Impulse Response Priors
for Discriminating
Structural Vector Autoregressions

Mark Dwyer

Working Paper Number 780
Department of Economics
University of California, Los Angeles
Bunche 2263
Los Angeles, CA 90095-1477
August, 1998

Impulse Response Priors for Discriminating Structural Vector Autoregressions

MARK DWYER*

Department of Economics, UCLA
Los Angeles, CA 90095-1477,USA

first version: October 1997
this version: August 1998

Keywords: Structural Vector Autoregression, Exact Identification, Impulse Responses, Priors, Bayes Factors, Importance Sampling, Monetary Policy.

JEL: C11, C32, C52, E50.

*8286 Bunche Hall, phone: 310-206-4610, fax: 310-825-9528, e-mail: dwyer@econ.ucla.edu. The research support of the UCLA Academic Senate is gratefully acknowledged. I am indebted to Carmen Fernández, Guido Imbens, Frank Kleibergen, Adrian Pagan, Mark Steel, Harald Uhlig, and seminar participants at Tilburg University, UCSD, and Camp Econometrics V for helpful comments. Some of this work was completed while visiting CentER, Tilburg University.

Abstract

The structural vector autoregression (SVAR) has become a central tool for research in empirical macroeconomics. Because the vast majority of these models are exactly identified, researchers have traditionally relied upon the informal use of prior information to compare alternative specifications. This paper surveys some of the structural dynamic restrictions used to evaluate SVARs. I provide a method for constructing prior distributions that incorporates this information on impulse responses. Based upon these Impulse Response Priors (IRPs) I employ a formal Bayesian model selection procedure for comparing SVAR specifications. I use this procedure to compare several alternative, six variable SVAR models of the interaction of real and monetary sectors of the U.S. economy. I make these comparisons under a variety of assumptions regarding the nature of the money supply rule, and lag length.

Emprically, I find strong evidence in favor of interpreting shocks to the federal funds rate as monetary policy shocks, as opposed to shocks to nonborrowed reserves. The most favored identification is one in which monetary policy reacts to contemporaneous movements in real variables and the price level. There is less evidence that monetary policy reacts as quickly to fluctuations in money demand.

1 Introduction

This paper offers some inferential techniques for one of the most popular tools in modern empirical macroeconomics — the structural vector autoregression (SVAR). Far from being atheoretical, SVARs are frequently derived from economic theory. Most researchers employing SVARs however, do not believe that the theoretical models from which they are derived are sufficiently rich to impose plausible overidentifying restrictions on the data. As a consequence, SVARs are largely employed as a means of *interpreting* data from a given theoretical perspective, rather than as a means of *restricting* data. In this paper, I assume that there are a number of competing economic models, which impose different identifying assumptions on the data.

Until now, the evaluation and comparison of alternative models represented by SVARs has been largely heuristic and qualitative. The absence of a method for comparing SVARs based upon statistical inference arises from a limitation of classical statistics. Since alternative SVARs are just identified, they differ only in the way that they parameterize the (usually normal) likelihood function. Alternative parameterizations are related by 1:1 mappings. Consequently, the invariance of the maximum likelihood estimator guarantees that alternative SVARs possess no discriminating restrictions from a classical viewpoint.

Nevertheless, authors working with SVARs frequently do make comparisons between alternative specifications. These are based upon an assessment of the resulting impulse responses of each model. The quality of these responses is judged informally relative to some prior assumptions. Usually, these assessments combine priors for several responses, in some unspecified manner.

In this paper, I introduce procedures for formally specifying these priors over the impulse responses of macroeconomic variables to various structural shocks. This allows researchers to be much more specific in their comparisons of SVARs, and to explicitly communicate the way in which they combine multidimensional assessments. This paper takes issue with the prevailing perspective of the literature, that has been recently articulated by Leeper, Sims, and Zha (1996). These authors propose that there is little to be gained by formalizing and quantifying the criteria we use for comparing alternative SVAR specifications. They state:

> ... [W]e use informal restrictions on the reasonableness of the impulse responses, ... in that we focus attention on results that do not produce implausible impulse responses. Our criterion for plausibility is loose. ... Our informal use of this sort of identifying information may give the impression of undisciplined data mining. We could have accomplished the same, at much greater computational cost, by imposing our beliefs about the forms of impulse responses as precise mathematical restrictions, but this would not have been any more "disciplined."

The purpose of this paper is to point out that formalizing such restrictions is not computationally costly, and that it is in fact substantially more disciplined, particularly because SVARs are compared across several dimensions. Informal, multidimensional assessments are difficult both to communicate and to replicate.

Two related papers are those of Faust (1998), and Uhlig (1997). Like the present paper, both of these authors formalize the impulse response restrictions that form the basis for SVAR comparisons. In contrast to the present paper, neither of these papers attempts to select among completely identified models. Faust examines the robustness of the claim that the variance share of money supply shocks in output fluctuations is small. He searches over all possible identifications for a counterexample, consistent with a set of explicit sign restrictions on impulse responses. Uhlig imposes sign restrictions via a penalty function on multiple responses to a money supply shock. Uhlig's work seeks to identify a money supply shock, without attempting to compare completely identified models.

Like the other papers, I formalize a multiple impulse response criteria.[1] Unlike the other work, I use a Bayesian prior for this formalization. However, a more substantive difference between this work and the others is that I focus on completely identified models. This orientation makes this paper substantially more amenable to macroeconomic theory. The presumption is that competing economic models give rise to alternative identifications, and that we aim to make comparisons between underlying theories. The work of Faust and Uhlig is not designed to make such comparisons, and is inherently more "atheoretical."

[1] Unlike Uhlig's work, responses to both money supply and money demand shocks are used to construct the criterion for model evaluation.

Paper Outline The remainder of this section establishes notation, and reviews the basic features of SVAR models. Section 2 examines some of the ways SVARs are identified and assessed in the recent literature. Subsection 2.2 specifies a candidate set of seven response criteria from the SVAR literature that focus on responses to money supply and money demand shocks.

Section 3 introduces the main framework for formalizing impulse response priors (IRPs). Subsection 3.1 specifies the general construction of the IRP. Subsection 3.2 covers certain technical issues related to ensuring that the IRP is a proper prior. Subsection 3.3 is also primarily technical. It describes how marginal likelihoods based upon these IRPs are calculated. The relative magnitudes of marginal likelihoods indicates the evidence in favor of one specification versus another. These two, technical subsections can be skipped by readers primarily interested in the application of IRPs to using monetary responses for selecting among SVAR identifications.

Section 4 applies these methods to selecting a preferred structural model for the interraction of real and monetary sectors of the U.S. economy. Subsection 4.2 gives the details of constructing an IRP for this application, and discusses some of the properties of the prior. Subsection 4.3 provides the model comparison results. Section 5 concludes. Appendices constitute section 6.

1.1 The SVAR Framework

To begin, consider the standard SVAR specification. This assumes that the n vector of variables $\mathbf{y}$ of interest are related by a simultaneous equation model (SEM) of the form:

$$\mathbf{A}(L)\mathbf{y}_t = A_0\mathbf{y}_t - \sum_{j=1}^{p} A_j\mathbf{y}_{t-j} = \boldsymbol{\varepsilon}_t, \tag{1.1}$$

where

$$\boldsymbol{\varepsilon}_t \sim \mathcal{NID}(0, I).$$

Two features of this SVAR specification specialize it from the traditional SEM specification.[2]

1. The structural errors $\boldsymbol{\varepsilon}_t$ are assumed to be contemporaneously uncorrelated.
2. The structural specification implies no restrictions on the elements of $A_1, \dots, A_p$.

The origin of the two significant distinctions between SVAR and SEM models noted in the previous section are attributed to Sims (1980). Sims argued that economic theory imposes much more structure over the contemporaneous interaction of variables, than it does with respect to their dynamic interactions. This is why SVARs generally avoid using conventional exclusion restrictions on lagged terms.

Contemporaneous restrictions are insufficient to identify structural models. In the absence of restrictions on predetermined variables, restrictions on the covariance of structural errors must be employed.[3] Sims added the assumption that the input shocks to impulse response analysis are contemporaneously uncorrelated. Indeed, without this restriction, impulse responses would be very limited guides to the "typical" response of a system to a given shock.

One argument to support such orthogonalized innovations is definitional: structural shocks should be uncorrelated as part of the objective of specifying a structural model. This requires the structural model to account for all covariations, both intertemporal, and intratemporal. Throughout the remainder of this paper, I maintain this assumption of orthogonal structural errors as well.

Two inessential differences between the conventional SEM and (1.1) are *(a)* the normalization of the structural variances to unity, and *(b)* the absence of strictly exogenous variables. With respect to the latter, they have been omitted for simplicity, since in the SVAR framework, they too are not subject to any exclusion restrictions.

[2] Cooley and LeRoy (1985) emphasize that the SVAR method of macroeconomic analysis can only be justified as a special case of the SEM approach.

[3] In the absence of both predetermined variable restrictions, and covariance matrix restrictions. the only identified models are those that lack any simultaneity.

The simultaneity in the above specification arises because A_0 is not in general diagonal. The associated reduced form specification is given by:

$$\mathbf{B}(L)\mathbf{y}_t = \mathbf{y}_t - \sum_{j=1}^{p} B_j \mathbf{y}_{t-j} = \mathbf{u}_t, \tag{1.2}$$

where $B_j = A_0^{-1}A_j$, $\mathbf{u}_t \sim \mathcal{NID}(0, \Sigma)$, with $\Sigma = A_0^{-1}A_0^{-1'}$. This reduced form is usually estimated on the data. SVARs are identified by providing a set of restrictions such that an estimator of A_0 can be derived from the relations above. Most SVARs are exactly identified, meaning that the mapping from (1.2) to (1.1) is 1:1. Thus while two alternative structures will differ in their implied structural form coefficients, $A_0, \ldots, A_p$, they will share the same reduced form parameters: $B_1, \ldots, B_p$ and Σ.

Since SVARs involve many coefficients, and since the matrices A_j, $j \geq 1$ are generally unrestricted by theory, researchers have adopted alternative methods for representing their estimation results. A standard tool here is impulse response analysis, which graphs the Wold Representation of the system. These impulse responses also have corresponding reduced and structural forms.

Under the assumption that the roots of $\mathbf{A}(z)$ are outside the unit circle, we have the Wold Representation:

$$\mathbf{y}_t = \sum_{j=0}^{\infty} C_j \mathbf{u}_{t-j} = \mathbf{B}^{-1}(L)\mathbf{u}_t, \tag{1.3}$$

where $C_0 = I$. Again for two alternative, exactly identified SVARs, this form will be identical. Corresponding to the different structural forms however are distinct ***structural*** impulse responses:

$$\mathbf{y}_t = \sum_{j=0}^{\infty} R_j \boldsymbol{\epsilon}_{t-j} = \mathbf{A}^{-1}(L)\boldsymbol{\epsilon}_t = \mathbf{B}^{-1}(L)A_0^{-1}\boldsymbol{\epsilon}_t. \tag{1.4}$$

As I describe in the next section, SVARs are generally compared in terms of these structural impulse responses. This paper aims to provide a straightforward means of formalizing SVAR model comparisons based upon such impulse responses.

1.2 Variations: Accommodating Nonstationarity and Cointegration

Above it was assumed that the system was covariance stationary. However, we can dispense with this assumption with only minimal changes to the basic definitions above.

To see this, it is useful to begin with the reduced form (1.2). Assuming that the elements of $\mathbf{y}$ are either I(0) or I(1), the first differences $\Delta\mathbf{y}_t$ have a the Wold Representation:

$$\Delta\mathbf{y}_t = \tilde{D}(L)\mathbf{u}_t = \sum_{j=0}^{\infty} \tilde{D}_j \mathbf{u}_{t-j}. \tag{1.5}$$

Recall too that $\mathbf{u}_t = A_0^{-1}\boldsymbol{\epsilon}_t$. Thus we can express this Wold Representation in terms of the structural innovations, by defining $D_j = \tilde{D}_j A_0^{-1}$.

Conditional on the initial value $\mathbf{y}_0$, the level of the series $\mathbf{y}_t$ is given by the partial sum:

$$\mathbf{y}_t = \sum_{j=1}^{t-1} \Delta\mathbf{y}_{t-j} + \mathbf{y}_0. \tag{1.6}$$

It is then possible to ask what the effect of a given structural shock $\boldsymbol{\epsilon}_t$ on $\mathbf{y}_{t+k}$ is for any given horizon k. Specifically

$$\begin{aligned} R_k \equiv \frac{\partial \mathbf{y}_{t+k}}{\partial \boldsymbol{\epsilon}_t} &= D_k + D_{k-1} + \cdots + D_1 + D_0 \\ &= \sum_{j=0}^{k} D_j, \end{aligned} \tag{1.7}$$

where $D_0 = A_0^{-1}$. This provides a formal definition of the conditional, structural impulse responses R_k, of (1.4). Note that unlike (1.4), the argument above does not assume that $\mathbf{B}(1)$ is invertible. When $\mathbf{y}_t$ has I(1) components, both $\mathbf{A}(1)$ and $\mathbf{B}(1)$ will be singular, with rank h, where h is the number of linearly independent cointegrating relations. Correspondingly, both $\mathbf{D}(1)$ and $\tilde{\mathbf{D}}(1)$ are singular with rank $r = n - h$, the number of unit roots for $|\mathbf{B}(z)|$.

2 A Brief Survey of Methods for Formulating and Evaluating SVARs

Several recent surveys of SVARs and their applications are available. See Christiano and Eichenbaum (1992a), Pagan and Robertson (1995), Pagan and Robertson (1996), and Leeper, Sims, and Zha (1996). The point of this section is to categorize the standard types of arguments used for identifying and comparing SVARs, and to indicate the types of restrictions that I will be accommodating within response priors.

2.1 Identifying SVARs

2.1.1 Impact Restrictions — Wold Causal Orderings

Since identification in (1.1) amounts to determining A_0, the most direct form of prior information consists in restrictions on A_0 itself. The most popular form for such restrictions to take goes by the name of a *Wold Causal Ordering* (WCO). WCOs order the variables in $\mathbf{y}_t$ to justify a lower triangular structure for A_0. Under this assumption, an estimate of A_0 can be recovered from the Cholesky decomposition of the reduced form covariance matrix Σ. This approach to identification in SVARs characterizes the surveys above by Christiano and Eichenbaum (1992a), Pagan and Robertson (1995), and Leeper, Sims, and Zha (1996).[4] The empirical section (4) focuses on alternative WCO identifications.

2.1.2 Alternative Impact Restrictions

While WCOs are the most popular approach to identification there are alternatives within the SVAR framework. What is required to identify A_0 is just a set of $n(n-1)/2$ linearly independent restrictions on the elements of A_0, ensuring a unique solution to the set of equations $A_0 \Sigma A_0' = I$. Gordon and Leeper (1994), and Leeper, Sims, and Zha (1996) provide examples of alternative zero restrictions on A_0 that can be used to identify SVARs. While the application in this paper is to WCOs, the methods extend immediately to these alternative impact restrictions.

2.1.3 Long Run Identifying Information

Another approach to identifying SVARs does not rely completely on contemporaneous restrictions on A_0, but consider long run restrictions as well. This approach has been used by Blanchard and Quah (1989), Galí (1992) and Lastrapes and Selgin (1995) among others. While such restrictions might in principle be expressed in terms of $\mathbf{A}(1)$, researchers have found it more convenient to impose restrictions on the long run, structural impulse responses of the first differences, namely via $\mathbf{D}(1)$. (Linear) long run restrictions are of the form

$$Q_1 vec(\mathbf{D}(1)) = q_2, \tag{2.1}$$

where Q_1, and q_2 are known. Such restrictions can aid in identification, since

$$\mathbf{D}(1) = \tilde{\mathbf{D}}(1) A_0^{-1}. \tag{2.2}$$

[4]These authors examine the work of Bernanke and Blinder (1992), Bernanke and Mihov (1995), Christiano and Eichenbaum (1992b), Christiano, Eichenbaum, and Evans (1996), Eichenbaum (1992), Eichenbaum and Evans (1995), Leeper and Gordon (1992), Sims (1986), Sims (1992), Sims (1993), Strongin (1995), among others, all of whom use WCOs for identification.

Since $\tilde{\mathbf{D}}(1)$ is estimable from the reduced form, long run restrictions imply that the contemporaneous effects in A_0 satisfy

$$(Q_1(I \otimes \tilde{\mathbf{D}}(1)))vec(A_0^{-1}) = q_2. \tag{2.3}$$

For example, Lastrapes and Selgin assume that $\mathbf{D}(1)$ is lower triangular, which, like the WCO assumption above, is sufficient to identify A_0. Galí combines zero restrictions on $\mathbf{D}(1)$ with zero restrictions on A_0 to identify an SVAR. These approaches to identification, like the impact restrictions above are covered by the techniques of this present paper. So long as alternative identifications do not also affect the reduced form specification, the approach to model comparison of this paper applies.

Stationarity Restrictions For systems with nonstationary variables, imposing cointegration specifications on the data does not, *in itself* contribute to structural identification. Cointegration imposes that $\mathbf{D}(1)$ is of reduced rank $r < n$. But from (2.2), it should be clear that this translates into a reduced rank restriction on $\tilde{\mathbf{D}}(1)$.

A_0 can be any arbitrary matrix, provided that $\tilde{\mathbf{D}}(1)$ has this appropriate reduced rank structure. As a consequence, cointegrating restrictions are testable, and cannot help in identifying A_0. To emphasize that cointegration restrictions impact the reduced form, I refer to them as *reduced form cointegrating restrictions.*

Nevertheless, researchers frequently *do* invoke cointegration assumptions in identifying structural models. This is possible, because these researchers impose more than just a particular cointegration rank for the system. These authors impose particular zero restrictions on $\mathbf{D}(1)$, that serve to identify particular structural shocks as permanent, and others as transitory. These restrictions can be expressed as

$$\mathbf{D}(1) = [\mathbf{D}(1)^*,\ \mathbf{0}], \tag{2.4}$$

where $\mathbf{D}(1)^*$ is $(n \times r)$. This restriction not only imposes cointegration, but specifies that the first r shocks are permanent, while the remaining h shocks are transitory. King, Plosser, Stock, and Watson (1991) impose precisely this type of restriction. They show that, in a three variable system with two cointegrating vectors, such a structural reduced rank specification is sufficient to identify A_0. A similar result obtains in a bivariate VAR for Blanchard and Quah (1989).

For larger systems, such structural reduced rank restrictions may be insufficient to completely identify A_0. Again King et al. illustrate this point in a six variable system, where they impose additional restrictions on $\mathbf{D}^*(1)$ to achieve identification. I refer to cointegrating restrictions of the form (2.4) as *structural cointegration restrictions.*

There have been two responses to the problem of incorporating cointegration into SVAR analysis in the literature. By and large, most research focuses on the level VAR specification. This constitutes the unrestricted specification, in that explicitly accounting for nonstationarity places restrictions on this level VAR. A motivation for emphasizing the level VAR specification may be that our priors are uninformative regarding the stationary properties of the data. This is the approach adopted in this paper.

A second approach found in the literature advocates pretesting a VAR to determine the nonstationary features of the system. Some work, like that of Galí , and King et al. employ economic reasoning to define candidate cointegrating vectors for the pretest stage. Of course, despite the economic foundations for these imposed cointegrating vectors, the very nature of pretesting implies that they may be examining the responses of misspecified models.

A more satisfactory approach to cointegration might be to maintain a given reduced form cointegrating specification, and to distinguish between models that invoke the structural cointegration restriction (2.4) and those that use alternative identifying assumptions. Imposing a particular reduced rank on $\mathbf{D}(1)$ then becomes another auxiliary assumption, such as imposing a particular lag length. This approach has the advantage in that the statistical representation of the data is maintained across alternative structural models. This after all, appears to be one of the underlying constraints of the SVAR literature. Put differently, just because a system appears to exhibit cointegration, we do not necessarily need to use that information for identification. Instead, we can ask whether a given structural cointegrating interpretation of the reduced form cointegrating restrictions is more or less consistent than alternative structural information, with what prior information we have regarding the impulse responses of the model. The extension of the methods of this paper to this case is left for future research.

2.2 Evaluating SVARs

An interesting feature of the identifying restrictions above is that they can be seen to impose information regarding either the contemporaneous ($t = 0$) or the long run ($t \to \infty$) impulse responses of the system. In contrast, intermediate horizon responses constitute the crucible in which alternative identifications are compared. Typically, researchers examine responses of alternative models and divide identifications in "acceptable" versus "unacceptable" identifications, based upon the degree to which each model's responses accord with some informal criteria.

As an example, here are some of the criteria that arise in SVAR model evaluation. These descriptions apply to a model with both real and monetary sectors. Unless otherwise noted, these priors are taken from Sims (1986). Many subsequent authors make either direct or indirect reference to the prior assessments in this paper. Here, I am ***not*** concerned with the source of this prior information, but rather on formalizing this information, to make it readily communicable as to how it is used in model comparisons.

1. **Output Response to Money Supply Shocks** This prior maintains that that output should respond positively, but temporarily to expansionary money supply shocks.

2. **Price Response to Money Supply Shocks** Sims (1992) observed that when innovations to an interest rate are treated as the monetary policy measure, there is an anomalous price increase following increases in the interest rate. Instead, since an interest rate hike is generally perceived to be contractionary, the prior is that real demand should fall. This phenomenon has come to be called the "price puzzle." In commenting on Sims' work, Eichenbaum (1992) focuses on narrow measures of money (nonborrowed reserves) as the policy measure, precisely to avoid this effect. Other authors introduce world commodity prices (in quarterly data) to this end.

3. **Money Market (Demand) Response to Money Supply Shocks** Much of the work in monetary SVARs focuses on determining whether or not expansionary monetary policy shocks cause initially opposite movements in money stock, and interest rates. When shocks to a monetary aggregate are interpreted as supply shocks, the short term interest rate is viewed as reflecting the demand response. In this case, the liquidity effect is generally assumed to be short lived.

   In contrast, when models propose interest rate shocks as primarily reflecting policy interventions, an opposite movement in the money stock, again interpreted as a liquidity effect, may have a more persistent response.[5]

4. **Money Stock Response to Money Demand Shocks** Increases in money demand should be accommodated (at higher rates), and consequently, money stock should increase, perhaps permanently.

5. **Output Response to Money Demand Shocks** Output should decline, potentially persistently, to unanticipated increases in money demand.

6. **Price Response to Money Demand Shocks** Initially, increased demand for money should drive prices down. In the long run, prices may rise, to offset increases in the nominal money stock (holding real balances fixed in the long run).

7. **Interest Rate Response to Money Demand Shocks** This maintains that increases in money demand should bid up the price of money (i.e. interest rates), at least temporarily.

The elements of this list are not meant to be exhaustive or even necessary for evaluating SVARs. I have emphasized assumptions found in the literature relating to money stock, interest rates, output, and prices because this is the application I consider below. For models with additional or alternative dimensions, this list is suitably modified. My point here is that even if we refrain both from imposing overidentifying restrictions on the data, and from working with a formal, stylized economic model, we can and should nevertheless use (Bayesian) statistical inference in incorporating what we believe we *do* know about the underlying economic structure. I now turn to translating the above example restrictions into an informative prior.

[5] These alternative formulations of the liquidity effect are laid out in Leeper, Sims, and Zha (1996).

3 Implementing Impulse Response Priors

3.1 Translating Response Criteria Into Response Priors

To begin, I need to define some features of what constitutes a model. Alternative models $M_1, \ldots, M_{k_m}$ imply alternative identifying restrictions on A_0. One should note, that while a structural model identifies A_0, it does so in order to interpret the structural shocks ε_t. Thus I define a model M as an interpretation of structural shocks, based upon a set of identifying restrictions. Specifically, I consider as distinct structures two models which imply the same WCO, but which identify the money supply shock with different structural shocks (money stock versus interest rate shocks). Following convention, I refer to identifications that associate money stock shocks with monetary policy innovations as "M-Rule" identifications, and identifications that associate interest rate shocks with monetary policy shocks as "R-Rule" identifications.

I formalize information over the intermediate horizon responses of the system by assuming that the researcher can construct a prior distribution over some subset of response horizons. Let l denote the number of such horizons $j_1 < \ldots < j_l$, where $j_1 > 0$. Suppose then that we have prior information regarding $q \leq n^2$ dynamic responses. Denote the l-vector of horizon values for the jth dynamic response as r_j. Then let $\mathbf{R}_{lq}$ denote the $(l \times q)$ matrix of responses $[r_1, \ldots, r_q]$, over which a formal prior is constructed. In the empirical example that follows, $l = 6$, where I have chosen to specify a prior over the 1, 6, 12, 24, 60, and 120 month horizons. I use information in the $q = 7$, qualitative response criteria from the preceding section to construct the impulse response prior.

I specify independent trinomial priors for these l response horizons. The aim here is to maintain information regarding the sign of anticipated structural responses, at particular horizons. I use a trinomial specification to capture both positive-negative, and zero value response restrictions. This prior for $\mathbf{R}_{lq}$ is of the form

$$\pi_{IR}(\mathbf{R}_{lq}) = \prod_{i=1}^{l}\prod_{j=1}^{q} \pi(r_{ij}), \quad \pi(r_{ij}) = \begin{cases} \pi_{ij}^{+}, & \text{for } r_{ij} > z_{ij}, \\ \pi_{ij}^{0}, & \text{for } z_{ij} \geq r_{ij} \geq -z_{ij}, \\ \pi_{ij}^{-}, & \text{for } r_{ij} < -z_{ij}. \end{cases} \tag{3.1}$$

Here $\pi_{ij}^{+} + \pi_{ij}^{0} + \pi_{ij}^{-} = 1$. The hyperparameters z_{ij} define significant positive and negative responses, for response j at horizon i. The criteria described in the previous section indicate the anticipated signs of responses at various horizons. This is not enough information for determining the z_{ij}. I employ the estimated reduced form impulse responses: $[c_1, \ldots, c_q]$, corresponding to the q informative structural responses $[r_1, \ldots, r_q]$. For each response, I measure its sample standard deviation, across horizons: $\hat{\sigma}_{c_j}$, $j = 1, \ldots, q$. I use these values to determine the z_{ij} hyperparameters, setting $z_{ij} = \hat{\sigma}_{c_j}/2$, for all horizons $i = 1, \ldots, l$.

The specification in (3.1) maintains independence with respect to the identification of A_0, that is to variation across structural models M_j. The motivation for this invariance is simply that the SVAR literature appears to use common response criteria when comparing alternative structural identifications. This prior invariance holds[6] with respect to the structural shocks denoted as "money demand" and "money supply." Note however that the association of these shocks with a particular response varies across models, specifically M-Rule models and R-Rule models.

In certain contexts, one might calibrate the impulse response priors to the variances of $\mathbf{y}$. There are two reasons why I do not pursue this approach here. First, I do not wish to impose stationarity on the data. Without stationarity, impulse responses are no longer associated with a variance decomposition of $\mathbf{y}$, because $\mathbf{y}$ may not have a finite variance. Second, even under a stationarity assumption, I would need to impose prior information regarding the structural variance decomposition of $\mathbf{y}$, which I do not have, at least based on the evaluative criteria in subsection (2.2) above.

Auxiliary Assumptions and Structural Comparisons SVAR specification involves a host of auxiliary assumptions, such as trend and seasonal dummy dependence, lag length selection, and reduced form cointegration specification. In principle, it should be possible to formulate priors over these auxiliary assumptions, so that structural model comparisons can be averaged over these specifications. With the

[6]See footnote (13).

exception of lag length, this paper adopts a more limited approach, of holding auxiliary assumptions fixed in comparing identifications.

Let ψ denote the reduced form parameters $(B_1, B_2, \ldots, B_p, \Sigma)$ of (1.2). I now turn to translating the prior over $\mathbf{R}_{lq}$ to one over these reduced form parameters ψ.

3.2 VAR Dimension and Prior Combination

Let $\mathcal{R}_{lq}(\psi|M_j) = \mathbf{R}_{lq}$ denote the mapping from reduced form parameters, to impulse responses, under the identifying restrictions of model M_j. This mapping is well defined. What concerns us here however, is the inverse of this mapping, when it exists, and where it is unique. In order for $\pi_{IR}(\mathbf{R}_{lq})$ to induce a prior over Ψ, it must be the case that $\mathcal{R}_{lq}(\cdot)$ is 1:1, which means at a minimum that the dimension of the IR prior lq must be as large as that of the VAR parameterization ψ, which is $n^2p + n(n+1)/2$. There are two reasons why I choose to keep the dimension of the IR prior much smaller than this. First, I want a prior which ultimately can be used to compare identifications over alternative auxiliary assumptions, such as lag length. If I had to match the dimension of the prior to the VAR dimension, I would not have an invariant prior. Second, imposing such a high dimension prior based on impulse responses effectively imposes stronger information regarding these responses than I actually have. Rather than specifying locations for impulse responses at a handful of horizons, matching dimensions would require specifying priors at virtually all response horizons, which is more information than I want my prior to incorporate.[7]

Instead, I propose to augment the IR prior $\pi_{IR}(\mathbf{R}_{lq})$ by a relatively uninformative prior that scales with the dimension of the system.[8] In doing so, the dimension of the IR prior can be kept constant and modest. I use a natural conjugate prior for this purpose, which I denote as $\pi_0(\psi)$. The prior that I induce over the reduced form parameters is of the form:

$$\pi_{IR,\psi}(\psi|M_j) \propto \pi_0(\psi) \cdot \pi_{IR}(\mathcal{R}_{lq}(\psi|M_j)). \tag{3.2}$$

Note that this induced prior is *not* a simple reparameterization of a prior over impulse responses, in that two Jacobians have been averaged out (or omitted). The first Jacobian corresponds to the parameterization of π_0 in terms of the complete impulse response paths $\mathbf{R}_{l\cdot n^2}$, of which $\mathbf{R}_{lq}$ is only a finite subset. The second Jacobian arises in translating this prior back into the ψ parameterization. Since the resulting prior is not a straightforward variable transformation, it is important to evaluate the features of the composite prior. This will be taken up in the empirical section 4.2.1 below.

Based upon these induced priors, I can now construct Bayes factors for structural model comparisons.

3.3 Importance Sampling and Model Comparison via Bayes Factors

Given two structural models M_1 and M_2, the Bayes factor in favor of model M_1 is given by the ratio of the marginal likelihoods

$$\mathcal{B}(M_1, M_2) = \frac{m(\mathbf{y}|M_1)}{m(\mathbf{y}|M_2)} = \frac{\int \mathcal{L}(\psi|\mathbf{y}, M_1))d\pi_{IR,\psi}(\psi|M_1)}{\int \mathcal{L}(\psi|\mathbf{y}, M_2))d\pi_{IR,\psi}(\psi|M_2)}. \tag{3.3}$$

Here $\mathcal{L}(\cdot)$ denotes the likelihood function with respect to the reduced form parameters ψ. In order to denote the integrating constants omitted in (3.2) explicitly, let

$$\pi_{IR,\psi}(\psi|M_j) = \mathfrak{a}(M_j)\pi_0(\psi)\pi_{IR}(\mathcal{R}_{lq}(\psi|M_j)). \tag{3.4}$$

Attempting to evaluate the integrals of (3.3) by direct monte carlo integration faces two significant obstacles. First, we do not know the integrating constants $\mathfrak{a}(M_j)$, so that we cannot draw from $\pi_{IR,\psi}(\psi|M_j)$ directly. Second, even if I could draw from the IR priors, they are likely to be much more diffuse than the likelihood,

[7]For example, for the twenty-four lag model considered below, using the seven response criteria from the previous section would require priors over the first one hundred and twenty-five periods — more than ten years of horizon specifications.

[8]See Appendix (6.2) for a formal interpretation of this composite prior construction method.

and as a consequence, one would be sampling nonnegligible contributions of the integrands quite inefficiently. Consequently, I take an alternative approach to calculating the marginal likelihoods for each specification.

The solution is to use importance sampling.[9] This provides a means of sampling the likelihood in the region where it is significant, weighting by the appropriate values of the prior. To this end I use a benchmark prior $\pi_0(\psi)$ over the reduced form parameters, which is sufficiently diffuse that the posterior (that has an analytical form) largely reflects the features of the likelihood function. The benchmark prior I use here is the same natural conjugate prior that appears in (3.2).[10] This results in certain simplifications that will be outlined below. Because this prior is diffuse, draws from this posterior reflect the central tendency of the likelihood. Let $m_0(y)$ be the associated marginal likelihood. Now consider the Bayes factor[11]

$$\mathcal{B}(M_j, M_0) = \frac{m(\mathbf{y}|M_j)}{m_0(\mathbf{y})} = \frac{\mathcal{L}(\psi, \mathbf{y})\pi_{IR,\psi}(\psi|M_j)/p_{IR,\psi}(\psi|\mathbf{y}, M_j)}{\mathcal{L}(\psi, \mathbf{y})\pi_0(\psi)/p_0(\psi|\mathbf{y})}, \ \forall\psi \in \Psi.$$

In this expression, $p_{IR,\psi}(\cdot)$ and $p_0(\cdot)$ are the posteriors associated with $\pi_{IR,\psi}(\cdot)$ and $\pi_0(\cdot)$ respectively. Cancelling the likelihoods, and rearranging, this yields

$$m(\mathbf{y}|M_j)p_{IR,\psi}(\psi|\mathbf{y}, M_j) = m_0(\mathbf{y})\frac{\pi_{IR,\psi}(\psi|M_j)}{\pi_0(\psi)}p_0(\psi|\mathbf{y}), \ \forall\psi \in \Psi.$$

Now we don't know the posterior $p_{IR,\psi}(\psi|\mathbf{y}, M_j)$. We don't need to however, if we integrate:

$$m(\mathbf{y}|M_j) = m_0(\mathbf{y})\int \frac{\pi_{IR,\psi}(\psi|M_j)}{\pi_0(\psi)}p_0(\psi|\mathbf{y})d\psi. \tag{3.5}$$

Thus by generating ψ from the posterior $p_0(\psi|\mathbf{y})$ and averaging, we can estimate $m(\mathbf{y}|M_j)$ up to the proportionality factor $m_0(\mathbf{y})$. We don't even need to know this factor, for if we are really interested in comparing the models M_1 and M_2, these proportionality factors will cancel. Thus the estimated Bayes factor favoring M_1 over M_2 becomes

$$\begin{aligned}\hat{\mathcal{B}}(M_1, M_2) = \frac{\hat{m}(\mathbf{y}|M_1)}{\hat{m}(\mathbf{y}|M_2)} &= \left.\frac{N^{-1}\sum_{\iota=1}^{N}\pi_{IR,\psi}(\psi^\iota|M_1)/\pi_0(\psi^\iota)}{N^{-1}\sum_{\iota=1}^{N}\pi_{IR,\psi}(\psi^\iota|M_2)/\pi_0(\psi^\iota)}\right|_{\psi^\iota\sim p_0(\psi|\mathbf{y})}. \\ &= \left.\frac{\sum_{\iota=1}^{N}\mathfrak{a}(M_1)\pi_{IR}(\mathcal{R}_{lq}(\psi^\iota|M_1))}{\sum_{\iota=1}^{N}\mathfrak{a}(M_2)\pi_{IR}(\mathcal{R}_{lq}(\psi^\iota|M_2))}\right|_{\psi^\iota\sim p_0(\psi|\mathbf{y})}.\end{aligned} \tag{3.6}$$

Here the notation $(\cdot)|_{\psi^\iota\sim p_0(\psi|\mathbf{y})}$ indicates that the ψ^ι are drawn from the benchmark posterior $p_0(\psi|\mathbf{y})$. Note that the last line of (3.6) exploits the simplification that the benchmark priors cancel with the corresponding terms within each composite prior $\pi_{IR,\psi}(\psi^\iota|M_j)$.[12]

4 Application: Money Transmission Mechanisms

4.1 Variable Definitions, and Model Descriptions

Here I consider modeling a six variable VAR for the U.S. I am using monthly data on the Industrial Production Index (y), the Consumer Price Index (p), a world commodity-price index (cp), the Federal Funds rate (ff), the level of Non-Borrowed Reserves (nbr), and the level of total reserves (tr). All data are in logs except the interest rate. This choice of variables follows that of Leeper, Sims, and Zha (1996), and many of the papers they discuss. There are 420 observations, from 1959:01 through 1993:12. The series are shown in Figure 1.

[9] See Geweke (1989) for a general discussion of importance sampling in econometrics.

[10] See Harkema (1971) for an explicit discussion of natural conjugate priors and posteriors for multivariate regression. Appendix (6.3) provides explicit expressions for this prior and posterior, along with the specific values of hyperparameters used, and a discussion of alternative prior specifications that might be used.

[11] Here M_0 does not actually represent a structural model, but the "hypothesis" corresponding to the unrestricted, benchmark prior $\pi_0(\psi)$.

[12] Appendix (6.1) details the evaluation of the integrating constants $\mathfrak{a}(M_j)$.

Alternative Structural VARs and the Limitations of WCOs I examine six alternative WCO orderings. These constitute six alternative SVAR specifications. I treat the "goods sector" consisting of $g = [y, p, cp]$ as a block, as well as the "money sector" $m = [nbr, tr]$. These intrasectoral orderings maintain that commodity prices incorporate more contemporaneous information than output or consumer prices, and that total reserves reflect contemporaneous movements in non-borrowed reserves.

The alternative orderings I consider are between these blocks g and m, and the interest rate $r = ff$. This reduces the number of orderings considered to the six permutations of $\{g, m, r\}$. When the g sector occurs first, output and price variables are assumed to respond sluggishly to financial markets. This position is maintained by Strongin (1995), Bernanke and Mihov (1995), and Christiano, Eichenbaum, and Evans (1996). In addition, the early ordering of the g sector implies that financial variables do incorporate contemporaneous real activity and price information. Moving the g sector down in the causal ordering means both that these real activity and price variables (g) *do* respond within the month to financial markets (r, m), and that these other (preceding) sectors *do not* incorporate contemporaneous production/price information.

The fact that the ordering has two implications reflects a general feature of WCOs. Variables that enter into the ordering early are both "leading variables" — those whose information is incorporated into subsequent variables' information sets, and variables that have more limited information sets. In order to avoid this dual implication of the ordering, one might consider block diagonal structures. Such models lead to overidentifying restrictions, which this literature generally ignores. Frequently, one finds justifications for an ordering based upon only one implication. For instance, in placing the g sector first, Leeper, Sims, and Zha (1996) emphasize the sluggish response of production/price variables to the financial variables, rather than the ability of financial markets to incorporate contemporaneous production/price information. Similarly, authors who place the m sector first emphasize the limitations of the Federal Reserve's information set, rather than the ability of the g sector to incorporate financial information.

4.2 Description of Impulse Response Prior

Lag Lengths, Model and Prior Dimensions I consider marginal likelihoods for alternative lag length specifications of 6,12, 18 and 24 months. The dimension of these specifications vary from 231 to 879 parameters. For the seven informative response priors, I specify the probabilities at the six monthly horizons of 1, 6, 12, 24, 60,and 120 months.

Impulse Response Prior Scaling My informative priors are constructed relative to the seven example priors described in section 2. These assumptions are specified in terms of money supply shocks (policy shocks) and money demand shocks. There are two possible mappings of these responses to identification, depending upon whether we maintain an M-Rule interpretation, or an R-Rule interpretation of monetary policy. This simply means we need to take a stand on whether shocks to m or to r will be considered policy shocks. With one exception,[13] I have chosen to specify these response priors in a symmetric fashion, so that the only change between M- and R- Rule interpretations is the definition of an expansionary shock. Under an M-Rule, positive innovations to m are expansionary, while under an R-Rule, negative innovations to r are expansionary.

As described in section 3, I used the standard deviations of the reduced form responses to set the scaling for the impulse response priors.[14] These standard deviations are given in Table 1. Under an M-Rule interpretation, the relevant reduced form responses are

$$\text{M-Rule:} \quad \frac{\partial y}{\partial u_{nbr}}, \frac{\partial p}{\partial u_{nbr}}, \frac{\partial ff}{\partial u_{nbr}}, \frac{\partial tr}{\partial u_{ff}}, \frac{\partial y}{\partial u_{ff}}, \frac{\partial p}{\partial u_{ff}}, \frac{\partial ff}{\partial u_{ff}}.$$

Under an R-Rule interpretation, the relevant reduced form responses are

$$\text{R-Rule:} \quad \frac{\partial y}{\partial u_{ff}}, \frac{\partial p}{\partial u_{ff}}, \frac{\partial nbr}{\partial u_{ff}}, \frac{\partial tr}{\partial u_{nbr}}, \frac{\partial y}{\partial u_{nbr}}, \frac{\partial p}{\partial u_{nbr}}, \frac{\partial ff}{\partial u_{nbr}}.$$

[13] The one exception is that the prior assumes a more persistent liquidity effect under an R-Rule interpretation.

[14] The responses used are those of a level VAR with an intercept, and twelve lags.

I have imposed a constant width zero region $z_{ij} = z_j$, $i = 1, \ldots, l$ across horizons for a given response, but there is no need to maintain this restriction in general. For simplicity, I have constructed response priors that vary monotonically with horizon. Tables 2 and 3 give the precise parameterization of the IR prior used. I now turn to the explicit assignment of probabilities I used to construct these priors.

1. **Prior-Contrary Events** I have chosen five percent for the prior probability of structural response horizons that conflict with the dominant sign restriction of the prior. Thus with respect to money supply shocks, negative output responses have a five percent prior probability at all horizons, as do negative price responses, and positive money demand responses. Similarly, with respect to money demand shocks, negative money stock responses have a five percent prior probability at all horizons, as do positive output responses, and negative interest rate responses.

2. **Prior-Consistent Events** I have chosen eighty percent for the prior probability of limit horizons (that is, either the first or the last horizon) satisfying the relevant sign restrictions. Thus for M-Rule liquidity effects, the probability of a negative, first period, interest rate response is set to eighty percent, as is the probability of positive, first month responses of output to money supply shocks; as well as the first month, negative price, and positive interest rate responses to money demand shocks. Similarly, at the ten year horizon, the positive price response to money supply shocks; as well as the positive money stock, positive price, and negative output responses to money demand shocks are all assigned an eighty percent prior probability.

3. **Complimentary Horizon Specifications** For each of the responses that are characterized by either an initial effect, or a long run effect, I chose the complimentary horizon probability of the prior-consistent sign restriction to be twenty percent. Thus for the priors that imply positive (negative) long run responses, the first month positive (negative) response probability is twenty percent. Similarly for the priors that imply positive (negative) short run responses, the last horizon (10 year) positive (negative) response probability is twenty percent.

4. **Dual Horizon Priors** Finally, there are two responses, that of prices to money demand, and the R-Rule specification of the liquidity effect, where both a short run and a long run sign restriction are posited. In these cases, eighty percent specification was used at both the one month, and the 10 year horizons.

Figures 2 and 3 present these seven, impulse response priors, along with the informative structural responses, for the six alternative identifications considered. Figure 2 depicts the priors and responses under an M-Rule regime, while Figure 3 depicts the same under an R-Rule regime. Each structural response is plotted with bands on either side of the origin reflecting the zero region of the associated response prior. The bar plots represent the relative probabilities of positive region, zero region, and negative region responses, as indicated by the prior. For each bar plot, the size of the zero region probability is scaled to the $\pm 1/2$ standard deviation line.

These figures illustrate a number of issues.

1. The first feature of these figures is that there is relatively little variation in responses across the various WCOs, given a money supply rule. Thus we might expect the model comparisons to have difficulty in choosing a complete identification scheme. The comparisons may still reveal some preference for R-rule vs. M-rule models however. This relative uniformity is frequently invoked as evidence of the "robustness" of structural VAR conclusions. But such arguments are based on only a handful of model comparisons. As Faust (1998) makes clear, such claims of robustness are largely refuted when broader classes of identifying restrictions are evaluated.

2. Second, it becomes apparent where each money supply rule has difficulty in generating responses consistent with prior information. For M-rule specifications, increases in the federal funds rate (interpreted as increases in money demand) are not accommodated. Instead they seem to lead to a persistent fall in total reserves. Prices also respond anomalously to interest rate increases under M-rules, only falling with a delay in the face of positive shocks to money demand, and then persisting in being negative at long horizons.

A final discrepancy between the M-Rule specifications and the maintained IR prior is in the money demand response to money supply shocks. While M-Rule orderings possess a short lived liquidity effect, the monotonicity assumption of the IR prior I employ here does not accommodate the significantly positive responses of the federal funds rate to non-borrowed reserves, at intermediate and long horizons.

3. For R-rule specifications, output expands persistently in response to increases in money demand. A second weakness arises for some orderings, in that output declines persistently in the face of money supply contractions. Taken together, these models appear to reflect long run nonneutrality.

4.2.1 Properties of the Impulse Response Prior

This subsection provides some qualitative implications of the IR prior described above. Since this paper focuses on comparing models in terms of their impulse responses, I have chosen to present the second moments of the impulse response functions, associated with the composite prior $\pi_{IR,\psi}(\psi|M_j)$. The question I want to address here is: to what extent does the IR prior dominate the responses of the data? If the prior generates responses with a smaller effective support than those found in the data under alternative identifications, then comparisons between these identifications are likely to be inconclusive. If on the other hand, the responses implied by the IR prior are relatively diffuse, then more substantive comparisons may be made between models.

I compute these moments via importance sampling, relative to the benchmark prior $\pi_0(\psi)$. Let $\mathcal{R}(\psi|M_j)$ denote the general mapping from reduced form parameters ψ to structural responses under model M_j. The estimate of the *kth* sample moment of the responses is generated by drawing $\psi \sim \pi_0(\psi)$:

$$\begin{aligned} \hat{E}\left(\mathcal{R}(\psi|M_j)^k|\pi_{IR,\psi}(\psi|M_j)\right) &= N^{-1}\sum_{\iota=1}^{N}\mathcal{R}(\psi^\iota|M_j)^k\frac{\pi_{IR,\psi}(\psi^\iota|M_j)}{\pi_0(\psi^\iota)}|_{\psi^\iota\sim\pi_0(\psi)} \\ &= N^{-1}\sum_{\iota=1}^{N}\mathcal{R}(\psi^\iota|M_j)^k\mathfrak{a}(M_j)\pi_{IR}(\psi^\iota|M_j)|_{\psi^\iota\sim\pi_0(\psi)}. \end{aligned} \tag{4.1}$$

Response Prior Standard Errors

1. Figure 4 plots the means and standard deviations of the informative responses, for twelve-lag, M-Rule specifications.[15] Twenty horizons are plotted. The reason for only plotting twenty horizons is that the standard errors are explosive. This is hardly surprising. The diffuse natural conjugate prior $\pi_0(\psi)$ incorporates no information that the system is dynamic. Thus generating a draw of the 432 random mean parameters of the six-variable, twelve-lag specification virtually assures that some of the implied roots of the system will be unstable.

2. Given this instability, the assumption that the prior response moments $E\left(\mathcal{R}(\psi|M_j)^k|\pi_{IR,\psi}(\psi|M_j)\right)$ even exist should be questioned. Additional evidence regarding the instability of these mean and variance estimates arises as we look across each row in Figure 4. Priors are constant within a given row, as they are in Figures 2 and 3. All of the response moments are based upon a common set of six-thousand draws from the benchmark prior $\pi_0(\psi)$. Thus the variation in orderings in this figure simply permutes these parameter draws. Since $\pi_0(\psi)$ is completely symmetric in ψ, these alternative parameter permutations act as additional prior sampling runs. The variation across a given row can then be interpreted as a measure of the instability of the standard error estimates.

3. All of this prior response instability however is good news. The scale of the forty month response deviations lie between 10^6 and 10^{10}, implying that the impulse responses generated by the prior are vastly more diffuse than those seen in the data, under the alternative identifications of interest. This is particularly reassuring, since the IR priors were calibrated to the reduced form. One reason that the IR priors preserve this diffuse nature is the use of the trinomial, that weights responses by their sign, but (apart from the zero region) not by the magnitude of their deviations from zero.

[15]Similar results for the R-Rule specifications have been omitted.

Response Prior Interquartile Ranges The standard error calculations above suggest that the IR priors generate extremely diffuse priors. To verify that these wide bounds are not due to occasional outlying responses that might by generated by heavy tailed response distributions, I attempted to compute interquartile ranges using Rejection/Acceptance (R/A) sampling. R/A sampling generates an approximately *iid* sample of structural responses, which can then be used to construct interquartile ranges.[16]

Unfortunately, R/A sampling does not work in this context. To understand why, note that based on the parameterizations in Tables 2 and 3,. the upper and lower bounds on $\pi_{IR}(\mathcal{R}(\psi))$ are $\pi^*_{IR} = 4.21 \times 10^{-8}$ and $\pi_{IR\,*} = 1.46 \times 10^{-53}$ respectively. As the appendix explains, in order for R/A sampling to work, I need to place an upper bound on $\pi_{IR}(\mathcal{R}(\psi))$, that is sample-independent. The natural choice is π^*_{IR}. Then a given response $\mathcal{R}(\psi)$ will be accepted as an *iid* realization with probability $\pi_{IR}(\mathcal{R}(\psi))/\pi^*_{IR}$. If all draws of ψ from $\pi_0(\cdot)$ imply a very small value for this ratio, then it will take extremely many iterations to generate a draw of **R**. Figure 5 plots the mean responses $\mathcal{R}(\psi)$ via importance sampling. The large, frequent sign fluctuations in these responses are at odds with the primarily monotonic sign restrictions imposed by the prior. Consequently the IR prior gives such responses extremely low probabilities, typically on the order of 10^{-20}. Combining this with π^*_{IR} implies typical acceptance probabilities on the order of 10^{-12}.

4.3 Results

4.3.1 Convergence of Marginal Likelihoods

Figure 6 depicts how a particular log and corresponding level marginal likelihood calculation varies with the length of the importance sampler run. The marginal likelihood does not appear to stabilize until around four thousand draws. To ensure stable estimates, the computations in this paper rely upon six thousand sampler iterations. Appendix (6.4) reports computation times, and some sampler design issues.

4.3.2 Marginal Likelihood Comparisons

1. Figure 7 displays log marginal likelihoods for each WCO, under the two money supply rules, for six, twelve, eighteen, and twenty-four lag specifications. It is immediately apparent that on average, the response prior configuration favors the R-Rule specifications. As mentioned above, neither M-Rule nor R-Rule formulations are completely consistent with the impulse response priors. But the M-Rule deviations from the prior restrictions appear to be more significant. Table 4 averages across WCOs. This imposes a prior that equally weights the alternative structural models. Here the evidence in favor of R-Rule specifications increases dramatically with lag length.

   When is a (log) Bayes factor significant? In citing Jeffreys (1961), Kass and Raftery (1995) propose that log-Bayes factors between one and three be considered "substantial" evidence, while factors between three and five imply "strong" evidence, and factors greater than five constitute "decisive" evidence.[17] With this convention, Table 4 implies that the evidence in favor of an R-Rule is substantial for the six-lag and twelve-lag specifications with log-Bayes factors of 1.09 and 4.76 respectively; and decisive for the eighteen, and twenty-four-lag specifications, with log-Bayes factors of 7.00, and 7.66 respectively.

2. It is also clear that in comparing alternative WCOs within a given monetary rule, most of the pairwise comparisons provide no more than substantial evidence between orderings.[18] One pattern that appears to be robust to money supply rule and lag length is the preference for placing the real (G) and/or monetary (M) sectors ahead of the interest rate (F). Of the eight lag length / money rule configurations, the only case that violates this regularity is the six lag, M-Rule, where the F-G-M ordering is preferred. This case aside, the results say that when the IR priors are filtered through the data, we are led to favor the federal funds rate as one of the faster reacting component of the model, incorporating more contemporaneous information than other sectors. The corollary to this of course is that the interest rate only affects these other sectors with a lag.

[16]See Tanner (1993), section 3.3.3. Appendix (6.5) provides the details of this sampling algorithm.

[17]As a basis of comparison, log marginal likelihood differences on the order of four and five approximately correspond to Bayes factors of 50:1 and 150:1 respectively.

[18]The sole decisive comparison is between GFM-R to FMG-R, within the twenty-four lag specification.

What feature of the IR prior generates this conclusion? We might get some idea by looking across models (rows) for deviations between the response priors and the structural responses in Figures 2 and 3. In Figure 2, the relative weak performance of the FGM-FMG (and GFM) orderings under a twelve lag, M-Rule specification may be attributable to the negative own response of the federal funds rate for horizons over forty months in these specifications. The IR prior here puts significant weight on zero or positive responses under an M-Rule, for the federal funds rate. Figure 3 suggests that the low marginal likelihood weighting of the FMG and FGM orderings under a twelve lag, R-Rule specification may be due to the price response to money demand (here non-borrowed reserves) at intermediate (5 year) horizons. The IR prior places greater weight on this response being positive, while for these orderings the response is within the zero region. If this graphical analysis is correct, neither of these inter-ordering comparisons would be robust to variations in the size of the "zero region."[19] The M-Rule result may weaken with widening the zero region width, while the R-Rule result may weaken with narrowing of the zero region width.

3. It is important to note that Figure 7 plots *log* marginal likelihoods. Varying the lag length has a substantial effect on the orderings, which seems to make it difficult to favor one ordering over another. This is consistent with Cooley and Dwyer (1998), who find similar evidence of a substantial impact of lag length variation on impulse response dynamics.

   But in the present context, this difficulty is more apparent than real. If we formally average over alternative lag lengths, some relatively stable rankings emerge. Table 5 uses three different priors over lag lengths to compare alternative orderings. The first column imposes an equal prior weighting on each lag length. The second column (Alt. 1) places greater weight on eighteen and twenty-four lag lengths, favoring them 2:1 over the six and twelve lag specifications. The last column (Alt. 2) places linearly increasing prior weight on longer lag specifications.[20]

   The same ordering dominates under all prior specifications. Under an M-Rule, the F-G-M ordering is preferred uniformly, while under an R-Rule, the G-F-M ordering is preferred. The second most preferred ordering under the R-Rule specification is M-G-F. These three orderings share the feature that the money supply reacts to contemporaneous fluctuations in real activity and prices. The two interest rate rule orderings differ as to whether monetary policy is capable of reacting to contemporaneous movements in money demand.

4. In accounting for lag length uncertainty, the priors above highlight the degradation in the relative performance of M-Rules for longer lag lengths. Table 6 presents the log Bayes factors, and level Bayes factors under the equal weighting, and two alternative priors used above.

   The evidence in favor of an R-rule specification increases as we place higher weight on the larger models. Even with equal weighting, the R-Rule specification is over one hundred times more consistent with the impulse response priors.

4.3.3 Mean Responses Implied by Dynamic Restrictions Under Alternative Orderings

In this section, I report estimated posterior mean impulse responses under the six alternative orderings examined above. Consistent with the Bayes factor comparisons, I limit attention to R-Rule specifications. I also maintain the twelve lag specification here.[21] For the seven informative responses, Figure 8 reports these posterior mean responses, together with standard errors of these estimates, and the conventional estimates of the structural responses (dashed lines). Note that these standard errors are *not* measures of the posterior standard deviation. Rather they are measures of the precision with which the posterior mean is estimated by the importance sampler.

[19]Recall that for the IR prior described in section (3), the width of each zero region is the standard deviation of the corresponding reduced form response.

[20]These alternative priors that place greater weight on larger models reflect my prior belief that the responses of smaller models tend to be more sensitive to the particular lag length chosen.

[21]These mean responses are computed using importance sampling of six thousand draws.

1. First note that the least squares estimates of the VAR are the posterior means, under the natural conjugate prior assumption. Thus the deviations here between the conventional response estimates and the sampler averaged estimates reflect the inclusion of the impulse response priors $\pi_{IR}(\psi)$.

2. The standard errors imply that the posterior means appear to be rather imprecisely determined. Using $\pm$ one standard error as a sensitive definition of the significance of the discrepancy between the conventional response estimates, and those based on the informative response priors, I find that many of the responses are not significantly altered by the informative priors.

3. Nevertheless, some responses do show substantial alterations. This arises where the impulse response priors are most at odds with the conventional response estimates, as is the case (under R-Rules) for the output response to money demand shocks. Here the weighting on negative responses has significantly pulled down the intermediate response across orderings. A second response which is substantially affected by the priors is the price response to money demand. These posterior means rise faster and then remain flat, relative to the conventional response estimates. This arises because the six and twelve month prior horizons attribute significant probability to positive responses, and the longer horizon components of the prior are consistent with the data, and so do not have much effect.

5 Conclusions and Extensions

5.1 Extensions

Here are a number of areas where these techniques might be fruitfully extended.

Prior Sensitivity It is straightforward to vary hyperparameters of the response priors, to assess the sensitivity of Bayes factors. It is also straightforward to examine how model comparisons vary as alternative aspects of the response priors are removed, revealing those aspects of the priors that have the most significant effects. The magnitudes of the sign weightings, as well as the selection of horizons for evaluation, could be varied to examine prior sensitivity. Another dimension along which the results of this paper might be examined is in disaggregating the real (G) and monetary (M) components of the models.

Alternative Response Priors The priors presented in subsection 2.2 are meant to be illustrative of the type of criteria used in assessing SVARs. Other criteria are available. For example, Galí (1992) proposes a set of sign restrictions on impulse responses that he views as being common to all "Keynesian" models. More formally, impulse responses could also be derived from a specific, linearized, calibrated general equilibrium model. Such exercises tend to generate testable overidentifying restrictions, that are generally not taken very seriously in this literature. Limiting the role of such theoretical models to creating response criteria for the comparison of alternative models may be deemed more appropriate.

Alternative SVAR Identifications While the example of this paper limited comparisons to Wold causal orderings (WCOs), the same techniques can be applied to more general model identifications, including those derived from linearized, stochastic, dynamic general equilibrium models.

Bayes Factors and Composite Models While this paper has focused on model comparisons (testing), it is also possible to exploit Bayes factors for the purpose of pooling alternative models. In the context of measuring monetary policy, this could reveal more sophisticated, composite measures of monetary policy. Combining several simple identifications could also lead to richer measures of responses to structural shocks. See Kass and Raftery (1995) for a general discussion of the use of Bayes factors for pooling statistical models.

Accommodating Structural Shifts Some authors have argued that the Bayesian framework should incorporate historical announcements of changes in regimes by large economic agents. In models with monetary or financial components, examples of such regime changes are the Federal Reserve Chairmanship, and announced changes in Fed operating procedures. By restricting attention to such subsamples, the methods of this paper provide a Bayesian approach to comparisons between M-Rules and R-Rules over subsamples, analogous to the classical approach of (Bernanke and Mihov (1995)).

Combining policy rule variation across subsamples provides another metric for comparing alternative identifications.

Cointegration Restrictions This paper has not attempted to impose cointegrating information. Most, but not all of the SVAR literature generally avoids this issue. There are two features of cointegration worth distinguishing. One affects the reduced form estimation of the model, in terms of restricting the likelihood function (the reduced rank restriction). The second feature of cointegration is its affect on identification, when combined with additional structural assumptions regarding the economic identities of permanent and transitory shocks. Making comparisons between alternative identifications which do and do not impose cointegration involves comparing models that do and do not impose some testable restrictions. This departs from the class of models covered by this paper, in that the priors vary across models. I have argued that it may make sense to maintain the likelihood restrictions of cointegration across model comparisons that do and do not exploit this information for identification. This is the subject of future research.

5.2 Conclusions

Empirical Conclusions An examination of the responses generated by the composite prior showed that they were quite diffuse. The priors imposed contradicted *all* of the structural models examined here in one way or another. Indeed, had they not, they would not be very useful for making model comparisons.

Among the WCOs examined here, R-rules for money supply shocks show significantly greater coherence with the impulse response priors. This preference for R-rules appears to be robust to variations in lag length.

Comparisons between individual orderings showed that under either an M-Rule, or an R-Rule specification, that placing the policy variable last in the ordering was quite consistent with the response priors. This suggests that monetary policy incorporates contemporaneous information on production, prices, and money demand. Under the preferred R-Rule specifications, there is somewhat weaker evidence that monetary policy fully incorporates contemporaneous movements in money demand. The ranking of alternative orderings was robust across three different priors over lag lengths. Incorporating the information in the response priors alters the point estimates (i.e. posterior means) of the structural responses for each of the models in predictable directions, with the most significant changes occuring in responses that are most at odds with the response prior.

Methodological Conclusions The informative component of the response priors advocated here assigns probability weights to sign restrictions on impulse responses of interest, at selected horizons. Since VARs contain many hundreds of parameters, this impulse response prior is combined with an uninformative, natural conjugate prior to create a proper prior. Via importance sampling, model comparisons based upon marginal likelihoods were straightforward to conduct. It was also shown to be straightforward to examine the effects of response priors on posterior mean responses.

The motivation for this exercise is to formally incorporate prior information regarding the shape of structural impulse response functions into comparisons of alternative, exactly identified models. While response based comparisons have been conducted informally in the structural VAR literature as a matter of course, little has been done to make formal comparisons. In focusing on exactly identified models, I follow the SVAR literature, which generally avoids imposing or testing overidentifying restrictions.

This paper has been silent as to *where* this prior information regarding impulse responses comes from. Whatever its source, this information is taken seriously by researchers working with structural VARs. I have not questioned the validity of this information, but only the way that this information has been used.

I have been critical here of the informal approach to model comparison and selection. Using straightforward techniques, I have shown that we can make statistically meaningful comparisons between alternative SVAR identifications. The priors I propose are quite simple, and they are easy to communicate. This is in contrast to the vague, qualitative comments that have so far dominated the debate over the identification of SVARs. It is rather ironic that a research program that has embraced Bayesian inference for the purposes of computing error bands for impulse responses, and for parameter reduction in large dimension VARs, has yet to embrace proper priors over the implicit impulse response assessments it employs. [22] This paper has

[22] For papers advocating Bayesian methods for these problems, see Sims and Zha (1995) and Sims and Zha (1996).

shown that such an application of Bayesian statistics is direct and immediate, and will only strengthen the structural VAR research program.

6 Appendix

6.1 Constants of Integration

A prior sampling procedure gives us a means of evaluating the integrating constant for the composite prior. Since

$$\int \pi_{IR,\psi}(\psi|M_j)d\psi = \int \frac{\mathfrak{a}(M_j)\pi_0(\psi)\pi_{IR}(\mathcal{R}(\psi|M_j))}{\pi_0(\psi)}\pi_0(\psi|)d\psi = 1, \tag{6.1}$$

a simulation-consistent estimator of the integrating constant is given by

$$\hat{\mathfrak{a}}(M_j) = \left[N^{-1}\sum_{\iota=1}^{N}\pi_{IR}(\mathcal{R}(\psi^{\iota}|M_j))\right]^{-1}_{\psi^{\iota}\sim\pi_0(\psi)}. \tag{6.2}$$

6.2 Combination of Priors

Assume that we have two distinct priors on the parameters ψ of interest, $\pi_1(\psi|\mu)$ and $\pi_2(\psi|\nu)$. These priors depend on their respective hyperparameters μ and ν. Given the first prior, we want to update our beliefs once we have received the information in the second prior. To do this, we consider the hyperparameters of the second prior, namely ν to represent the signal inherent in this prior. Thus we want to construct a likelihood function corresponding to this signal. Via Bayes' Rule we have

$$\pi_2(\psi|\nu) = \mathcal{L}(\nu|\psi)\pi_0(\psi)/\pi_\nu(\nu). \tag{6.3}$$

This says that after rearranging, we can express a likelihood for this signal ν in terms of the prior π_2, and unconditional priors over both ψ and ν :

$$\mathcal{L}(\nu|\psi) = \pi_2(\psi|\nu)\pi_\nu(\nu)/\pi_0(\psi). \tag{6.4}$$

Combining this then with our initial prior π_1 yields an updated posterior distribution which reflects the prior information in both priors.

$$\begin{aligned} p(\psi|\nu,\mu) &\propto \mathcal{L}(\nu|\psi)\pi_1(\psi|\mu) \\ &\propto \pi_2(\psi|\nu)\pi_1(\psi|\mu)/\pi_0(\psi). \end{aligned} \tag{6.5}$$

Here the constant of proportionality is a function of the hyperparameters only, namely μ, ν. The final simplification we make is to assume that the underlying diffuse prior on ψ is proportional to a constant. In this sense, we are formally justified in applying products of priors to construct an overall prior that we then take to the data.

6.3 Natural Conjugate Benchmark Priors

The natural conjugate prior/posterior are applied to the reduced form parameterization $\mathbf{B}(z), H$, where $H = \Sigma^{-1}$. Letting $\mathbf{x}_t$ denote the np vector of lagged values of $\mathbf{y}_t$, the reduced form (1.2) can be expressed as

$$\mathbf{Y} = \mathbf{X}B + \mathbf{U}, \tag{6.6}$$

where

$$B = \begin{bmatrix} B_1' \\ \vdots \\ B_p' \end{bmatrix}. \tag{6.7}$$

Then let $\beta = vec(B)$, so that

$$vec(\mathbf{Y}) = (I_n \otimes \mathbf{X})\beta + vec(\mathbf{U}). \tag{6.8}$$

For $\mathbf{u}_t \sim \mathcal{NID}(0, H^{-1})$, the (Normal-Wishart) natural conjugate prior specifies that

$$\pi(\beta|H) = \mathcal{N}(\underline{\beta}, H^{-1} \otimes \underline{W}^{-1}), \tag{6.9}$$

$$\pi(H) = \mathcal{W}(\underline{G}, \underline{\nu}). \tag{6.10}$$

When combined with the multivariate normal likelihood for (6.6), this yields the Normal Wishart posterior

$$p(\beta, H|\mathbf{Y}) = \mathcal{N}_\beta(\bar{\beta}, H^{-1} \otimes \bar{W}^{-1})\mathcal{W}_H(\bar{G}, \bar{\nu}), \tag{6.11}$$

where

$$\begin{aligned} \bar{W} &= \underline{W} + \mathbf{X}'\mathbf{X}, \\ \bar{\beta} &= vec(\bar{B}), \text{ where } \bar{B} = \bar{W}^{-1}(\underline{W}\,\underline{B} + \mathbf{X}'\mathbf{X}\hat{B}), \\ \bar{\nu} &= \underline{\nu} + T, \\ (\bar{G}^{-1} + \bar{B}'\bar{W}\bar{B}) &= (\underline{G}^{-1} + \underline{B}'\underline{W}\,\underline{B}) + (\hat{\mathbf{U}}'\hat{\mathbf{U}} + \hat{B}'\mathbf{X}'\mathbf{X}\hat{B}). \end{aligned} \tag{6.12}$$

The choice of the prior hyperparameters for the natural conjugate distribution was motivated by several concerns. First, I want to make this prior "uninformative," especially for the mean parameters β. To do this, I set $\underline{\beta} = 0$, and I require that $\underline{W}$ is small relative to $\mathbf{X}'\mathbf{X}$. For the specification with twelve lags, I set $\underline{W} \propto 1/1000$.

Second, since the dimension of $\underline{W}$ is a function of the number of lags, I vary the scale of this parameter as the number of lags grows to ensure that $|\underline{W}| \neq 0$. I choose $\underline{W}$ so that $|\underline{W}| = |\underline{W}|^*$ is constant across specifications. This ensures that I will be able to invert $\underline{W}$, which is necessary for evaluating the prior density.

Third, if the variability of the posterior is too large, I will generate many draws of the mean parameters β that imply explosive roots. To mitigate this effect, I choose $\underline{G}$ sufficiently large, specifically $\underline{G} \propto 1000$.

Alternative specifications are found in the literature. The standard "Minnesota Prior" (Doan, Litterman, and Sims (1984)) sets the prior means of first lags of own coefficients equal to unity rather than zero, to reflect the prior knowledge that the data is likely to be I(1). Also, this same prior weights higher lag coefficients with a smaller variance (corresponding to a larger element in $\underline{W}$). This restriction has been found to improve the forecasting properties based upon the resulting posterior distribution. Sims and Zha (1996) propose an elaboration on the Minnesota prior. Such priors might be useful in the current context, in that they would help to distinguish the model from a standard multivariate regression specification. This topic is left for future research.

6.4 Computation Times and Sampler Design

All of the structural models compared in this paper share a common reduced form specification. The sampling used to evaluate marginal likelihoods is based upon draws from the common reduced form prior and posterior. I cut computing times in half (for the twelve lag case) by making these prior and posterior draws once, and using the same realizations to compute all marginal likelihoods. On a Pentium 120MHz pc, running Gauss for OS/2 (ver. 4), the six thousand draws of prior and posterior realizations for the baseline (twelve lag) model took 1.23 hours. From these realizations, computation of the marginal likelihood of a given structural model, on a Pentium 90MHz pc, took approximately 4 hours (again running Gauss for OS/2, version 4).

The marginal likelihood evaluation depends upon examining the structural impulse responses for each parameter realization. These responses are computed recursively (via a "DO-loop"). Consequently the marginal likelihood evaluation time depends crucially on the maximum impulse response horizon considered in the informative prior. For the baseline case, this was one hundred and twenty, corresponding to a 10 year horizon.

6.5 Rejection/Acceptance Sampling of Impulse Responses

Exploiting a theorem of Ripley (1987), the following two, readily verifiable conditions are sufficient for applying R/A sampling.

1. Conditional on a given model M_j and the reduced form parameters ψ, the conditional density for $\mathbf{R}$ is simply an indicator function for $\mathcal{R}(\psi)$. That is, $\mathbf{R} = \mathcal{R}(\psi|M_j)$ is a deterministic function of ψ. Drawing $\mathbf{R}$ from this conditional density means computing $\mathcal{R}(\psi|M_j)$.

2. The following inequality holds:

$$\frac{\pi_{IR,\psi}(\psi|M_j)}{\pi_0(\psi)} \leq \alpha(M_j)\pi^*_{IR}$$

   where π^*_{IR} is the maximum realizable value of the informative prior $\pi_{IR}(\mathcal{R}_{I_q}(\psi))$. For the specific prior used in this paper, $\pi^*_{IR} = 4.21 \times 10^{-8}$.

Under these conditions, the R/A sampler works as follows.

1. Draw ψ^* from $\pi_0(\psi)$.
2. Draw u^* from a uniform $[0,1]$ distribution.
3. If

$$u^* \leq \frac{\pi_{IR}(\mathcal{R}(\psi^*))}{\pi^*_{IR}},$$

   evaluate and keep $\mathcal{R}(\psi^*)$. Otherwise return to the first step.
4. Repeat these steps to draw additional realizations from implied IR prior distribution for $\mathbf{R}$.

References

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 5145.

BLANCHARD, O., AND D. QUAH (1989): "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, 79(4), 655–673.

CHRISTIANO, L. J., AND M. EICHENBAUM (1992a): "Identification and the Liquidity Effect of a Monetary Policy Shock," in *Political Economy, Growth, and Business Cycles*, ed. by Z. H. Alex Cukierman, and L. Leiderman, chap. 12, pp. 335–370. MIT Press.

——— (1992b): "Liquidity Effects, Monetary Policy, and the Business Cycle," Federal Reserve Bank of Minneapolis, Institute for Empirical Macroeconomics Discussion Paper 70.

CHRISTIANO, L. J., M. EICHENBAUM, AND C. EVANS (1996): "The Effects of Monetary Shocks: Evidence from the Flow of Funds," *Review of Economics and Statistics*, 78(1), 16–34.

COOLEY, T. F., AND M. DWYER (1998): "Business Cycle Analysis Without Much Theory: A Look at Structural VARs," *Journal of Econometrics*, 83, 57–88.

COOLEY, T. F., AND S. F. LEROY (1985): "Atheoretical Macroeconometrics: A Critique," *Journal of Monetary Economics*, 16, 283–308.

DOAN, T., R. LITTERMAN, AND C. SIMS (1984): "Forecasting and Conditional Projection Using Realistic Prior Distributions," *Econometric Reviews*, 3, 1–144.

EICHENBAUM, M. (1992): "Comments on 'Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy,' by Christopher Sims," *European Economic Review*, 36, 1001–1011.

EICHENBAUM, M., AND C. L. EVANS (1995): "Some Empirical Evidence on the Effects of Shocks to Monetary Policy on Exchange Rates," *Quarterly Journal of Economics*, 110(4), 975–1009.

FAUST, J. (1998): "The Robustness of Identified VAR Conclusions About Money," Board of Governors of the Federal Reserve System, International Finance Discussion Paper no. 610.

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

GEWEKE, J. (1989): "Bayesian Inference In Econometric Models Using Monte Carlo Integration," *Econometrica*, 57(6), 1317–1339.

GORDON, D. B., AND E. M. LEEPER (1994): "The Dynamic Impacts of Monetary Policy: An Exercise in Tentative Identification," *Journal of Political Economy*, 102(6), 1228–1247.

HARKEMA, R. (1971): *Simultaneous Equations, A Bayesian Approach*. Rotterdam University Press.

JEFFREYS, H. (1961): *Theory of Probability*. Oxford University Press, 3rd edn.

KASS, R., AND A. RAFTERY (1995): "Bayes Factors," *Journal of the American Statistical Association*, 90, 773–795.

KING, R. G., C. I. PLOSSER, J. H. STOCK, AND M. W. WATSON (1991): "Stochastic Trends and Economic Fluctuations," *American Economic Review*, 81, 819–840.

LASTRAPES, W. D., AND G. SELGIN (1995): "The Liquidity Effect: Identifying Short-Run Interest Rate Dynamics Using Long-Run Restrictions," *Journal of Macroeconomics*, 17, 387–404.

LEEPER, E. M., AND D. B. GORDON (1992): "In Search of the Liquidity Effect," *Journal of Monetary Economics*, 29(3), 341–369.

LEEPER, E. M., C. A. SIMS, AND T. ZHA (1996): "What Does Monetary Policy Do?," *Brookings Papers on Economic Activity*, (2), 1–63.

PAGAN, A., AND J. ROBERTSON (1996): "Structural Models of the Liquidity Effect," Australian National University working paper.

PAGAN, A. R., AND J. C. ROBERTSON (1995): "Resolving the Liquidity Effect," *Federal Reserve Bank of St. Louis Review*, 77(3), 33–54.

RIPLEY, B. (1987): *Stochastic Simulation.* Wiley, New York.

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

——— (1993): "A Nine-Variable Probabilistic Macroeconomic Forecasting Model," in *Business Cycles, Indicators and Forecasting*, ed. by J. H. Stock, and M. W. Watson, chap. 4, pp. 179–204. University of Chicago Press for the NBER.

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

——— (1986): "Are Forecasting Models Usable for Policy Analysis?," *Federal Reserve Bank of Minneapolis Quarterly Review*, Winter, 2–16.

SIMS, C. A., AND T. ZHA (1995): "Error Bands for Impulse Responses," FRB of Atlanta working paper #95-6.

——— (1996): "Bayesian Methods for Dynamic Multivariate Models," Federal Reserve Bank of Atlanta working paper 96-13.

STRONGIN, S. (1995): "The Identification of Monetary Policy Disturbances: Explaining the Liquidity Puzzle," *Journal of Monetary Economics*, 35(3), 463–497.

TANNER, M. A. (1993): *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions.* Springer-Verlag, New York.

UHLIG, H. (1997): "What Are The Effects of Monetary Policy? Results From An Agnostic Identification Procedure," CentER, Tilburg University working paper.

Table 1: Standard Deviations of Reduced Form Responses[a]

	u_y	u_p	u_{cp}	u_{ff}	u_{nbr}	u_{tr}
y	0.0020	0.0013	0.0024	**0.0007**	**0.0017**	0.0016
p	0.0007	0.0010	0.0015	**0.0018**	**0.0022**	0.0023
cp	0.0008	0.0032	0.0105	0.0016	0.0023	0.0024
ff	0.0440	0.0567	0.0788	**0.1087**	**0.0676**	0.0652
nbr	0.0016	0.0015	0.0032	**0.0014**	0.0034	0.0023
tr	0.0012	0.0016	0.0036	**0.0014**	**0.0023**	0.0022

[a] Values used in constructing the IR priors are indicated in **bold**.

Table 2: Trinomial Prior Response Parameterization[a]

M-Rule Prior Specifications

		Horizon (Months)					
M-Rule		1	6	12	24	60	120
NB⟹FF	+	0.05	0.05	0.05	0.05	0.05	0.05
	0	0.15	0.27	0.39	0.51	0.63	0.75
	−	0.80	0.68	0.56	0.44	0.32	0.20
NB⟹PP	+	0.20	0.32	0.44	0.56	0.68	0.80
	0	0.75	0.63	0.51	0.39	0.27	0.15
	−	0.05	0.05	0.05	0.05	0.05	0.05
NB⟹YY	+	0.80	0.68	0.56	0.44	0.32	0.20
	0	0.15	0.27	0.39	0.51	0.63	0.75
	−	0.05	0.05	0.05	0.05	0.05	0.05
FF⟹TR	+	0.20	0.32	0.44	0.56	0.68	0.80
	0	0.75	0.63	0.51	0.39	0.27	0.15
	−	0.05	0.05	0.05	0.05	0.05	0.05
FF⟹PP	+	0.10	0.24	0.38	0.52	0.66	0.80
	0	0.10	0.10	0.10	0.10	0.10	0.10
	−	0.80	0.66	0.52	0.38	0.24	0.10
FF⟹YY	+	0.05	0.05	0.05	0.05	0.05	0.05
	0	0.75	0.63	0.51	0.39	0.27	0.15
	−	0.20	0.32	0.44	0.56	0.68	0.80
FF⟹FF	+	0.80	0.68	0.56	0.44	0.32	0.20
	0	0.15	0.27	0.39	0.51	0.63	0.75
	−	0.05	0.05	0.05	0.05	0.05	0.05

[a] Entries are prior probabilities associated with given sign restriction (+ 0 −) at given horizon, for each response.

Table 3: Trinomial Prior Response Parameterization[a]

R-Rule Prior Specifications

		Horizon (Months)					
R-Rule		1	6	12	24	60	120
FF⟹NB	+	0.05	0.05	0.05	0.05	0.05	0.05
	0	0.15	0.15	0.15	0.15	0.15	0.15
	−	0.80	0.80	0.80	0.80	0.80	0.80
FF⟹PP	+	0.05	0.05	0.05	0.05	0.05	0.05
	0	0.75	0.63	0.51	0.39	0.27	0.15
	−	0.20	0.32	0.44	0.56	0.68	0.80
FF⟹YY	+	0.05	0.05	0.05	0.05	0.05	0.05
	0	0.15	0.27	0.39	0.51	0.63	0.75
	−	0.80	0.68	0.56	0.44	0.32	0.20
NB⟹TR	+	0.20	0.32	0.44	0.56	0.68	0.80
	0	0.75	0.63	0.51	0.39	0.27	0.15
	−	0.05	0.05	0.05	0.05	0.05	0.05
NB⟹PP	+	0.10	0.24	0.38	0.52	0.66	0.80
	0	0.10	0.10	0.10	0.10	0.10	0.10
	−	0.80	0.66	0.52	0.38	0.24	0.10
NB⟹YY	+	0.05	0.05	0.05	0.05	0.05	0.05
	0	0.75	0.63	0.51	0.39	0.27	0.15
	−	0.20	0.32	0.44	0.56	0.68	0.80
NB⟹FF	+	0.80	0.68	0.56	0.44	0.32	0.20
	0	0.15	0.27	0.39	0.51	0.63	0.75
	−	0.05	0.05	0.05	0.05	0.05	0.05

[a] Entries are prior probabilities associated with given sign restriction (+ 0 −) at given horizon, for each response.

Table 4: Logarithms of WCO-Averaged Marginal Likelihoods

	M-Rule	R-Rule	Difference
6 lags	0.414	1.500	1.086
12 lags	-2.835	1.925	4.760
18 lags	-3.842	3.153	6.995
24 lags	-2.532	5.132	7.664

Table 5: Log Marginal Likelihoods Averaged Over Lag Specifications:

Alternative Prior Lag Weightings[a]

	Equal	Alt. 1	Alt. 2
GMF-M	-1.993	-1.911	-1.824
GFM-M	-1.020	-1.375	-1.787
MFG-M	-3.155	-3.333	-3.195
MGF-M	-2.755	-2.751	-2.664
FMG-M	-2.655	-3.029	-3.469
FGM-M	0.608	0.203	-0.305
GMF-R	3.115	3.246	3.291
GFM-R	5.378	5.658	5.823
MFG-R	2.935	3.149	3.299
MGF-R	3.701	3.948	4.039
FMG-R	0.027	0.120	0.152
FGM-R	1.891	2.129	2.256

[a] "Equal" refers to equal weighting of lag lengths. "Alt. 1" refers to a weighting $\{0.167, 0.167, 0.333, 0.333\}$ for lags $\{6, 12, 18, 24\}$ respectively. "Alt. 2" refers to a weighting of $\{0.1, 0.2, 0.3, 0.4\}$.

Table 6: Bayes Factors in Favor of an R-Rule Specification:

Alternative Prior Lag Weightings[a]

	Equal	Alt. 1	Alt. 2
log BF	4.804	5.410	5.941
level BF	121.987	223.698	380.382

[a] "log BF" refers to log Bayes factors, while "level BF" refers to the (level) Bayes Factor. For the definition of the lag priors used, see footnote to Table 5.

Figure 1: Data for Six Variable Vector Autoregression

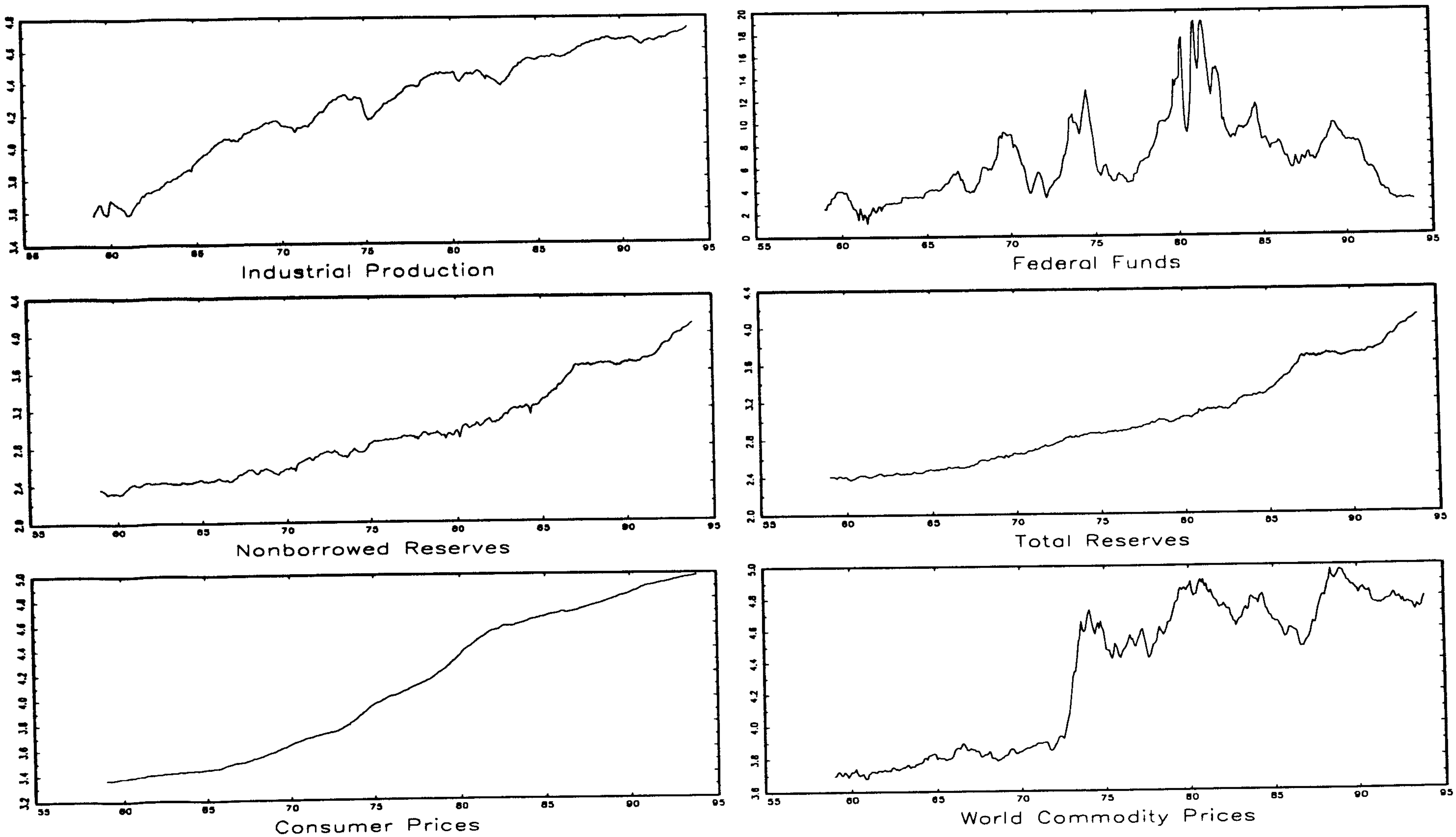


Figure 2: Structural Responses and Impulse Response Priors, M-Rule, 12 Lags

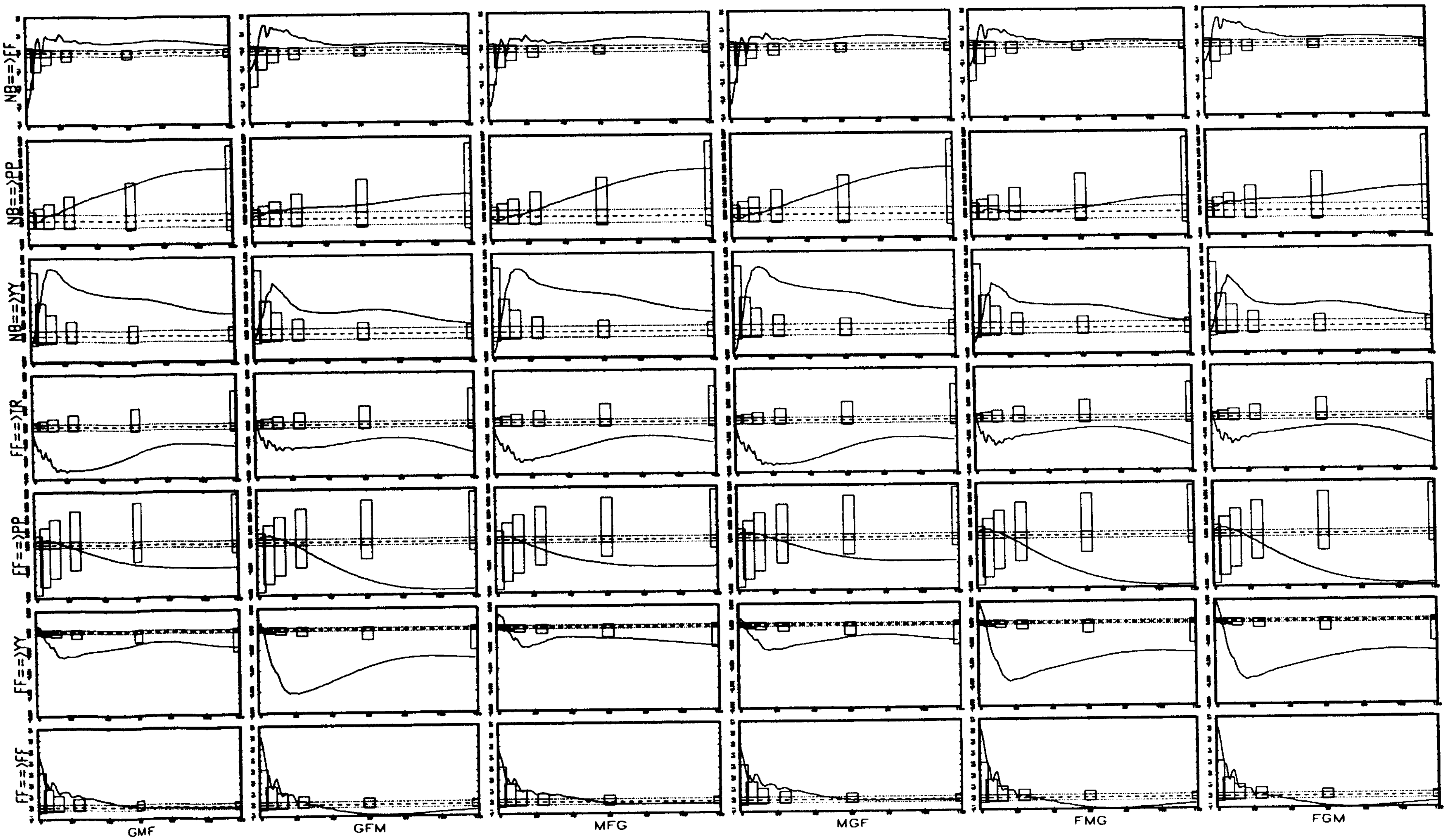


Figure 3: Structural Responses and Impulse Response Priors, R-Rule, 12 Lags

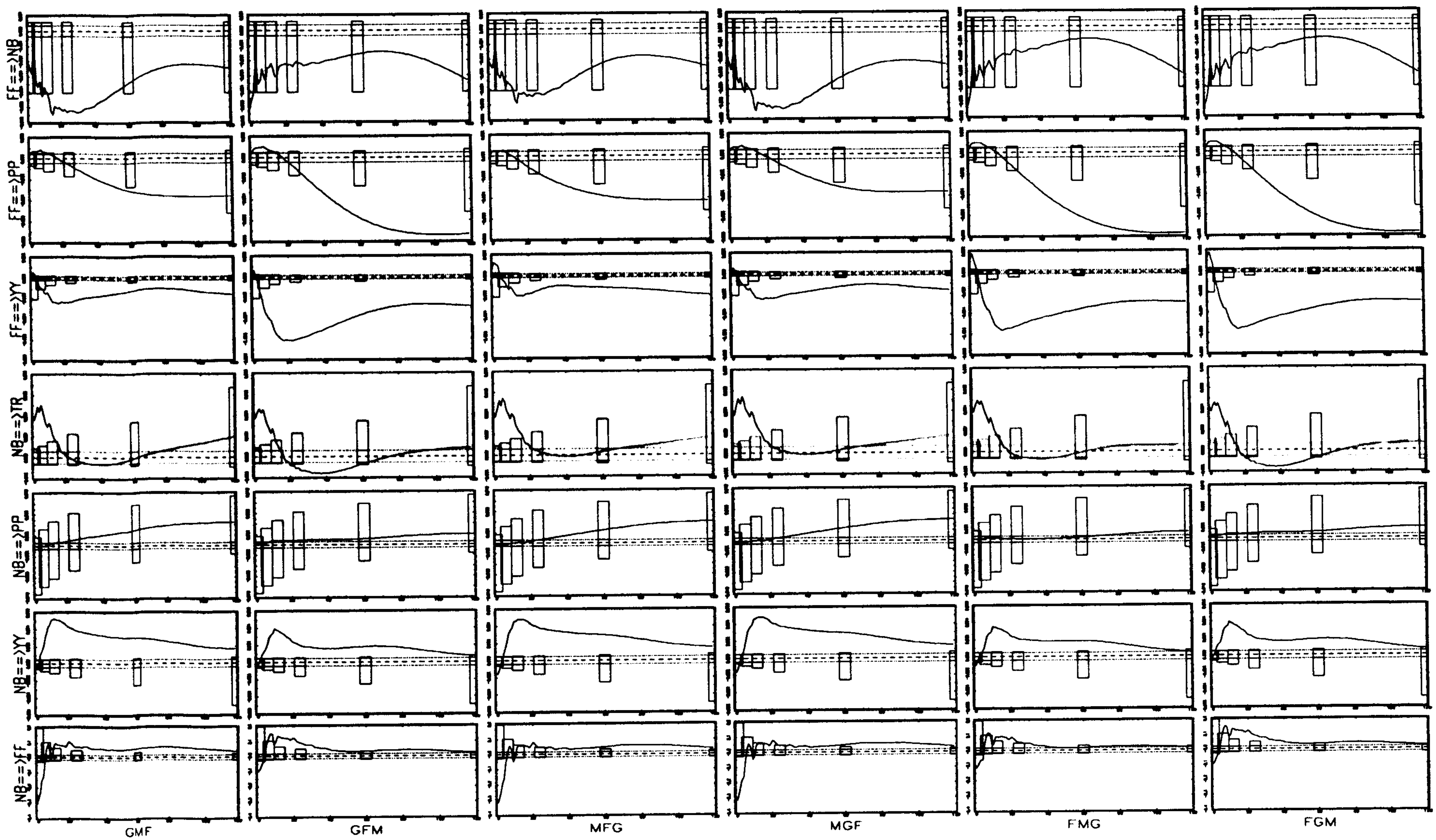


Figure 4: Standard Deviations of IR Prior Impulse Responses, M-Rule

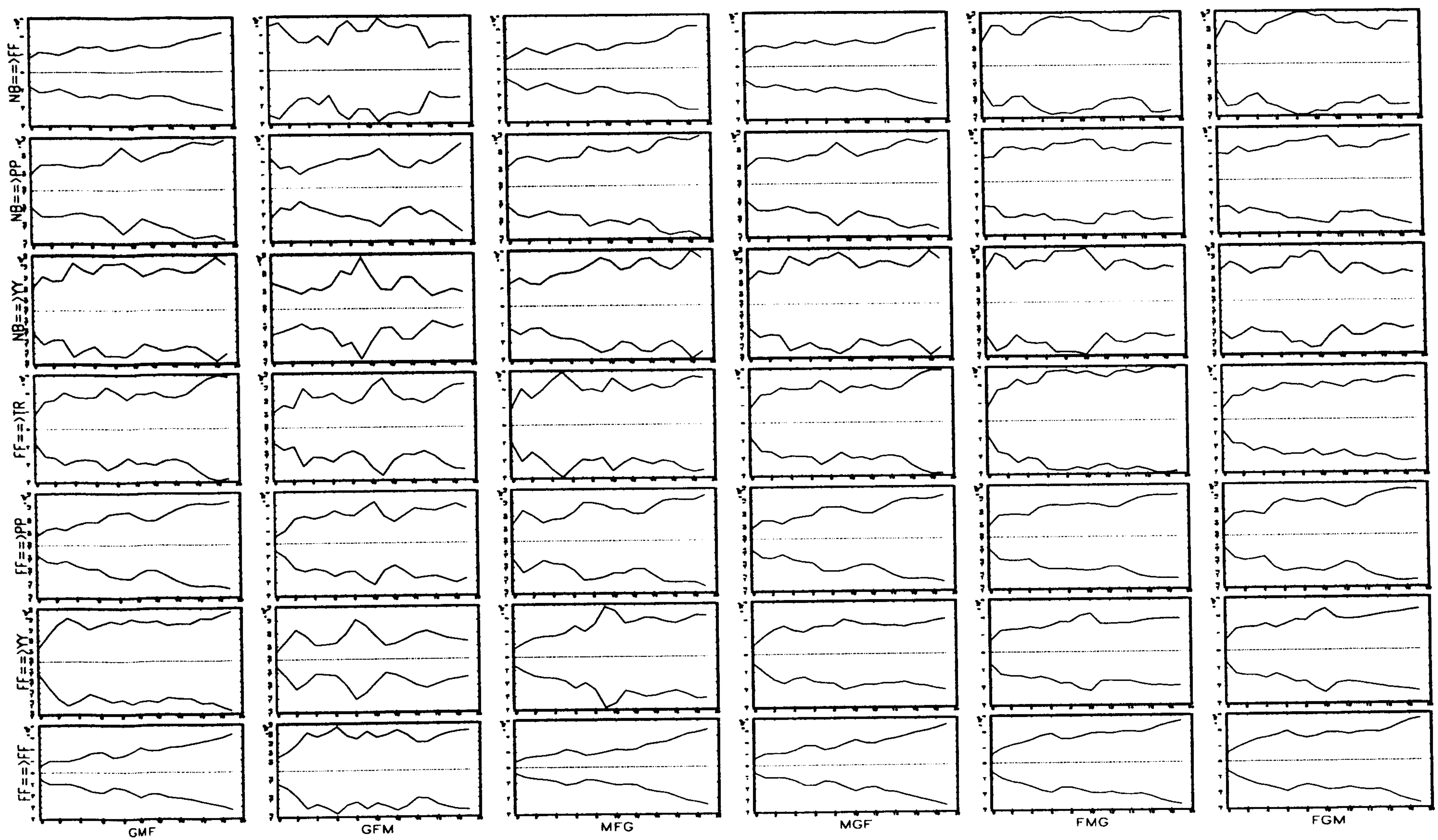


Figure 5: Means of IR Prior Impulse Responses, M-Rule

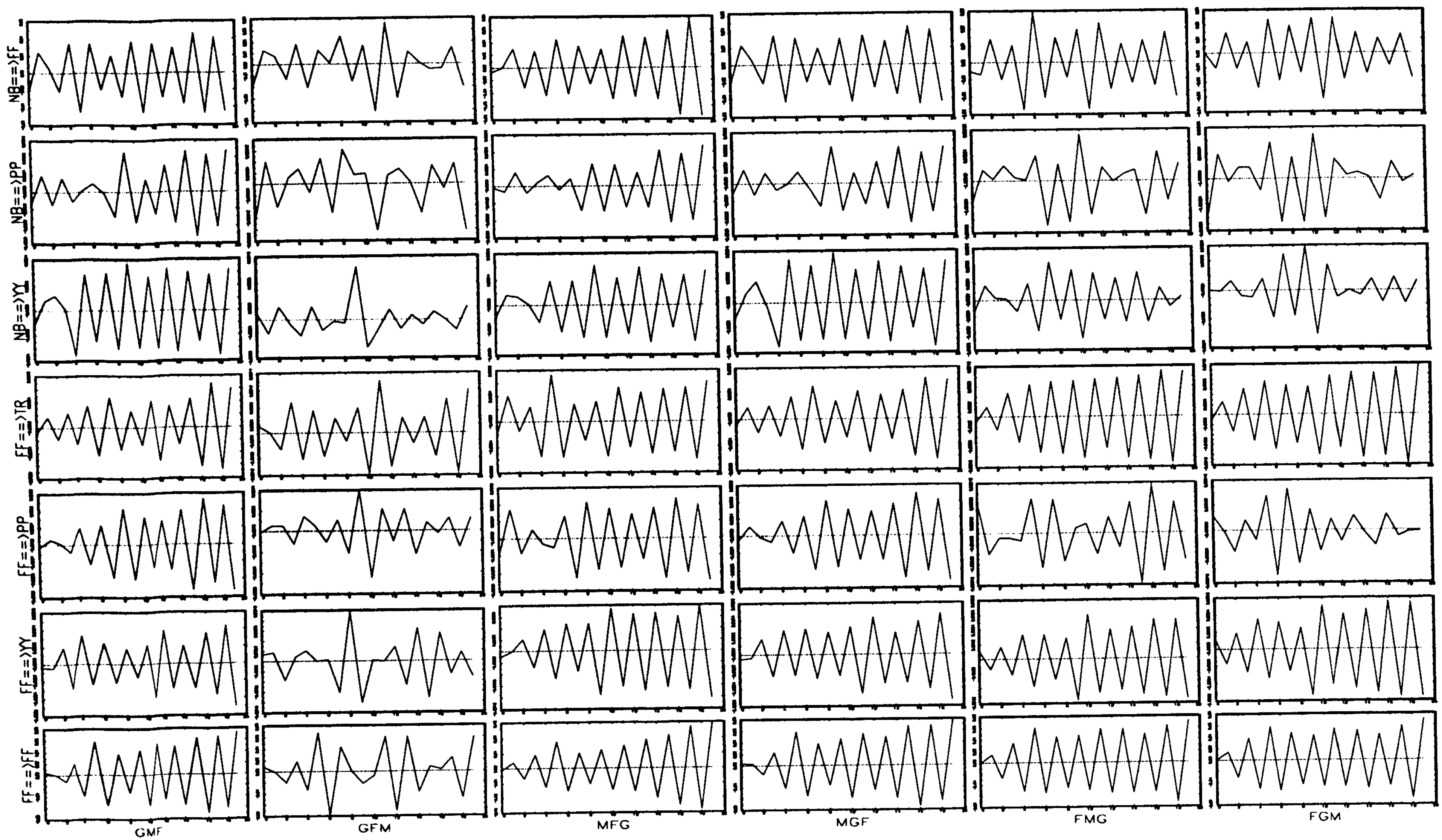


Figure 6: Marginal Likelihood Convergence via Importance Sampling

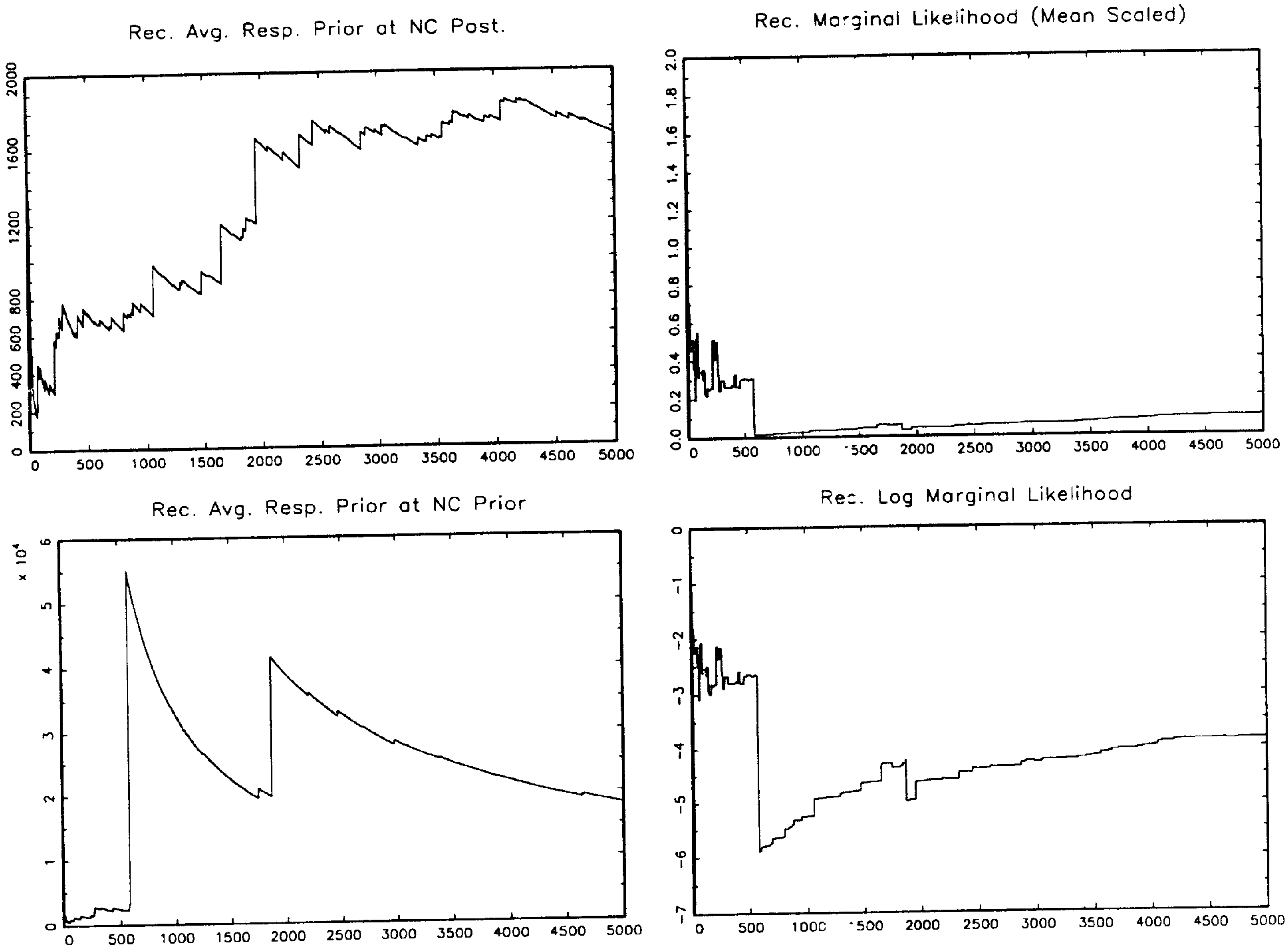
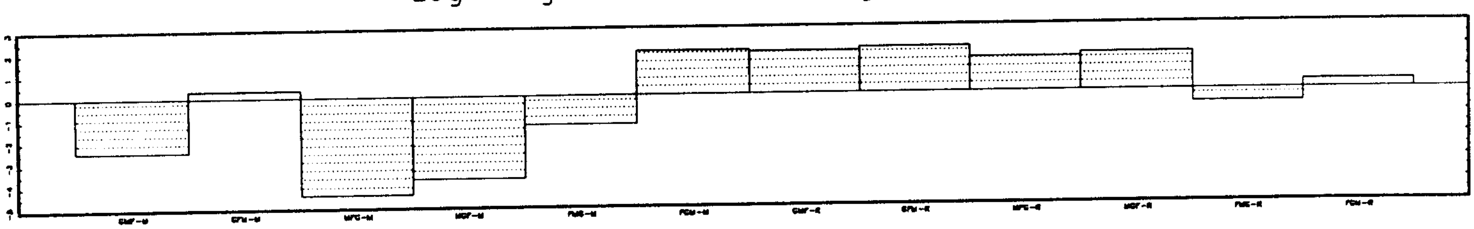
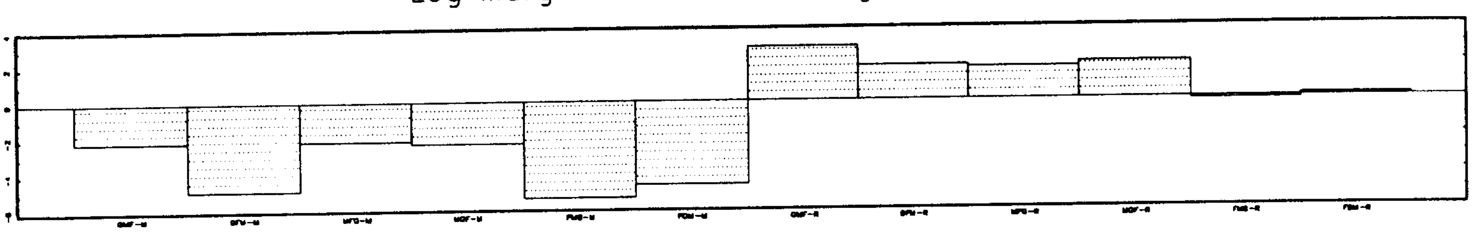


Figure 7: Marginal Likelihood Comparisons for Alternative Orderings & Rules

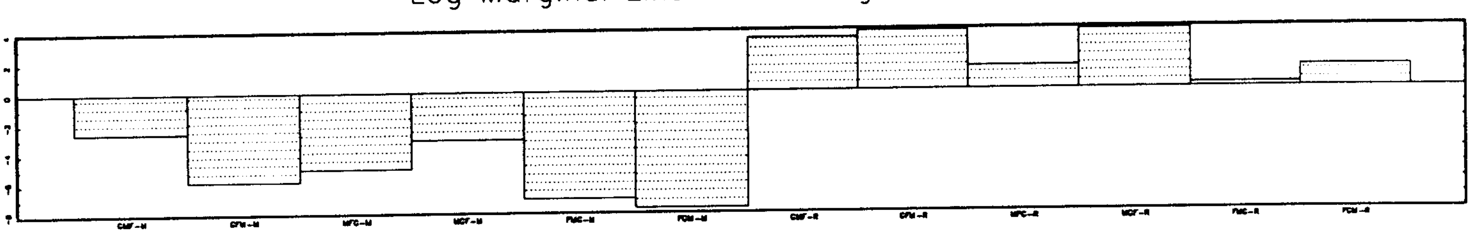
Log Marginal Likelihoods: Lag=06, 6K Runs



Log Marginal Likelihoods: Lag=12, 6K Runs



Log Marginal Likelihoods: Lag=18, 6K Runs



Log Marginal Likelihoods: Lag=24, 6K Runs

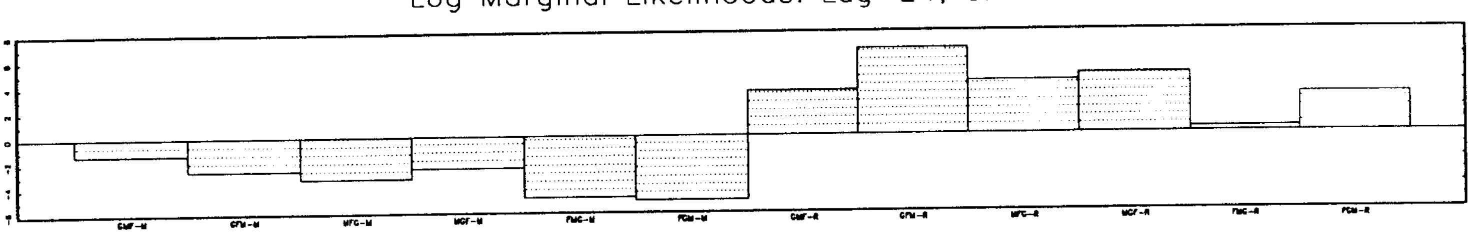


Figure 8: Posterior Mean and Importance Sampling
Standard Errors for Structural Responses
R-Rule, 12 Lags