Inference on Impulse Response Functions in Structural VAR Models*

Atsushi Inoue† North Carolina State University

Lutz Kilian‡ University of Michigan

May 19, 2011

Abstract

Skepticism toward traditional identifying assumptions based on exclusion restrictions has led to a surge in the use of structural VAR models in which structural shocks are identified by restricting the sign of the responses of selected macroeconomic aggregates to these shocks. Researchers commonly report the vector of pointwise posterior medians of the impulse responses as a measure of central tendency of the estimated response functions, along with pointwise 68 percent posterior error bands. It can be shown that this approach cannot be used to characterize the central tendency of the structural impulse response functions. We propose an alternative method of summarizing the evidence from sign-identified VAR models designed to enhance their practical usefulness. Our objective is to characterize the most likely admissible model(s) within the set of structural VAR models that satisfy the sign restrictions. We show how the set of most likely structural response functions can be computed from the posterior mode of the joint distribution of admissible models both in the fully identified and in the partially identified case, and we propose a highest-posterior density credible set that characterizes the joint uncertainty about this set. Our approach can also be used to resolve the long-standing problem of how to conduct joint inference on sets of structural impulse response functions in exactly identified VAR models. We illustrate the differences between our approach and the traditional approach for the analysis of the effects of monetary policy shocks and of the effects of oil demand and oil supply shocks.

JEL: C32, C52; E37

KEYWORDS: Vector Autoregression; Simultaneous Inference; Impulse Responses; Sign Restrictions; Median; Mode; Credible Set.

*We thank Harald Uhlig for providing the data used in one of the empirical applications. We also thank Christiane Baumeister, Alastair Hall, Sujit Ghosh, Dan Phaneuf, Frank Schorfheide and Tao Zha for helpful discussions.

†Department of Agricultural and Resource Economics, Box 8109, North Carolina State University, Raleigh, NC 27695-8109. E-mail: atsushi@ncsu.edu.

‡University of Michigan, Department of Economics, 238 Lorch Hall, Ann Arbor, MI 48109-1220. E-mail: lkilian@umich.edu.

1 Introduction

One of the most widely studied questions in empirical macroeconomics is to what extent an unanticipated monetary tightening reduces real output. It is widely recognized that answering this question requires the estimation of structural models in which cause and effect are clearly differentiated. Much of the literature since the 1980s has relied on structural vector autoregressive models in which exclusion restrictions on the instantaneous feedback from monetary policy shocks to macroeconomic aggregates ensure the identification of the policy shock. Skepticism toward these traditional identifying assumptions in recent years has made increasingly popular an alternative class of structural VAR models in which policy shocks are identified by restricting the sign of the responses of selected macroeconomic aggregates to policy shocks. For example, Uhlig (2005) postulated that an unexpected monetary policy contraction is associated with an increase in the federal funds rate, the absence of price increases and the absence of increases in nonborrowed reserves for some time following the policy shock. This approach is considerably more agnostic than traditional identification approaches in some dimensions, while more restrictive in others. Uhlig showed that sign-identified models may produce substantially different results from conventional structural VAR models.

Although the original applications of this approach were to models of monetary policy, sign-identified VAR models have become increasingly popular in other areas as well and are now part of the mainstream of empirical macroeconomics. They have been used to study fiscal shocks (e.g., Canova and Pappa 2007; Mountford and Uhlig 2009, Pappa 2009; Caldara 2011), technology shocks (e.g., Dedola and Neri 2007; Peersman and Straub 2009), and various other shocks in open economies (e.g., Canova and De Nicolo 2002; Scholl and Uhlig 2008), in oil markets (e.g., Baumeister and Peersman 2009; Kilian and Murphy 2010, 2011), and in labor markets (e.g., Fujita 2011), for example.

In all these applications, the cost of remaining agnostic about the structural model is that the data are potentially consistent with a wide range of structural models that are all admissible in that they satisfy the identifying restrictions. An unresolved question in the literature is how to represent the results of such agnostic identification procedures when the set of admissible models includes a range of models with conflicting interpretations. One early approach, exemplified by Faust (1998), has been to focus on the model that is most favorable to the hypothesis of interest. This allows us to establish the extent to which this hypothesis could potentially explain the data. It may also help us to rule out a hypothesized explanation, if none of the admissible models supports this hypothesis. The problem is that this approach is not informative about whether any one of the admissible models is a more likely explanation of the data than some other model.

More commonly, researchers have reported the vector of pointwise posterior medians of the impulse responses as a measure of the central tendency of the impulse response functions, along

with pointwise 68% posterior error bands. This approach suffers from two distinct shortcomings. The first shortcoming is that the vector of pointwise posterior median responses will have no structural economic interpretation unless the pointwise posterior medians of all impulse response coefficients in the VAR system correspond to the same structural model, which is highly unlikely a priori (see, e.g., Fry and Pagan 2011). In other words, in practice, none of the models in the set of admissible structural models constructed by the researcher will exhibit the impulse response dynamics embodied in the *median response function* obtained by connecting the dots between the pointwise posterior medians. The second shortcoming is that median response functions are not a valid statistical summary of the admissible set of impulse response functions. It is well known that the median of a vector variable is not the vector of the medians, rendering the vector of pointwise medians inappropriate as a statistical measure of the central tendency of the impulse response functions (e.g., Chauduri 1996; Koltchinskii 1997; Liu, Parelius and Singh 1999). This means that even if there were an admissible structural model with the same impulse response function as the median response function, there would be no compelling reason to focus on this model in interpreting the evidence.[1]

Similar problems arise in the construction of pointwise impulse response error bands in sign-identified models based on the quantiles of the marginal posterior distributions of the impulse responses. Moreover, these error bands do not take account of the dependence of the impulse response estimates across horizons and across response functions and hence may overstate or understate the true uncertainty about the dynamics of the system. The latter problem is well known, but there are no alternative methods in the literature that address these limitations, leaving researchers with little choice, but to rely on potentially misleading measures of uncertainty.

In this paper, we propose a new method of summarizing the evidence from sign-identified VAR models that addresses these shortcomings and is designed to enhance the practical usefulness of sign-identified models. Our objective is to identify the most likely admissible model(s) within the set of structural VAR models that satisfy the sign restrictions. A structural VAR model is defined by the set of structural impulse response functions associated with a given set of reduced-form VAR parameters and a given structural impact multiplier matrix. There is a one-to-one mapping from the joint posterior density of these model parameters to the joint posterior density of the corresponding set of structural impulse response functions, allowing us to derive the latter density analytically by the change-of-variable method. This enables us to assign a posterior density value to each structural model. The most likely or modal model by

[1] This second point is also relevant for a recent proposal by Fry and Pagan (2011) designed to overcome the lack of structural interpretation of median response functions. Their idea is to search for the admissible structural model with impulse response functions closest to the median response functions. This proposal deals with the first shortcoming, but not the second. Because the median vector is not well defined statistically, there is no compelling reason to focus on a structural model with response functions close to the vector of medians.

construction is the admissible model that maximizes the joint posterior density of the admissible structural VAR models. A $100(1-\alpha)\%$ highest posterior density credible set of admissible models may be formed by ranking the admissible models based on the value of this joint density.

In practice, we proceed in two steps. Under the conventional assumption of a diffuse Gaussian-inverse Wishart prior, we begin by generating repeated draws from the joint posterior distribution of the reduced-form VAR parameters and of the rotation matrices used in constructing the structural impact multiplier matrix. For each candidate structural VAR model, we first compute the posterior density value associated with that model; we then evaluate the set of implied structural impulse response functions. We discard the structural models that are inadmissible in that they do not satisfy the identifying sign restrictions. We then rank the remaining structural models by the value of their posterior density, making it straightfoward to determine the most likely admissible model and to characterize its impulse response dynamics. The set of structural impulse response functions associated with the modal admissible model by construction will be economically interpretable and statistically well defined, addressing the two main critiques of traditional median response functions.

This baseline procedure is designed for fully identified structural VAR models. Many structural VAR models, however, are only partially identified in that only a subset of the structural shocks are identified. Such models are sometimes also referred to as semi-structural VAR models. For example, in the model of Uhlig (2005) only the responses to monetary policy shocks are identified. In this case, responses to unidentified shocks become irrelevant in constructing the modal model. Instead the mode and credible set must be based on the marginal posterior density of the subset of impulse response functions of interest. We propose a modification of our baseline procedure that accomplishes this task. Marginalizing the joint density requires Monte Carlo integration, which renders this procedure computationally more challenging than the baseline procedure for fully identified models.

Although our approach was designed to aid in the interpretation of impulse response dynamics in sign-identified models, essentially the same approach can also be used to resolve the long-standing problem of how to conduct joint inference on sets of structural impulse response functions in exactly identified models. It is well known that the pointwise error bands commonly attached to the structural impulse response functions in exactly identified VAR models fail to convey the true uncertainty surrounding these impulse response functions. This problem has been long recognized, but few practical alternatives have been proposed in the literature and none of these alternative methods are applicable to sign-identified models.[2] Our final contribution is to show how a simplified version of our proposed procedure may be used to construct

[2] Sims and Zha (1999), for example, caution against connecting the dots representing pointwise error bands and discuss possible alternative strategies for exactly identified models. Related work based on joint asymptotic approximations for exactly identified models includes Lütkepohl (1990) and Jorda (2009).

credible sets for the structural impulse response functions for exactly identified structural VAR models, providing a convenient alternative to traditional pointwise error bands that is easy to implement. As in the earlier analysis, our approach can accommodate both fully identified and partially identified models.

Table 1 summarizes the four distinct types of structural VAR models considered in this paper. The remainder of the paper is organized as follows. In section 2, we describe the proposed procedure and its implementation in more detail. Although Uhlig's model focuses on identifying only the monetary policy shock, many related studies have used sign restrictions to identify simultaneously a variety of macroeconomic shocks (e.g., Canova and Paustian 2007; Gambetti, Pappa and Canova 2008). Our analysis allows for that situation. Our analysis also allows for refinements of the sign restriction approach in the form of additional bounds on elasticities, for example, or of bounds on cross-correlations (e.g., Canova and De Nicolo 2002; Kilian and Murphy 2011). These modifications do not affect the substance of our method and can be easily incorporated.

In section 3, we contrast the proposed modal model responses with traditional median responses. We demonstrate that median response functions can be quite misleading in practice. We illustrate how this problem may arise in practice and provide a simple analytical example which shows that even in simple settings the median response is likely to be biased away from the responses associated with the most likely or modal model. The median responses not only tend to be biased relative to the modal response, but, in models with richer dynamics, the median response function may have the opposite sign of the response function in the modal model. This observation suggests that there is a need to reexamine the findings of earlier studies based on median responses from sign-identified VAR models.

The empirical relevance of these insights is demonstrated in section 4. In section 4a, we pursue this question in the context of a fully identified structural VAR model. We focus on the example of a model of the global market for crude oil in the tradition of Baumeister and Peersman (2009) and Kilian and Murphy (2011). We show that the modal model generates economically plausible responses. The 68 percent joint credible sets tend to be wider than conventional 68 percent pointwise impulse response error bands, but many responses are fairly precisely estimated nevertheless with credible sets that exclude the zero line at some or all horizons. We demonstrate that the responses in the modal model can be substantially different from conventional median response functions. The bias in the median response functions can be upward or downward. In many cases, the responses of the modal model are outside of the conventional 68 percent pointwise posterior error bands.

In section 4b, we consider a partially identified structural VAR model, building on the analysis of U.S. monetary policy in Uhlig (2005). We explore in particular the question of what

the effect is of an unanticipated monetary policy tightening on real U.S. output. We show that the method of summarizing the evidence matters. For example, Uhlig reported a peak median output response of 0.15 for this model. For the same data, we obtain a peak response of almost 0.5 based on the modal model. Moreover, that modal model estimate is near the upper end of the credible set and outside the conventional pointwise posterior error band.

Both the median estimate and the response estimate based on the modal model are counter-intuitive in that a monetary tightening would be expected to cause a decline in real output over time rather than an increase. This outcome reflects the fact that the identifying assumptions are not overly informative. Even in Uhlig's original analysis, there was substantial pointwise probability mass on both negative and positive responses of real output. Our 68 percent credible set further widens the set of probable responses. The explicit reason that Uhlig (2005) did not impose further restrictions is that he wished to be as agnostic as possible about the response of real output. It has been shown that this approach is valid only to the extent that we view models in which real output increases in response to a monetary tightening as equally plausible a priori as models in which real output declines (see Kilian and Murphy 2011). In an effort to relax that assumption, we also considered an alternative set of models that impose an additional sign restriction on the response of real GDP after 6 months (and only at that horizon). This identifying assumption leaves the short-run as well as the longer-run response of real output unrestricted.

The resulting modal model produces substantially different and more economically plausible results, including a cumulative drop in real GDP of -0.3 percentage points in the second quarter. The response estimate for the modal model is at the lower end of the credible set and again outside the conventional pointwise posterior error band. It also is substantially different from the response estimate obtained from the traditional Cholesky decomposition. Even in this alternative model, however, the 68 percent credible set includes many positive real output responses, suggesting that the data are not very informative about the response of real output. We conclude that there remains substantial uncertainty about the effects of monetary policy shocks on real output, whereas there is strong evidence of the effects of oil demand shocks on the real price of oil in the earlier example.

In section 5, we show that our approach of constructing joint credible sets for the structural impulse response functions can also be adapted to exactly identified structural VAR models, providing a convenient alternative to traditional pointwise error bands that is easy to implement. For standard semi-structural monetary policy VAR models of the type considered by Uhlig (2005) as a benchmark, we show that properly accounting for the joint uncertainty about all impulse responses renders the impulse response estimates less informative than conventional pointwise error bands suggest. For example, evidence of the price puzzle vanishes. On the other

hand, several of the response functions including that of real GDP remain precisely enough estimated for the VAR model to be economically informative. In addition, we illustrate that, for the reasons already discussed, commonly used posterior median response functions may differ from the response functions in the modal posterior model. The concluding remarks are in section 6. Some technical details of the proposed procedure can be found in the technical appendix.

2 Evaluating the Posterior of Sign-Identified VAR Models

2.1 Preliminaries

Consider the n-variate reduced-form VAR(p) model:

$$y_t \;=\; c + B_1 y_{t-1} + B_2 y_{t-2} + \cdots + B_p y_{t-p} + e_t, \tag{1}$$

for $t = 1, ..., T$, where $e_t \overset{iid}{\sim} N(0_{n\times 1}, \Sigma)$ and Σ is positive definite. Write (1) as

$$Y \;=\; XB + e \tag{2}$$

where $Y = [y_1\ y_2\ \cdots\ y_T]'$, $X = [X_1\ X_2\ \cdots\ X_T]'$, $X_t = [1\ y'_{t-1}\ \cdots\ y'_{t-p}]'$, $B = [c\ B_1\ \cdots\ B_p]'$, and $e = [e_1\ e_2\ \cdots\ e_T]'$.

Throughout this paper, we follow the conventional approach of specifying a normal-inverse Wishart prior distribution for the reduced-form VAR parameters and a Haar distribution for the rotation matrix. This approach ensures that all admissible models are a priori equally likely (see Uhlig 2005, p. 389):

$$\begin{aligned} \mathrm{vec}(B)|\Sigma &\sim N(\mathrm{vec}(\bar{B}_0), \Sigma \otimes N_0^{-1}), \\ \Sigma &\sim IW_n(\nu_0 S_0, \nu_0), \end{aligned}$$

where N_0 is an $np \times np$ positive definite matrix, S_0 is an $n \times n$ covariance matrix, and $\nu_0 > 0$. Then the posterior is given by

$$\mathrm{vec}(B)|\Sigma \;\sim\; N(\mathrm{vec}(\bar{B}_T), \Sigma \otimes N_T^{-1}), \tag{3}$$

$$\Sigma \;\sim\; IW_n(\nu_T S_T, \nu_T), \tag{4}$$

where $\nu_T = T + \nu$, $N_T = N_0 + X'X$, $\bar{B}_T = N_T^{-1}(N_0\bar{B}_0 + X'X\hat{B})$,

$$S_T \;=\; \frac{\nu_0}{\nu_T} S_0 + \frac{T}{\nu_T}\hat{\Sigma} + \frac{1}{\nu_T}(\hat{B} - \bar{B}_0)' N_0 N_T^{-1} X'X(\hat{B} - \bar{B}_0),$$

$\hat{B} = (X'X)^{-1}X'Y$ and $\hat{\Sigma} = (Y - X\hat{B})'(Y - X\hat{B})/T$.

Consider a set of $n \times n$ matrices consisting of orthonormal column vectors:

$$\mathcal{U} \;=\; \left\{U \;=\; [u_1\; u_2\; \cdots\; u_n]' \;:\; u_i \in\; \Re^{n\times 1},\; < u_i, u_j > =\; 1(i=j),\; \forall\, i,j = 1,2,...,n\right\},$$

where $< \cdot, \cdot >$ denotes the inner product and $1(\cdot)$ is the indicator function. Let U denote a draw from the uniform distribution over $\mathcal{U}$. By construction, $UU' = I_n$. Define $\overline{A} = AU$ where A is the Cholesky decomposition of Σ, such that $AA' = \Sigma$. Note that $\overline{AA}' = AUU'A' = AA' = \Sigma$. When all structural shocks in the model are identified, we say that the model is fully identified; when only a subset of the structural shocks is identified, we say that the model is partially identified. For further discussion see, e.g., Rubio-Ramirez, Waggoner and Zha (2010) and Fry and Pagan (2011). The discussion in section 2.2 focuses on sign-identified models in which all structural shocks are jointly identified (corresponding to case 1 in Table 1). Partially identified structural VAR models based on sign restrictions (corresponding to case 2 in Table 1) are discussed in section 2.3.

2.2 Fully Identified Models

2.2.1 The Posterior Mode of Sign-Identified Structural Impulse Responses

Let $vech(A)$ denote the $n(n+1)/2 \times 1$ vector that consists of the on-diagonal elements and the below-diagonal elements of A and let $veck(U)$ denote the $n(n-1)/2 \times 1$ vector that consists of above-diagonal elements of U. Ignoring the intercept for notational convenience, let $B = [B_1\; \cdots\; B_p]'$. As shown in the appendix, because there is a one-to-one mapping between B and the reduced-form vector moving average coefficient matrices Φ_i, $i = 1, 2, ..., p$ (see equation 10.1.19 of Hamilton, 1994, p. 260, for example) and because Σ is nonsingular and U is orthonormal, there is a one-to-one mapping between the first $p+1$ structural impulse responses $\tilde{\Theta} = [\overline{A},\; \Phi_1\overline{A},\; \Phi_2\overline{A}, \cdots,\; \Phi_p\overline{A}]'$ on the one hand and the tuple formed by the reduced-form VAR parameters and the rotation matrix, $(B, \text{vech}(A), \text{veck}(U))$, on the other. The nonlinear function $\widetilde{\Theta} = h(B, \text{vech}(A), \text{veck}(U))$ is known. Using the change-of-variables method, the posterior density of $\widetilde{\Theta}$ can be written as

$$\begin{aligned} f(\tilde{\Theta}) &= \left|\frac{\partial[\text{vec}(B)'\; \text{vech}(A)'\; \text{veck}(U)']}{\partial \text{vec}(\tilde{\Theta})}\right| f(B, A, U) \\ &= \left(\left|\frac{\partial \text{vec}(\tilde{\Theta})}{\partial[\text{vec}(B)'\; \text{vech}(A)'\; \text{veck}(U)']}\right|\right)^{-1} \left|\frac{\partial \Sigma}{\partial A}\right| f(B, \Sigma, U) \\ &= \left(\left|\frac{\partial \text{vec}(\tilde{\Theta})}{\partial[\text{vec}(B)'\; \text{vech}(A)'\; \text{veck}(U)']}\right|\right)^{-1} \left|\frac{\partial \Sigma}{\partial A}\right| f(B|\Sigma) f(\Sigma), \end{aligned} \tag{5}$$

where B, $\Sigma = AA'$, and U are the unique values that satisfy the nonlinear function $\widetilde{\Theta} = h(B, \text{vech}(A), \text{veck}(U))$. Here f's denote posterior densities whose conditioning on the data is omitted for notational simplicity. Because U is uniformly distributed on $\mathcal{U}$ the following result holds:

Proposition 1. The posterior density of $\tilde{\Theta}$ is

$$f(\tilde{\Theta}) \propto \left(\left| \frac{\partial \text{vec}(\tilde{\Theta})}{\partial [\text{vec}(B)' \ \text{vech}(A)' \ \text{veck}(U)']} \right| \right)^{-1} \left| \frac{\partial \Sigma}{\partial A} \right| f(B|\Sigma) f(\Sigma). \tag{6}$$

The Jacobian matrix and its construction are discussed in the technical appendix.

Let $\times$ denote the set of structural impulse responses $\tilde{\Theta}$ that satisfy the sign restrictions. The modal model by construction is the admissible model that maximizes the posterior density of the sign-identified structural impulse responses. Because the impulse responses that do not satisfy the sign restrictions are discarded, the posterior density of the sign-identified impulse responses can be written as

$$g(\tilde{\Theta}) = \begin{cases} \frac{f(\tilde{\Theta})}{P(\tilde{\Theta} \in \mathbf{\Theta})} & \text{if } \tilde{\Theta} \in \mathbf{\Theta} \\ 0 & \text{if } \tilde{\Theta} \notin \mathbf{\Theta} \end{cases} \tag{7}$$

where $P(\tilde{\Theta} \in \mathbf{\Theta})$ is the posterior probability that $\tilde{\Theta} \in \mathbf{\Theta}$. Because $P(\tilde{\Theta} \in \mathbf{\Theta})$ does not depend on Θ, finding the mode of the posterior of the sign-identified structural impulse responses reduces to finding the maximum of the right hand side of (6) subject to the sign restrictions. In particular, it is not necessary to reweight $P(\tilde{\Theta} \in \mathbf{\Theta})$ to account for draws from the posterior that have been rejected.

In practice we proceed as follows:

Step 1. Take a random draw, (B, Σ), from the posterior of the reduced-form VAR parameters.

Step 2. For each (B, Σ), consider N random draws of the rotation U, and for each combination (B, Σ, U) compute the set of implied structural impulse responses $\tilde{\Theta}$.

Step 3. If $\tilde{\Theta}$ satisfies the sign restrictions, store the value of $\tilde{\Theta}$ and the value of $f(\tilde{\Theta})$. Otherwise discard $\tilde{\Theta}$.

Step 4. Repeat steps 2 and 3 M times and find the element of Θ that maximizes (6).

2.2.2 Credible Sets for Structural Impulse Response Functions

Define the $100(1-\alpha)\%$ highest posterior density (HPD) credible set by

$$S = \{\tilde{\Theta} \in \mathbf{\Theta} : f(\tilde{\Theta}) \geq c_\alpha\} \tag{8}$$

where $f(\tilde{\Theta})$ is the posterior density of $\tilde{\Theta}$ and c_α is the largest constant such that

$$P(S) \;\geq\; 1-\alpha$$

(see Definition 5 of Berger, 1985, p. 140).

In practice, we compute the $100(1-\alpha)\%$ HPD credible set as follows:

Step 1. Take a random draw, (B, Σ), from the posterior distribution of the reduced-form VAR parameters.

Step 2. For each (B, Σ), consider N random draws of the rotation U, and for each combination (B, Σ, U) compute the set of implied structural impulse responses $\tilde{\Theta}$.

Step 3. If $\tilde{\Theta}$ satisfies the sign restrictions, store the value of $\tilde{\Theta}$ and the value of $f(\tilde{\Theta})$. Otherwise discard $\tilde{\Theta}$.

Step 4. Repeat Steps 2 and 3 M times and sort the pairs $\{(\tilde{\Theta}, f(\tilde{\Theta}))\}$ in descending order by the value of $f(\tilde{\Theta})$. The $100(1-\alpha)\%$ HPD credible set consists of the set of $\tilde{\Theta}$'s contained in the first $(1-\alpha)Q$ sorted pairs, where Q refers to the number of models among the $M \cdot N$ draws that satisfy the sign restrictions.

Credible sets differ from conventional error bands for impulse responses in that the elements of the credible set are vectors representing the impulse response functions up to some prespecified horizon. There is no reason for credible sets to be dense necessarily. Rather a plot of the credible set will typically exhibit a shot-gun pattern.

2.3 Partially Identified Models

A common situation in VAR models of monetary policy is that the structural model is only partially identified in that we are concerned with identifying the policy shock, but no other structural shocks. If we are concerned with a subset of impulse response functions only, what matters for constructing the posterior mode is not the joint impulse response distribution, but the marginalized distribution obtained by integrating out responses to shocks that are not identified. To simplify the exposition we will focus on the case in which only impulse responses to one structural shock are identified. The method proposed below can be modified to allow for impulse responses to more than one shocks.

The sign-identified structural impulse responses, $\theta_1 = \Phi_1 a$, ..., $\theta_p = \Phi_p a$, where $a = Au$, do not single out a unique value of $\Phi_1, ..., \Phi_p$. That is because any Φ_i that satisfies

$$\Phi_i a \;=\; \theta_i,\; i = 1, ..., p \tag{9}$$

is consistent with θ_i and there are infinitely many of such Φ_i. Given np restrictions of the form

(9), one therefore needs to integrate out the joint posterior distribution of B with respect to $(n-1)p$ of the parameters in B. Our approach exploits the following proposition.

Proposition 2:

$$\begin{aligned} f(\theta_0, \theta_1, ..., \theta_p) &\propto \int f(\theta_0, \theta_1, ..., \theta_p | \tilde{u}, \Phi^{(2)}, \tilde{\Sigma}) f(\Phi^{(2)}, \tilde{\Sigma}) d(\tilde{u}, \Phi^{(2)}, \tilde{\Sigma}) \\ &= \int |a_1|^{np} |\tilde{A}| \; f(\Phi^{(1)} | B^{(2)}, \tilde{\Sigma}) f(B^{(2)}, \tilde{\Sigma}) d(\tilde{u}, B^{(2)}, \tilde{\Sigma}) \\ &= \int |a_1|^{np} |\tilde{A}| \; f(B^{(1)} | B^{(2)}, \tilde{\Sigma}) f(B^{(2)} | \tilde{\Sigma}) f(\tilde{\Sigma}) d(\tilde{u}, B^{(2)}, \tilde{\Sigma}), \end{aligned} \tag{10}$$

A heuristic proof of Proposition 2 may be constructed as follows. Consider a random draw of $\tilde{\Sigma}$ from the posterior distribution of Σ and condition on its Cholesky decomposition, say $\tilde{A}$. Given $\tilde{A}$, the sign-identified impulse responses in the impact period, θ_0, uniquely pin down the value of $\tilde{u}$.[3] Conditional on $\tilde{\Sigma}$, we draw the second through last columns of B_i, $i = 1, ..., p$, from the unconstrained posterior distribution. Postmultiplying

$$\Phi_1 \;=\; B_1, \tag{11}$$

by a yields

$$\theta_1 \;=\; B_1 a \tag{12}$$

Because a has a continuous distribution under our assumptions, $a_1 \neq 0$ with probability one. Thus it follows from (12) that the first column of B_1 is obtained from

$$B_{1,j1} \;=\; \frac{\theta_1 - \sum_{k=2}^{n} a_j B_{1,jk}}{a_1} \quad \text{for } j = 1, ..., n \tag{13}$$

with probability one, where $B_{i,jk}$ and a_j denote the (j,k)th element of B_i and the jth element of a_j, respectively. We now have a value of B_1 that is consistent with (9).

Next, we postmultiply

$$\Phi_2 \;=\; B_1 \Phi_1 + B_2, \tag{14}$$

by a to obtain

$$\theta_2 \;=\; B_1 \theta_1 + B_2 a, \tag{15}$$

from which we obtain

$$B_{2,j1} \;=\; \frac{\phi_2 - \sum_{k=2}^{n} a_j B_{2,jk}}{a_1} \quad \text{for } j = 1, ..., n \tag{16}$$

[3] Because $\tilde{A}$ is nonsingular, $\tilde{u} = \tilde{A}^{-1}\theta_0$ is uniquely defined and satisfies $\tilde{u}'\tilde{u} = u'A'\tilde{A}^{-1\prime}\tilde{A}^{-1}Au = u'A'^{-1}Au = u'u = 1$ where A is the original Cholesky decomposition of Σ from which a is obtained.

where

$$\phi_2 \;=\; \theta_2 - B_1\theta_1.$$

This provides a value of the first column of B_2 that is consistent with (9). This process may be repeated recursively until we reach B_p. In the last step, we postmultiply

$$\Phi_p \;=\; B_1\Phi_{p-1} + B_2\Phi_{p-2} + \cdots + B_p, \tag{17}$$

by a to obtain

$$\theta_p \;=\; B_1\theta_{p-1} + B_2\theta_{p-2} + \cdots + B_p a, \tag{18}$$

from which we obtain the first column of B_p as:

$$B_{p,j1} \;=\; \frac{\phi_p - \sum_{k=2}^{n} a_j B_{p,jk}}{a_1} \;\text{ for } j = 1, ..., n \tag{19}$$

where

$$\phi_p \;=\; \theta_p - B_1\theta_{p-1} - \cdots - B_{p-1}\theta_1$$

Therefore, given $\tilde{\Sigma}$ and the second through last columns of B_i for $i = 1, 2, ..., p$, the value of θ_1,..., θ_p implies a unique value of $\tilde{u}$ and of the first columns of each B_i (and vice versa).

Let $B^{(1)}$ denote the $pn(n-1)$ column vector obtained by stacking the first columns of the B_i's, $i = 1, ..., p$, and let $B^{(2)}$ denote the corresponding second through last columns. Then the marginal posterior density of the subset of structural impulse responses of interest is

$$\begin{aligned} f(\theta_0, \theta_1, ..., \theta_p) \;&\propto\; \int f(\theta_0, \theta_1, ..., \theta_p | u, \Phi^{(2)}, \tilde{\Sigma}) f(\Phi^{(2)}, \tilde{\Sigma}) d(\Phi^{(2)}, \tilde{\Sigma}) \\ &=\; \int |a_1|^{np} |\tilde{A}| \; f(\Phi^{(1)} | B^{(2)}, \tilde{\Sigma}) f(B^{(2)}, \tilde{\Sigma}) d(B^{(2)}, \tilde{\Sigma}) \\ &=\; \int |a_1|^{np} |\tilde{A}| \; f(B^{(1)} | B^{(2)}, \tilde{\Sigma}) f(B^{(2)} | \tilde{\Sigma}) f(\tilde{\Sigma}) d(B^{(2)}, \tilde{\Sigma}), \end{aligned} \tag{20}$$

where the first equality follows because the distribution of u is uniform, the second equality follows from applying the change-of-variables method to (13), (16),...,(19), and the last equality follows from using the block diagonality of the Jacobian matrix and applying the change-of-variables method to (11), (14),...,(17), $|\tilde{A}|$ follows from $\theta_0 = \tilde{A}\tilde{u}$, (16),...,(19) and the last

equality follows from applying the change-of-variables method to (11), (14),...,(17).[4] Q.E.D.

Proposition 2 allows us to estimate the posterior density of $\theta_0, \theta_1, ..., \theta_p$ (up to scale) by Monte Carlo integration (see, e.g., Robert and Casella 2004).

To summarize:

Step 1. Generate M draws of (B, Σ), from the posterior distribution of the reduced-form VAR parameters with N independent draws each of the rotations U.

Step 2. For each of the $M \cdot N$ draws compute the set of sign-identified structural impulse responses of interest, $\theta_0, \theta_1, ..., \theta_p$.

Step 2a For each of these sets of structural impulse response functions use (13),(16),...,(19) to construct L draws of $B^{(2)}$ and Σ, from which L draws of $B^{(1)}$ are constructed. Evaluate the value of

$$|a_1|^{np}|\tilde{A}| \; f(B^{(1)}|B^{(2)}, \tilde{\Sigma}). \tag{21}$$

Step 2b Compute the average of (21) across the L draws considered in step 2a. This Monte Carlo integration yields (up to scale) an estimate of the density $f(\theta_0, \theta_1, ..., \theta_p)$.

Given the marginal posterior density of the structural response functions of interest, we may then compute the mode and credible sets as outlined earlier.

2.4 Implementation

In practice, we specify a diffuse Gaussian prior for the VAR slope parameters such that the posterior mode of the slope parameters equals the least-squares estimator. For expository purposes, we follow the literature in setting $\alpha = 0.32$. The mode and the credible set of course may be computed from the same loop. The procedures were implemented in MATLAB or FORTRAN, depending on the computational requirements.

3 Posterior Modes versus Posterior Medians

Before discussing the extent to which median response functions differ from response functions of the modal model in practice, it is useful to develop some intuition about the limitations of median response functions. Let $\Theta_{ij,h}$ denote the response of variable i to structural shock j at

[4] In the fully identified case, the interpretation of the density $f(\Theta_0, ..., \Theta_p)$ as a posterior density is immediate given that there is a one-to-one mapping from $(vec(B)', vech(A)', veck(U)')$ to $(\Theta_0, ..., \Theta_p)$. In the partially identified case, the argument that $f(\theta_0, \theta_1, ..., \theta_p)$ is a posterior density is more involved. Note that marginalizing the joint posterior of the reduced-form parameters with respect to $B^{(2)}$ and the $\widetilde{\Sigma}$ that satisfies the identifying restrictions yields the marginal posterior of $B^{(1)}$ conditional on the identifying restrictions being satisfied. Because the mapping between this $B^{(1)}$ and the impulse responses characterized in equation (20) is one-to-one conditional on $B^{(2)}$ and $\widetilde{\Sigma}$, it follows that equation (20) is the posterior distribution of $\theta_0, \theta_1, ..., \theta_p$.

horizon h. One situation in which the vector of pointwise posterior medians is misleading as a measure of the central tendency of a structural impulse response function is when posterior draws of $\Theta_{ij,h}$, $h = 1, ..., H$, for given i and j, cross one another. This situation is illustrated in Figure 1. The figure focuses on the response of real GDP to an unanticipated monetary policy shock for a horizon of up to 36 months. This example was constructed by plotting a randomly chosen subset of nine admissible response functions for the Uhlig (2005) model discussed in section 4.2. It is evident that, for different horizons, the pointwise posterior median responses coincide with responses of different admissible models. Specifically, the median response function coincides with the response function of model 1 at horizons 4, 12-13, and 17-25; with the response function of model 2 at horizon 9 and 15-16; with that of model 3 at horizon 16; with that of model 5 at horizons 5-8, 10 and 27; with that of model 6 at horizons 0-3, 11 and 15, and with that of model 9 at horizons 28-36. There is, in fact, no structural model in the admissible set that could replicate the response pattern implied by the posterior median response function, rendering this statistic economically meaningless. Note that it is not sufficient to show that there are no cross-overs between impulse response functions. Similar problems may arise even in the absence of cross-overs when the order of the models differs for two response functions at some horizon h. Verifying the absence of these problems is not practically feasible, given the thousands of admissible models implied by typical sign-identified VAR models. A key advantage of the approach discussed in section 2 is that it avoids both of these problems by construction. Moreover, as discussed in the introduction, median response functions are not appropriate measures of the central tendency of the vector of impulse responses, making them difficult to intepret not only from an economic, but even from a purely statistical point of view. Focusing instead on the impulse response functions of the most likely structural model therefore seems natural in our context.[5]

It is important to understand the implications of this choice. Even if we focus on one impulse response coefficient at a time, the posterior mode will differ systematically from the posterior median. To illustrate these differences, consider a scalar process $y_t = \rho y_{t-1} + \varepsilon_t$, $\varepsilon_t \overset{iid}{\sim} N(0, 1)$, where $t = 1, ..., T$. In population, $\rho = 0$. We generate a random draw of length $T = 200$ from this process and postulate a diffuse Gaussian prior for ρ. Consider the impulse response at horizon h defined as $\partial y_{t+h}/\partial \varepsilon_t = \rho^h$. We focus on $h = 1$ for illustrative purposes. The left panel of Figure 2 shows the posterior density of ρ in the absence of sign restrictions. The distribution is centered on the least-squares estimate, which for this sample is slightly positive.

[5] Our emphasis on the mode of the joint distribution of models is not without precedence. The same approach is used in classical maximum likelihood estimation, for example. Likewise, there is precedence for focusing on the peak of the posterior in Bayesian analysis (see, e.g., Rubio-Ramirez, Waggoner and Zha 2009). There clearly are situations in which the use of the mode may be problematic (such as for bimodal distributions with equally high peaks in both tails and little probability mass in the center), but in that case the median or the mean would not be adequate summary statistics either. In any case, our empirical analysis below suggests that such extreme examples are not practically relevant.

By construction, the least-squares estimate is both the mode and the median of the posterior distribution of ρ. The right panel shows the corresponding posterior density after imposing the restriction $\rho > 0$, which is equivalent to restricting all impulse responses to be positive. The dotted line is the posterior median; the dashed line is the mode. The right panel illustrates that the posterior median in the sign-identified model by construction implies a larger impulse response than the posterior mode, which is the most likely value of ρ by construction. In other words, the posterior median is a biased estimator of the response associated with the most likely model even when dealing with one impulse response at a time.

This intuition carries over to impulse responses in sign-identified VAR models. The problem of using median responses is not merely one of overstating (or understating) the magnitude of the impulse responses. As we will illustrate in section 4.2, in higher-dimensional models with richer dynamics it is even possible to find examples in which impulse responses of the most likely admissible model are of a different sign than the pointwise posterior median. Such examples suggest that there is a need more generally to reexamine the findings of earlier studies that were based on posterior median responses from sign-identified VAR models. In the next section, we consider one fully identified model and one partially identified model for illustrative purposes.

4 Sign-Identified Models

4.1 Fully Identified Case: Oil Demand and Supply Shocks

There is a growing literature of models of the global market for crude oil based on fully identified structural VAR models. Here we follow Kilian and Murphy (2011) in specifying a monthly VAR(24) model with intercept for 1973.2-2008.9. The set of variables consists of monthly data for the percent change in global oil production, a measure of global real activity (in deviations from trend), and the real price of crude oil. The variables are defined and discussed in detail in Kilian (2009). We combine some of the key identifying assumptions from the existing literature. We first impose sign restrictions on the impact responses of each variable to each structural shock. An unanticipated oil supply disruption causes oil production to fall, the real price of oil to increase, and global real activity to fall on impact. An unanticipated increase in the flow demand for oil driven by the global business cycle causes global oil production, global real activity and the real price of oil to increase on impact. Other positive demand shocks (such as shocks to oil inventory demand driven by forward looking behavior) cause oil production and the real price of oil to increase on impact and global real activity to fall. Second, we bound the impact price elasticity of oil supply by 0.025, as suggested by Kilian and Murphy (2011). This elasticity can be expressed as the ratio of two impact responses. This identifying restriction is consistent with widely held views among oil economists that the short-run price

elasticity of oil supply is close to zero. Very similar results would be obtained if we doubled that bound. Finally, we follow Baumeister and Peersman (2009) in restricting the real price of oil to be positive for the first year in response to unanticipated oil supply disruptions and in response to positive oil demand shocks. We construct the posterior distribution of the impulse responses estimates based on $M = 5000$ draws from the reduced-form posterior distribution with $N = 20,000$ rotations each.

Figure 3 displays the responses of each variable to each shock in the modal model along with the corresponding 68% credible sets. The responses have been normalized such that each structural shock implies an increase in the real price of oil. All structural response function estimates are consistent with standard economic intuition. For example, a negative flow supply shock is associated with a persistent decline in oil production, a modest increase in the real price of oil, and a gradual modest decline in global real economic activity. A positive flow demand shock is associated with a persistent and hump-shaped response in both global real activity and the real price of oil and with little response in global crude oil production. Other demand shocks (such as shocks to oil inventory demand) cause a temporary increase in the real price of oil, a temporary decline in global real activity for about 20 months and little response in global crude oil production. The corresponding credible sets indicate considerable uncertainty about the price responses and to a lesser extent for the responses in real activity, whereas the credible sets for oil production responses are quite narrow. Nevertheless, several response functions are precisely enough estimated to conclude that the response differs from zero.

There also are important differences between the most likely estimates provided by the modal model and the conventional median response functions. Figure 4 demonstrates that the median response function may be closer to zero or further away from zero than the responses of the modal model. For example, median response functions overestimate the magnitude of the price response to *other demand shocks*, but underestimate the response of global real activity to the same shock. Moreover, pointwise 68% posterior error bands provide little protection against mis-characterizing the impulse response dynamics, as shown in Figure 5. In many cases, the response functions of the modal model are outside the pointwise error bands for at least some horizon. Figure 5 also illustrates that pointwise intervals tend to understate the estimation uncertainty compared with credible sets that capture the joint uncertainty over all impulse response functions. This example illustrates that the way estimates of sign-identified VAR models are represented matters for the interpretation of the data.

4.2 Partially Identified Case: Monetary Policy Shocks

Whereas the preceding example dealt with a fully identified model, this section considers an example of a partially identified model. We focus on the model of U.S. monetary policy proposed

by Uhlig (2005). Our focus in this section is not so much on whether this specific model is an appropriate model of U.S. monetary policy, but whether the method of statistical evaluation makes a difference for the economic interpretation of the results. The central question in Uhlig (2005) is what the effects of an unanticipated monetary contraction are on real output. We follow Uhlig in constructing a VAR(12) model without intercept. The set of variables consists of monthly U.S. data for the log of interpolated real GDP of the US, the log of the interpolated GDP deflator, the log of a commodity price index, total reserves, non-borrowed reserves and the federal funds rate. The sample period is 1965.1-2003.12 to ensure compatibility with Uhlig's original analysis.[6] The numerical stability of the results requires a fairly large number of draws, especially for M and L. We construct the posterior distribution of the impulse responses estimates based on $M = 5,000$ draws from the reduced-form posterior distribution with $N = 500$ rotations each. We set $L = 20,000$.

Figure 6 demonstrates that there are important differences between the median response estimates of the response of real output and the response in the modal model. Whereas Uhlig reported a peak median output response of 0.15 percentage points, for the same data, we obtain a peak response of almost 0.5 percentage points based on the modal model. Moreover, that peak value is near the upper end of the credible set and outside the conventional pointwise posterior error band. It should be noted that both the median estimate and the response estimate based on the modal model are counterintuitive in that a monetary tightening would be expected to cause a decline in real output over time rather than an increase. This outcome reflects the fact that the identifying assumptions are not overly informative. Even in Uhlig's original analysis, there was substantial pointwise probability mass on both negative and positive responses of real output. Our 68% credible set further widens the set of probable response functions.

The explicit reason why Uhlig (2005) did not impose further restrictions is that he wished to be as agnostic as possible about the response of real output. It has been shown that this approach is appropriate only to the extent that we view models in which real output increases in response to a monetary tightening as equally plausible a priori as models in which real output declines (see Kilian and Murphy 2011). In an effort to relax that assumption, in Figure 7 we consider an alternative set of models that impose an additional sign restriction on the response of real GDP after 6 months (and only at that horizon). This identifying assumption leaves the short-run as well as the longer-run response of real output unrestricted. The resulting modal model produces substantially different and more economically plausible results, including a cumulative drop in real GDP of -0.3 percentage points in the second quarter. The response estimate for the modal model is at the lower end of the credible set and again outside the conventional pointwise posterior error band. It also is substantially different from the response estimate obtained

[6] For a more detailed description of the data the reader is referred to Uhlig (2005). The data set was provided by Harald Uhlig.

from the traditional Cholesky decomposition. One difference is that the reduction in real GDP in Figure 7 is temporary, whereas traditional Cholesky models imply a much more persistent decline in real GDP. Even in this alternative model, however, the 68% credible set includes many positive real output responses, suggesting that the data are not informative about the response of real output. Likewise the other response functions are estimated only very imprecisely. We conclude that there remains substantial uncertainty about the effects of monetary policy shocks on real output, whereas there is strong evidence of the effects of oil demand shocks on the real price of oil in the earlier example.

Figures 8 and 9 elaborate further on the results in Figure 7. Figure 8 illustrates that there can be substantial differences between the median response function estimates and the response function estimates based on the modal model. For example, the decline in real GDP caused by an unanticipated monetary contraction is much larger in the modal model, at least in the short run. In some cases, the median and the modal response of real GDP differs not only in magnitude, but in sign. Finally, Figure 9 demonstrates that the modal model response may be outside the conventional pointwise 68% error bands. This is true in particular for the response of real GDP and to a lesser extent for the response of commodity prices and the own-response of the federal funds rate.

5 Exactly Identified Models

Our approach is not limited to sign-identified models. It can also be applied to exactly identified models. A case in point is a recursively identified model, in which the policy reaction function is ordered last in the system of equations. This type of model is commonly used in the monetary policy literature and indeed is the point of departure for the analysis in Uhlig (2005). The only difference to the earlier model is that the order of the variables matters. We follow Uhlig in ordering the variables of the VAR(12) model as real GDP, GDP deflator, commodity prices, federal funds rate, nonborrowed reserves and total reserves. The model in question is only partially identified. The object of interest are the responses to an orthogonalized federal funds rate shock.

Uhlig (2005) reports the posterior median impulse response functions. As in the sign-identified model, this measure of the central tendency of the structural impulse responses is valid only pointwise and need not correspond to the impulse response functions implied by the most likely structural model. Uhlig also reports pointwise 68% posterior error bands for this model. It is well known that these pointwise error bands fail to convey the true uncertainty surrounding these impulse response functions. This problem has been long recognized, but few practical alternatives have been proposed in the literature, which explains why these methods have remained the standard in the structural VAR literature (see, e.g., Lütkepohl 1990, Sims

and Zha 1999, Jorda 2009).

A solution to this problem is provided by a simplified version of our baseline procedure for sign-identified models. In the fully identified case, corresponding to case 3 in Table 1, it suffices to replace the rotation matrices in the procedure outlined in section 2.2 by the identity matrix and to modify the Jacobian of the transformation accordingly. We obtain

$$f(\tilde{\Theta}) \;\propto\; \left(\left| \frac{\partial[\text{vech}(\tilde{\Theta}_0)' \;\text{vec}(\left[\tilde{\Theta}_1, \tilde{\Theta}_2, ..., \tilde{\Theta}_p\right])']}{\partial[\text{vec}(B)' \;\text{vech}(A)']} \right| \right)^{-1} \left| \frac{\partial \Sigma}{\partial A} \right| f(B|\Sigma) f(\Sigma),$$

where $\tilde{\Theta}_0$ is the impulse response matrix in the impact period and $\tilde{\Theta}_1$, $\tilde{\Theta}_2$,...,$\tilde{\Theta}_p$ are the impulse response matrices at higher horizons. It follows from $\tilde{\Theta}_j = \Theta_j A$ for $j = 1, ..., p$, and equation (29) in the appendix that the determinant of the Jacobian reduces to $|A|^{np} \, |E_A'[(A \otimes I_n) + (I_n \otimes A) K_{n^2}] E_A|$, where E_A is defined in the appendix. On the basis of this result, we can evaluate the posterior as discussed earlier. This simplified procedure could be applied, for example, to the fully structural oil market VAR model in Kilian (2009).

The partially point-identified case considered by Uhlig (2005) corresponding to case 4 in Table 1 is more involved. Suppose that only impulse responses to the kth shock are considered (e.g., monetary policy shocks). Because the impulse responses in the impact period correspond to the kth column of the Cholesky decomposition, we need to draw Σ conditional on the kth column of its Cholesky decomposition when marginalizing the joint posterior density. We write the Cholesky decomposition as

$$A \;=\; \begin{bmatrix} A_{11} & 0 & 0 \\ A_{21} & A_{22} & 0 \\ A_{31} & A_{32} & A_{33} \end{bmatrix}.$$

where A_{11}, A_{21}, A_{22}, A_{31}, A_{32} and A_{33} are $n_1 \times n_1$, 1×1, $n_2 \times n_1$, $n_2 \times 1$ and $n_2 \times n_2$ matrices, $n_1 + 1 + n_2 = n$, and $[0 \;\; A_{22}' \;\; A_{32}']'$ corresponds to the kth column of A. Then

$$\Sigma \;=\; \begin{bmatrix} A_{11}A_{11}' & A_{11}A_{21}' & A_{11}A_{31}' \\ A_{21}A_{11}' & A_{21}A_{21}' + A_{22}A_{22}' & A_{21}A_{31}' + A_{22}A_{32}' \\ A_{31}A_{11}' & A_{31}A_{21}' + A_{32}A_{22}' & A_{31}A_{31} + A_{32}A_{32}' + A_{33}A_{33}' \end{bmatrix}.$$

Because

$$\begin{aligned} &\begin{bmatrix} A_{21}A_{21}' + A_{22}A_{22}' & A_{21}A_{31}' + A_{22}A_{32}' \\ A_{31}A_{21}' + A_{32}A_{22}' & A_{31}A_{31} + A_{32}A_{32}' + A_{33}A_{33}' \end{bmatrix} - \begin{bmatrix} A_{22}A_{22}' & A_{22}A_{32}' \\ A_{32}A_{22}' & A_{32}A_{32}' + A_{33}A_{33}' \end{bmatrix} \\ = &\begin{bmatrix} A_{21}A_{21}' & A_{21}A_{31}' \\ A_{31}A_{21}' & A_{31}A_{31} + A_{33}A_{33}' \end{bmatrix} \end{aligned} \tag{22}$$

is positive definite and A_{21}, A_{31} and A_{33} are unconstrained, one can draw $\widetilde{\Sigma}$ from the posterior distribution of $\widetilde{\Sigma}$ conditional on the kth Cholesky decomposition by drawing $\widetilde{\Sigma}$ from the unconditional posterior distribution of Σ and retaining the draws in which (22) is positive definite.

We draw B's that are consistent with the impulse responses $\theta_1, \theta_2, ..., \theta_p$ in the same way as in Proposition 2 except that we condition on the nth columns of B_i rather than the first column of B_i. This is because the first element of θ_0 is zero (unless the first shock is considered). Let $B^{(1)}$ denote the last columns of B_i's and let $B^{(2)}$ denote their first through second-to-last columns. Then we can summarize our approach in the following proposition:

Proposition 3:

$$\begin{aligned} f(\theta_0, \theta_1, ..., \theta_p) &\propto \int f(\theta_0, \theta_1, ..., \theta_p | \tilde{u}, \Phi^{(2)}, \tilde{\Sigma}) f(\Phi^{(2)}, \tilde{\Sigma}) d(\Phi^{(2)}, \tilde{\Sigma}) \\ &= \int |a_n|^{np} \; f(\Phi^{(1)} | B^{(2)}, \widetilde{\Sigma}) f(B^{(2)}, \tilde{\Sigma}) d(B^{(2)}, \tilde{\Sigma}) \\ &= \int |a_n|^{np} \; f(B^{(1)} | B^{(2)}, \tilde{\Sigma}) f(B^{(2)} | \tilde{\Sigma}) f(\widetilde{\Sigma}) d(B^{(2)}, \tilde{\Sigma}), \end{aligned} \tag{23}$$

where the integration is taken over $B^{(2)}$ and the $\widetilde{\Sigma}$ whose $(n_2+1) \times (n_2+1)$ lower-right submatrix satisfies the restriction (22).

The resulting procedure allows the user to construct credible sets for the structural impulse response functions for exactly identified structural VAR models that account for the joint uncertainty in the set of structural impulse response functions. It provides a convenient alternative to traditional pointwise error bands that is easy to implement. We illustrate this point for the responses to an unanticipated monetary tightening in the (partially) recursively identified VAR model used as a benchmark in Uhlig (2005). Uhlig's Figure 4 reported the pointwise median response functions and pointwise 68% posterior error bands for this model. We instead report the response functions of the modal model and the corresponding 68% joint credible set. All results are based on $M = 5,000$ and $L = 20,000$. Figure 10 shows that, even after accounting for the full uncertainty about the impulse response dynamics, the response functions of real GDP, of the federal funds rate and of nonborrowed reserves are precisely enough estimated to be economically informative at least at some horizons. The price puzzle and the puzzling initial increase in real GDP in response to an unanticipated monetary tightening, in contrast, can be attributed to estimation uncertainty. This result highlights the importance of simultaneous inference for all structural impulse responses. A user of pointwise 68% posterior error bands would have concluded that these puzzles cannot be explained merely by estimation uncertainty.

Figure 10 also illustrates that the median response functions may differ substantially from the response functions of the modal model. For example, the response of total reserves in the modal model differs not only in magnitude, but also in sign. The negative response of commodity

prices doubles in magnitude compared with the results in Uhlig (2005) and the response of the GDP deflator turns negative after only 24 months rather than 48 months. These differences highlight that accounting for the dependence across impulse response estimates is important even in point-identified models.

6 Concluding Remarks

Conventional approaches to summarizing the evidence from sign-identified impulse response functions based on quantiles of the pointwise posterior distribution of impulse responses lack a clear economic interpretation and fail to convey the uncertainty about the structural responses functions. We proposed an alternative approach based on a characterization of the most likely models in the set of admissible structural models. The approach discussed in this paper is explicitly Bayesian in nature. The use of Bayesian methods facilitates the interpretation of sign-identified VAR models and is standard in this literature. In fact, it is not clear how to extend our approach to evaluating sign-identified VAR models to frequentist settings.[7]

For exactly identified VAR models, in contrast, one could construct joint asymptotic normal approximations of the distribution of the impulse responses (see, e.g., Lütkepohl 1990, Mittnik and Zadrozny 1993). This allows joint inference based on the Bonferroni principle, although that method is impractical given its low power (see Lütkepohl 1990). As an alternative, Jorda (2009) recently proposed the construction of joint confidence intervals based on Scheffé's method. Likewise, our approach in this paper could be adapted to provide joint confidence sets for structural impulse response functions in exactly identified models based on the conventional asymptotic normal approximation.

Finally, in related work, Sims and Zha (1999) proposed a joint impulse response error band based on the mean and variance of the joint posterior distrubution of the structural impulse responses in exactly identified models. Their Bayesian error band could also be computed for impulse response functions in sign-identified models, but has no obvious rationale in the absence of normality. The reason is that only under normality the first two moments provide an adequate characterization of the distribution. It is well known that the finite-sample posterior distribution of impulse responses tends to be highly nonnormal. In the case of exactly identified models, the joint impulse response distribution is at least aymptotically normal. In the sign-identified model, in contrast, asymptotic normality breaks down, as shown by Moon, Schorfheide, Granziera, and Lee (2009), making this approach less appealing. In either case, the probability content of the 68% joint error band of Sims and Zha (1999) will be at best approximately 68% in finite

[7] Moon, Schorfheide, Granziera, and Lee (2009) recently proposed pointwise frequentist confidence intervals for impulse response estimates obtained from sign-identified VAR models. They showed that Bayesian and classical inference do not coincide even asymptotically in sign-identified VAR models. Moon et al. do not address the question of how to construct joint confidence regions or the question of which response estimates are most likely.

samples. In contrast, our credible set has exactly 68% probability content in finite samples by construction. An interesting question for future research would be to see how well the error band proposed by Sims and Zha approximates its nominal probability content in the exactly identified and in the sign-identified case.

Two empirical examples illustrated that the way information from structural VAR model estimates is represented matters. Responses based on the modal model and the associated credible sets can generate very different assessments of the evidence than traditional methods. It may seem that there could be alternatives to the use of the posterior mode in characterizing the structural impulse responses of the most likely model. We motivated our analysis by observing that posterior median response functions are not appropriate for characterizing the central tendency of structural impulse response functions. The use of the mode of the posterior is not without precedent in Bayesian analysis (see, e.g, Rubio-Ramirez, Waggoner and Zha 2010, p. 684). One obvious alternative statistic might seem to be the vector of posterior means. The posterior mean is statistically well-defined in the vector case and does not suffer from the second shortcoming discussed in the introduction. Given a finite set of admissible posterior draws, the posterior mean in general will not correspond to any one structural model in the set of admissible models, however, making it vulnerable to the first shortcoming. One way of addressing this concern would be to search for the admissible structural model that produces impulse responses closest to those of the posterior mean response functions, building on an idea in Fry and Pagan (2011).[8] We do not pursue this idea in this paper. One disadvantage of the posterior mean approach compared with focusing on the most likely or modal model is that there is no natural way of constructing joint credible sets in the sign-identified model. In contrast, our approach in this paper has the advantage of allowing for a unified treatment of estimation and inference in both the exactly identified and the sign-identified VAR model.

[8] Fry and Pagan (2011) suggested to minimize the distance to the median response function. That idea is not advisable because the posterior median response function is not a well defined statistical object, as discussed earlier, but the same idea could be applied without problems to the posterior mean response function.

Technical Appendix

Proof of the result that the mapping between the first $p+1$ structural impulse responses $\tilde{\Theta} = [\bar{A}, \Phi_1\bar{A}, \cdots, \Phi_p\bar{A}]'$ and $[B, \text{vech}(A), \text{veck}(U)]$ is one-to-one: Because there is a one-to-one mapping between B and the reduced-form vector moving average coefficient matrices Φ_i, $i = 1, 2, ..., p$ (see equation 10.1.19 of Hamilton, 1994, p.260, for example), establishing this result comes down to showing that the mapping between $\bar{A}$ and the pair vech(A) and veck(U) is one-to-one. We prove this result in two steps. Recall that - because the orthonormality restriction $UU' = I_n$ imposes $n(n+1)/2$ restrictions on the $n \times n$ matrix U - one can recover all elements of U from its upper diagonal elements, $veck(U)$. First, suppose that $\bar{A}_1 \neq \bar{A}_2$ but $A_1 = A_2$. Then $\bar{A}_j\bar{A}_j' = A_jU_jU_j'A_j' = A_jA_j'$ regardless of the value of U_j because $U_jU_j' = I_n$ for $j = 1, 2$. Because $A_1 = A_2$ this implies $\bar{A}_1\bar{A}_1' = \bar{A}_2\bar{A}_2'$ which contradicts the assumption that $\bar{A}_1 \neq \bar{A}_2$. Hence, $\bar{A}_1 \neq \bar{A}_2$ implies vech$(A)_1 \neq$ vech$(A)_2$. Second, suppose that $[A_1, U_1] \neq [A_2, U_2]$ but $\bar{A}_1 = \bar{A}_2$. Because $A_1A_1' = \bar{A}_1\bar{A}_1 = \bar{A}_2\bar{A}_2' = A_2A_2'$ and the Cholesky decomposition is uniquely determined for positive definite matrix Σ, it has to be the case that $A_1 = A_2$ and $U_1 \neq U_2$. Because A_j is nonsingular for $j = 1, 2$, however, $U_1 = A_1^{-1}\bar{A}_1 = A_2^{-1}\bar{A}_2 = U_2$, which is a contradiction. Hence, $[A_1, U_1] \neq [A_2, U_2]$ implies $\bar{A}_1 \neq \bar{A}_2$.

Derivation of the Jacobian matrix in Proposition 1: Let D_n denote the $n^2 \times n(n+1)/2$ duplication matrix of zeros and ones such that vec$(M) = D_n$vech(M) for $n \times n$ symmetric matrix M (see Magnus and Neudecker, 1999, pp.49). D_n^+ denotes the Moore-Penrose inverse of D_n so that we can write vech$(M) = D_n^+$vec(M). K_n denotes the $n^2 \times n^2$ communication matrix such that vec$(M') = K_n$vec(M) for $n \times n$ matrix M (see Magnus and Neudecker, 1999, pp.46–47).

Let $\Phi = [\Phi_1' \ \Phi_2' \ \cdots \ \Phi_p']'$ and $\overline{A}_1 = [\overline{A}'\Phi_1' \ \overline{A}'\Phi_2' \ \cdots \ \overline{A}'\Phi_p']'$ where Φ_i is the ith reduced-form vector moving average coefficient matrix. Because $\overline{A}_1 = \Phi AU$,

$$\frac{\partial \begin{bmatrix} \text{vec}(\overline{A}) \\ \text{vec}(\overline{A}_1) \end{bmatrix}}{\partial[\text{vec}(A)' \ \ \text{vec}(U)' \ \ \text{vec}(\Phi)']} = \begin{bmatrix} U' \otimes I_n & I_n \otimes A & O_{n^2 \times n^2 p} \\ U' \otimes \Phi & I_n \otimes \Phi A & A \otimes I_{np} \end{bmatrix}. \tag{24}$$

We need to replace the partial derivatives with respect to vec(A) and vec(U) in (24) with those with respect to vech(A) and veck(U). It follows from $UU' = I_n$ that

$$[(U \otimes I_n) + (I_n \otimes U)K_n]dvec(U) \ = \ 0_{n^2 \times 1}, \tag{25}$$

from which we obtain

$$D_n^+[(U \otimes I_n) + (I_n \otimes U)K_n]dvec(U) \ = \ 0_{n(n+1)/2 \times 1}. \tag{26}$$

Let E_h and E_k denote the $(n^2 \times n(n+1)/2)$ and $(n^2 \times n(n-1)/2)$ matrices of zeros and ones such that

$$vec(U) \;=\; [E_h \; E_k] \begin{bmatrix} \text{vech}(U) \\ \text{veck}(U) \end{bmatrix}.$$

Then (26) can be written as

$$D_n^+[(U \otimes I_n) + (I_n \otimes U)K_n]E_h \text{vech}(U) + D_n^+[(U \otimes I_n) + (I_n \otimes U)K_n]E_k \text{veck}(U) \;=\; 0_{n(n+1)/2 \times 1}. \tag{27}$$

Applying the implicit function theorem to (27), the Jacobian of vec(U) with respect to veck(U) can be written as

$$J_U \;=\; E_k - E_h \left\{ D_n^+[(U \otimes I_n) + (I_n \otimes U)K_n]E_h \right\}^{-1} D_n^+[(U \otimes I_n) + (I_n \otimes U)K_n]E_k \tag{28}$$

Thus, it follows from (24) and (28) that

$$J_1 \;\equiv\; \frac{\partial \begin{bmatrix} \text{vec}(\overline{A}) \\ \text{vec}(\tilde{\theta}) \end{bmatrix}}{\partial[\text{vech}A' \;\; \text{veck}(U)' \;\; \text{vec}(\Phi)']} \;=\; \begin{bmatrix} (U' \otimes I_n)D_n & (I_n \otimes A)J_U & O_{n^2 \times n^2 p} \\ (U' \otimes \Phi)D_n & (I_n \otimes \Phi A)J_U & A' \otimes I_{np} \end{bmatrix}. \tag{29}$$

Because (29) is block-diagonal, its determinant is given by the product of determinants:

$$\begin{aligned} |J_1| \;&=\; |(U' \otimes I_n)D_n \;\; (I_n \otimes A)J_U| \; |A' \otimes I_{np}| \\ &=\; |(U' \otimes I_n)D_n \;\; (I_n \otimes A)J_U| \; |A|^{np}. \end{aligned} \tag{30}$$

Because of the recursive relationships (11), (14),...,(17) between B and Φ, the Jacobian matrix of θ with respect to B is block-diagonal and each diagonal block has unit determinant. Thus

$$|J_2| \;\equiv\; \left| \frac{\partial \text{vec}(\Phi)}{\partial \text{vec}(B)} \right| \;=\; 1. \tag{31}$$

Since the Jacobian of vec(Σ) with respect to vec(A) is

$$[(A \otimes I_n) + (I_n \otimes A)K_n], \tag{32}$$

the determinant of the Jacobian of vech(Σ) with respect to vech(A) is given by

$$|J_3| \;\equiv\; |D_n'[(A \otimes I_n) + (I_n \otimes A)K_n]D_n| \tag{33}$$

Therefore it follows from (30), (31) and (33) that the determinant of the Jacobian in (6) is given by the product of $|J_1|$ in (30) and $|J_3|$ in (33) subject to sign.

References

1. Baumeister, C., and G. Peersman (2009), "Sources of the Volatility Puzzle in the Crude Oil Market," mimeo, Ghent University.
2. Berger, J. (1985), *Statistical Decision Theory and Bayesian Analysis*, Springer-Verlag: New York.
3. Caldara, D. (2011), "The Analytics of SVARs: A Unified Framework to Measure Fiscal Multipliers," mimeo, Stockholm
4. Canova, F. and D. De Nicolo (2002), "Monetary Disturbances Matter for Business Cycle Fluctuations in the G-7," *Journal of Monetary Economics*, 49, 1131-1159.
5. Canova, F., and E. Pappa (2007), "Price Dispersions in Monetary Unions: The Role of Fiscal Shocks," *Economic Journal*, 117, 713-737.
6. Canova, F., and M. Paustian (2007), "Measurement with Some Theory: Using Sign Restrictions to Evaluate Business Cycle Models," mimeo, Pompeu Fabra.
7. Chauduri, P. (1996), "On a Geometric Notion of Quantiles for Multivariate Data," *Journal of the American Statistical Association*, 91, 862-872.
8. Dedola, N. and S. Neri (2007), "What Does a Technology Shock Do? A VAR Analysis with Model-Based Sign Restrictions" *Journal of Monetary Economics*, 54, 512-549.
9. Faust, J. (1998), "On the Robustness of Identified VAR Conclusions about Money," *Carnegie-Rochester Conference Series on Public Policy* 49, 207-244.
10. Fry, R. and A. Pagan (2011), "Sign Restrictions in Structural Vector Autoregressions: A Critical Review," forthcoming: *Journal of Economic Literature*.
11. Fujita, S. (2011), "Dynamics of Worker Flows and Vacancies: Evidence from the Sign Restriction Approach," *Journal of Applied Econometrics*, 26, 89-121.
12. Gambetti, L, Pappa, E., and F. Canova (2008), "The Structural Dynamics of U.S. Output and Inflation: What Explains the Changes?" *Journal of Money, Credit and Banking*, 40, 369-388.
13. Hamilton, J.D. (1994), *Time Series Analysis,* Princeton University Press: Princeton, NJ.
14. Jordà, Ò. (2009), "Simultaneous Confidence Regions for Impulse Responses," *Review of Economics and Statistics*, 91, 629-647.
15. Kilian, L. (2009), "Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market," *American Economic Review*, 99, 1053-1069.
16. Kilian, L., and D. Murphy (2010), "The Role of Inventories and Speculative Trading in the Global Market for Crude Oil," mimeo, University of Michigan.

17. Kilian, L., and D. Murphy (2011), "Why Agnostic Sign Restrictions Are Not Enough: Understanding the Dynamics of Oil Market VAR Models," forthcoming: *Journal of the European Economic Association.*

18. Koltchinskii, V.I. (1997), "M-Estimation, Convexity and Quantiles," *Annals of Statistics*, 25, 435-477.

19. Liu, R.Y., J.M. Parelius, and K. Singh (1999), "Multivariate Analysis by Data Depth: Descriptive Statistics, Graphics, and Inference," *Annals of Statistics*, 27, 783-858.

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

21. Magnus, J.R., and H. Neudecker (1999), *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Wiley: Chichester, UK.

22. Mittnik, S., and P.A. Zadrozny (1993), "Asymptotic Distributions of Impulse Responses, Step Responses, and Variance Decompositions of Estimated Linear Dynamic Models," *Econometrica*, 61, 857-870.

23. Moon, H.R., Schorfheide, F., Granziera, E., and M. Lee (2009), "Inference for VARs Identified with Sign Restrictions," mimeo, University of Southern California.

24. Mountford, A., and H. Uhlig (2009), "What are the Effects of Fiscal Policy Shocks?" *Journal of Applied Econometrics*, 24, 960-992.

25. Pappa, E. (2009), "The Effects of Fiscal Shocks on Employment and Real Wages," *International Economic Review*, 50, 217-244.

26. Peersman, G., and R. Straub (2009) "Technology shocks and Robust Sign Restrictions in a Euro Area SVAR," *International Economic Review*, 50, 727-750.

27. Robert, C.P., and G. Casella (2004), *Monte Carlo Statistical Methods*, 2nd ed., Springer: New York, NY.

28. Rubio-Ramirez, J., Waggoner, D., and T. Zha. (2010) "Structural Vector Autoregressions: Theory for Identification and Algorithms for Inference," *Review of Economic Studies*, 77, 665-696.

29. Scholl, A. and H. Uhlig (2008), "New Evidence on the Puzzles: Results from Agnostic Identification on Monetary Policy and Exchange Rate," *Journal of International Economics*, 76, 1-13.

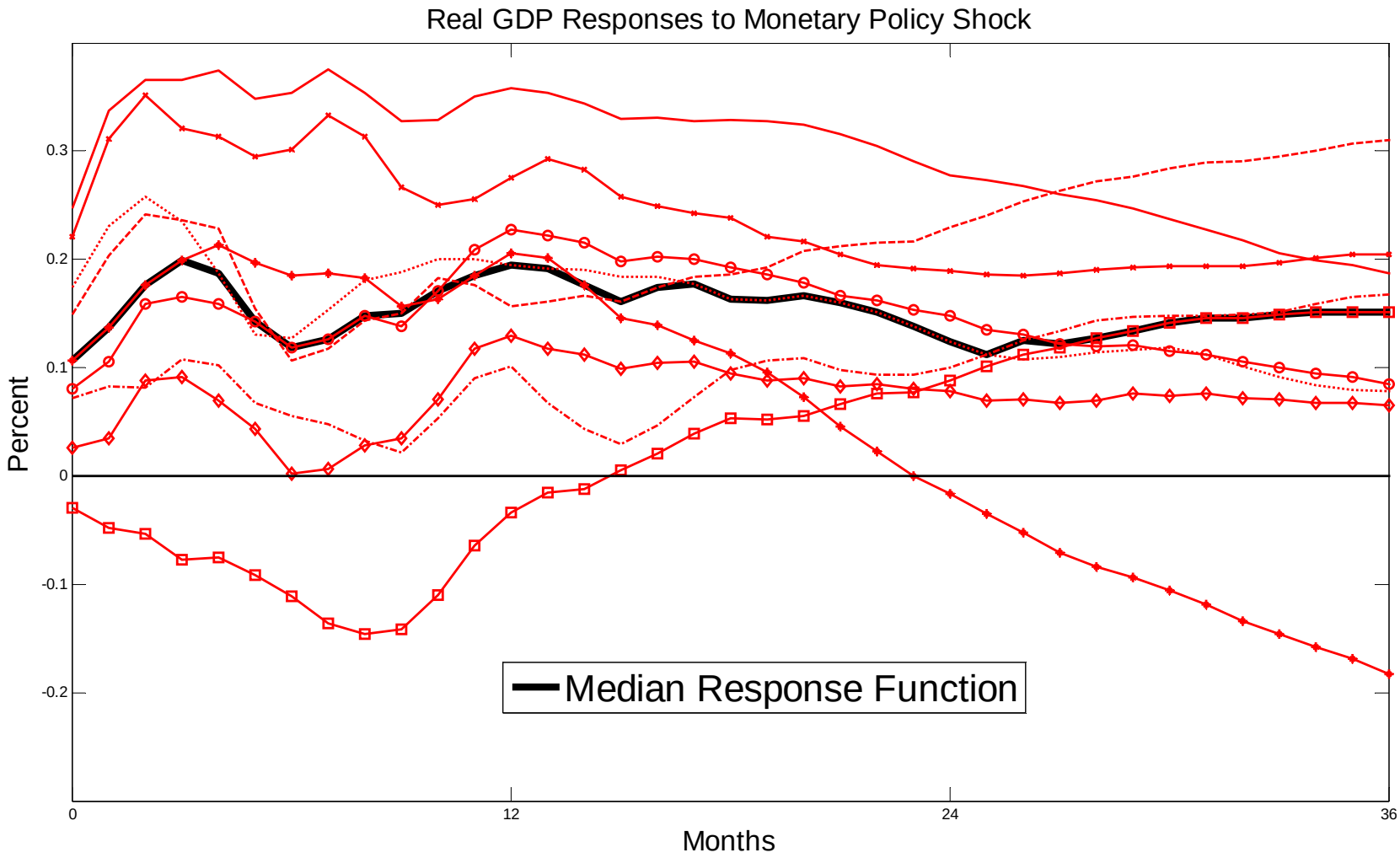
30. Sims, C.A., and T. Zha (1999), "Error Bands for Impulse Responses," *Econometrica*, 67, 1113-1155.

31. Uhlig, H. (2005), "What are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure," *Journal of Monetary Economics,* 52, 381–419.

Table 1: Four Types of Structural VAR Models

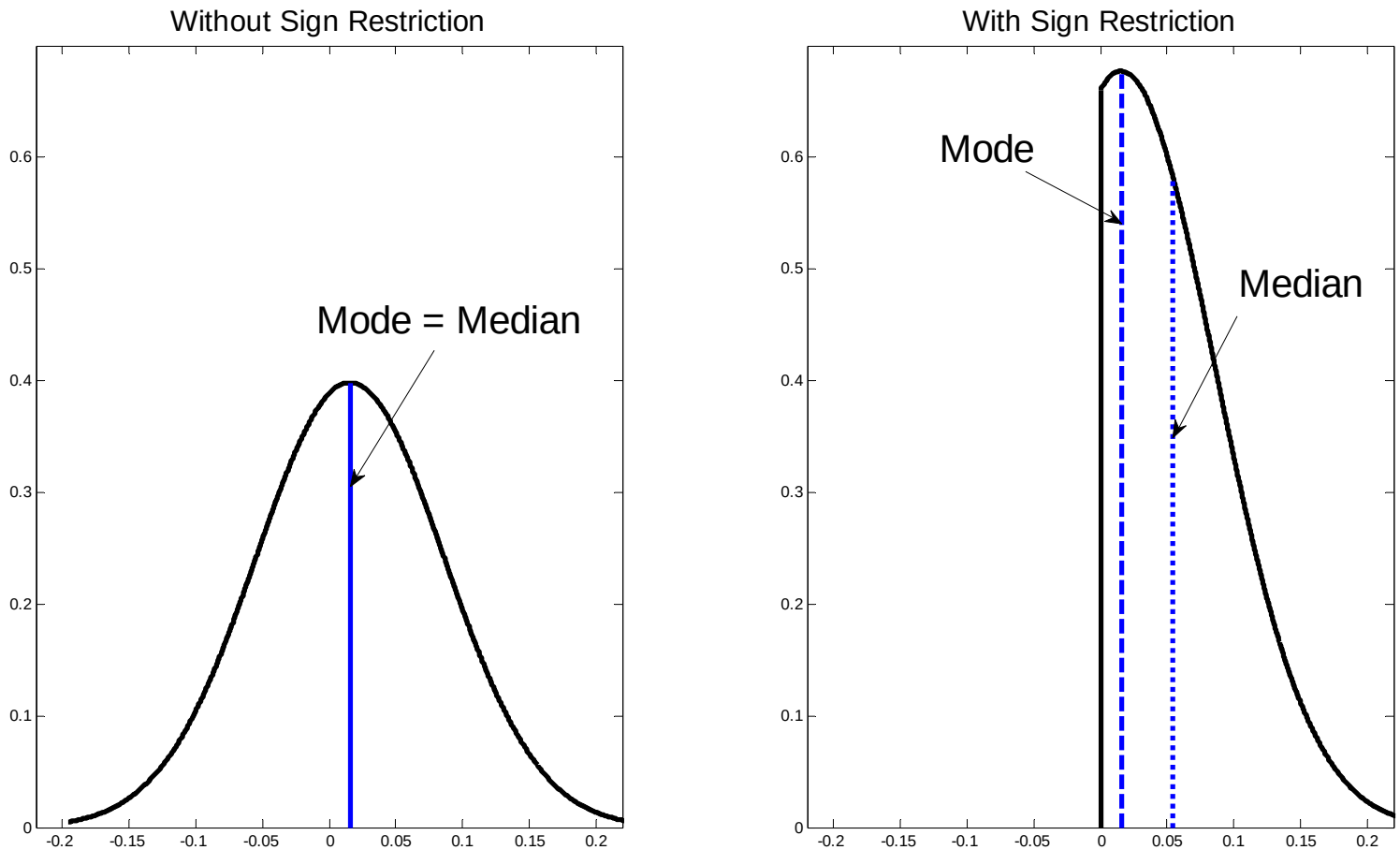
	Fully Identified	Partially Identified
Sign-Identified	Case 1	Case 2
Exactly Identified	Case 3	Case 4

Figure 1: Randomly Selected Response Functions from a Sign-Identified VAR Model



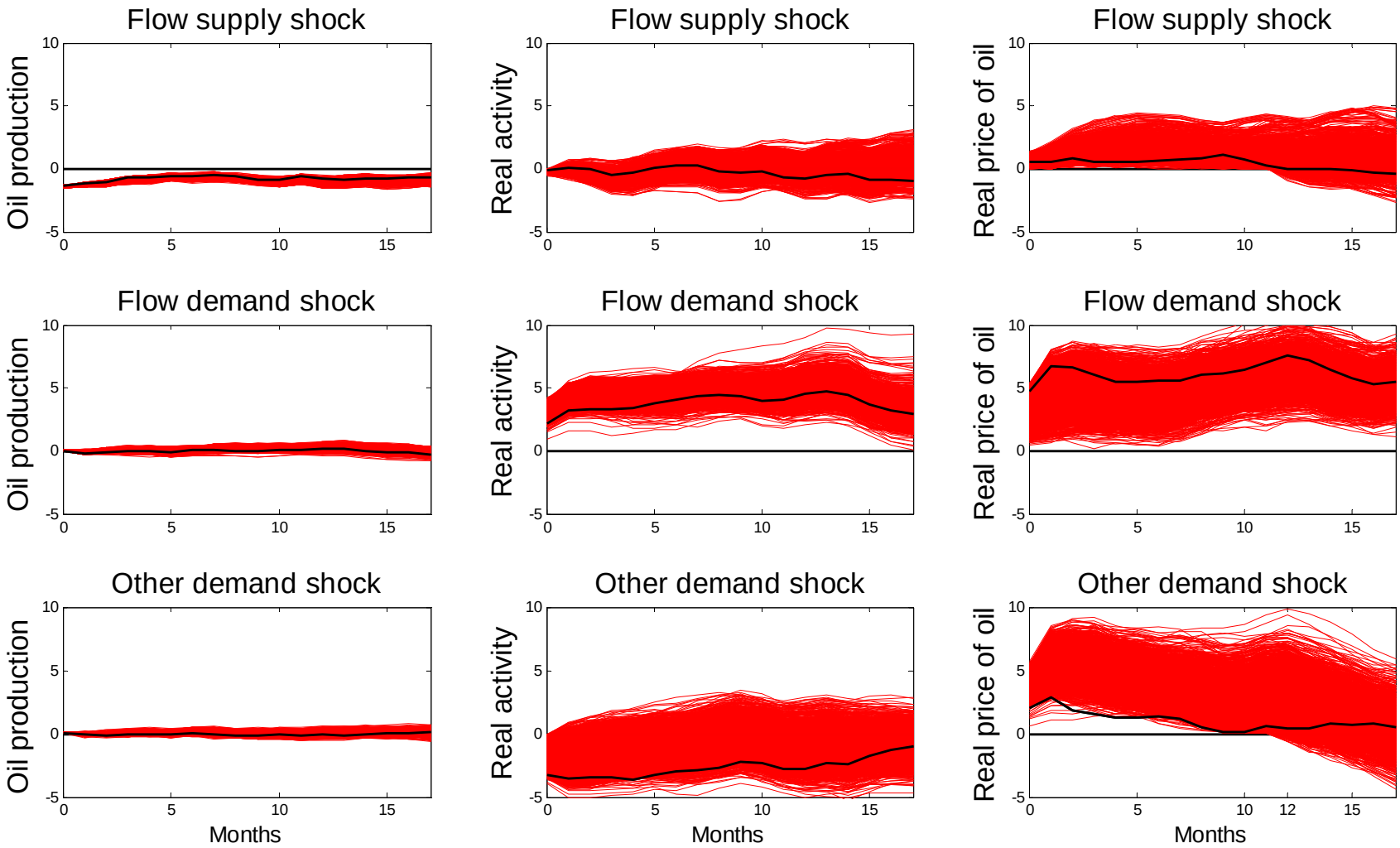
NOTES: Based on nine randomly selected responses from the posterior of the model used as an empirical example in section 4.2. The median response function is constructed from the pointwise posterior medians. It coincides with responses from six different admissible structural models depending on the horizon.

Figure 2: Posterior Mode versus Posterior Median for a Scalar Impulse Response Coefficient Posterior Density of the AR(1) Slope Parameter



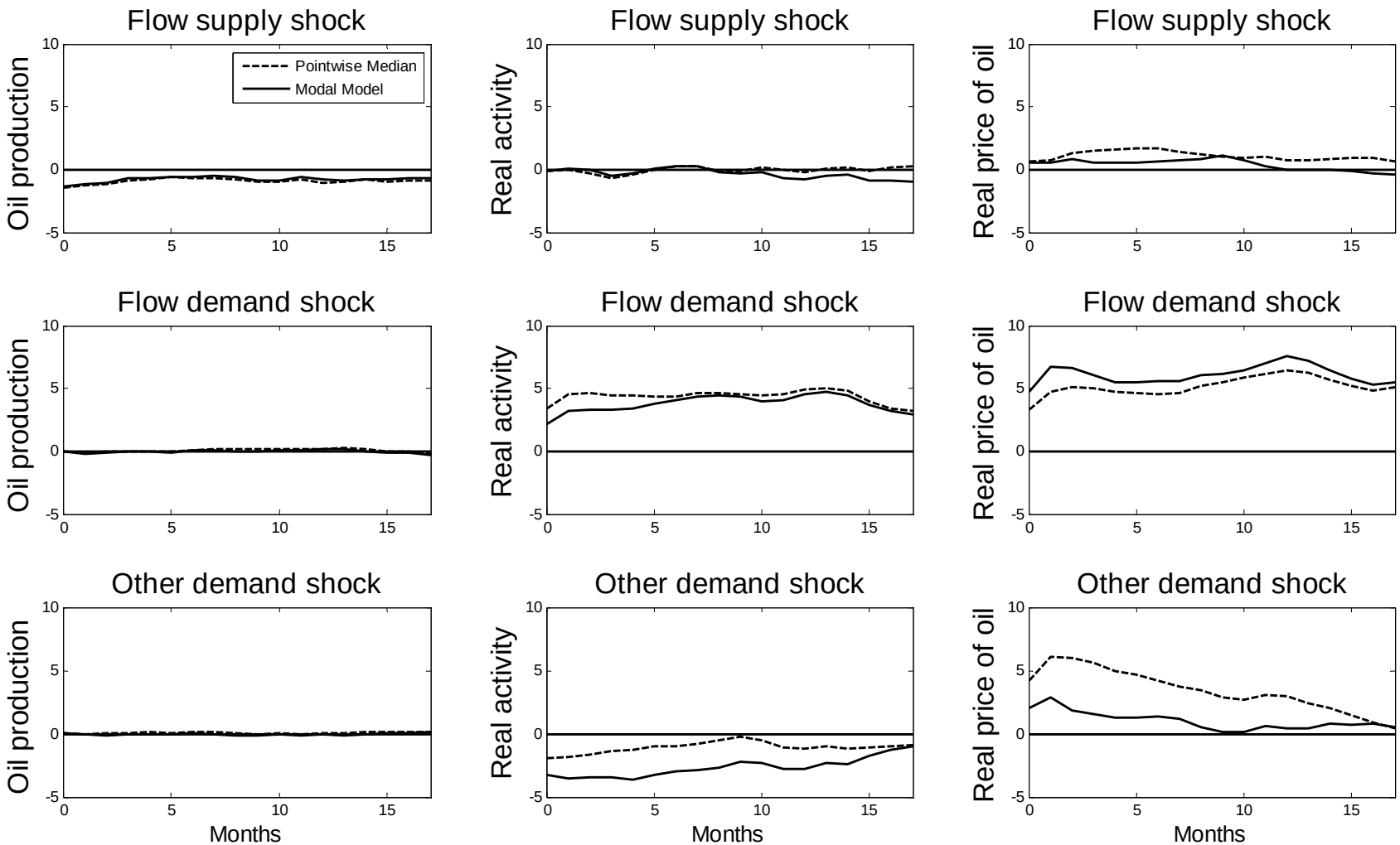
NOTES: Based on a draw of length $T = 200$ from the process $y_t = \rho y_{t-1} + \varepsilon_t, t = 1,...,T,$ where $\varepsilon_t \overset{iid}{\sim} N(0,1)$ and $\rho = 0.$ The posterior of ρ is based on a conventional diffuse Gaussian prior for $\rho.$ $\partial y_{t+h} / \partial \varepsilon_t = \rho^h.$ We focus on horizon 1 for expository purposes.

Figure 3: Structural Responses in the Sign-Identified Oil Market Model
Response Functions in the Modal Model and 68% Joint Regions of High Posterior Density



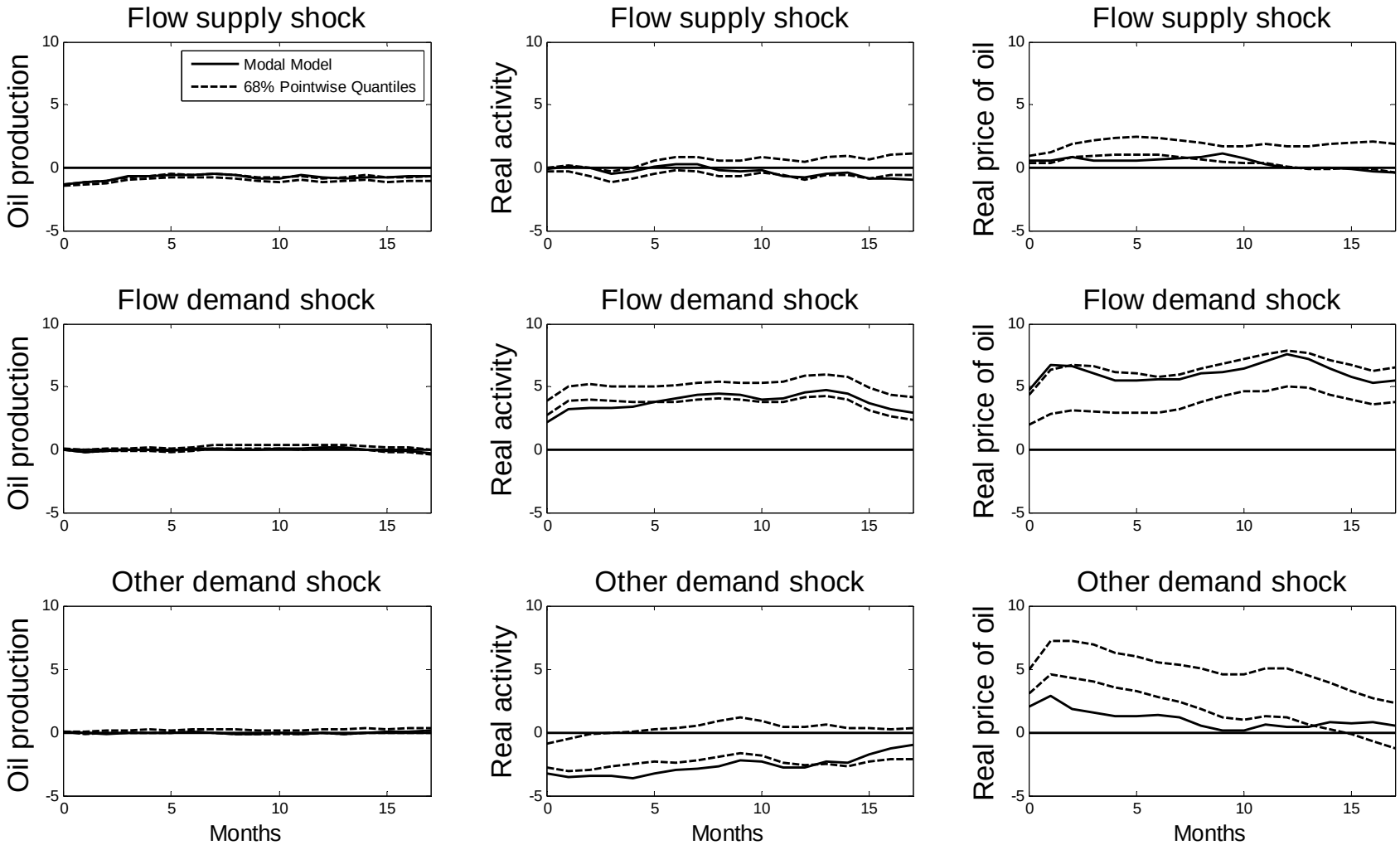
NOTES: The data and model are described in the text. All shocks have been normalized to imply an increase in the real price of oil.

Figure 4: Structural Responses in the Sign-Identified Oil Market Model
Response Functions in the Modal Model and Pointwise Median Responses



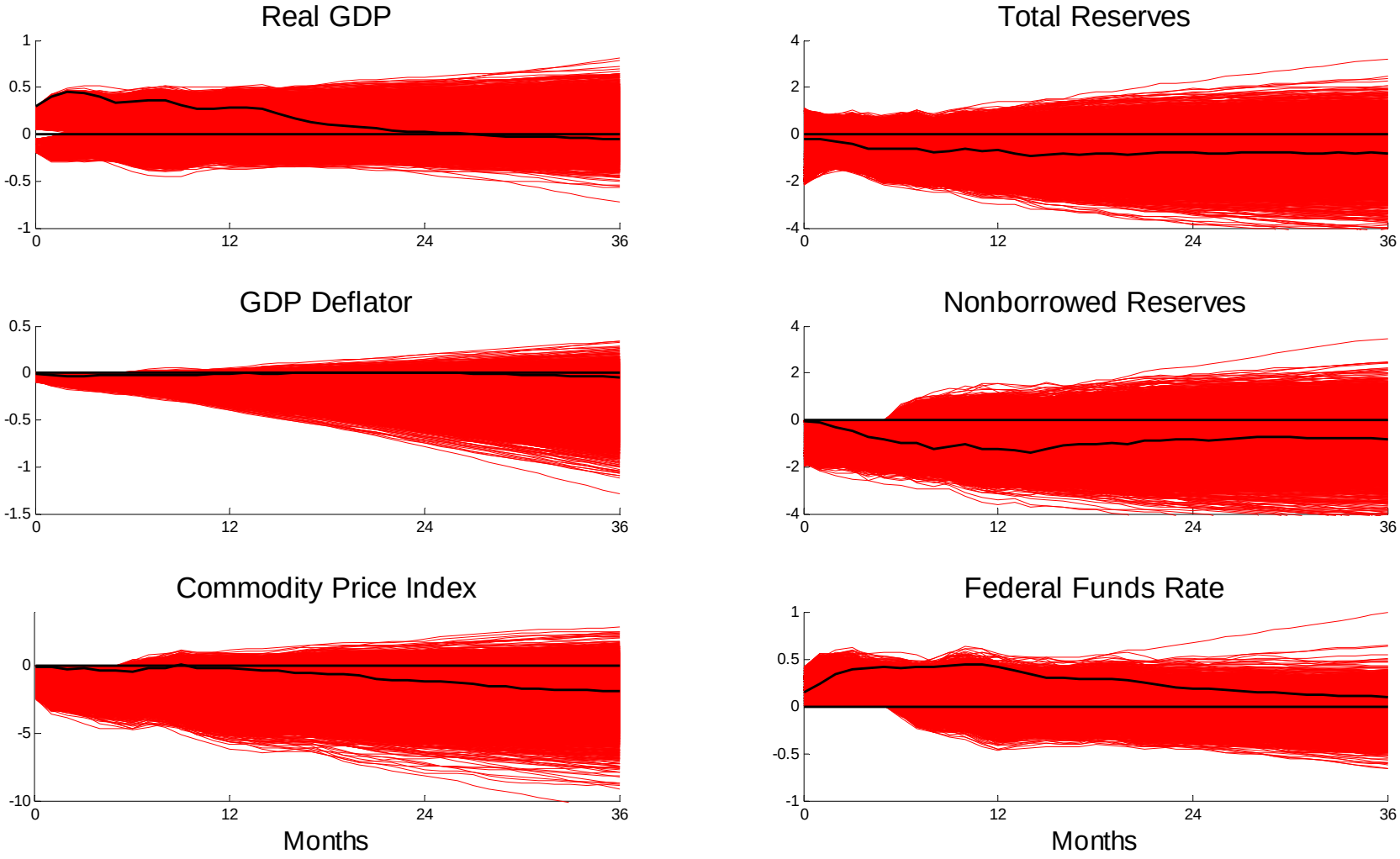
NOTES: See Figure 3.

Figure 5: Structural Responses in the Sign-Identified Oil Market Model
Response Functions in the Modal Model and Pointwise 68% Posterior Error Bands



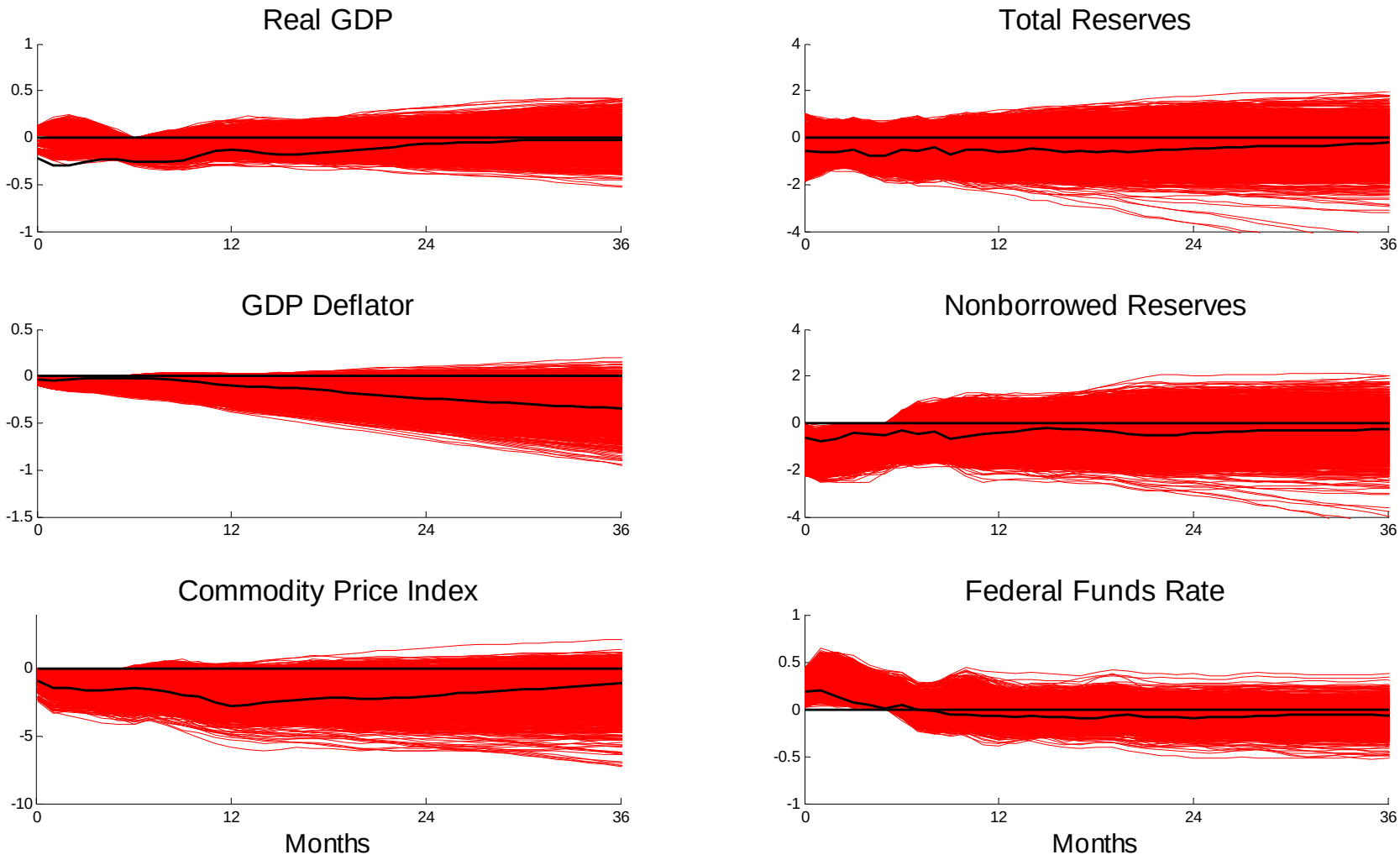
NOTES: See Figure 3.

Figure 6: Responses to a Monetary Policy Tightening in the Uhlig (2005) Sign-Identified Model Response Functions in the Modal Model and 68% Joint Regions of High Posterior Density



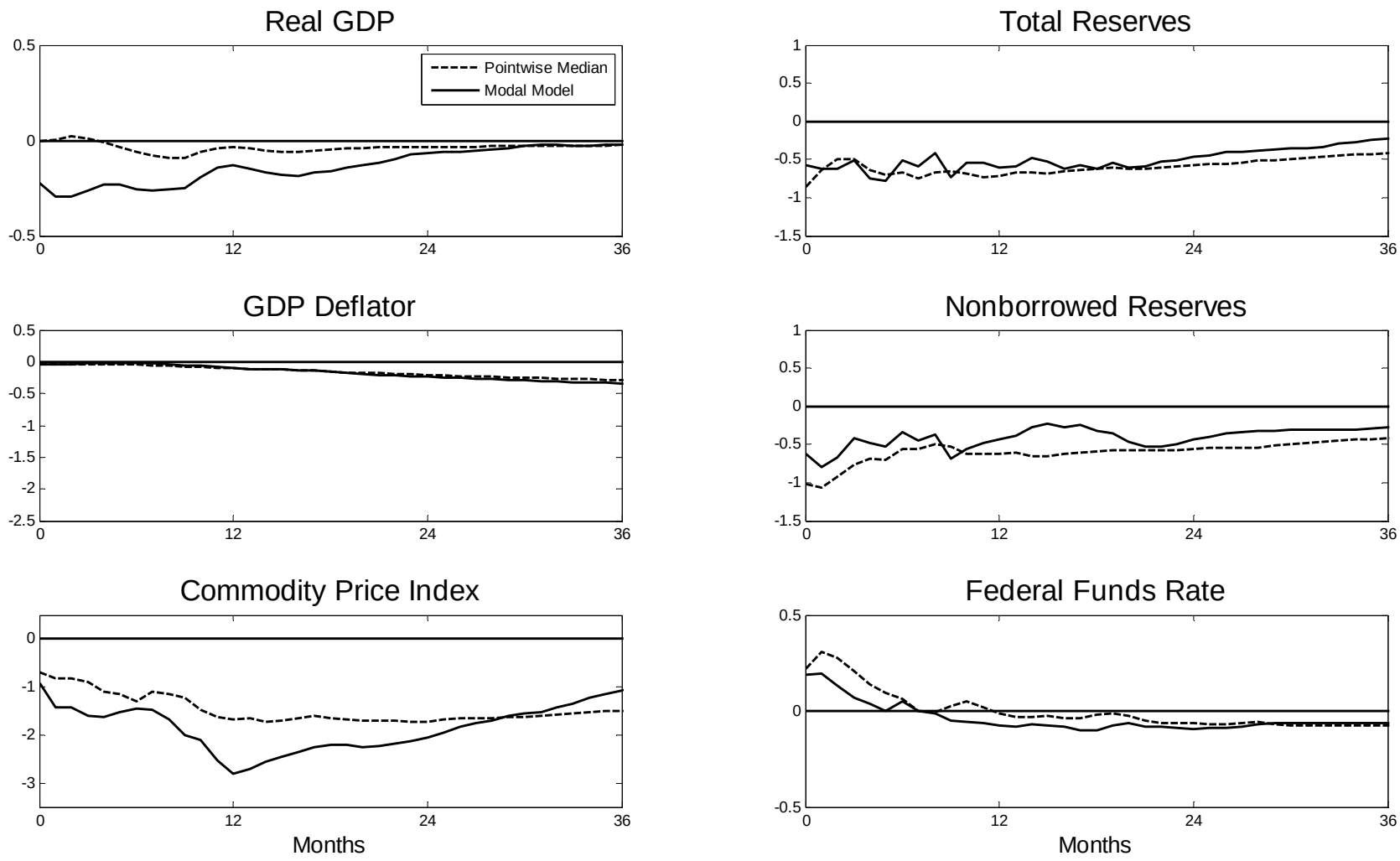
NOTES: The model specification and sign restrictions are the same as in Figure 6 of Uhlig (2005).

Figure 7: Responses to a Monetary Policy Tightening in the Modified Sign-Identified Model Response Functions in the Modal Model and 68% Joint Regions of High Posterior Density



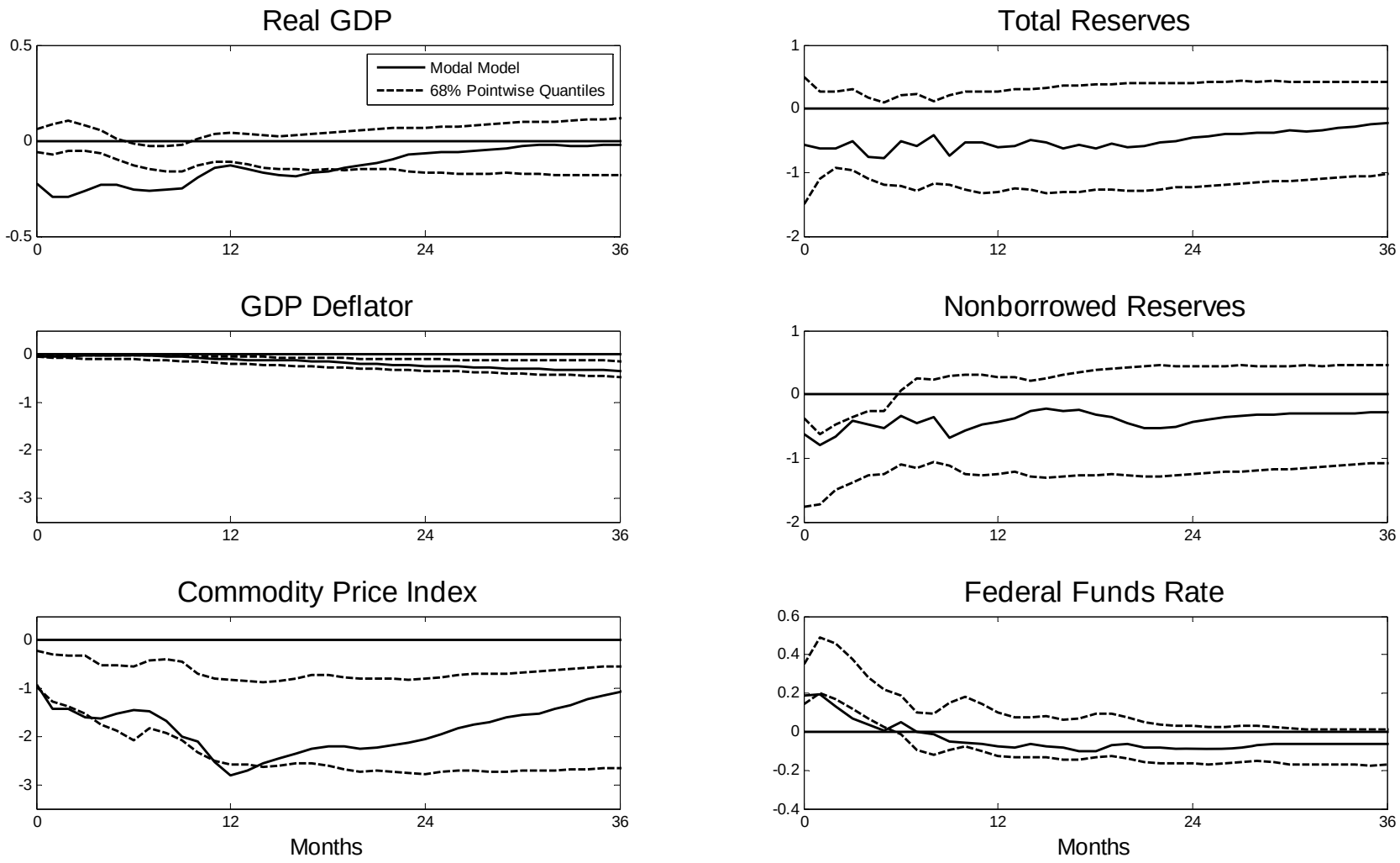
NOTES: Relative to the model in Figure 5, an additional scalar restriction has been imposed that real GDP must have declined 6 months after the monetary policy tightening.

Figure 8: Responses to a Monetary Policy Tightening in the Modified Sign-Identified Model Response Functions in the Modal Model and Pointwise Median Responses



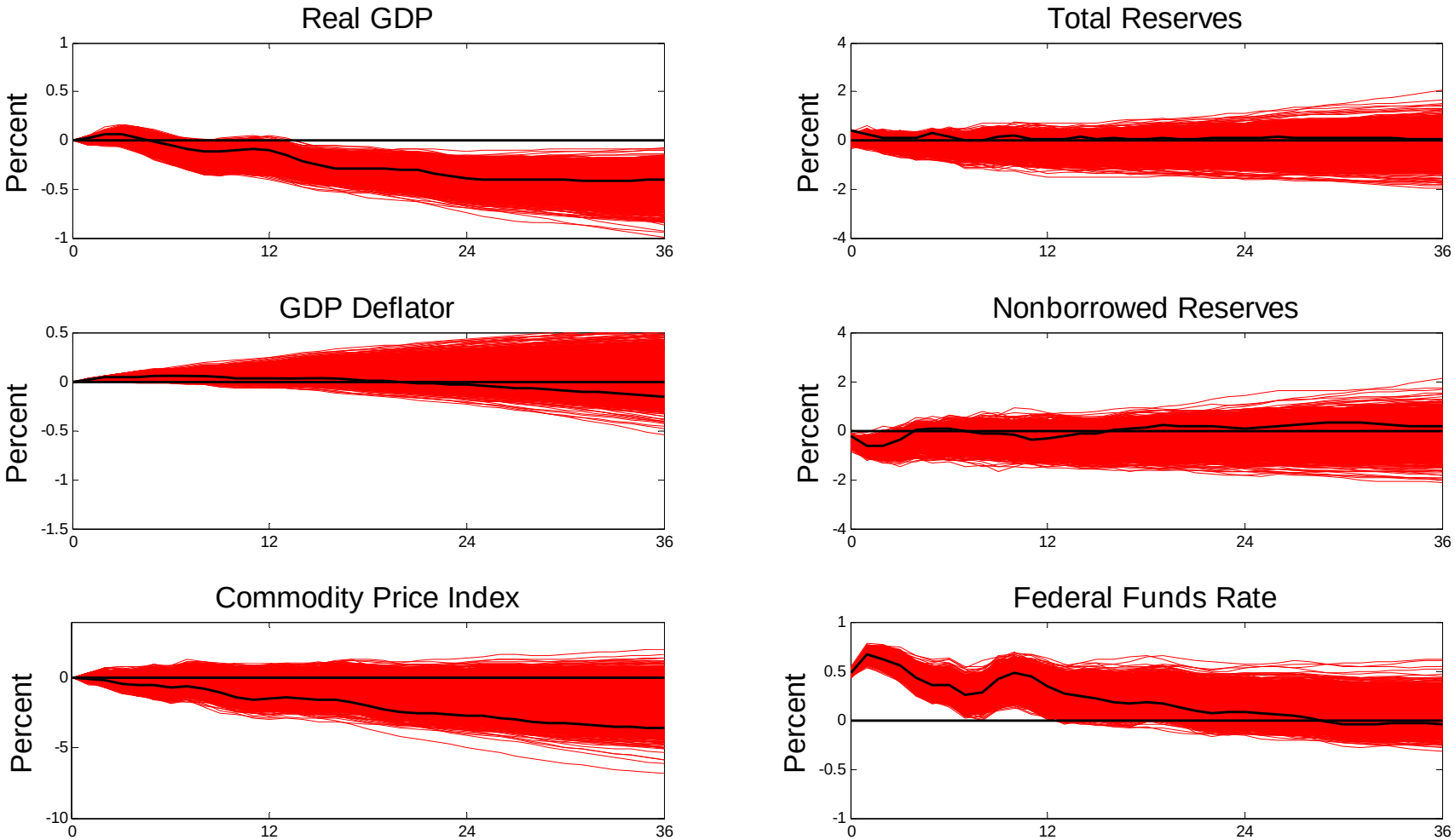
NOTES: See Figure 7.

Figure 9: Responses to a Monetary Policy Tightening in the Modified Sign-Identified Model Response Functions in the Modal Model and Pointwise 68% Posterior Error Bands



NOTES: See Figure 7.

Figure 10: Responses to a Monetary Policy Tightening in the Partially Identified Cholesky Model
Response Functions in the Modal Model and 68% Joint Regions of High Posterior Density



NOTES: The model specification and identifying restrictions are the same as in Figure 5 of Uhlig (2005).