(\mathbf{x}_t\mathbf{x}_t')]^{-1} \otimes \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & 2\mathbf{D}_n^{+}(\mathbf{\Sigma} \otimes \mathbf{\Sigma})\mathbf{D}_n^{+\prime} \end{bmatrix},$$

$\mathbf{x}_t = \left[\mathbf{y}_{t-1}', \mathbf{y}_{t-2}', \cdots, \mathbf{y}_{t-p}'\right]'$, and $\mathbf{D}_n^{+}$ is the Moore-Penrose inverse of $\mathbf{D}_n$, as defined in Appendix A. Refer to Hamilton (1994) for its derivation and extended discussion.

Due to the block diagonal property of $\mathbf{\Sigma}_\theta$, the asymptotic distribution of the second-step estimator of $\mathbf{\Phi}_0$ depends on $\mathbf{\Sigma}_\sigma$ only, so that one may use the asymptotic variance of $\hat{\mathbf{\Phi}}_0$ or $\mathbf{\Sigma}_\sigma$ for asymptotic distributions of impulse responses and forecast-error variance decomposition. For example, Giannini (1992) uses the asymptotic variance of $\hat{\mathbf{\Phi}}_0$ , while Lütkepohl (1990) uses $\mathbf{\Sigma}_\sigma$. When partially identified models are considered, it is convenient to use $\mathbf{\Sigma}_\sigma$, because the transformation matrix used for ML estimation is a function of $\boldsymbol{\sigma}$. Write

$$\boldsymbol{\Phi}_0 = \begin{bmatrix} \underset{(n_1\times n_1)}{\boldsymbol{\Phi}_{11}} & \underset{(n_1\times n_2)}{\mathbf{0}} & \underset{(n_1\times n_3)}{\mathbf{0}} \\ \underset{(n_2\times n_1)}{\boldsymbol{\Phi}_{21}} & \underset{(n_2\times n_2)}{\boldsymbol{\Phi}_{22}} & \underset{(n_2\times n_3)}{\mathbf{0}} \\ \underset{(n_3\times n_1)}{\boldsymbol{\Phi}_{31}} & \underset{(n_3\times n_2)}{\boldsymbol{\Phi}_{32}} & \underset{(n_3\times n_3)}{\boldsymbol{\Phi}_{33}} \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} \underset{(n_1\times n_1)}{\boldsymbol{\Sigma}_{11}} & \underset{(n_1\times n_2)}{\boldsymbol{\Sigma}_{12}} & \underset{(n_1\times n_3)}{\boldsymbol{\Sigma}_{13}} \\ \underset{(n_2\times n_1)}{\boldsymbol{\Sigma}_{21}} & \underset{(n_2\times n_2)}{\boldsymbol{\Sigma}_{22}} & \underset{(n_2\times n_3)}{\boldsymbol{\Sigma}_{23}} \\ \underset{(n_3\times n_1)}{\boldsymbol{\Sigma}_{31}} & \underset{(n_3\times n_2)}{\boldsymbol{\Sigma}_{32}} & \underset{(n_3\times n_3)}{\boldsymbol{\Sigma}_{33}} \end{bmatrix}$$

and denote $\boldsymbol{\sigma}^* = (\boldsymbol{\sigma}'_{11}, \boldsymbol{\sigma}'_{21}, \boldsymbol{\sigma}'_{31}, \boldsymbol{\sigma}'_{22}, \boldsymbol{\sigma}'_{32}, \boldsymbol{\sigma}'_{33})'$. Let $\boldsymbol{\Sigma}_{\sigma^*}$ be the asymptotic variance of $\hat{\boldsymbol{\sigma}}^*$, where $\boldsymbol{\sigma}_{ii} = \text{vech}(\boldsymbol{\Sigma}_{ii})$ and $\boldsymbol{\sigma}_{ij} = \text{vec}(\boldsymbol{\Sigma}_{ij})$ if $i \neq j$. It is particularly interesting to derive the asymptotic property of the second-step estimator of $\bar{\boldsymbol{\Phi}}_0 = [\mathbf{0} \vdots \boldsymbol{\Phi}'_{22} \vdots \boldsymbol{\Phi}'_{32}]'$. From $\boldsymbol{\Phi}_0 = \mathbf{M}^{-1}\mathbf{B}_d^{-1}\mathbf{F}$, we can show that

$$\bar{\boldsymbol{\Phi}}_0 = \tilde{\mathbf{M}}\boldsymbol{\Phi}_{22}, \tag{3.1}$$

where $\tilde{\mathbf{M}}$ is the second-column block of $\mathbf{M}^{-1}$. Since $\mathbf{M}$ is lower block triangular, we can write $\tilde{\mathbf{M}} = [\mathbf{0} \vdots \mathbf{I}_{n_2} \vdots -\mathbf{M}'_{32}]'$, where $\mathbf{M}_{32}$ is defined in Lemma 2.1. For computational purposes, define $\bar{\boldsymbol{\Sigma}} = \bar{\mathbf{M}}\boldsymbol{\Sigma}\bar{\mathbf{M}}' = \boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}'_{21}$, $\tilde{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}_{32} - \boldsymbol{\Sigma}_{31}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}'_{21}$, and write $\mathbf{M}_{32} = -\tilde{\boldsymbol{\Sigma}}\bar{\boldsymbol{\Sigma}}^{-1}$ in a compact form. In addition, denote $\bar{\boldsymbol{\sigma}} = \text{vech}(\bar{\boldsymbol{\Sigma}})$ and $\tilde{\boldsymbol{\sigma}} = \text{vec}(\tilde{\boldsymbol{\Sigma}})$. Since $\bar{\boldsymbol{\Phi}}_0$ in (3.1) depends on $\tilde{\mathbf{M}}$ and $\boldsymbol{\Phi}_{22}$ we derive the asymptotic variance of $\hat{\tilde{\mathbf{M}}}$ in Lemma 3.1 below and that of $\hat{\boldsymbol{\Phi}}_{22}$ later in Theorem 3.1.

**Lemma 3.1.** *(Asymptotic distributions of the first-step estimators)*

*(a)* $\sqrt{T}(\hat{\boldsymbol{\sigma}}^* - \boldsymbol{\sigma}^*) \xrightarrow{d} N(\mathbf{0}, \mathbf{G}_{\sigma^*\sigma}\boldsymbol{\Sigma}_\sigma\mathbf{G}'_{\sigma^*\sigma})$, *where*

$$\mathbf{G}_{\sigma^*\sigma} = \begin{bmatrix} \mathbf{D}_{n_1}^+ & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{2}\mathbf{K}_{n_1n_2} & \mathbf{0} & \frac{1}{2}\mathbf{I}_{n_2n_1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{1}{2}\mathbf{K}_{n_1n_3} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \frac{1}{2}\mathbf{I}_{n_3n_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_{n_2}^+ & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \frac{1}{2}\mathbf{K}_{n_2n_3} & \mathbf{0} & \frac{1}{2}\mathbf{I}_{n_3n_2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_{n_3}^+ \end{bmatrix} \begin{bmatrix} \mathbf{K}_{nn_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{nn_2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{K}_{nn_3} \end{bmatrix} \mathbf{D}_n.$$

*(b)* $\sqrt{T}(\hat{\bar{\boldsymbol{\sigma}}} - \bar{\boldsymbol{\sigma}}) \xrightarrow{d} N(\mathbf{0}, \bar{\mathbf{G}}_{\sigma\sigma^*}\boldsymbol{\Sigma}_{\sigma^*}\bar{\mathbf{G}}'_{\sigma\sigma^*})$, *where*

$$\bar{\mathbf{G}}_{\sigma\sigma^*} = \left[\mathbf{D}_{n_2}^+(\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} \otimes \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1})\mathbf{D}_{n_1} \vdots -2\mathbf{D}_{n_2}^+(\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} \otimes \mathbf{I}_{n_2}) \vdots \mathbf{0} \vdots \mathbf{I}_{\frac{n_2(n_2+1)}{2}} \vdots \mathbf{0} \vdots \mathbf{0}\right].$$

*(c)* $\sqrt{T}(\hat{\tilde{\boldsymbol{\sigma}}} - \tilde{\boldsymbol{\sigma}}) \xrightarrow{d} N(\mathbf{0}, \tilde{\mathbf{G}}_{\sigma\sigma^*}\boldsymbol{\Sigma}_{\sigma^*}\tilde{\mathbf{G}}'_{\sigma\sigma^*})$*, where*

$$\tilde{\mathbf{G}}_{\sigma\sigma^*} = \left[(\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} \otimes \boldsymbol{\Sigma}_{31}\boldsymbol{\Sigma}_{11}^{-1})\mathbf{D}_{n_1} \vdots -\mathbf{K}_{n_2 n_3}(\boldsymbol{\Sigma}_{31}\boldsymbol{\Sigma}_{11}^{-1} \otimes \mathbf{I}_{n_2}) \vdots -\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} \otimes \mathbf{I}_{n_3} \vdots\ \mathbf{0} \vdots\ \mathbf{I}_{n_3 n_2} \vdots\ \mathbf{0}\right].$$

*(d)* $\sqrt{T}(\hat{\bar{\mathbf{m}}} - \bar{\mathbf{m}}) \xrightarrow{d} N(\mathbf{0}, \bar{\mathbf{G}}_{m\sigma^*}\boldsymbol{\Sigma}_{\sigma^*}\bar{\mathbf{G}}'_{m\sigma^*})$*, where*

$$\bar{\mathbf{G}}_{m\sigma^*} = \begin{bmatrix} (\boldsymbol{\Sigma}_{11}^{-1} \otimes \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1})\mathbf{D}_{n_1} & -\boldsymbol{\Sigma}_{11}^{-1} \otimes \mathbf{I}_{n_2} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0}_{n_2^2 \times \frac{n_1(n_1+1)}{2}} & \mathbf{0}_{n_2^2 \times n_2 n_1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0}_{n_2 n_3 \times \frac{n_1(n_1+1)}{2}} & \mathbf{0}_{n_2 n_3 \times n_2 n_1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

*(e)* $\sqrt{T}(\hat{\tilde{\mathbf{m}}} - \tilde{\mathbf{m}}) \xrightarrow{d} N(\mathbf{0}, \tilde{\mathbf{G}}_{m\sigma^*}\boldsymbol{\Sigma}_{\sigma^*}\tilde{\mathbf{G}}'_{m\sigma^*})$*, where*

$$\tilde{\mathbf{G}}_{m\sigma^*} = \mathbf{K}_{n_2 n} \begin{bmatrix} \mathbf{0}_{n_1 n_2 \times \frac{n(n+1)}{2}} \\ \mathbf{0}_{n_2^2 \times \frac{n(n+1)}{2}} \\ \mathbf{K}_{n_3 n_2}\left[(\bar{\boldsymbol{\Sigma}}^{-1} \otimes \mathbf{I}_{n_3})\tilde{\mathbf{G}}_{\sigma\sigma^*} - (\bar{\boldsymbol{\Sigma}}^{-1} \otimes \tilde{\boldsymbol{\Sigma}}\bar{\boldsymbol{\Sigma}}^{-1})\mathbf{D}_{n_2}\bar{\mathbf{G}}_{\sigma\sigma^*}\right] \end{bmatrix}.$$

*Proof.* See Appendix B. ∎

As $\tilde{\mathbf{M}}$ depends on $\bar{\boldsymbol{\Sigma}}$ and $\tilde{\boldsymbol{\Sigma}}$, we first derive the asymptotic variances of $\hat{\bar{\boldsymbol{\Sigma}}}$ and $\hat{\tilde{\boldsymbol{\Sigma}}}$ in Lemma 3.1 (b) and (c), respectively, then we provide the asymptotic variance of $\hat{\tilde{\mathbf{M}}}$ in terms of $\boldsymbol{\Sigma}_{\sigma^*}$ in Lemma 3.1 (e). The result of Lemma 3.1 (d) is used to derive the asymptotic variance of $\hat{\bar{\mathbf{B}}}_{\mathbf{0}}$ in Theorem 3.2 below. With the result of Lemma 3.1 (a), we can express the asymptotic variances of $\hat{\bar{\mathbf{M}}}$ and $\hat{\tilde{\mathbf{M}}}$ in terms of $\boldsymbol{\Sigma}_\sigma$ using $\sqrt{T}(\hat{\bar{\mathbf{m}}} - \bar{\mathbf{m}}) \xrightarrow{d} N(\mathbf{0}, \bar{\mathbf{G}}_{m\sigma}\boldsymbol{\Sigma}_\sigma\bar{\mathbf{G}}'_{m\sigma^*})$ and $\sqrt{T}(\hat{\tilde{\mathbf{m}}} - \tilde{\mathbf{m}}) \xrightarrow{d} N(\mathbf{0}, \tilde{\mathbf{G}}_{m\sigma}\boldsymbol{\Sigma}_\sigma\tilde{\mathbf{G}}'_{m\sigma^*})$, where $\bar{\mathbf{G}}_{m\sigma} = \bar{\mathbf{G}}_{m\sigma*}\mathbf{G}_{\sigma*\sigma}$ and $\tilde{\mathbf{G}}_{m\sigma} = \tilde{\mathbf{G}}_{m\sigma*}\mathbf{G}_{\sigma*\sigma}$.

Now, we express the asymptotic variance of $\hat{\boldsymbol{\Phi}}_{22}$ in terms of $\boldsymbol{\Sigma}_\sigma$. This requires three steps. First, we compute the total derivative of $\boldsymbol{\Phi}_{22}$ in terms of the second-step ML estimator $\hat{\boldsymbol{\lambda}}_s$ in Lemma 3.2. Then, we derive the asymptotic variance of $\hat{\boldsymbol{\lambda}}_s$ in Theorem 3.1, which is equivalent to inverse of the information matrix of $\hat{\boldsymbol{\lambda}}_s$ given in Theorem 2.1. By writing the asymptotic variance of $\hat{\boldsymbol{\lambda}}_s$ in terms of $\boldsymbol{\Sigma}_\sigma$, we can simplify the expression of the asymptotic distribution of impulse responses in the next section. Finally, the asymptotic variance of $\hat{\hat{\mathbf{\Phi}}}_0$ is derived in terms of $\mathbf{\Sigma}_\sigma$ in Theorem 3.2.

**Lemma 3.2.** *(Total derivatives of the second-step inferred estimators)*

*(a)* $d\mathbf{k}_{22} = \mathbf{G}_{k\lambda} d\boldsymbol{\lambda}_s$, *where*

$$\mathbf{G}_{k\lambda} = \left[\mathbf{I}_{n_2} \otimes \mathbf{F}_{22}^{-1} \vdots -\mathbf{B}'_{22}\mathbf{F}'^{-1}_{22} \otimes \mathbf{F}_{22}^{-1}\right] \mathbf{S}_\lambda.$$

*(b)* $d\boldsymbol{\phi}_{22} = \mathbf{G}_{\phi\lambda} d\boldsymbol{\lambda}_s$, *where*

$$\mathbf{G}_{\phi\lambda} = \left[-\mathbf{F}'_{22}\mathbf{B}'^{-1}_{22} \otimes \mathbf{B}_{22}^{-1} \vdots \mathbf{I}_{n_2} \otimes \mathbf{B}_{22}^{-1}\right] \mathbf{S}_\lambda.$$

*(c)* $d\text{vech}(\mathbf{B}_{22}^{-1}\mathbf{F}_{22}\mathbf{F}'_{22}\mathbf{B}'^{-1}_{22}) = \mathbf{G}_{\phi\phi\lambda} d\boldsymbol{\lambda}_s$, *where*

$$\mathbf{G}_{\phi\phi\lambda} = 2\mathbf{D}_{n_2}^{+} \left[-\mathbf{B}_{22}^{-1}\mathbf{F}_{22}\mathbf{F}'_{22}\mathbf{B}'^{-1}_{22} \otimes \mathbf{B}_{22}^{-1} \vdots \mathbf{B}_{22}^{-1}\mathbf{F}_{22} \otimes \mathbf{B}_{22}^{-1}\right] \mathbf{S}_\lambda.$$

*Proof.* See Appendix B. ∎

The results of Lemma 3.2 (a) and (b) are used to derive the asymptotic variance of $\hat{\mathbf{K}}_{22}$ and $\hat{\mathbf{\Phi}}_{22}$, respectively, in Theorem 3.2 below. Note, however, that these results are expressed in terms of the asymptotic variance of the ML estimator, $\mathbf{\Sigma}_{\lambda_s}$, rather than $\mathbf{\Sigma}_\sigma$. To express the asymptotic variance of $\hat{\boldsymbol{\lambda}}_s$ in terms of $\mathbf{\Sigma}_\sigma$, we use the following property

$$\bar{\mathbf{M}}\mathbf{\Sigma}\bar{\mathbf{M}}' = \mathbf{B}_{22}^{-1}\mathbf{F}_{22}\mathbf{F}'_{22}\mathbf{B}'^{-1}_{22},$$

where the asymptotic variance of the left-hand side is a function of $\mathbf{\Sigma}_\sigma$, while the right-hand side is a function of $\mathbf{\Sigma}_{\lambda_s}$ from Lemma 3.2 (c). From these results, we can state the asymptotic distributions of the ML estimator $\hat{\boldsymbol{\lambda}}_s$ in terms of $\mathbf{\Sigma}_\sigma$ in the following theorem:

**Theorem 3.1.** *(Asymptotic distributions of the second-step ML estimators)* $\sqrt{T}(\hat{\boldsymbol{\lambda}}_s - \boldsymbol{\lambda}_s) \xrightarrow{d} N(\mathbf{0}, \mathbf{G}_{\lambda\sigma}\boldsymbol{\Sigma}_\sigma\mathbf{G}'_{\lambda\sigma})$, *where*

$$\mathbf{G}_{\lambda\sigma} = \mathbf{G}^{+}_{\phi\phi\lambda}\bar{\mathbf{G}}_{\sigma\sigma^*}\mathbf{G}_{\sigma^*\sigma}$$

*Proof.* It follows from $\bar{\boldsymbol{\Sigma}} = \mathbf{B}_{22}^{-1}\mathbf{F}_{22}\mathbf{F}'_{22}\mathbf{B}'^{-1}_{22}$ that $d\bar{\boldsymbol{\sigma}} = d\text{vech}(\mathbf{B}_{22}^{-1}\mathbf{F}_{22}\mathbf{F}'_{22}\mathbf{B}'^{-1}_{22})$. Therefore, $d\bar{\boldsymbol{\sigma}} = \mathbf{G}_{\phi\phi\lambda}d\boldsymbol{\lambda}_s$ from Lemma 3.2 (c). Since $\mathbf{G}_{\phi\phi\lambda}$ has a full column rank, we get $d\boldsymbol{\lambda}_s = \mathbf{G}^{+}_{\phi\phi\lambda}d\bar{\boldsymbol{\sigma}}$, where $d\bar{\boldsymbol{\sigma}} = \bar{\mathbf{G}}_{\sigma\sigma^*}d\boldsymbol{\sigma}^* = \bar{\mathbf{G}}_{\sigma\sigma^*}\mathbf{G}_{\sigma^*\sigma}d\boldsymbol{\sigma}$. Finally, the delta method yields $\boldsymbol{\Sigma}_{\lambda_s} = \mathbf{G}_{\lambda\sigma}\boldsymbol{\Sigma}_\sigma\mathbf{G}'_{\lambda\sigma}$. ∎

Let $\mathbf{G}_{b\sigma} = [\ \mathbf{I}_{n_{b_s}} \vdots\ \mathbf{0}\ ]\mathbf{G}_{\lambda\sigma}$, where $n_{b_s}$ is the number of free parameters in $\mathbf{B}_{22}$. With this notation, the asymptotic distributions of $\hat{\bar{\boldsymbol{\Phi}}}_0$ and other inferred estimators, such as $\hat{\bar{\mathbf{B}}}_{\mathbf{0}}$ and $\hat{\bar{\mathbf{K}}}_{\mathbf{0}}$, are given in the next theorem.

**Theorem 3.2.** *(Asymptotic distributions of second-step inferred estimators)*

*(a)* $\sqrt{T}(\hat{\mathbf{k}}_{22} - \mathbf{k}_{22}) \xrightarrow{d} N(\mathbf{0}, \mathbf{G}_{k\sigma}\boldsymbol{\Sigma}_\sigma\mathbf{G}'_{k\sigma})$, *where*

$$\mathbf{G}_{k\sigma} = \mathbf{G}_{k\lambda}\mathbf{G}_{\lambda\sigma}.$$

*(b)* $\sqrt{T}(\hat{\boldsymbol{\phi}}_{22} - \boldsymbol{\phi}_{22}) \xrightarrow{d} N(\mathbf{0}, \mathbf{G}_{\phi\sigma}\boldsymbol{\Sigma}_\sigma\mathbf{G}'_{\phi\sigma})$, *where*

$$\mathbf{G}_{\phi\sigma} = \mathbf{G}_{\phi\lambda}\mathbf{G}_{\lambda\sigma}.$$

*(c)* $\sqrt{T}(\hat{\bar{\mathbf{b}}}_0 - \bar{\mathbf{b}}_0) \xrightarrow{d} N(\mathbf{0}, \bar{\mathbf{G}}_{b\sigma}\boldsymbol{\Sigma}_\sigma\bar{\mathbf{G}}'_{b\sigma})$, *where*

$$\bar{\mathbf{G}}_{b\sigma} = (\bar{\mathbf{M}}' \otimes \mathbf{I}_{n_2})\mathbf{S}_b\mathbf{G}_{b\sigma} + (\mathbf{I}_n \otimes \mathbf{B}_{22})\bar{\mathbf{G}}_{m\sigma^*}\mathbf{G}_{\sigma^*\sigma}.$$

*(d)* $\sqrt{T}(\hat{\bar{\mathbf{k}}}_0 - \bar{\mathbf{k}}_0) \xrightarrow{d} N(\mathbf{0}, \bar{\mathbf{G}}_{k\sigma}\boldsymbol{\Sigma}_\sigma\bar{\mathbf{G}}'_{k\sigma})$, *where*

$$\bar{\mathbf{G}}_{k\sigma} = (\bar{\mathbf{M}}' \otimes \mathbf{I}_{n_2})\mathbf{G}_{k\sigma} + (\mathbf{I}_n \otimes \mathbf{F}_{22}^{-1}\mathbf{B}_{22})\bar{\mathbf{G}}_{m\sigma^*}\mathbf{G}_{\sigma^*\sigma}.$$

*(e)* $\sqrt{T}(\hat{\bar{\boldsymbol{\phi}}}_0 - \bar{\boldsymbol{\phi}}_0) \xrightarrow{d} N(\mathbf{0}, \bar{\mathbf{G}}_{\phi\sigma}\boldsymbol{\Sigma}_\sigma\bar{\mathbf{G}}'_{\phi\sigma})$, *where*

$$\bar{\mathbf{G}}_{\phi\sigma} = (\mathbf{I}_{n_2} \otimes \tilde{\mathbf{M}})\mathbf{G}_{\phi\sigma} + (\boldsymbol{\Phi}'_{22} \otimes \mathbf{I}_n)\tilde{\mathbf{G}}_{m\sigma^*}\mathbf{G}_{\sigma^*\sigma}.$$

*Proof.* (a) From Lemma 3.2 and Theorem 3.1, we obtain $d\mathbf{k}_{22} = \mathbf{G}_{k\lambda} d\boldsymbol{\lambda}_s = \mathbf{G}_{k\lambda}\mathbf{G}_{\lambda\sigma} d\boldsymbol{\sigma}$.

(b) Similarly, $d\boldsymbol{\phi}_{22} = \mathbf{G}_{\phi\lambda} d\boldsymbol{\lambda}_s = \mathbf{G}_{\phi\lambda}\mathbf{G}_{\lambda\sigma} d\boldsymbol{\sigma}$.

(c) From $\bar{\mathbf{B}}_0 = \mathbf{B}_{22}\bar{\mathbf{M}}$, we can show that $d\bar{\mathbf{b}}_0 = d\text{vec}(\mathbf{B}_{22}\bar{\mathbf{M}}) = (\bar{\mathbf{M}}' \otimes \mathbf{I}_{n_2}) d\text{vec}(\mathbf{B}_{22}) + (\mathbf{I}_n \otimes \mathbf{B}_{22}) d\text{vec}(\bar{\mathbf{M}})$. From Lemma 3.1 and Theorem 3.1, we get $d\bar{\mathbf{b}}_0 = (\bar{\mathbf{M}}' \otimes \mathbf{I}_{n_2})\mathbf{S}_b d\mathbf{b}_s + (\mathbf{I}_n \otimes \mathbf{B}_{22})\bar{\mathbf{G}}_{m\sigma^*} d\boldsymbol{\sigma}^* = \left[(\bar{\mathbf{M}}' \otimes \mathbf{I}_{n_2})\mathbf{S}_b \mathbf{G}_{b\sigma} + (\mathbf{I}_n \otimes \mathbf{B}_{22})\bar{\mathbf{G}}_{m\sigma^*}\mathbf{G}_{\sigma^*\sigma}\right] d\boldsymbol{\sigma}$.

(d) From $\bar{\mathbf{K}}_0 = \mathbf{K}_{22}\bar{\mathbf{M}}$, it follows that $d\bar{\mathbf{k}}_0 = d\text{vec}(\mathbf{K}_{22}\bar{\mathbf{M}}) = (\bar{\mathbf{M}}' \otimes \mathbf{I}_{n_2}) d\text{vec}(\mathbf{K}_{22}) + (\mathbf{I}_n \otimes \mathbf{K}_{22}) d\text{vec}(\bar{\mathbf{M}})$. From Lemma 3.1 and 3.2, we obtain $d\bar{\mathbf{k}}_0 = [(\bar{\mathbf{M}}' \otimes \mathbf{I}_{n_2})\mathbf{G}_{k\sigma} + (\mathbf{I}_n \otimes \mathbf{F}_{22}^{-1}\mathbf{B}_{22})\bar{\mathbf{G}}_{m\sigma^*}\mathbf{G}_{\sigma^*\sigma}] d\boldsymbol{\sigma}$.

(e) Similar to the proof of (d). The delta method completes the proof. ∎

*Remark* 3.1. When the model is just-identified, the result of Theorem 3.2 (e) can be stated as $\sqrt{T}(\hat{\bar{\boldsymbol{\phi}}}_0 - \bar{\boldsymbol{\phi}}_0) \xrightarrow{d} N(\mathbf{0}, \bar{\mathbf{G}}_{\phi\sigma}\boldsymbol{\Sigma}_\sigma \bar{\mathbf{G}}'_{\phi\sigma})$, where

$$\bar{\mathbf{G}}_{\phi\sigma} = (\mathbf{F}_{22}^{-1}\mathbf{B}_{22}\bar{\mathbf{M}} \otimes \mathbf{I}_n)\mathbf{D}_n + (\mathbf{I}_{n_2} \otimes \boldsymbol{\Sigma})\mathbf{K}_{n_2 n}\bar{\mathbf{G}}_{k\sigma}$$

from the relation $\bar{\boldsymbol{\Phi}}_0 = \boldsymbol{\Sigma}\bar{\mathbf{K}}_0'$. Among others, King, Plosser, Stock, and Watson (1991) used the relation $\bar{\boldsymbol{\Phi}}_0 = \boldsymbol{\Sigma}\bar{\mathbf{K}}_0'$ for just-identified models. The first-step OLS estimator $\hat{\boldsymbol{\Sigma}}$ can be used to obtain $\hat{\bar{\boldsymbol{\Phi}}}_0 = \hat{\boldsymbol{\Sigma}}\hat{\bar{\mathbf{K}}}_{\mathbf{0}}'$ because $\boldsymbol{\Sigma}_{ML} = \boldsymbol{\Sigma}_{OLS}$ for just-identified models, where $\boldsymbol{\Sigma}_{OLS} = \hat{\boldsymbol{\Sigma}}$ and $\boldsymbol{\Sigma}_{ML} = \hat{\boldsymbol{\Phi}}_0\hat{\boldsymbol{\Phi}}_0'$. Note that we need to estimate $\bar{\boldsymbol{\Phi}}_0$ using (3.1) for over-identified models in which $\boldsymbol{\Sigma}_{ML} \neq \boldsymbol{\Sigma}_{OLS}$.

# 4 Asymptotics of Impulse Responses and Forecast-Error Variance Decompositions

This section provides the asymptotic distributions of impulse responses and forecast-error variance decomposition, which are widely used as standard tools for economic analysis in the applied VAR literature (see, e.g., Baillie, 1987; Runkle, 1987). Given that $\mathbf{y}_t$ is covariance stationary, there exists a Wold representation

$$\mathbf{y}_t = \boldsymbol{\Psi}(L)\boldsymbol{\epsilon}_t$$

and the corresponding structural moving average representation

$$\mathbf{y}_t = \boldsymbol{\Phi}(L)\mathbf{e}_t,$$

where $\boldsymbol{\Psi}(L) = \mathbf{A}(L)^{-1} = \sum_{i=0}^{\infty} \boldsymbol{\Psi}_i L^i$, $\boldsymbol{\Psi}_0 = \mathbf{I}_n$, $\boldsymbol{\Phi}(L) = \mathbf{B}(L)^{-1}\mathbf{F} = \sum_{i=0}^{\infty} \boldsymbol{\Phi}_i L^i$, and $\boldsymbol{\Phi}_i = \boldsymbol{\Psi}_i\boldsymbol{\Phi}_0$. In particular, $i$-step impulse responses to the second set of structural shocks, $\mathbf{e}_{2t}$ are given by

$$\bar{\boldsymbol{\Phi}}_i = \boldsymbol{\Psi}_i\bar{\boldsymbol{\Phi}}_0.$$

It is often of interest to trace the accumulated responses

$$\boldsymbol{\Psi}_{ci} = \sum_{j=0}^{i} \boldsymbol{\Psi}_j, \quad \bar{\boldsymbol{\Phi}}_{ci} = \boldsymbol{\Psi}_{ci}\bar{\boldsymbol{\Phi}}_0$$

and the total accumulated responses

$$\boldsymbol{\Psi}(1) = \sum_{j=0}^{\infty} \boldsymbol{\Psi}_j = \mathbf{A}(1)^{-1}, \quad \bar{\boldsymbol{\Phi}}(1) = \boldsymbol{\Psi}(1)\bar{\boldsymbol{\Phi}}_0.$$

Let $\boldsymbol{\ell}_p$ be the $p$-dimensional vector with ones and denote

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \cdots & \mathbf{A}_{p-1} & \mathbf{A}_p \\ \mathbf{I}_n & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_n & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I}_n & \mathbf{0} \end{bmatrix} \text{ and } \mathbf{J}_{np} = \begin{bmatrix} \mathbf{I}_n \vdots \mathbf{0}_{n\times n(p-1)} \end{bmatrix}.$$

With this notation, the asymptotic distributions of the impulse responses are given in the next theorem.

**Theorem 4.1.** *Suppose* $\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma}_\theta)$ *and* $\boldsymbol{\lambda}_{22} = \mathbf{S}_\lambda\boldsymbol{\lambda}_s + \mathbf{s}_\lambda$. *Then* *(a)* $$\sqrt{T}\text{vec}(\hat{\mathbf{\Psi}}_i - \mathbf{\Psi}_i) \xrightarrow{d} N(\mathbf{0}, \mathbf{G}_{\Psi ai}\mathbf{\Sigma}_a\mathbf{G}'_{\Psi ai}), \quad i = 1, 2, \cdots,$$

*where*

$$\mathbf{G}_{\Psi ai} = \frac{\partial \text{vec}(\mathbf{\Psi}_i)}{\partial \mathbf{a}'} = \sum_{j=0}^{i-1} \mathbf{J}_{np}(\mathbf{A}')^{i-1-j} \otimes \mathbf{\Psi}_j;$$

*(b)* $$\sqrt{T}\text{vec}(\hat{\mathbf{\Psi}}_{ci} - \mathbf{\Psi}_{ci}) \xrightarrow{d} N(\mathbf{0}, \mathbf{G}_{\Psi cai}\mathbf{\Sigma}_a\mathbf{G}'_{\Psi cai}), \quad i = 1, 2, \cdots,$$

*where*

$$\mathbf{G}_{\Psi cai} = \frac{\partial \text{vec}(\mathbf{\Psi}_{ci})}{\partial \mathbf{a}'} = \sum_{j=0}^{i} \mathbf{G}_{\Psi aj};$$

*(c)* $$\sqrt{T}\text{vec}(\hat{\mathbf{\Psi}}(1) - \mathbf{\Psi}(1)) \xrightarrow{d} N(\mathbf{0}, \mathbf{G}_{\Psi 1a}\mathbf{\Sigma}_a\mathbf{G}'_{\Psi 1a})$$

*where*

$$\mathbf{G}_{\Psi 1a} = \frac{\partial \text{vec}(\mathbf{\Psi}(1))}{\partial \mathbf{a}'} = \boldsymbol{\ell}_p{}' \otimes \mathbf{\Psi}(1)' \otimes \mathbf{\Psi}(1);$$

*(d)* $$\sqrt{T}\text{vec}(\hat{\bar{\mathbf{\Phi}}}_i - \bar{\mathbf{\Phi}}_i) \xrightarrow{d} N(\mathbf{0}, \bar{\mathbf{G}}_{\Phi ai}\mathbf{\Sigma}_a\bar{\mathbf{G}}'_{\Phi ai} + \bar{\mathbf{G}}_{\Phi\sigma i}\mathbf{\Sigma}_\sigma\bar{\mathbf{G}}'_{\Phi\sigma i}), \quad i = 0, 1, 2, \cdots,$$

*where*

$$\begin{aligned} \bar{\mathbf{G}}_{\Phi ai} &= \frac{\partial \text{vec}(\bar{\mathbf{\Phi}}_i)}{\partial \mathbf{a}'} = \begin{cases} \mathbf{0}, & i = 0 \\ (\bar{\mathbf{\Phi}}_0' \otimes \mathbf{I}_n)\mathbf{G}_{\Psi ai}, & i = 1, 2, \cdots \end{cases} \quad \textit{and} \\ \bar{\mathbf{G}}_{\Phi\sigma i} &= \frac{\partial \text{vec}(\bar{\mathbf{\Phi}}_i)}{\partial \boldsymbol{\sigma}'} = (\mathbf{I}_{n_2} \otimes \mathbf{\Psi}_i)\bar{\mathbf{G}}_{\phi\sigma}; \end{aligned}$$

*(e)* $$\sqrt{T}\text{vec}(\hat{\bar{\mathbf{\Phi}}}_{ci} - \bar{\mathbf{\Phi}}_{ci}) \xrightarrow{d} N(\mathbf{0}, \bar{\mathbf{G}}_{\Phi cai}\mathbf{\Sigma}_a\bar{\mathbf{G}}'_{\Phi cai} + \bar{\mathbf{G}}_{\Phi c\sigma i}\mathbf{\Sigma}_\sigma\bar{\mathbf{G}}'_{\Phi c\sigma i}), \quad i = 0, 1, 2, \cdots,$$

*where*

$$\begin{aligned} \bar{\mathbf{G}}_{\Phi cai} &= \frac{\partial \text{vec}(\bar{\mathbf{\Phi}}_{ci})}{\partial \mathbf{a}'} = \sum_{j=0}^{i} \bar{\mathbf{G}}_{\Phi aj} \quad \textit{and} \\ \bar{\mathbf{G}}_{\Phi c\sigma i} &= \frac{\partial \text{vec}(\bar{\mathbf{\Phi}}_{ci})}{\partial \boldsymbol{\sigma}'} = \sum_{j=0}^{i} \bar{\mathbf{G}}_{\Phi\sigma j}; \end{aligned}$$

*(f)* $$\sqrt{T}\text{vec}(\hat{\bar{\mathbf{\Phi}}}(1) - \bar{\mathbf{\Phi}}(1)) \xrightarrow{d} N(\mathbf{0}, \bar{\mathbf{G}}_{\Phi 1a}\mathbf{\Sigma}_a\bar{\mathbf{G}}'_{\Phi 1a} + \bar{\mathbf{G}}_{\Phi 1\sigma}\mathbf{\Sigma}_\sigma\bar{\mathbf{G}}'_{\Phi 1\sigma}),$$

*where*

$$\begin{aligned}\bar{\mathbf{G}}_{\Phi 1a} &= \frac{\partial \text{vec}(\bar{\mathbf{\Phi}}_i)}{\partial \mathbf{a}'} = (\bar{\mathbf{\Phi}}'_0 \otimes \mathbf{I}_n)\mathbf{G}_{\Psi 1a} \quad \textit{and} \\ \bar{\mathbf{G}}_{\Phi 1\sigma} &= \frac{\partial \text{vec}(\bar{\mathbf{\Phi}}_i)}{\partial \boldsymbol{\sigma}'} = (\mathbf{I}_{n_2} \otimes \mathbf{\Psi}(1))\bar{\mathbf{G}}_{\phi\sigma}.\end{aligned}$$

*Proof.* (a)–(c) See Lütkepohl (1990) Proposition 1.

(d) It follows from $\bar{\mathbf{\Phi}}_i = \mathbf{\Psi}_i\bar{\mathbf{\Phi}}_0$ and Theorem 3.2 that $d\text{vec}(\bar{\mathbf{\Phi}}_i) = (\bar{\mathbf{\Phi}}'_0 \otimes \mathbf{I}_n)d\text{vec}(\mathbf{\Psi}_i) + (\mathbf{I}_{n_2} \otimes \mathbf{\Psi}_i)d\text{vec}(\bar{\mathbf{\Phi}}_0) = (\bar{\mathbf{\Phi}}'_0 \otimes \mathbf{I}_n)\mathbf{G}_{\Psi ai}d\mathbf{a} + (\mathbf{I}_{n_2} \otimes \mathbf{\Psi}_i)\bar{\mathbf{G}}_{\phi\sigma}d\boldsymbol{\sigma}$.

(e) Immediate from (d).

(f) Similar to the proof of (d). ∎

The structural model considered in this paper includes general classes of VAR models. The following corollary shows that the result of Lütkepohl (1990) is a special case when the model is fully- and just-identified with recursive assumptions.

**Corollary 4.1.** *Suppose that the model is just-identified and $\mathbf{\Phi}_0$ is lower triangular. Then,*

$$\bar{\mathbf{G}}_{\Phi\sigma i} = (\mathbf{I}_n \otimes \mathbf{\Psi}_i)\mathbf{L}'_n\left[2\mathbf{L}_n\mathbf{N}_n(\mathbf{\Phi}_0 \otimes \mathbf{I}_n)\mathbf{L}'_n\right]^{-1}, \quad i = 0, 1, \cdots, s;$$

*which is equivalent to Lütkepohl's (1990) Proposition 1-(v).*

*Proof.* Consider a fully identified model with $n_2 = n$ in which $\mathbf{B}_0 = \mathbf{I}_n$, $\mathbf{F} = \mathbf{\Phi}_0$, and $\bar{\mathbf{M}} = \mathbf{I}_n$. From Theorem 4.1 (d), we obtain $\mathbf{G}_{\Phi\sigma i} = (\mathbf{I}_n \otimes \mathbf{\Psi}_i)\bar{\mathbf{G}}_{\phi\sigma} = (\mathbf{I}_n \otimes \mathbf{\Psi}_i)\mathbf{G}_{\phi\sigma} = (\mathbf{I}_n \otimes \mathbf{\Psi}_i)\mathbf{G}_{\phi\lambda}\mathbf{G}_{\lambda\sigma}$, where $\mathbf{G}_{\phi\lambda} = \mathbf{S}_\lambda$ and $\mathbf{G}_{\lambda\sigma} = \mathbf{G}^+_{\phi\phi\lambda} = [2\mathbf{D}^+_n(\mathbf{\Phi}_0 \otimes \mathbf{I}_n)\mathbf{S}_\lambda]^{-1}$ for just-identified models. Moreover, $\mathbf{S}_\lambda = \mathbf{L}'_n$ because $\text{vec}(\mathbf{\Phi}_0) = \mathbf{L}'_n\text{vech}(\mathbf{\Phi}_0)$ when $\mathbf{\Phi}_0$ is lower triangular. The property of $\mathbf{D}^+_n = \mathbf{L}_n\mathbf{N}_n$ completes the proof. ∎

We now derive the asymptotic distributions of forecast-error variance decompositions. Let $\bar{w}_{h,ij}$ be the contribution of the $j$-th shock in $\mathbf{e}_{2t}$ to the $h$-step forecast-error variance of the $i$-th variable, $y_{it}$, which is obtained using

$$\bar{w}_{h,ij} = \sum_{s=0}^{h-1} \bar{\phi}_{s,ij}^2 / MSE_i(h), \tag{4.1}$$

where $MSE_i(h) = \sum_{s=0}^{h-1} \boldsymbol{\iota}_i' \boldsymbol{\Psi}_s' \boldsymbol{\Sigma} \boldsymbol{\Psi}_s \boldsymbol{\iota}_i$ is the mean square error of the $h$-step forecast of $y_{it}$ and $\boldsymbol{\iota}_i$ is the $i$-th column of $\mathbf{I}_n$. For computational purposes, write the forecast-error variance components (4.1) using the following matrices

$$\bar{\mathbf{W}}_h = [\bar{w}_{h,ij}]_{n \times n_2} = \mathbf{W}_{h\Psi} \bar{\mathbf{W}}_{h\Phi}, \tag{4.2}$$

where $\mathbf{W}_{h\Psi} = \left[ (\sum_{s=0}^{h-1} \boldsymbol{\Psi}_s \boldsymbol{\Sigma} \boldsymbol{\Psi}_s') \odot \mathbf{I}_n \right]^{-1}$, $\bar{\mathbf{W}}_{h\Phi} = \sum_{s=0}^{h-1} (\bar{\boldsymbol{\Phi}}_s \odot \bar{\boldsymbol{\Phi}}_s)$, and $\odot$ is the Hadamard operator that just takes the product of corresponding pairs of entries. See Appendix A for details. Following Giannini (1992), let $D_v(\mathbf{M}) = diag(\text{vec}(\mathbf{M}))$ be the square matrix with $\text{vec}(\mathbf{M})$ on the diagonal. Using this notation, the asymptotic distributions of the forecast-error variance components are given in the next theorem.

**Theorem 4.2.** *Suppose* $\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma}_\theta)$ *and* $\boldsymbol{\lambda}_{22} = \mathbf{S}_\lambda \boldsymbol{\lambda}_s + \mathbf{s}_\lambda$. *Then,*

$$\sqrt{T}\text{vec}(\hat{\bar{\mathbf{W}}}_h - \bar{\mathbf{W}}_h) \xrightarrow{d} N(\mathbf{0}, \bar{\mathbf{G}}_{wah} \boldsymbol{\Sigma}_a \bar{\mathbf{G}}_{wah}' + \bar{\mathbf{G}}_{w\sigma h} \boldsymbol{\Sigma}_\sigma \bar{\mathbf{G}}_{w\sigma h}'),$$

*where*

$$\begin{aligned}\bar{\mathbf{G}}_{wah} = \frac{\partial \text{vec}(\bar{\mathbf{W}}_h)}{\partial \mathbf{a}'} &= 2(\mathbf{I}_{n_2} \otimes \mathbf{W}_{h\Psi}) \sum_{s=0}^{h-1} D_v(\bar{\mathbf{\Phi}}_s)\bar{\mathbf{G}}_{\Phi as} \\ &- 2(\bar{\mathbf{W}}'_{h\Phi}\mathbf{W}'_{h\Psi} \otimes \mathbf{W}_{h\Psi})D_v(\mathbf{I}_n)\mathbf{N}_n \sum_{s=0}^{h-1}(\mathbf{\Psi}_s\mathbf{\Sigma} \otimes \mathbf{I}_n)\mathbf{G}_{\Psi as} \text{ and} \\ \bar{\mathbf{G}}_{w\sigma h} = \frac{\partial \text{vec}(\bar{\mathbf{W}}_h)}{\partial \boldsymbol{\sigma}'} &= 2(\mathbf{I}_{n_2} \otimes \mathbf{W}_{h\Psi}) \sum_{s=0}^{h-1} D_v(\bar{\mathbf{\Phi}}_s)\bar{\mathbf{G}}_{\Phi\sigma s} \\ &- (\bar{\mathbf{W}}'_{h\Phi}\mathbf{W}'_{h\Psi} \otimes \mathbf{W}_{h\Psi})D_v(\mathbf{I}_n) \sum_{s=0}^{h-1}(\mathbf{\Psi}_s \otimes \mathbf{\Psi}_s)\mathbf{D}_n.\end{aligned}$$

*Proof.* Write $d\text{vec}(\bar{\mathbf{W}}_h) = (\bar{\mathbf{W}}'_{h\Phi} \otimes \mathbf{I}_n)d\text{vec}(\mathbf{W}_{h\Psi}) + (\mathbf{I}_{n_2} \otimes \mathbf{W}_{h\Psi})d\text{vec}(\bar{\mathbf{W}}_{h\Phi})$, where

$$\begin{aligned}d\text{vec}(\mathbf{W}_{h\Psi}) &= -(\mathbf{W}'_{h\Psi} \otimes \mathbf{W}_{h\Psi})d\text{vec}((\sum_{s=0}^{h-1} \mathbf{\Psi}_s\mathbf{\Sigma}\mathbf{\Psi}'_s) \odot \mathbf{I}_n) \\ &= -(\mathbf{W}'_{h\Psi} \otimes \mathbf{W}_{h\Psi})D_v(\mathbf{I}_n) \sum_{s=0}^{h-1} d\text{vec}(\mathbf{\Psi}_s\mathbf{\Sigma}\mathbf{\Psi}'_s) \\ &= -(\mathbf{W}'_{h\Psi} \otimes \mathbf{W}_{h\Psi})D_v(\mathbf{I}_n) \sum_{s=0}^{h-1} [2\mathbf{N}_n(\mathbf{\Psi}_s\mathbf{\Sigma} \otimes \mathbf{I}_n)\mathbf{G}_{\Psi as}d\mathbf{a} + (\mathbf{\Psi}_s \otimes \mathbf{\Psi}_s)\mathbf{D}_n d\boldsymbol{\sigma}] \text{ and} \\ d\text{vec}(\bar{\mathbf{W}}_{h\Phi}) &= \sum_{s=0}^{h-1} d\text{vec}(\bar{\mathbf{\Phi}}_s \odot \bar{\mathbf{\Phi}}_s) = \sum_{s=0}^{h-1} 2D_v(\bar{\mathbf{\Phi}}_s)d\text{vec}(\bar{\mathbf{\Phi}}_s) = \sum_{s=0}^{h-1} 2D_v(\bar{\mathbf{\Phi}}_s)(\bar{\mathbf{G}}_{\Phi as}d\mathbf{a} + \bar{\mathbf{G}}_{\Phi\sigma s}d\boldsymbol{\sigma}).\end{aligned}$$

The delta method after rearrangement completes the proof. ∎

Similar to Corollary 4.1, the asymptotic distribution of the forecast-error variance decomposition in Lütkepohl (1990) is a special case when the model is fully- and just-identified with recursive assumptions, as shown in the next corollary.

**Corollary 4.2.** *Suppose that the model is just-identified and $\mathbf{\Phi}_0$ is lower triangular. Then, the asymptotic distribution of the $(i, j)_{th}$ component of $\mathbf{W}_h$ follows*

$$\sqrt{T}(\hat{w}_{h,ij} - w_{h,ij}) \xrightarrow{d} N(0, \mathbf{g}_{wah,ij}\mathbf{\Sigma}_a\mathbf{g}'_{wah,ij} + \mathbf{g}_{w\sigma h,ij}\mathbf{\Sigma}_\sigma\mathbf{g}'_{w\sigma h,ij})$$

*which is equivalent to Lütkepohl's (1990) Proposition 1-(v), where*

$$\mathbf{g}_{wah,ij} = \begin{cases} \mathbf{0}, & h=1 \\ 2\sum_{s=1}^{h-1}[MSE_i(h)(\boldsymbol{\iota}_i'\boldsymbol{\Psi}_s\boldsymbol{\Phi}_0\boldsymbol{\iota}_j)(\boldsymbol{\iota}_j'\boldsymbol{\Phi}_0'\otimes\boldsymbol{\iota}_i')\mathbf{G}_{\Psi as} & \\ -(\boldsymbol{\iota}_i'\boldsymbol{\Psi}_s\boldsymbol{\Phi}_0\boldsymbol{\iota}_j)^2\sum_{m=1}^{h-1}(\boldsymbol{\iota}_i'\boldsymbol{\Psi}_m\boldsymbol{\Sigma}\otimes\boldsymbol{\iota}_i')\mathbf{G}_{\Psi am}]/MSE_i(h)^2, & h>1 \end{cases}$$

$$\mathbf{g}_{w\sigma h,ij} = \begin{array}{l} \sum_{s=0}^{h-1}[2MSE_i(h)(\boldsymbol{\iota}_i'\boldsymbol{\Psi}_s\boldsymbol{\Phi}_0\boldsymbol{\iota}_j)(\boldsymbol{\iota}_j'\otimes\boldsymbol{\iota}_i')\mathbf{G}_{\Phi\sigma s} \\ -(\boldsymbol{\iota}_i'\boldsymbol{\Psi}_s\boldsymbol{\Phi}_0\boldsymbol{\iota}_j)^2\sum_{m=0}^{h-1}(\boldsymbol{\iota}_i'\boldsymbol{\Psi}_m\otimes\boldsymbol{\iota}_i'\boldsymbol{\Psi}_m)\mathbf{D}_n]/MSE_i(h)^2, \qquad h\geq 1 \end{array}$$

*Proof.* Consider a fully identified model with $n = n_2$. From $\bar{w}_{h,ij} = \boldsymbol{\iota}_i'\bar{\mathbf{W}}_h\boldsymbol{\iota}_j = (\boldsymbol{\iota}_j'\otimes\boldsymbol{\iota}_i')\mathrm{vec}(\bar{\mathbf{W}}_h)$, it follows that

$$\begin{aligned} \mathbf{g}_{wah,ij} = (\boldsymbol{\iota}_j'\otimes\boldsymbol{\iota}_i')\mathbf{G}_{wah} &= 2(\boldsymbol{\iota}_j'\otimes\boldsymbol{\iota}_i')(\mathbf{I}_{n_2}\otimes\mathbf{W}_{h\Psi})\sum_{s=0}^{h-1}D_v(\bar{\boldsymbol{\Phi}}_s)\bar{\mathbf{G}}_{\Phi as} \\ &- 2(\boldsymbol{\iota}_j'\otimes\boldsymbol{\iota}_i')(\bar{\mathbf{W}}_{h\Phi}'\mathbf{W}_{h\Psi}'\otimes\mathbf{W}_{h\Psi})D_v(\mathbf{I}_n)\mathbf{N}_n\sum_{s=0}^{h-1}(\boldsymbol{\Psi}_s\boldsymbol{\Sigma}\otimes\mathbf{I}_n)\mathbf{G}_{\Psi as} \textit{ and} \\ \mathbf{g}_{w\sigma h,ij} = (\boldsymbol{\iota}_j'\otimes\boldsymbol{\iota}_i')\mathbf{G}_{w\sigma h} &= 2(\mathbf{I}_{n_2}\otimes\mathbf{W}_{h\Psi})\sum_{s=0}^{h-1}D_v(\bar{\boldsymbol{\Phi}}_s)\bar{\mathbf{G}}_{\Phi\sigma s} \\ &- (\bar{\mathbf{W}}_{h\Phi}'\mathbf{W}_{h\Psi}'\otimes\mathbf{W}_{h\Psi})D_v(\mathbf{I}_n)\sum_{s=0}^{h-1}(\boldsymbol{\Psi}_s\otimes\boldsymbol{\Psi}_s)\mathbf{D}_n. \end{aligned}$$

We can show that $(\boldsymbol{\iota}_j'\otimes\boldsymbol{\iota}_i')(\mathbf{I}_n\otimes\mathbf{W}_{h\Psi})\sum_{s=0}^{h-1}D_v(\bar{\boldsymbol{\Phi}}_s) = \sum_{s=0}^{h-1}\frac{\phi_{s,ij}}{MSE_i(h)}(\boldsymbol{\iota}_j'\otimes\boldsymbol{\iota}_i')$ and $(\boldsymbol{\iota}_j'\otimes\boldsymbol{\iota}_i')(\bar{\mathbf{W}}_{h\Phi}'\mathbf{W}_{h\Psi}'\otimes\mathbf{W}_{h\Psi})D_v(\mathbf{I}_n) = \sum_{s=0}^{h-1}\left[\sum_{m=0}^{h-1}\frac{\phi_{m,ij}^2}{MSE_m(h)^2}(\boldsymbol{\iota}_i'\otimes\boldsymbol{\iota}_i')\right]$. After incorporating the property of $(\boldsymbol{\iota}_i'\otimes\boldsymbol{\iota}_i')\mathbf{N}_n = (\boldsymbol{\iota}_i'\otimes\boldsymbol{\iota}_i')$, the delta method completes the proof.

∎

# 5 Application

To implement the generalized two-step ML estimation described in Section 2, we extend Bernanke and Mihov (1998) to investigate the effects of monetary policy shocks to exchange rates in an open economy. As Bernanke and Blinder (1992) proposed, consider a structural VAR model

$$\mathbf{B}_0\mathbf{y}_t = \sum_{i=1}^{p}\mathbf{B}_i\mathbf{y}_{t-i} + \mathbf{F}\mathbf{e}_t,$$

where $\mathbf{B}_0$ is lower block triangular and $\mathbf{F}$ is block diagonal as defined in Assumptions 2.1 and 2.2, respectively. We consider three sets of variables in $\mathbf{y}_t$: a set of non-policy variables, $\mathbf{y}_{1t}$, which are not affected by monetary policy shocks contemporaneously, a set of policy indicators, $\mathbf{y}_{2t}$, which describe the stance of the Fed's monetary policy, and a set of macroeconomic variables, $\mathbf{y}_{3t}$, which are influenced by monetary policy shocks contemporaneously. In our application, $\mathbf{y}_{1t}$ includes the industrial production index, consumer price index, and world commodity price index; $\mathbf{y}_{2t}$ includes total reserves, non-borrowed reserves, and federal funds rates, as chosen by Bernanke and Mihov (1998).[3] We extend the model by considering exchange rates and term structures in $\mathbf{y}_{3t}$. We may extend our method to an open economy model in which exchange rates and term structures are considered as policy indicators (see, e.g., Fung and Yuan, 1999). In such a case, the model falls into the same structure as Bernanke and Mihov (1998) because there are no variables in $\mathbf{y}_{3t}$. For methodological purposes, we consider exchange rates and term structures as non-policy variables to shed light on the generality of our model.

Under Assumption 2.1, consider the second block of the structural VAR model

$$\mathbf{B}_{22}\bar{\mathbf{M}}\mathbf{y}_t = \sum_{i=1}^{p} \bar{\mathbf{B}}_i \mathbf{y}_{t-i} + \mathbf{F}_{22}\mathbf{e}_{2t},$$

and describe the market for bank reserves using

$$\mathbf{B}_{22}\bar{\boldsymbol{\epsilon}}_{2t} = \mathbf{F}_{22}\mathbf{e}_{2t}, \tag{5.1}$$

where $\bar{\boldsymbol{\epsilon}}_{2t} = \bar{\mathbf{M}}\boldsymbol{\epsilon}_t$, which is the second set of innovations orthogonalized to the first set of innovations, $\boldsymbol{\epsilon}_{1t}$. Following Bernanke and Mihov (1998), we denote the innovation

[3]Bernanke and Mihov (1998) used real GDP and the GDP deflator at the monthly frequency using interpolation.

in the demand for total reserves $\epsilon_{TR}$, the innovation in the demand for borrowed reserves $\epsilon_{BR}$, the innovation in the demand for non-borrowed reserves $\epsilon_{NBR}$, and the innovation in the federal funds rate $\epsilon_{FFR}$. Bernanke and Mihov (1998) assumed that the market is described by the following set of equations:[4]

$$\begin{array}{lrcl} \text{(Demand for total reserves)} & \bar{\epsilon}_{TR} & = & -\alpha\bar{\epsilon}_{FFR} + \eta_d e_d \\ \text{(Demand for borrowed reserves)} & \bar{\epsilon}_{BR} & = & \beta\bar{\epsilon}_{FFR} + \eta_b e_b \\ \text{(Demand for nonborrowed reserves)} & \bar{\epsilon}_{NBR} & = & \phi_d e_d + \phi_b e_b + \eta_s e_s, \end{array}$$

where $e_d$ is a demand disturbance, $e_b$ is a disturbance to the borrowing function, and $e_s$ is the shock to monetary policy that we want to identify. Note that $\bar{\epsilon}_{BR} = \bar{\epsilon}_{TR} - \bar{\epsilon}_{NBR}$ and all the equations are expressed in orthogonalized innovation forms. For ML estimation, write the equations in the form of (5.1)

$$\begin{bmatrix} 1 & 0 & \alpha \\ 0 & 1 & 0 \\ 1 & -1 & -\beta \end{bmatrix} \begin{bmatrix} \bar{\epsilon}_{TR} \\ \bar{\epsilon}_{NBR} \\ \bar{\epsilon}_{FFR} \end{bmatrix} = \begin{bmatrix} \eta_d & 0 & 0 \\ \phi_d & \eta_s & \phi_b \\ 0 & 0 & \eta_b \end{bmatrix} \begin{bmatrix} e_d \\ e_s \\ e_b \end{bmatrix}.$$

The model is unidentified because seven unknowns need to be estimated from six relations in $\mathbf{B}_{22}\bar{\mathbf{M}}\mathbf{\Sigma}\bar{\mathbf{M}}'\mathbf{B}_{22}' = \mathbf{F}_{22}\mathbf{F}_{22}'$. Following Bernanke and Mihov (1998), we consider four alternative models regarding restrictions on the monetary policy shock

$$e_s = \frac{1}{\eta_s}\left[-(\frac{\phi_d}{\eta_d} + \frac{\phi_b}{\eta_b})\bar{\epsilon}_{TR} + (1 + \frac{\phi_b}{\eta_b})\bar{\epsilon}_{NBR} - (\alpha\frac{\phi_d}{\eta_d} - \beta\frac{\phi_b}{\eta_b})\bar{\epsilon}_{FFR}\right]. \tag{5.2}$$

The alternative models are summarized as follows:

i) Bernanke and Blinder (1992) model (BB): $\frac{\phi_d}{\eta_d} = 1$, $\frac{\phi_b}{\eta_b} = -1$, $e_s = -\frac{1}{\eta_s}(\alpha + \beta)\bar{\epsilon}_{FFR}$;

ii) Christiano and Eichenbaum (1991) model (CE): $\frac{\phi_d}{\eta_d} = 0$, $\frac{\phi_b}{\eta_b} = 0$, $e_s = \frac{1}{\eta_s}\bar{\epsilon}_{NBR}$;

iii) Strongin (1995) model (ST): $\alpha = 0$, $\frac{\phi_b}{\eta_b} = 0$, $e_s = -\frac{1}{\eta_s}(\frac{\phi_d}{\eta_d}\bar{\epsilon}_{TR} + \bar{\epsilon}_{NBR})$; and

[4] We consider a simplified model in which the innovation to the discount rate is zero.

iv) Just-identified model (JI): $\alpha = 0$, $e_s = \frac{1}{\eta_s}\left[-(\frac{\phi_d}{\eta_d} + \frac{\phi_b}{\eta_b})\bar{\epsilon}_{TR} + (1 + \frac{\phi_b}{\eta_b})\bar{\epsilon}_{NBR} + \beta\frac{\phi_b}{\eta_b}\bar{\epsilon}_{FFR}\right]$. See Bernanke and Mihov (1998) for details.

The unrestricted VAR model is estimated over the sample period from January 1970 to June 2001 using monthly data obtained from the Federal Reserve Bank of St. Louis. The industrial production index, the consumer price index, the world commodity price index, and exchange rates (U.S/Canada) are taken as differences using logarithms. The total and non-borrowed reserves are normalized using the 36-month moving average of the total reserves. The ten-year treasury bill rate less three-month treasury bill rate is used for the term structure. Time plots of data for levels and differences are shown in Figure 4.1. See Appendix C for a description of the data. We choose 12 months as the lag length, although choosing shorter lag lengths does not alter our main results.

The second-step ML estimates are given in Table 4.1. We begin with the over-identification tests in Panel A. Bernanke-Blinder model and Strongin model are not rejected at the 5% significance level, while Christiano-Eichenbaum model is rejected at the 1% significance level. The ML estimates of $\mathbf{B}_{22}$ and $\mathbf{F}_{22}$ for the alternative models are listed in Panel B. All the free parameters are significantly different from zero at the 5% significance level. The long-run neutrality of money can be tested using the estimates of $\mathbf{\Phi}(1)_{12}$ in the fifth column of Panel E. The long-run neutrality of money is not rejected for the Bernanke-Blinder, Christiano-Eichenbaum, or just-identified models, while it is rejected for the Strongin model at the 5% significance level.

Figure 4.2 shows the estimated dynamic responses of macroeconomic variables and policy indicators to expansionary monetary policy shocks for the alternative models. These results are robust for the model selected. The dynamic responses of output have the usual humped shape. Output increases in the short run and starts to decrease after one or two years. The effects vanish in the long run, implying the long-run neutrality of money. The dynamics of price are in sharp contrast with the results of Bernanke and Mihov (1998). The results are subject to the 'price puzzle', in which an expansionary monetary policy shock is followed by a subsequent fall in price as pointed out by Sims (1992). The fall in price is significant for the first nine months and is insignificant in the long horizon in the Bernanke-Blinder and just-identified models, while it is insignificant for every horizon in the Christiano-Eichenbaum and Strongin models. Although the world commodity price index is incorporated in the Fed's information set, as suggested by Sims (1992) and Leeper, Sims, and Zha (1996), the price puzzle does not disappear when the industrial production index and the consumer price index are used to measure output and overall price, respectively.[5] The dynamic responses of total reserves, non-borrowed reserves, and federal funds rates show the liquidity effects in which expansionary monetary policy shocks are accompanied by an increase in non-borrowed reserves and a fall in federal funds rates. Exchange rates exhibit over-shooting behavior in the Bernanke-Blinder and the just-identified models. The dynamic responses of exchange rates are in sharp contrast with Eichenbaum and Evans (1995), who found such evidence only with a twenty-month delay. In addition, Jang and Ogaki (2004) and Kalyvitis and Michaelides (2001) gave evidence for instantaneous overshooting. The Christiano-Eichenbaum and Strongin models yield

[5] The price puzzle disappears when we use the quarterly real GDP and the GDP deflator for the measure of output and overall price. As Bernanke and Mihov (1995) noted, it is "difficult to defend [applying] the identification assumption of no feedback from policy to" the non-policy variables in $\mathbf{y}_{1t}$ at a quarterly frequency. One solution is to use interpolated monthly GDP data, as suggested by Bernanke and Mihov (1995).

the depreciation of the U.S. dollar after an expansionary monetary policy shock, but the dynamic responses do not exhibit the overshooting behavior and are insignificant over all time horizons. The dynamic responses of term structures show a humped shape. An expansionary monetary policy shock yields an increase in term structures for the first twenty months in the Bernanke-Blinder and just-identified models, while the responses are insignificant in the Christiano-Eichenbaum and Strongin models. The increase in term structures due to an expansionary policy shock is consistent with the literature. See Evans and Marshall (1997) for an example.

Table 4.2 shows the forecast-error variance decompositions in the four alternative models. The policy indicator that includes the largest fraction of the forecast-error variance attributed to the monetary policy shock varies across the four alternative models, as implied by (5.2). Note that 98% of the federal funds rates in the Bernanke-Blinder model, 96% of the non-borrowed reserves in the Christiano-Eichenbaum model, and 77% of the non-borrowed reserves in the Strongin model are attributed to the monetary policy shock in the first month. In the just-identified model, 55% of the non-borrowed reserves and 61% of the federal funds rates are attributed to the monetary policy shock. The fraction of the exchange rates forecast-error variance attributed to the monetary policy shock is relatively small, ranging from 1% to 3% in the first month. The fraction after six months in the Bernanke-Blinder model is at most 7%. As Faust and Rogers (2000) pointed out, it is not attractive to exclude the exchange rates from the Fed's information set. This result is consistent with Jang and Ogaki (2004), who found that the fraction is relatively small in recursive VAR models, while it is relatively large in vector error-correction models with long-run restrictions.

# 6 Concluding Remarks

This paper generalizes the existing VAR literature. First, it generalizes Giannini (1992) to consider VAR models that are not necessarily fully identified. It shows that partially identified models can be estimated using generalized two-step ML estimation with a transformation matrix that diagonalizes the model. Second, generalizing Lütkepohl (1990), this paper also derives the asymptotic distributions of impulse responses and forecast-error variance decomposition of general classes of VAR models. In particular, it shows that the result of Lütkepohl (1990) is a special case when the model is fully- and just-identified with recursive assumptions. Finally, as an application, we extend Bernanke and Mihov (1998) to an open economy. We find that exchange rates tend to overshoot and term structures show hump-shaped responses to monetary policy shocks. One possible extension of this paper would be two-step ML estimation of partially identified models with both short- and long-run restrictions.

## References


Baillie, R. T. (1987): "Inference in Dynamic Models Containing 'Surprise' Variables," *Journal of Econometrics*, 35, 101–117.

Bernanke, B. S. (1986): "Alternative Explanations of the Money-Income Correlation," *Carnegie-Rochester Conference Series on Public Policy*, 25, 49–100.

Bernanke, B. S., and A. S. Blinder (1992): "The Federal Funds Rate and the Channels of Monetary Transmission," *American Economic Review*, 82(4), 901–921.

Bernanke, B. S., and I. Mihov (1995): "Measuring Monetary Policy," NBER Working Paper No. 5145.

——— (1998): "Measuring Monetary Policy," *Quarterly Journal of Economics*, 113(3), 869–902.

Blanchard, O. J. (1989): "A Traditional Interpretation of Macroeconomic Fluctuations," *American Economic Review*, 79(5), 1146–1164.

Blanchard, O. J., and D. Quah (1989): "The Dynamic Effects of Aggregate Supply and Demand Disturbances," *American Economic Review*, 77, 655–673.

Blanchard, O. J., and M. W. Watson (1986): "Are Business Cycles All Alike?," in *The American Business Cycle: Continuity and Change*, ed. by R. J. Gordon, vol. 25 of *National Bureau of Economic Research Studies in Business Cycles*, pp. 123–182. University of Chicago Press, Chicago.

Christiano, L. J., and M. Eichenbaum (1991): "Identification and the Liquidity Effect of a Monetary Policy Shock," NBER Working Paper No. 3920.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (1999): "Monetary Policy Shocks: What Have We Learned and to What End?," in *Handbook of Macroeconomics*, ed. by J. Taylor, and M. Woodford, vol. 1, chap. 2, pp. 65–148. Elsevier Science.

Eichenbaum, M., and C. L. Evans (1995): "Some Empirical Evidence on the Effects of Shocks to Monetary Policy on Exchange Rate," *Quarterly Journal of Economics*, 110, 975–1009.

Evans, C. L., and D. A. Marshall (1997): "Monetary Policy and the Term Structure of Nominal Interest Rates: Evidence and Theory," Federal Reserve Bank of Chicago, Working Paper, Macroeconomic Issues No. WP-97-10.

Faust, J., and J. H. Rogers (2000): "Monetary Policy's Role in Exchange Rate Behavior," Manuscript.

Fung, B. S., and M. Yuan (1999): "Measuring the Stance of Monetary Policy," Money, Monetary Policy, and Transmission Mechanisms, Proceedings of a conference held by the Bank of Canada.

Gali, J. (1992): "How Well Does the IS-LM Model Fit Postwar U.S. Data?," *Quarterly Journal of Economics*, 107(2), 709–738.

Giannini, C. (1992): *Topics in Structural VAR Econometrics.* Apringer Verlag, New York.

Hamilton, J. D. (1994): *Time Series Analysis.* Princeton University Press, Princeton.

Jang, K., and M. Ogaki (2004): "The Effects of Monetary Policy Shocks on Exchange Rates: A Structural Vector Error Correction Model Approach," *Journal of the Japanese and International Economies*, 18(1), 99–114.

Kalyvitis, S., and A. Michaelides (2001): "New Evidence on the Effects of US Monetary Policy on Exchange Rates," *Economics Letters*, 71(2), 255–263.

Keating, J. W. (1999): "Structural Inference With Long-Run Recursive Empirical Models," Working Paper No. 1999-3, University of Kansas, forthcoming in *Macroeconomic Dynamics.*

King, R. G., C. I. Plosser, J. H. Stock, and M. W. Watson (1991): "Stochastic Trends and Economic Fluctuations," *American Economic Review*, 81(4), 810–840.

Leeper, E. M., C. A. Sims, and T. Zha (1996): "What Dose Monetary Policy Do?," *Brookings Papers on Economic Activity*, 2, 1–78.

Lütkepohl, H. (1990): "Asymptotic Distributions of Impulse Response Functions and Forecast Error Variance Decompositions of Vector Autoregressive Models," *Review of Economics and Statistics*, 72(1), 116–125.

Magnus, J. R. (1988): *Linear Structures.* Charles Griffin & Company Limited, London, U.K.

Runkle, D. E. (1987): "Vector Autoregressions and Reality," *Journal of Business and Economic Statistics*, 5, 437–442.

Sims, C. A. (1980): "Macroeconomics and Reality," *Econometrica*, 48, 1–48.

——— (1992): "Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy," *European Economic Review*, 36, 975–1000.

Strongin, S. (1995): "The Identification of Monetary Policy Disturbances Explaining the Liquidity Puzzle," *Journal of Monetary Economics*, 35, 463–497.

# Appendix

## A Definitions and properties

We follow the definitions used by Magnus (1988).

**Definition A.1.** *(The commutation matrix)* $\mathbf{K}_{mn} vec(\mathbf{A}) = vec(\mathbf{A}')$ *for any* $m \times n$ *matrix* $\mathbf{A}$.

**Definition A.2.** *(The matrix* $\mathbf{N}_n$*)* $\mathbf{N}_n \text{vec}(\mathbf{A}) = \frac{1}{2}\text{vec}(\mathbf{A} + \mathbf{A}')$ *for any* $n \times n$ *matrix* $\mathbf{A}$.

**Definition A.3.** *(The duplication matrix* $\mathbf{D}_n$*)* $\mathbf{D}_n vech(\mathbf{A}) = vec(\mathbf{A})$ *for any symmetric* $n \times n$ *matrix* $\mathbf{A}$.

**Definition A.4.** *(The elimination matrix* $\mathbf{L}_n$*)* $\mathbf{L}_n vec(\mathbf{A}) = vech(\mathbf{A})$ *for any lower triangular* $n \times n$ *matrix* $\mathbf{A}$.

**Definition A.5.** *(The Moore-Penrose inverse matrix* $\mathbf{A}^+$*) An* $n \times m$ *matrix* $\mathbf{A}^+$ *is the Moore-Penros inverse of a real* $m \times n$ *matrix if* $\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}, \mathbf{A}^+\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+, (\mathbf{A}\mathbf{A}^+)' = \mathbf{A}\mathbf{A}^+$*, and* $(\mathbf{A}^+\mathbf{A})' = \mathbf{A}^+\mathbf{A}$.

**Definition A.6.** *(The Hadamard operator* $\odot$*)* $\mathbf{A} \odot \mathbf{B} = [a_{ij}b_{ij}]$ *for any* $m \times n$ *matrices* $\mathbf{A}$ *and* $\mathbf{B}$.

**Property A.1.** *(The commutation property)*

*(i)* $\mathbf{K}'_{mn} = \mathbf{K}^{-1}_{mn} = \mathbf{K}_{nm}$.

*(ii)* $\mathbf{K}_{pm}(\mathbf{A} \otimes \mathbf{B}) = (\mathbf{B} \otimes \mathbf{A})\mathbf{K}_{qn}$ *for any* $m \times n$ *matrix* $\mathbf{A}$ *and* $p \times q$ *matrix* $\mathbf{B}$.

*(iii)* $\mathbf{K}_{n1} = \mathbf{K}_{1n} = \mathbf{I}_n$.

**Property A.2.** *(The property of* $\mathbf{N}_n$*)*

*(i)* $\mathbf{N}_n = \frac{1}{2}(\mathbf{I}_{n^2} + \mathbf{K}_{nn})$.

*(ii)* $\mathbf{N}_n = \mathbf{N}'_n = \mathbf{N}^2_n$.

*(iii)* $\mathbf{N}_n\mathbf{K}_{nn} = \mathbf{N}_n = \mathbf{K}_{nn}\mathbf{N}_n$.

*(iv)* $\mathbf{N}_n(\mathbf{A} \otimes \mathbf{B})\mathbf{N}_n = \mathbf{N}_n(\mathbf{B} \otimes \mathbf{A})\mathbf{N}_n$.

*(v)* $\mathbf{N}_n(\mathbf{A} \otimes \mathbf{A})\mathbf{N}_n = \mathbf{N}_n(\mathbf{A} \otimes \mathbf{A}) = (\mathbf{A} \otimes \mathbf{A})\mathbf{N}_n$.

**Property A.3.** *(The duplication property)*

*(i)* $\mathbf{K}_{nn}\mathbf{D}_n = \mathbf{D}_n = \mathbf{N}_n\mathbf{D}_n$.

*(ii)* $\mathbf{D}_n^+\mathbf{K}_{nn} = \mathbf{D}_n^+ = \mathbf{D}_n^+\mathbf{N}_n$.

*(iii)* $\mathbf{D}_n\mathbf{D}_n^+ = \mathbf{N}_n$.

**Property A.4.** *(The elimination property)*

*(i)* $\mathbf{L}_n\mathbf{L}_n' = \mathbf{I}_{\frac{n(n+1)}{2}}$.

*(ii)* $\mathbf{L}_n^+ = \mathbf{L}_n'$.

*(iii)* $\mathbf{L}_n\mathbf{D}_n = \mathbf{I}_{\frac{n(n+1)}{2}}$.

*(iv)* $\mathbf{D}_n\mathbf{L}_n\mathbf{N}_n = \mathbf{N}_n$.

*(v)* $\mathbf{D}_n^+ = \mathbf{L}_n\mathbf{N}_n$.

**Property A.5.** *(The Moore-Penrose inverse matrix)*

*(i)* $\mathbf{A}^+ = (\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$ *if* $\mathbf{A}$ *has full-column rank.*

*(ii)* $\mathbf{A}^+ = \mathbf{A}'(\mathbf{A}\mathbf{A}')^{-1}$ *if* $\mathbf{A}$ *has full-row rank.*

**Property A.6.** *(The Hadamard operator)*

*(i)* $\mathbf{A} \odot \mathbf{B} = \mathbf{B} \odot \mathbf{A}$.

*(ii)* $\mathbf{A} \odot \mathbf{I}_n = diag(\mathbf{A})$.

*(iii)* $\text{vec}(\mathbf{A} \odot \mathbf{B}) = \text{vec}(\mathbf{A}) \odot \text{vec}(\mathbf{B}) = D_v(\mathbf{A})\text{vec}(\mathbf{B}) = D_v(\mathbf{B})\text{vec}(\mathbf{A})$.

## B Proof of Lemma

### Lemma 2.1

*Proof.* Each block of the reduced-form VAR model becomes mutually orthogonalized when the transformation matrix $\mathbf{M}$ is multiplied. Therefore, we can concentrate on the second block of the model to estimate $\mathbf{B}_{22}$ and $\mathbf{F}_{22}$ from $\mathbf{B}_{22}\bar{\mathbf{M}}\boldsymbol{\Sigma}\bar{\mathbf{M}}'\mathbf{B}_{22}' = \mathbf{F}_{22}\mathbf{F}_{22}'$, where $\bar{\mathbf{M}} = [-\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} \vdots \mathbf{I}_{n_2} \vdots \mathbf{0}]$. Therefore, $\mathbf{B}_{22}$ and $\mathbf{F}_{22}$ are two-step ML estimable. Finally, $\mathbf{B}_{21}$ is obtained by $\mathbf{B}_{21} = \mathbf{B}_{22}\mathbf{M}_{21}$. ∎

### Lemma 3.1

*Proof.* (a) Write $\text{vec}(\boldsymbol{\Sigma}) = \text{vec}(\boldsymbol{\Sigma}_{\cdot 1}, \boldsymbol{\Sigma}_{\cdot 2}, \boldsymbol{\Sigma}_{\cdot 3}) = [\text{vec}(\boldsymbol{\Sigma}_{\cdot 1})', \text{vec}(\boldsymbol{\Sigma}_{\cdot 2})', \text{vec}(\boldsymbol{\Sigma}_{\cdot 3})']'$, where

$$\text{vec}(\boldsymbol{\Sigma}_{\cdot 1}) = \mathbf{K}_{n_1 n}\text{vec}(\boldsymbol{\Sigma}_{\cdot 1}') = \mathbf{K}_{n_1 n}\begin{bmatrix} \text{vec}(\boldsymbol{\Sigma}_{11}) \\ \text{vec}(\boldsymbol{\Sigma}_{21}') \\ \text{vec}(\boldsymbol{\Sigma}_{31}') \end{bmatrix} = \mathbf{K}_{n_1 n}\begin{bmatrix} \mathbf{D}_{n_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{n_2 n_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{K}_{n_3 n_1} \end{bmatrix}\begin{bmatrix} \boldsymbol{\sigma}_{11} \\ \boldsymbol{\sigma}_{21} \\ \boldsymbol{\sigma}_{31} \end{bmatrix}.$$

Therefore,

$$\begin{bmatrix} \boldsymbol{\sigma}_{11} \\ \boldsymbol{\sigma}_{21} \\ \boldsymbol{\sigma}_{31} \end{bmatrix} = \begin{bmatrix} \mathbf{D}_{n_1}^{+} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{n_1 n_2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{K}_{n_1 n_3} \end{bmatrix}\mathbf{K}_{n n_1}\text{vec}(\boldsymbol{\Sigma}_{\cdot 1}).$$

Similarly,

$$\begin{bmatrix} \boldsymbol{\sigma}_{21} \\ \boldsymbol{\sigma}_{22} \\ \boldsymbol{\sigma}_{32} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{n_2 n_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{n_2}^{+} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{K}_{n_2 n_3} \end{bmatrix}\mathbf{K}_{n n_2}\text{vec}(\boldsymbol{\Sigma}_{\cdot 2}) \quad \text{and} \quad \begin{bmatrix} \boldsymbol{\sigma}_{31} \\ \boldsymbol{\sigma}_{32} \\ \boldsymbol{\sigma}_{33} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{n_3 n_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n_3 n_2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{D}_{n_3} \end{bmatrix}\mathbf{K}_{n n_3}\text{vec}(\boldsymbol{\Sigma}_{\cdot 3}).$$

Therefore, from

$$\begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix}\begin{bmatrix} \boldsymbol{\sigma}_{11} \\ \boldsymbol{\sigma}_{21} \\ \boldsymbol{\sigma}_{31} \\ \boldsymbol{\sigma}_{22} \\ \boldsymbol{\sigma}_{32} \\ \boldsymbol{\sigma}_{33} \end{bmatrix} = \begin{bmatrix} \mathbf{D}_{n_1}^{+} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{n_1 n_2} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{K}_{n_1 n_3} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{n_2 n_1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_{n_2}^{+} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{K}_{n_2 n_3} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{n_3 n_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{n_3 n_2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_{n_3}^{+} \end{bmatrix}\begin{bmatrix} \mathbf{K}_{n n_1}\text{vec}(\boldsymbol{\Sigma}_{\cdot 1}) \\ \mathbf{K}_{n n_2}\text{vec}(\boldsymbol{\Sigma}_{\cdot 2}) \\ \mathbf{K}_{n n_3}\text{vec}(\boldsymbol{\Sigma}_{\cdot 3}) \end{bmatrix}$$

it follows that

$$d\boldsymbol{\sigma}^* = \begin{bmatrix} \mathbf{D}_{n_1}^{+} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{2}\mathbf{K}_{n_1 n_2} & \mathbf{0} & \frac{1}{2}\mathbf{I}_{n_2 n_1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{1}{2}\mathbf{K}_{n_1 n_3} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \frac{1}{2}\mathbf{I}_{n_3 n_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_{n_2}^{+} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \frac{1}{2}\mathbf{K}_{n_2 n_3} & \mathbf{0} & \frac{1}{2}\mathbf{I}_{n_3 n_2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_{n_3}^{+} \end{bmatrix}\begin{bmatrix} \mathbf{K}_{n n_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{n n_2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{K}_{n n_3} \end{bmatrix}\mathbf{D}_n d\boldsymbol{\sigma}.$$

(b) Write $d\text{vec}(\bar{\boldsymbol{\Sigma}}) = d\text{vec}(\boldsymbol{\Sigma}_{22} - \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{21}')$, where $d\text{vec}(\bar{\boldsymbol{\Sigma}}) = \mathbf{D}_{n_2} d\bar{\boldsymbol{\sigma}}$ and $d\text{vec}(\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{21}') = -(\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} \otimes \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1})\mathbf{D}_{n_1} d\boldsymbol{\sigma}_{11} + 2\mathbf{N}_{n_2}(\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} \otimes \mathbf{I}_{n_2}) d\boldsymbol{\sigma}_{21}$. Therefore,

$$\begin{aligned} d\bar{\boldsymbol{\sigma}} &= \mathbf{D}_{n_2}^{+}(\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} \otimes \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1})\mathbf{D}_{n_1} d\boldsymbol{\sigma}_{11} - 2\mathbf{D}_{n_2}^{+}(\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} \otimes \mathbf{I}_{n_2}) d\boldsymbol{\sigma}_{21} + d\boldsymbol{\sigma}_{22} \\ &= \left[\mathbf{D}_{n_2}^{+}(\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} \otimes \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1})\mathbf{D}_{n_1} \vdots -2\mathbf{D}_{n_2}^{+}(\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} \otimes \mathbf{I}_{n_2}) \vdots \mathbf{0} \vdots \mathbf{I}_{\frac{n_2(n_2+1)}{2}} \vdots \mathbf{0} \vdots \mathbf{0}\right] d\boldsymbol{\sigma}^*. \end{aligned}$$

(c) It follows from $d\text{vec}(\tilde{\boldsymbol{\Sigma}}) = d\text{vec}(\boldsymbol{\Sigma}_{32} - \boldsymbol{\Sigma}_{31}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{21}')$ that

$$\begin{aligned} d\tilde{\boldsymbol{\sigma}} &= (\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} \otimes \boldsymbol{\Sigma}_{31}\boldsymbol{\Sigma}_{11}^{-1})\mathbf{D}_{n_1} d\boldsymbol{\sigma}_{11} - \mathbf{K}_{n_2n_3}(\boldsymbol{\Sigma}_{31}\boldsymbol{\Sigma}_{11}^{-1} \otimes \mathbf{I}_{n_2}) d\boldsymbol{\sigma}_{21} - (\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} \otimes \mathbf{I}_{n_3}) d\boldsymbol{\sigma}_{31} + d\boldsymbol{\sigma}_{32} \\ &= \left[(\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} \otimes \boldsymbol{\Sigma}_{31}\boldsymbol{\Sigma}_{11}^{-1})\mathbf{D}_{n_1} \vdots -\mathbf{K}_{n_2n_3}(\boldsymbol{\Sigma}_{31}\boldsymbol{\Sigma}_{11}^{-1} \otimes \mathbf{I}_{n_2}) \vdots -\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1} \otimes \mathbf{I}_{n_3} \vdots \mathbf{0} \vdots \mathbf{I}_{n_3n_2} \vdots \mathbf{0}\right] d\boldsymbol{\sigma}^*. \end{aligned}$$

(d) Write $d\text{vec}(\bar{\mathbf{M}}) = \left[-d\text{vec}(\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1})', \mathbf{0}, \mathbf{0}\right]'$, where $d\text{vec}(\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}) = -(\boldsymbol{\Sigma}_{11}^{-1} \otimes \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1})\mathbf{D}_{n_1} d\boldsymbol{\sigma}_{11} + (\boldsymbol{\Sigma}_{11}^{-1} \otimes \mathbf{I}_{n_2}) d\boldsymbol{\sigma}_{21}$. Therefore,

$$d\bar{\mathbf{m}} = \begin{bmatrix} (\boldsymbol{\Sigma}_{11}^{-1} \otimes \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1})\mathbf{D}_{n_1} & -\boldsymbol{\Sigma}_{11}^{-1} \otimes \mathbf{I}_{n_2} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0}_{n_2^2 \times \frac{n_1(n_1+1)}{2}} & \mathbf{0}_{n_2^2 \times n_2n_1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0}_{n_2n_3 \times \frac{n_1(n_1+1)}{2}} & \mathbf{0}_{n_2n_3 \times n_2n_1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} d\boldsymbol{\sigma}^*.$$

(e) Write $d\text{vec}(\tilde{\mathbf{M}}) = \mathbf{K}_{n_2n} d\text{vec}(\tilde{\mathbf{M}}')$, where

$$d\text{vec}(\tilde{\mathbf{M}}') = \begin{bmatrix} \mathbf{0}_{n_1n_2 \times 1} \\ \mathbf{0}_{n_2^2 \times 1} \\ \mathbf{K}_{n_3n_2} d\text{vec}(\tilde{\boldsymbol{\Sigma}}\bar{\boldsymbol{\Sigma}}^{-1}) \end{bmatrix}.$$

It follows from (c), (d), and $d\text{vec}(\tilde{\boldsymbol{\Sigma}}\bar{\boldsymbol{\Sigma}}^{-1}) = (\bar{\boldsymbol{\Sigma}}^{-1} \otimes \mathbf{I}_{n_3}) d\tilde{\boldsymbol{\sigma}} - (\bar{\boldsymbol{\Sigma}}^{-1} \otimes \tilde{\boldsymbol{\Sigma}}\bar{\boldsymbol{\Sigma}}^{-1})\mathbf{D}_{n_2} d\bar{\boldsymbol{\sigma}}$ that

$$d\tilde{\mathbf{m}} = \mathbf{K}_{n_2n} \begin{bmatrix} \mathbf{0}_{n_1n_2 \times \frac{n(n+1)}{2}} \\ \mathbf{0}_{n_2^2 \times \frac{n(n+1)}{2}} \\ \mathbf{K}_{n_3n_2}\left[(\bar{\boldsymbol{\Sigma}}^{-1} \otimes \mathbf{I}_{n_3})\tilde{\mathbf{G}}_{\sigma\sigma^*} - (\bar{\boldsymbol{\Sigma}}^{-1} \otimes \tilde{\boldsymbol{\Sigma}}\bar{\boldsymbol{\Sigma}}^{-1})\mathbf{D}_{n_2}\bar{\mathbf{G}}_{\sigma\sigma^*}\right] \end{bmatrix} d\boldsymbol{\sigma}^*.$$

The delta method completes the proof. ∎

**Lemma 3.2**

*Proof.* (a) Write $d\text{vec}(\mathbf{K}_{22}) = d\text{vec}(\mathbf{F}_{22}^{-1}\mathbf{B}_{22}) = (\mathbf{I}_{n_2} \otimes \mathbf{F}_{22}^{-1}) d\text{vec}(\mathbf{B}_{22}) + (\mathbf{B}_{22}' \otimes \mathbf{I}_{n_2}) d\text{vec}(\mathbf{F}_{22}^{-1})$, where $d\text{vec}(\mathbf{B}_{22}) = \mathbf{S}_b d\mathbf{b}_s$ and $d\text{vec}(\mathbf{F}_{22}^{-1}) = -(\mathbf{F}_{22}'^{-1} \otimes \mathbf{F}_{22}^{-1})\mathbf{S}_f d\mathbf{f}_s$. Therefore,

$$\begin{aligned} d\mathbf{k}_{22} &= \left[\mathbf{I}_{n_2} \otimes \mathbf{F}_{22}^{-1} \vdots -\mathbf{B}_{22}'\mathbf{F}_{22}'^{-1} \otimes \mathbf{F}_{22}^{-1}\right] \begin{bmatrix} \mathbf{S}_b & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_f \end{bmatrix} \begin{bmatrix} d\mathbf{b}_s \\ d\mathbf{f}_s \end{bmatrix} \\ &= \left[\mathbf{I}_{n_2} \otimes \mathbf{F}_{22}^{-1} \vdots -\mathbf{B}_{22}'\mathbf{F}_{22}'^{-1} \otimes \mathbf{F}_{22}^{-1}\right] \mathbf{S}_\lambda d\boldsymbol{\lambda}_s. \end{aligned}$$

(b) It follows from $\boldsymbol{\Phi}_{22} = \mathbf{K}_{22}^{-1}$ that

$$\begin{aligned} d\boldsymbol{\phi}_{22} &= -(\mathbf{K}_{22}'^{-1} \otimes \mathbf{K}_{22}^{-1}) d\mathbf{k}_{22} \\ &= \left[ -\mathbf{F}_{22}'\mathbf{B}_{22}'^{-1} \otimes \mathbf{B}_{22}^{-1} \,\vdots\, \mathbf{I}_{n_2} \otimes \mathbf{B}_{22}^{-1} \right] \mathbf{S}_\lambda d\boldsymbol{\lambda}_s. \end{aligned}$$

(c) Write $d\text{vec}(\mathbf{B}_{22}^{-1}\mathbf{F}_{22}\mathbf{F}_{22}'\mathbf{B}_{22}'^{-1}) = d\text{vec}(\boldsymbol{\Phi}_{22}\boldsymbol{\Phi}_{22}')$, where $d\text{vec}(\boldsymbol{\Phi}_{22}\boldsymbol{\Phi}_{22}') = (\boldsymbol{\Phi}_{22} \otimes \mathbf{I}_{n_2})d\boldsymbol{\phi}_{22} + (\mathbf{I}_{n_2} \otimes \boldsymbol{\Phi}_{22})\mathbf{K}_{n_2 n_2} d\boldsymbol{\phi}_{22} = 2\mathbf{N}_{n_2}(\boldsymbol{\Phi}_{22} \otimes \mathbf{I}_{n_2})d\boldsymbol{\phi}_{22}$. Therefore,

$$\begin{aligned} d\text{vech}(\mathbf{B}_{22}^{-1}\mathbf{F}_{22}\mathbf{F}_{22}'\mathbf{B}_{22}'^{-1}) &= 2\mathbf{D}_{n_2}^{+}(\boldsymbol{\Phi}_{22} \otimes \mathbf{I}_{n_2})\mathbf{G}_{\phi\lambda} d\boldsymbol{\lambda}_s \\ &= 2\mathbf{D}_{n_2}^{+} \left[ -\mathbf{B}_{22}^{-1}\mathbf{F}_{22}\mathbf{F}_{22}'\mathbf{B}_{22}'^{-1} \otimes \mathbf{B}_{22}^{-1} \,\vdots\, \mathbf{B}_{22}^{-1}\mathbf{F}_{22} \otimes \mathbf{B}_{22}^{-1} \right] \mathbf{S}_\lambda d\boldsymbol{\lambda}_s. \end{aligned}$$

The delta method completes the proof. ∎

# C Data

Monthly data from January 1970 to June 2001 are used. The world price index was obtained from the International Financial Statistics CD-ROM and website. Other data were obtained from the Federal Reserve Bank of St. Louis.

- y: U.S. industrial production index (1997 =100). Seasonally adjusted. Log difference × 1200.
- p: U.S. consumer price index for all urban consumers. All items (1982–84=100). Log difference × 1200.
- pc: The world non-fuel primary commodities price index (1995=100). The world price index from January 1980 to June 2001 was obtained from the IFS website (series 00176NFDZF). The world price index from January 1970 to December 1979 was constructed by backward recursion using the growth rate of series 00176AXDZF from the IFS CD-ROM. Log difference × 1200.
- TR: Board of governors' total reserves (billions). Adjusted for changes in reserve requirements. Normalized using the 36-month moving average of total reserves.
- NBR: Non-borrowed reserves of depository institutions (billions). Normalized using the 36-month moving average of total reserves.
- FFR: Effective federal funds rates (%).
- er: U.S./Canada foreign exchange rates. Log difference × 1200.
- TS: Term structures. The 10-year treasury constant maturity rate less the 3-month treasury bill secondary market rate.

Figure 4.1: Plots of levels and differences in variables

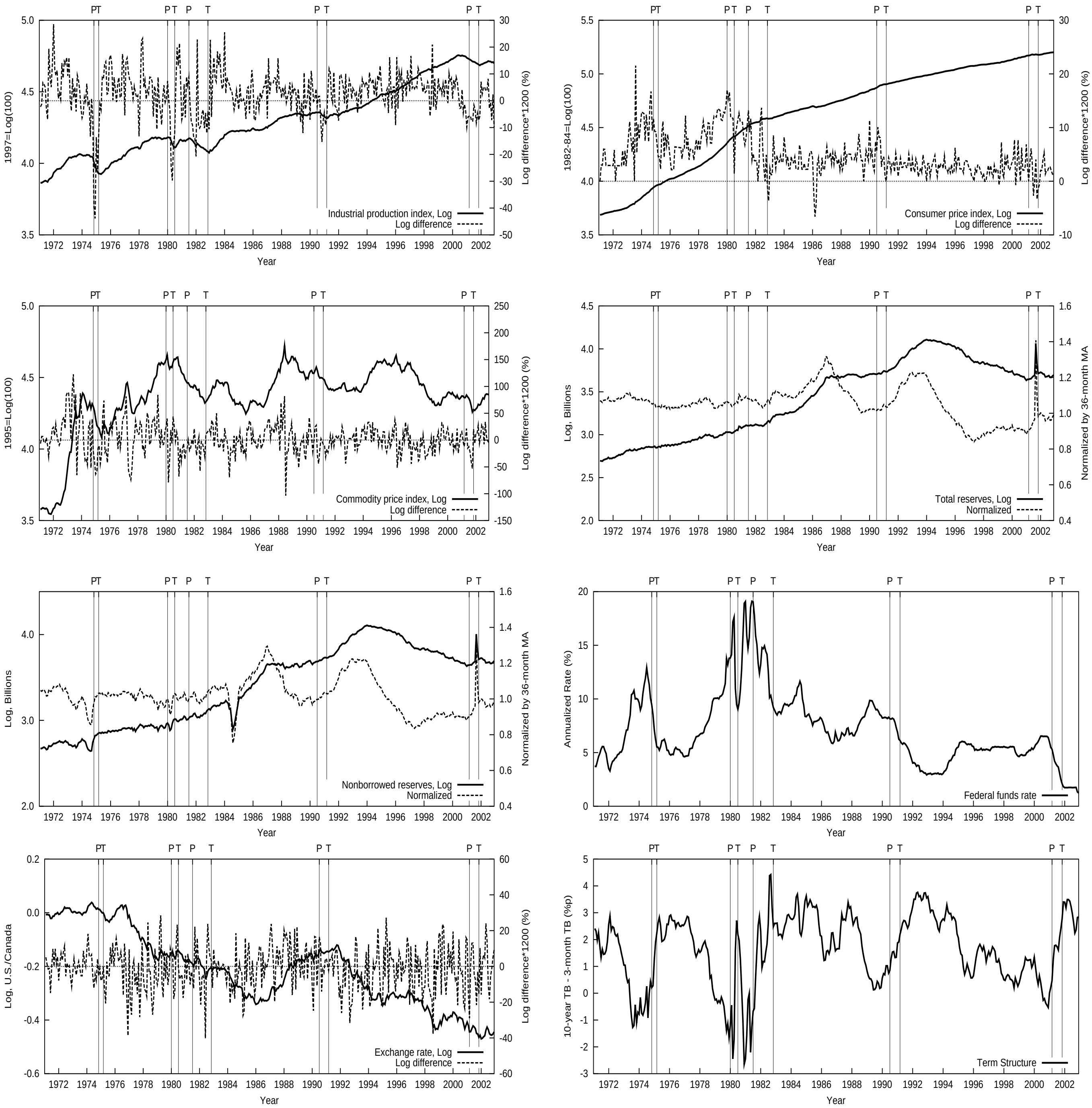
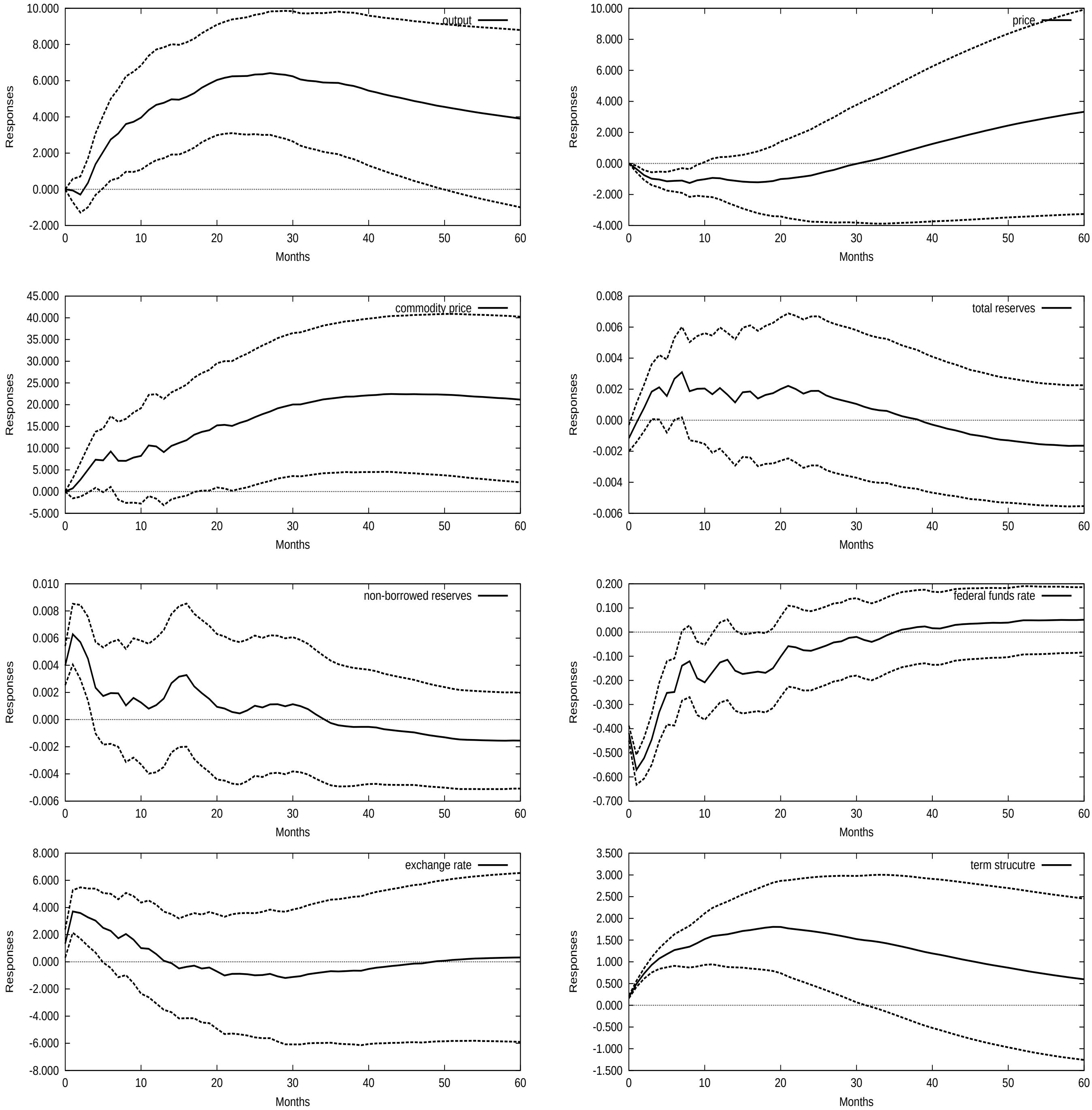


Figure 4.2: Impulse responses to monetary policy shock in alternative models

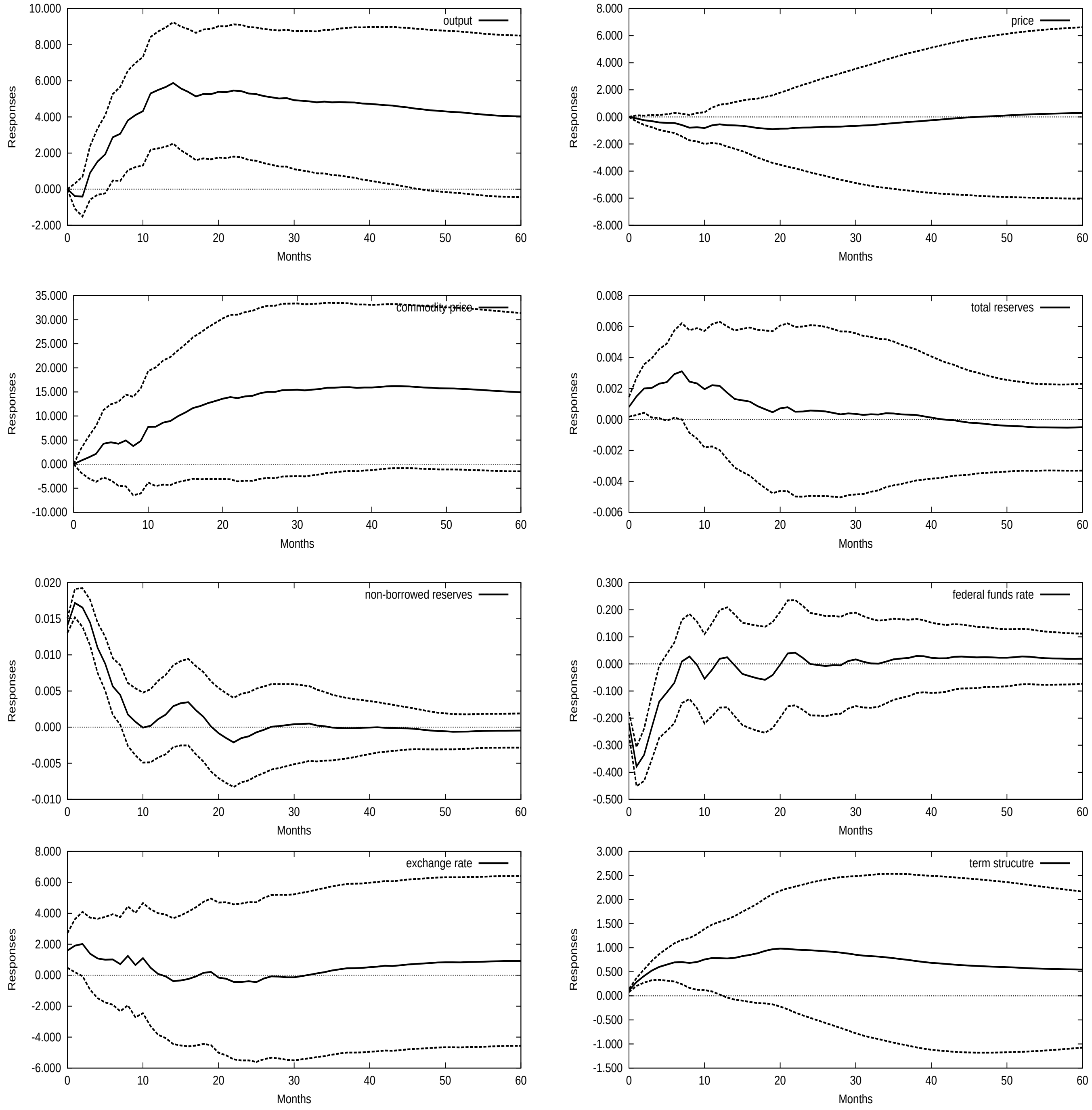
A. Bernanke-Blinder Model




Note: We chose 12 as the lag length of VAR. Upper and lower bounds are calculated by 95% confidence levels.

Figure 4.2: (Continued)

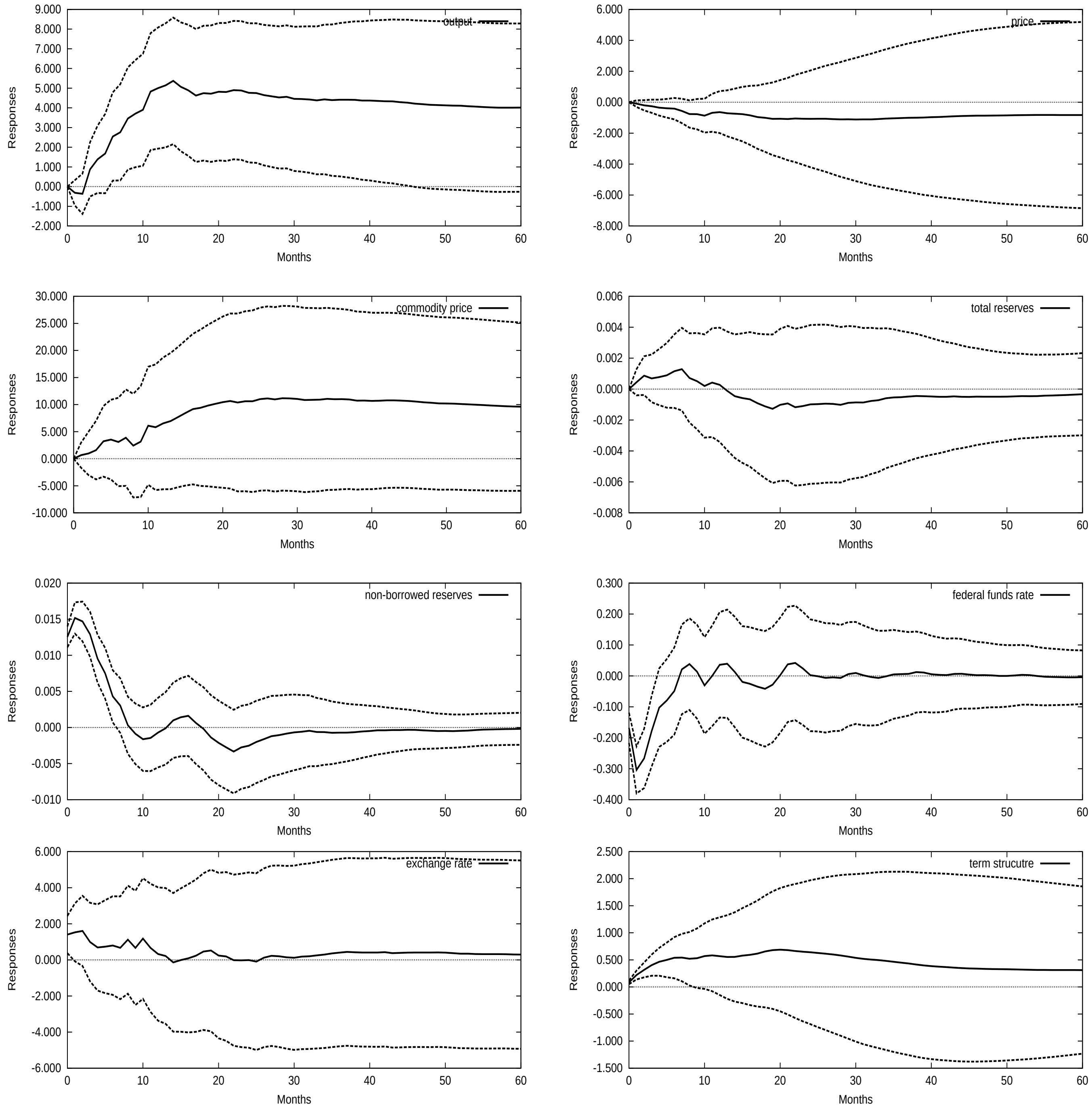
B. Christiano-Eichenbaum Model




Note: We chose 12 as the lag length of VAR. Upper and lower bounds are calculated by 95% confidence levels.

Figure 4.2: (Continued)

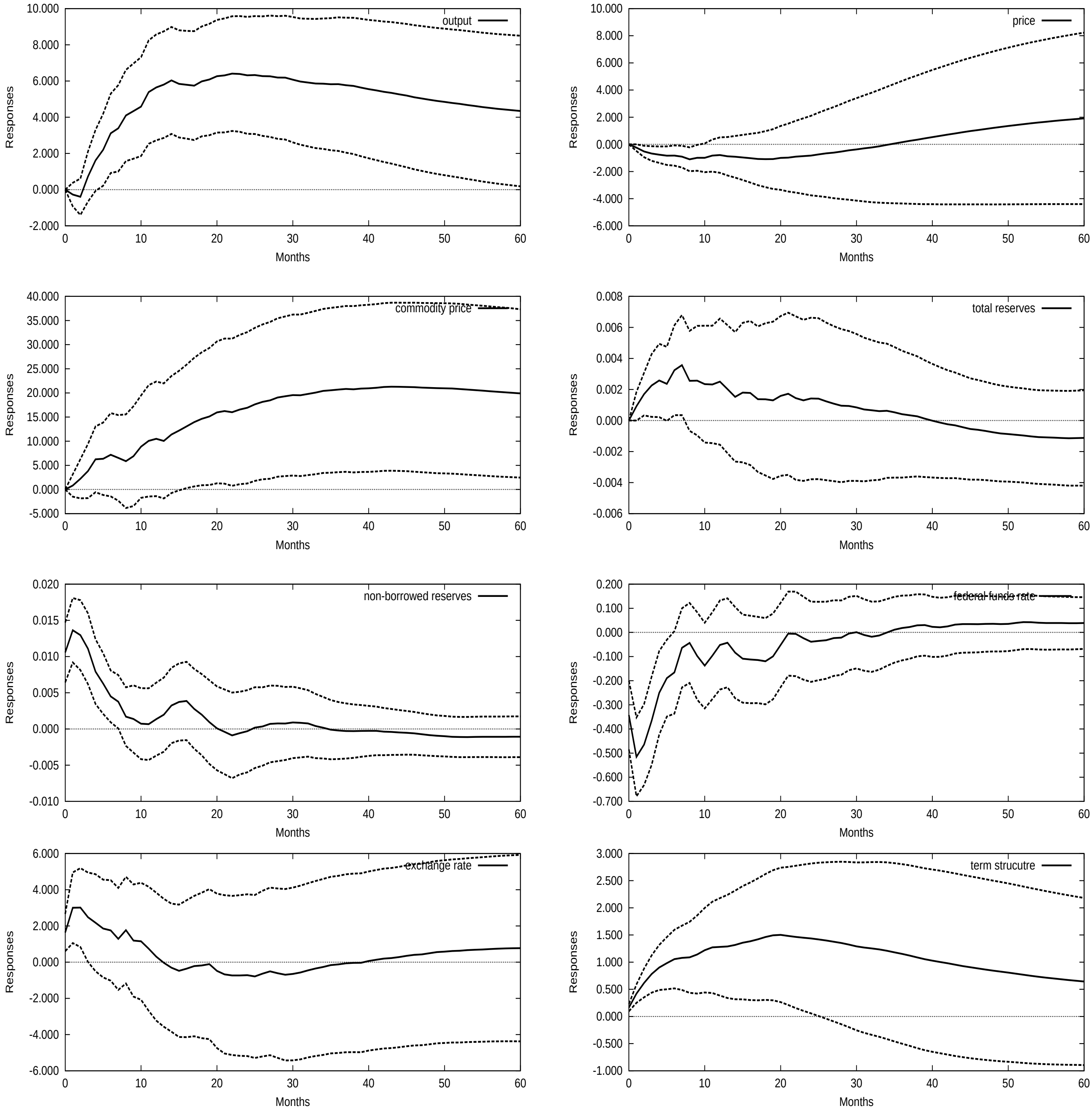
C. Strongin Model




Note: We chose 12 as the lag length of VAR. Upper and lower bounds are calculated by 95% confidence levels.

Figure 4.2: (Continued)

D. Just-identified Model




Note: We chose 12 as the lag length of VAR. Upper and lower bounds are calculated by 95% confidence levels.

## Table 4.1: Parameter estimates

**A. Over-identification test statistics**

| Model | LR test ($\chi^2_{(1)}$) | p-value |
|---|---|---|
| Bernanke-Blinder Model (BB) | 3.2909 | 0.0697 |
| Christiano-Eichenbaum Model (CE) | 75.2612** | 0.0000 |
| Strongin Model (ST) | 3.0162 | 0.0824 |

**B. ML estimates ($\mathbf{B}_{22}, \mathbf{F}_{22}$).**

| $\mathbf{B}_{22}$ | (4,4) | (5,4) | (6,4) | (4,5) | (5,5) | (6,5) | $\alpha$ | (5,6) | $\beta$ |
|---|---|---|---|---|---|---|---|---|---|
| BB | 1.0000 | 0.0000 | 1.0000 | 0.0000 | 1.0000 | -1.0000 | -0.0028** | 0.0000 | -0.0123** |
| | (.) | (.) | (.) | (.) | (.) | (.) | (0.0010) | (.) | (0.0015) |
| CE | 1.0000 | 0.0000 | 1.0000 | 0.0000 | 1.0000 | -1.0000 | 0.0037** | 0.0000 | -0.0603** |
| | (.) | (.) | (.) | (.) | (.) | (.) | (0.0012) | (.) | (0.0054) |
| ST | 1.0000 | 0.0000 | 1.0000 | 0.0000 | 1.0000 | -1.0000 | 0.0000 | 0.0000 | -0.0748** |
| | (.) | (.) | (.) | (.) | (.) | (.) | (.) | (.) | (0.0090) |
| JI | 1.0000 | 0.0000 | 1.0000 | 0.0000 | 1.0000 | -1.0000 | 0.0000 | 0.0000 | -0.0309* |
| | (.) | (.) | (.) | (.) | (.) | (.) | (.) | (.) | (0.0125) |

| $\mathbf{F}_{22}$ | $\eta^d$ | $\phi^d$ | (6,4) | (4,5) | $\eta^s$ | (6,5) | (4,6) | $\phi^b$ | $\eta^b$ |
|---|---|---|---|---|---|---|---|---|---|
| BB | 0.0081** | 0.0081** | 0.0000 | 0.0000 | 0.0040** | 0.0000 | 0.0000 | -0.0116** | 0.0116** |
| | (0.0003) | (0.0003) | (.) | (.) | (0.0008) | (.) | (.) | (0.0004) | (0.0004) |
| CE | 0.0085** | 0.0000 | 0.0000 | 0.0000 | 0.0140** | 0.0000 | 0.0000 | 0.0000 | 0.0232** |
| | (0.0004) | (.) | (.) | (.) | (0.0005) | (.) | (.) | (.) | (0.0021) |
| ST | 0.0082** | 0.0063** | 0.0000 | 0.0000 | 0.0125** | 0.0000 | 0.0000 | 0.0000 | 0.0286** |
| | (0.0003) | (0.0007) | (.) | (.) | (0.0005) | (.) | (.) | (.) | (0.0036) |
| JI | 0.0082** | 0.0063** | 0.0000 | 0.0000 | 0.0106** | 0.0000 | 0.0000 | -0.0068* | 0.0140** |
| | (0.0003) | (0.0007) | (.) | (.) | (0.0021) | (.) | (.) | (0.0033) | (0.0030) |

**C. Inferred parameter estimates ($\mathbf{K}_{22}, \mathbf{\Phi}_{22}$).**

| $\mathbf{K}_{22}$ | (4,4) | (5,4) | (6,4) | (4,5) | (5,5) | (6,5) | (4,6) | (5,6) | (6,6) |
|---|---|---|---|---|---|---|---|---|---|
| BB | 123.8419** | 0.0000 | 86.2670** | 0.0000 | 0.0000 | -86.2670** | -0.3430** | -2.3897** | -1.0587** |
| | (4.9363) | (.) | (3.7719) | (.) | (.) | (3.7719) | (0.1279) | (0.0899) | (0.1354) |
| CE | 117.4153** | 0.0000 | 43.0950** | 0.0000 | 71.1827** | -43.0950** | 0.4346** | 0.0000 | -2.5990** |
| | (5.6098) | (.) | (5.8491) | (.) | (2.6215) | (5.8491) | (0.1436) | (.) | (0.1052) |
| ST | 122.5857** | -61.7820** | 34.9436** | 0.0000 | 79.7121** | -34.9436** | 0.0000 | 0.0000 | -2.6137** |
| | (8.7615) | (18.8947) | (6.6455) | (.) | (4.6738) | (6.6455) | (.) | (.) | (0.1074) |
| JI | 122.5857** | -27.5065 | 71.6292** | 0.0000 | 48.7950* | -71.6292** | 0.0000 | -1.4151* | -2.2108** |
| | (4.6136) | (23.2405) | (15.1959) | (.) | (22.3573) | (15.1959) | (.) | (0.6898) | (0.4416) |

| $\mathbf{\Phi}_{22}$ | (4,4) | (5,4) | (6,4) | (4,5) | (5,5) | (6,5) | (4,6) | (5,6) | (6,6) |
|---|---|---|---|---|---|---|---|---|---|
| BB | 0.0081** | 0.0081** | 0.0000 | -0.0012** | 0.0040** | -0.4185** | 0.0000 | -0.0116** | 0.0000 |
| | (0.0003) | (0.0003) | (.) | (0.0004) | (0.0007) | (0.0157) | (.) | (0.0005) | (.) |
| CE | 0.0080** | 0.0000 | 0.1331** | 0.0008* | 0.0140** | -0.2195** | 0.0013** | 0.0000 | -0.3625** |
| | (0.0004) | (.) | (0.0133) | (0.0003) | (0.0005) | (0.0211) | (0.0004) | (.) | (0.0195) |
| ST | 0.0082** | 0.0063** | 0.0245 | 0.0000 | 0.0125** | -0.1677** | 0.0000 | 0.0000 | -0.3826** |
| | (0.0006) | (0.0013) | (0.0210) | (.) | (0.0007) | (0.0250) | (.) | (.) | (0.0157) |
| JI | 0.0082** | 0.0063** | 0.0595** | 0.0000 | 0.0106** | -0.3423** | 0.0000 | -0.0068* | -0.2332* |
| | (0.0003) | (0.0007) | (0.0222) | (.) | (0.0021) | (0.0726) | (.) | (0.0033) | (0.1069) |

Note: Standard errors are in parentheses. Statistics significantly different from zeros are denoted by * or ** at a 5% or 1% significance level, respectively.

Table 4.1: (Continued)

D. Inferred parameter estimates $(\mathbf{B}_{21}, \mathbf{K}_{21})$.

| $\mathbf{B}_{21}$ | (4,1) | (5,1) | (6,1) | (4,2) | (5,2) | (6,2) | (4,3) | (5,3) | (6,3) |
|---|---|---|---|---|---|---|---|---|---|
| BB | 0.0000 | 0.0004** | -0.0001 | -0.0002 | 0.0007* | -0.0010** | 0.0000 | 0.0000 | 0.0000 |
| | (0.0001) | (0.0001) | (0.0001) | (0.0002) | (0.0004) | (0.0003) | (0.0000) | (0.0000) | (0.0000) |
| CE | -0.0001 | 0.0004** | 0.0008** | -0.0002 | 0.0007* | -0.0009 | 0.0000* | 0.0000 | -0.0001 |
| | (0.0001) | (0.0001) | (0.0002) | (0.0002) | (0.0004) | (0.0006) | (0.0000) | (0.0000) | (0.0001) |
| ST | 0.0000 | 0.0004** | 0.0011** | -0.0002 | 0.0007* | -0.0009 | 0.0000 | 0.0000 | -0.0001* |
| | (0.0001) | (0.0001) | (0.0004) | (0.0002) | (0.0004) | (0.0007) | (0.0000) | (0.0000) | (0.0001) |
| JI | 0.0000 | 0.0004** | 0.0003 | -0.0002 | 0.0007* | -0.0010** | 0.0000 | 0.0000 | 0.0000 |
| | (0.0001) | (0.0001) | (0.0003) | (0.0002) | (0.0004) | (0.0004) | (0.0000) | (0.0000) | (0.0000) |

| $\mathbf{K}_{21}$ | (4,1) | (5,1) | (6,1) | (4,2) | (5,2) | (6,2) | (4,3) | (5,3) | (6,3) |
|---|---|---|---|---|---|---|---|---|---|
| BB | 0.0046 | 0.0483** | -0.0106 | -0.0288 | 0.0018 | -0.0838** | 0.0030 | -0.0060* | 0.0008 |
| | (0.0088) | (0.0086) | (0.0088) | (0.0259) | (0.0259) | (0.0262) | (0.0024) | (0.0024) | (0.0024) |
| CE | -0.0110 | 0.0251** | 0.0366** | -0.0278 | 0.0531* | -0.0403 | 0.0048* | -0.0006 | -0.0048* |
| | (0.0089) | (0.0085) | (0.0087) | (0.0259) | (0.0260) | (0.0265) | (0.0024) | (0.0024) | (0.0024) |
| ST | -0.0023 | 0.0293** | 0.0399** | -0.0287 | 0.0739** | -0.0323 | 0.0039 | -0.0026 | -0.0052* |
| | (0.0084) | (0.0087) | (0.0088) | (0.0260) | (0.0270) | (0.0267) | (0.0024) | (0.0024) | (0.0024) |
| JI | -0.0023 | 0.0463** | 0.0181 | -0.0287 | 0.0439 | -0.0686* | 0.0039 | -0.0048 | -0.0027 |
| | (0.0084) | (0.0103) | (0.0166) | (0.0259) | (0.0336) | (0.0300) | (0.0024) | (0.0025) | (0.0029) |

E. Inferred parameter estimates $(\mathbf{\Phi}(1)_{12})$.

| $\mathbf{\Phi}(1)_{12}$ | (1,4) | (2,4) | (3,4) | (1,5) | (2,5) | (3,5) | (1,6) | (2,6) | (3,6) |
|---|---|---|---|---|---|---|---|---|---|
| BB | -3.8435 | 3.3318 | 18.4644 | 3.0351 | 3.8829 | 17.8095* | -3.4561 | 2.7872 | -1.6416 |
| | (4.6629) | (5.3453) | (15.9874) | (2.4312) | (2.7705) | (8.3285) | (2.4027) | (2.7431) | (8.1972) |
| CE | -6.5147 | 3.4611 | 10.9385 | 4.0590 | 0.0876 | 14.0809 | -0.2138 | 5.7240 | 19.9722 |
| | (4.9011) | (5.6180) | (16.7990) | (2.1539) | (2.4823) | (7.4671) | (3.2081) | (3.6598) | (10.9493) |
| ST | -4.4858 | 3.4791 | 17.2370 | 4.1220* | -0.7735 | 9.6853 | 0.8707 | 5.1163 | 18.0410 |
| | (4.8215) | (5.4624) | (16.3082) | (2.0593) | (2.3740) | (7.0881) | (2.8041) | (3.2151) | (9.6221) |
| JI | -4.5653 | 3.0121 | 15.5904 | 3.9293 | 2.0374 | 17.6389* | -1.4856 | 4.7447 | 10.0388 |
| | (4.6780) | (5.3625) | (16.0491) | (2.0831) | (2.7330) | (7.6031) | (3.0677) | (3.2856) | (11.0858) |

F. Inferred parameter estimates $(\mathbf{\Phi}(1)_{22})$.

| $\mathbf{\Phi}(1)_{22}$ | (4,4) | (5,4) | (6,4) | (4,5) | (5,5) | (6,5) | (4,6) | (5,6) | (6,6) |
|---|---|---|---|---|---|---|---|---|---|
| BB | 0.5767* | 0.4574* | 8.4964 | 0.0587 | 0.0522 | -2.6179 | 0.0921 | 0.0047 | 0.7201 |
| | (0.2923) | (0.1976) | (10.0252) | (0.1537) | (0.1045) | (5.1902) | (0.1497) | (0.1011) | (5.1398) |
| CE | 0.5789 | 0.4197* | 9.2849 | 0.0485 | 0.1035 | -0.5335 | 0.2644 | 0.1805 | 0.5602 |
| | (0.3076) | (0.2079) | (10.5245) | (0.1372) | (0.0929) | (4.5871) | (0.2013) | (0.1361) | (6.8652) |
| ST | 0.5865 | 0.4558* | 8.7607 | -0.0266 | 0.0430 | -1.2119 | 0.1666 | 0.1097 | -0.9869 |
| | (0.3007) | (0.2028) | (10.2455) | (0.1284) | (0.0865) | (4.3844) | (0.1752) | (0.1183) | (6.0167) |
| JI | 0.5713 | 0.4458* | 8.8508 | 0.0652 | 0.0938 | -1.5394 | 0.1553 | 0.0696 | -0.1814 |
| | (0.2929) | (0.1980) | (10.0452) | (0.1350) | (0.0880) | (4.3391) | (0.1769) | (0.1222) | (6.0555) |

Note: Standard errors are in parentheses. Statistics significantly different from zeros are denoted by * or ** at a 5% or 1% significance level, respectively.

Table 4.1: (Continued)

| G. Inferred parameter estimates ($\mathbf{\Phi}_{32}, \mathbf{\Phi}(1)_{32}$). | | | | | | |
|---|---|---|---|---|---|---|
| $\mathbf{\Phi}_{32}$ | (7,4) | (8,4) | (7,5) | (8,5) | (7,6) | (8,6) |
| BB | 0.1330 | 0.0185 | 1.3431* | 0.1899** | -0.9499 | -0.0083 |
| | (0.5244) | (0.0127) | (0.5268) | (0.0145) | (0.5260) | (0.0127) |
| CE | -0.8246 | -0.0474** | 1.5917** | 0.1104** | 0.7276 | 0.1657** |
| | (0.5761) | (0.0152) | (0.5679) | (0.0169) | (0.5409) | (0.0155) |
| ST | -0.0712 | 0.0063 | 1.4053** | 0.0847** | 0.8607 | 0.1726** |
| | (0.5553) | (0.0160) | (0.5267) | (0.0168) | (0.5253) | (0.0145) |
| JI | -0.1497 | -0.0094 | 1.6359** | 0.1621** | -0.0296 | 0.1004 |
| | (0.5292) | (0.0162) | (0.5256) | (0.0338) | (0.7289) | (0.0521) |
| $\mathbf{\Phi}(1)_{32}$ | (7,4) | (8,4) | (7,5) | (8,5) | (7,6) | (8,6) |
| BB | 6.8452 | 2.6273 | -0.3471 | 0.8138 | 0.2732 | 0.2134 |
| | (6.2844) | (1.7919) | (3.3059) | (0.9447) | (3.2682) | (0.9286) |
| CE | 6.7509 | 2.3499 | 0.7734 | 0.6948 | 1.8244 | 1.5748 |
| | (6.6018) | (1.8840) | (2.9522) | (0.8485) | (4.3325) | (1.2347) |
| ST | 6.9142 | 2.6127 | 0.0075 | 0.2843 | 0.6915 | 1.1752 |
| | (6.4208) | (1.8316) | (2.8078) | (0.8007) | (3.8055) | (1.0834) |
| JI | 6.8511 | 2.5054 | 0.3698 | 0.8571 | 0.5809 | 0.8408 |
| | (6.2958) | (1.7959) | (2.7854) | (0.8303) | (3.8190) | (1.1184) |

Note: Standard errors are in parentheses. Statistics significantly different from zeros are denoted by * or ** at a 5% or 1% significance level, respectively.

Table 4.2: Fraction of the forecast-error variance attributed to monetary policy shock

| Horizon | $y$ | $p$ | $pc$ | $TR$ | $NBR$ | $FFR$ | $er$ | $TS$ |
|---|---|---|---|---|---|---|---|---|
| | | | | A. Bernanke-Blinder Model | | | | |
| 1 | 0.00 | 0.00 | 0.00 | 0.02 | 0.08** | 0.91** | 0.02 | 0.38** |
| | (0.00) | (0.00) | (0.00) | (0.01) | (0.03) | (0.03) | (0.01) | (0.04) |
| 6 | 0.04* | 0.06** | 0.02 | 0.02 | 0.08** | 0.53** | 0.07** | 0.36** |
| | (0.02) | (0.02) | (0.02) | (0.02) | (0.04) | (0.07) | (0.03) | (0.07) |
| 12 | 0.05** | 0.05** | 0.04* | 0.02 | 0.04* | 0.35** | 0.06** | 0.28** |
| | (0.03) | (0.02) | (0.02) | (0.03) | (0.03) | (0.07) | (0.02) | (0.07) |
| 36 | 0.05** | 0.05** | 0.05** | 0.02 | 0.02 | 0.18** | 0.07** | 0.20** |
| | (0.02) | (0.02) | (0.02) | (0.03) | (0.02) | (0.07) | (0.02) | (0.06) |
| 60 | 0.05** | 0.07** | 0.05** | 0.02 | 0.02 | 0.15** | 0.07** | 0.20** |
| | (0.02) | (0.03) | (0.02) | (0.03) | (0.02) | (0.06) | (0.02) | (0.06) |
| | | | | B. Christiano-Eichenbaum Model | | | | |
| 1 | 0.00 | 0.00 | 0.00 | 0.01 | 0.96** | 0.25** | 0.02 | 0.13** |
| | (0.00) | (0.00) | (0.00) | (0.01) | (0.02) | (0.04) | (0.02) | (0.04) |
| 6 | 0.05** | 0.01 | 0.01 | 0.04 | 0.78** | 0.18** | 0.03 | 0.11** |
| | (0.02) | (0.01) | (0.01) | (0.03) | (0.08) | (0.05) | (0.02) | (0.05) |
| 12 | 0.08** | 0.02 | 0.02 | 0.03 | 0.40** | 0.10** | 0.03* | 0.08** |
| | (0.03) | (0.01) | (0.02) | (0.04) | (0.08) | (0.03) | (0.02) | (0.04) |
| 36 | 0.07** | 0.01 | 0.03 | 0.01 | 0.17** | 0.05** | 0.03* | 0.06** |
| | (0.03) | (0.01) | (0.02) | (0.02) | (0.05) | (0.02) | (0.02) | (0.03) |
| 60 | 0.07** | 0.02 | 0.03 | 0.01 | 0.15** | 0.04** | 0.03** | 0.06* |
| | (0.03) | (0.01) | (0.02) | (0.02) | (0.05) | (0.02) | (0.02) | (0.03) |
| | | | | C. Strongin Model | | | | |
| 1 | 0.00 | 0.00 | 0.00 | 0.00 | 0.77** | 0.15** | 0.02 | 0.08** |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.08) | (0.04) | (0.01) | (0.03) |
| 6 | 0.04* | 0.01 | 0.01 | 0.00 | 0.61** | 0.11** | 0.02 | 0.07* |
| | (0.02) | (0.01) | (0.01) | (0.01) | (0.09) | (0.04) | (0.01) | (0.04) |
| 12 | 0.07** | 0.02 | 0.02 | 0.00 | 0.31** | 0.06** | 0.02* | 0.05* |
| | (0.03) | (0.01) | (0.01) | (0.01) | (0.07) | (0.02) | (0.01) | (0.03) |
| 36 | 0.06** | 0.01 | 0.02 | 0.00 | 0.13** | 0.03** | 0.03* | 0.04* |
| | (0.02) | (0.01) | (0.01) | (0.01) | (0.04) | (0.01) | (0.01) | (0.02) |
| 60 | 0.06** | 0.01 | 0.02 | 0.00 | 0.12** | 0.02** | 0.03* | 0.04* |
| | (0.02) | (0.01) | (0.01) | (0.01) | (0.04) | (0.01) | (0.01) | (0.02) |
| | | | | D. Just-identified Model | | | | |
| 1 | 0.00 | 0.00 | 0.00 | 0.00 | 0.55** | 0.61** | 0.03 | 0.28** |
| | (0.00) | (0.00) | (0.00) | (0.00) | (0.23) | (0.25) | (0.02) | (0.11) |
| 6 | 0.05** | 0.03 | 0.02 | 0.04 | 0.46** | 0.39** | 0.04* | 0.26** |
| | (0.02) | (0.02) | (0.02) | (0.03) | (0.18) | (0.15) | (0.02) | (0.11) |
| 12 | 0.08** | 0.03* | 0.02 | 0.04 | 0.24** | 0.23** | 0.04* | 0.19** |
| | (0.03) | (0.02) | (0.02) | (0.04) | (0.09) | (0.11) | (0.02) | (0.10) |
| 36 | 0.07** | 0.03 | 0.03 | 0.02 | 0.10** | 0.11* | 0.05** | 0.14* |
| | (0.02) | (0.02) | (0.02) | (0.03) | (0.04) | (0.07) | (0.02) | (0.07) |
| 60 | 0.07** | 0.04 | 0.03 | 0.02 | 0.09** | 0.10 | 0.05** | 0.14* |
| | (0.02) | (0.03) | (0.02) | (0.03) | (0.04) | (0.06) | (0.02) | (0.07) |

Note: Standard errors are in parentheses. Statistics significantly different from zeros are denoted by * or ** at a 10% or 5% significance level, respectively.