Flexible prior beliefs on impulse responses in Bayesian vector autoregressive models*

Fabio Canova[†] Andrzej Kocięcki [‡] Michele Piffer[§]

November 25, 2023

Abstract

We develop a prior for the coefficients of a VAR that allows for flexible non-dogmatic beliefs on the shape of the structural impulse responses. We achieve this using an alternative setting of the moments of a Normal prior distribution on the autoregressive parameters of the model. The methodology we propose is computationally no more demanding than existing prior specifications; yet it offers a tool for Bayesian shrinkage over key outputs of the model. Introducing the prior belief that monetary policy shocks generate temporary but persistent effects leads to a hump-shaped response of GDP, with the peak response occurring between twelve and eighteen months after the shock.

JEL classification: C32, E52.

Keywords: Non-dogmatic beliefs, impulse responses, monetary policy, identification, structural shocks.

*We are thankful to Marta Bańbura, Luca Gambetti, Paolo Gelain, Michele Lenza, Christian Matthes, Giovanni Ricco, Giuseppe Ragusa, Herman van Dijk and Alessio Volpicella for helpful comments and suggestions. We also thank the participants of the 2022 IAAE annual congress and the 2022 Sailing the Macro workshop for comments on the early stage of this project.

[†]BI Norwegian Business School, Oslo, Norway. e-mail: fabio.canova@bi.no

[‡]University of Warsaw, Warsaw, Poland. e-mail: akociecki@wne.uw.edu.pl

[§]King's Business School, King's College London, UK. e-mail: m.b.piffer@gmail.com

1 Introduction

Impulse response functions (IRFs) are one of the most popular tools in modern macroeconomics and they have proved essential in exploring how structural shocks affect the economy. Applications of impulse response analysis include studying how the economy responds to policy interventions (Caldara and Kamps, 2017, Miranda-Agrippino and Ricco, 2021), to financial disruptions (Gilchrist and Zakrajšek, 2012), and to geopolitical and uncertainty-related risks (Piffer and Podstawski, 2018), among others.

Researchers typically hold views on what dynamic responses ought to be considered reasonable. Beliefs are commonly held on the timing, on the persistence of the responses, as well as on whether they are hump-shaped or not. For example, it is generally believed that a contractionary monetary policy shock should persistently decrease inflation, even though the effect may take time to materialize. The long standing debates about the 'price puzzle', the 'liquidity puzzle', and the 'exchange rate puzzle' provide other examples of the beliefs researchers may have about the likely path of structural responses (Ramey, 2016 Gourinchas and Tornell, 2004). We focus our analysis on non-dogmatic beliefs, that is beliefs that are not imposed for identification since, otherwise, one would introduce unnecessary restrictions on the data.

Unfortunately, macroeconomists are severely constrained in their ability to support the estimation of impulse response functions with non-dogmatic prior beliefs on their timing, persistence or shape. The computational convenience of working with Vector Autoregressive models has been acknowledged in the literature (Kilian and Lütkepohl, 2017). Yet, the existing practice of using a Minnesota-like prior in SVARs does not allow a researcher to non-dogmatically introduce, even indirectly, appropriate constraints on the impulse response functions. Beliefs on the shape of the impulse responses can be naturally introduced in Moving Average (MA) models (Plagborg-Møller, 2019). Yet, computationally, MA models are demanding to work with, and may require approximation techniques (Barnichon and Matthes, 2018). An alternative is to work with Local Projections models and impose priors directly on the coefficients of the relevant shock at all horizons. However, an unresolved challenge in working with the Bayesian version of Local Projections is the specification of a meaningful covariance structure for the residuals (Ferreira et al., 2023).

The contribution of the paper. This paper proposes a prior structure for the autoregressive parameters of SVAR models that contemporaneously achieves two goals.

First, the moments of the prior can be selected so that the implied distribution on IRFs is approximately centered around the values researchers select, with the second moment determining the width of the prior IRFs intervals. Second, because it has a Normal format, it retains the computational convenience of using highly tractable posterior samplings. Thus, one can introduce non-dogmatic beliefs about features of the impulse responses by simply replacing the Minnesota-like specification with our proposed specification, which is obtained analytically from the recursive computation of IRFs. Our prior is not designed for identification purposes and can be combined with existing identification strategies, for example, zero restrictions, sign restrictions or external instruments. It is also compatible with both the approach of Rubio-Ramirez et al. (2010) and of Baumeister and Hamilton (2015), because it remains agnostic about the prior specification on the contemporaneous impulse responses and/or the structural elasticities.

There are four main advantages of our settings. It has been acknowledged that Bayesian shrinkage is more naturally introduced on endogenous outcomes of the models rather than on its parameters (see, for instance, Van Dijk and Kloek, 1980 and Harvey et al., 2007). In the SVAR literature, priors on the endogenous outcomes of the model are currently viable only when considering the unconditional properties of the observables (see Villani, 2009 and Jarociński and Marcet, 2019). The first advantage of our method is to allow for direct shrinkage of the impulse responses, which are undoubtedly a key outcome of SVAR models (Kilian, 2022). A second advantage is that, by leading the data to produce particular shapes, it can sharpen inference, even when exact identification restrictions are not introduced. This offers an alternative to the current practice of progressively tightening identifying restrictions when the uncertainty about the IRFs remains large (Kilian and Murphy, 2012). The third advantage is that the specification is flexible enough to introduce a combination of tighter beliefs on some impulse responses and looser beliefs on others. Because the SVAR jointly estimates all impulse responses at once, sharper inference can then be obtained also for the dynamic responses of variables for which prior beliefs are loosely formulated. Finally, our prior can be applied to a wide class of models including linear Gaussian SVARs, linear non-Gaussian SVARs and nonlinear transition switching VARs when computing impulse responses within or across states.

We illustrate the properties of our methodology using data simulated from a standard three-variable New Keynesian model. The model features a very persistent effect of a government spending shock on the output gap. We show that in a sample of a realistic size, both a flat and a Minnesota prior, under either a white noise or a random walk specifications, lead to posterior IRFs that largely underestimate the degree of persistence of the output gap response. By contrast our prior, which allows for the belief that the effect is relatively persistent, leads to posterior IRFs that mimic the true responses, provided the prior variance is not too large. In addition, the prior imposed on the responses of the output gap to a government spending shock sharpens inference on the responses to all shocks, making credible sets uniformly smaller. We discuss how to select the parameters of the prior specification and make it operational in applied work. As expected, in a large sample the prior becomes irrelevant, and all prior approaches lead to the same impulse response outcome.

We study how US output responds to a US monetary policy interventions. This issue has received considerable attention over the last twenty years (Christiano et al., 1999, Uhlig, 2005) but the shape of the output dynamics still remains unsettled in the literature. We take a standard VAR model with six variables and identify the monetary policy disturbances using sign restrictions on the impact effect of the shock. The flat and the Minnesota prior lead to a posterior IRFs with exponentially declining shape and no hump. When we impose the belief that monetary shocks generate temporary but persistent effects on the variables of the model, a belief which is in line with a wide class of current New Keynesian macroeconomic models, the posterior distribution of real GDP responses displays hump-shaped dynamics, independently of degrees of persistence imposed with our prior. More importantly, this occurs even though our prior does not restrict responses to be hump-shaped.

Interestingly, our results indicate that it takes between one and one and a half year for the monetary shock to generate its largest GDP effect. Quantitatively speaking, a one standard deviation shock that increases the federal funds rate on impact by 20 basis points, leads to approximately a maximum decrease in real GDP of 0.20%. Thus, our results support the widely-held view that a central bank may affect real economic activity; but this occurs with long and variable lags, and the magnitude of the response is generally small (Buda et al., 2023).

The relationship with the literature. There is a large literature on Bayesian VAR models, see Koop and Korobilis (2010) and Miranda-Agrippino and Ricco (2019) for a detailed discussion. We build on Baumeister and Hamilton (2015, 2018), who raised academic attention on the benefit of explicitly introducing prior beliefs on the structural objects of interest. We extend their analysis by focusing on beliefs at longer horizons, and concentrating on non-dogmatic priors for autoregressive parameters rather than on identifying restrictions. The paper is also related to the work of Barnichon and Matthes (2018), who impose prior shape restrictions on the MA models, and to an earlier contribution of Kociecki (2010), who works with recursive identification and considers the case in which the prior on the impulse responses must be jointly Normal. Our approach has the same flavour as the ones of Villani (2009) and Andrle and Benes (2013), who construct priors for endogenous objects of a model. The need for tools that explicitly introduce non-dogmatic beliefs on impulse responses was acknowledged early on by Gordon and Boccanfuso (2001) and Dwyer (1998), who nevertheless do not proceed to construct a prior for impulse responses. Our method is also compatible with the approach by Arias et al. (2018), who provide a tool for combining sign and zero restrictions. While we do not follow the method of Giacomini and Kitagawa (2021), we acknowledge that a robust approach to sign restrictions can aid inference also with our methodology.

The rest of the paper is organized as follows. Section 2 discusses in detail the prior beliefs on impulse responses we propose in this paper. Section 3 illustrates the properties of our prior specification using simulated data. Section 4 studies how monetary policy shocks are transmitted to real output. Section 5 concludes. An Online Appendix contains the derivations of the prior distribution for autoregressive coefficients consistent with shape beliefs about impulse responses and the computational details mentioned in the paper.

2 The empirical methodology

This section explains the specification of our prior. We then show that our prior can be combined with commonly employed identification strategies to produce structural analyses, and demonstrate that it is straightforward to employ standard posterior algorithms to sample the objects of interest.

2.1 The model

We write the Structural Vector Autoregressive (SVAR) model as

$$\boldsymbol{y}_t = \sum_{l=1}^{p} \Pi_l \boldsymbol{y}_{t-l} + \boldsymbol{c} + B\boldsymbol{\epsilon}_t, \tag{1a}$$

$$\boldsymbol{\epsilon}_t \sim N(\boldsymbol{0}, I_k), \tag{1b}$$

where $\boldsymbol{y}_t$ is a $k \times 1$ vector of observables, Π_l is a $k \times k$ matrix of autoregressive reduced form coefficients at horizon $l = 1, .., p$, $\boldsymbol{c}$ is a $k \times 1$ vector of constants, and B is a $k \times k$ non-singular matrix. The vector $\boldsymbol{\epsilon}_t$ contains the structural shocks, whose covariance matrix is normalized to the identity matrix. The model can also be written in other ways. A useful alternative is given by:

$$\boldsymbol{y}_t = \sum_{l=1}^{p} \Pi_l \boldsymbol{y}_{t-l} + \boldsymbol{c} + \boldsymbol{u}_t, \tag{2a}$$

$$\boldsymbol{u}_t \sim N(\boldsymbol{0}, \Sigma), \tag{2b}$$

$$\boldsymbol{u}_t = B\boldsymbol{\epsilon}_t, \tag{2c}$$

$$\Sigma = BB', \tag{2d}$$

$$B = \chi(\Sigma)Q, \tag{2e}$$

where $\chi(.)$ is a function capturing any square root factorization of Σ, and Q an orthonormal matrix. Equations (1a)-(1b) and (2a)-(2e) define the same SVAR, while equation (2e) highlights the correspondence between reduced form and structural representations.

Let Ψ_h denote the impulse response function (IRF) h periods after the shocks, and let M be the maximum horizon of interest. The mapping between SVAR objects $(B, \Pi_1, .., \Pi_p)$ and IRF objects $(\Psi_0, \Psi_1, .., \Psi_M)$ can be obtained recursively, and for

$M \geq p$ it is given by (Kilian and Lütkepohl, 2017):

$$\Psi_0 = B, \tag{3a}$$

$$\Psi_1 = \Pi_1 \Psi_0, \tag{3b}$$

$$\Psi_2 = \Pi_1 \Psi_1 + \Pi_2 \Psi_0, \tag{3c}$$

$$\Psi_3 = \Pi_1 \Psi_2 + \Pi_2 \Psi_1 + \Pi_3 \Psi_0, \tag{3d}$$

$$\cdots$$

$$\Psi_p = \Pi_1 \Psi_{p-1} + \Pi_2 \Psi_{p-2} + \cdots + \Pi_p \Psi_0, \tag{3e}$$

$$\Psi_{p+1} = \Pi_1 \Psi_p + \Pi_2 \Psi_{p-1} + \cdots + \Pi_p \Psi_1, \tag{3f}$$

$$\cdots$$

$$\Psi_M = \Pi_1 \Psi_{M-1} + \Pi_2 \Psi_{M-2} + \cdots + \Pi_p \Psi_{M-p}. \tag{3g}$$

When B is non-singular, (3) provides a one-to-one mapping between the SVAR parameters $(B, \Pi_1, .., \Pi_p)$ and the IRF parameters $(\Psi_0, \Psi_1, .., \Psi_M)$, for any $M \geq p$. Thus, any prior beliefs on the SVAR parameters imply prior beliefs on the elements of the IRFs via the system of equations (3). Notice that the system is highly nonlinear, because of the cross products of Π_l and Ψ_h, $l = 1, .., p$, $h = 0, .., M$.

For the rest of the paper we make use of the following notation. Let $\Pi = [\Pi_1, .., \Pi_p]$ be a matrix of dimensions $k \times kp$. Vectorizing the matrices we have $\boldsymbol{b} = \text{vec}(B)$, $\boldsymbol{\pi} = \text{vec}(\Pi)$. The parameters of the SVAR are then given by $(\boldsymbol{\pi}, \boldsymbol{c}, B)$. Define $H \geq p$ the horizon until which the researcher is willing to introduce prior beliefs on the impulse response where, generally, $H \leq M$, since a researcher need not have prior beliefs stretching as far as the horizon of interest. Furthermore, let $\Psi_F = [\Psi_1, .., \Psi_H]$, $\Psi = [\Psi_0, \Psi_F]$, $\boldsymbol{\psi} = \text{vec}(\Psi)$, $\boldsymbol{\psi}_F = \text{vec}(\Psi_F)$ and $\boldsymbol{\psi}_h = \text{vec}(\Psi_h)$, for $h = 0, 1, .., H$. Vectorizing the first $H + 1$ equations of the system (3) one obtains:

$$\boldsymbol{\psi}_0 = \boldsymbol{b}, \tag{4}$$

$$\boldsymbol{\psi}_F = R\boldsymbol{\pi}, \tag{5}$$

where

$$R = R_H \otimes I_k \equiv \underbrace{\begin{bmatrix} \Psi_0' & 0 & 0 & \dots & 0 \\ \Psi_1' & \Psi_0' & 0 & .. & 0 \\ \Psi_2' & \Psi_1' & \Psi_0' & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Psi_{p-1}' & \Psi_{p-2}' & \Psi_{p-3}' & \dots & \Psi_0' \\ \Psi_p' & \Psi_{p-1}' & \Psi_{p-2}' & \dots & \Psi_1' \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Psi_{H-1}' & \Psi_{H-2}' & \Psi_{H-3}' & \dots & \Psi_{H-p}' \end{bmatrix}}_{R_H} \otimes I_k. \tag{6}$$

The matrix R_H is a function of Ψ, it is of dimension $Hk \times pk$, and is of full column rank as long as B is non-singular.

2.2 Our proposed approach

Bayesian inference in SVAR models typically involves formulating prior beliefs on $p(\boldsymbol{\pi}, \boldsymbol{c}, B) = p(B) \cdot p(\boldsymbol{\pi}|B) \cdot p(\boldsymbol{c}|B, \boldsymbol{\pi})$. Because researchers frequently entertain prior beliefs on Ψ, it is necessary to ensure that the selected $p(\boldsymbol{\pi}, \boldsymbol{c}, B)$ imply a $p(\Psi)$ in line with such beliefs. This is not currently possible under the standard approach that uses a Normal distribution for $\boldsymbol{\pi} \sim N(\boldsymbol{\mu}, V)$, where $(\boldsymbol{\mu}, V)$ are selected to produce a flat prior, a random walk or a white noise prior process. For example, it is standard in the literature to set $(\boldsymbol{\mu}, V)$ as follows:

$$E((\Pi_h)_{ij}) = \begin{cases} \delta_i, & j = i, h = 1 \\ 0 & \text{otherwise} \end{cases}, \qquad V((\Pi_h)_{ij}) = \begin{cases} \frac{\lambda^2}{h^2}, & j = i \\ \eta \frac{\lambda^2}{h^2} \frac{\sigma_i^2}{\sigma_j^2} & \text{otherwise} \end{cases}. \tag{7}$$

A flat prior is obtained choosing λ to be large, while the random walk and the white noise priors can obtained choosing δ_i to be 1 or 0, respectively (see Canova, 2007, Bańbura et al., 2010 and Koop and Korobilis, 2010 for popular selections of the remaining hyperparameters). The random walk and white noise priors are frequently referred to as the Minnesota-like prior for $\boldsymbol{\pi}$, following the contribution of Doan et al. (1984) and Litterman (1986).[1] While the forecasting properties of these specifications

[1]The original Minnesota prior postulated a known and diagonal matrix Σ. Later extensions combine the Normal prior on $\boldsymbol{\pi}$ specified as in equation (7) with an inverse-Wishart distribution for Σ, possibly adding a Kronecker structure. We will refer to the Minnesota or Minnesota-like prior only with reference to the prior for $\boldsymbol{\pi}$, while remaining intentionally silent about the priors on B or

have been well-documented, neither the flat nor the Minnesota-like prior for $\boldsymbol{\pi}$ allow for flexibility in the specification of $p(\Psi)$. The induced prior on Ψ can be derived analytically, but its shape is constrained by the combination of the Normal prior, the shrinkage assumptions (7), and the prior for B. We refer the reader to Section 3 and Section 4 for an illustration of the prior on impulse responses induced by the Minnesota specification.

Our approach focuses attention on $p(\boldsymbol{\pi}|B)$ while remaining agnostic about $p(B)$ and $p(\boldsymbol{c}|B,\boldsymbol{\pi})$. Let $\bar{\Psi}_h$, for $h = 0, 1, .., H$, be the prior mean value a researcher would like to select for the impulse responses up to horizon H. As we will see, the mapping provided by (3) allows for any $\bar{\Psi}_h$, $h = 0, 1, .., H$, to be transformed into a prior for the autoregressive parameters of the structural VAR, given any $p(B)$ and $p(\boldsymbol{c}|B,\boldsymbol{\pi})$. Let $\bar{\boldsymbol{\psi}}_h = \text{vec}(\bar{\Psi}_h)$, $\bar{\boldsymbol{\psi}} = (\bar{\boldsymbol{\psi}}_0', \bar{\boldsymbol{\psi}}_F')' = \text{vec}(\bar{\Psi})$, $\bar{\Psi}_F = [\bar{\Psi}_1, .., \bar{\Psi}_H]$, $\bar{\Psi} = [\bar{\Psi}_0, \bar{\Psi}_F]$. Assume that the prior $p(B)$ implies $E(B) = \bar{\Psi}_0$ (more on this below). Define the artificial random variables W_h, $h = 1, .., H$ as

$$W_1 = \Pi_1\bar{\Psi}_0 - \bar{\Psi}_1, \tag{8a}$$
$$W_2 = \Pi_1\bar{\Psi}_1 + \Pi_2\bar{\Psi}_0 - \bar{\Psi}_2, \tag{8b}$$
$$W_3 = \Pi_1\bar{\Psi}_2 + \Pi_2\bar{\Psi}_1 + \Pi_3\bar{\Psi}_0 - \bar{\Psi}_3, \tag{8c}$$
$$\vdots$$
$$W_p = \Pi_1\bar{\Psi}_{p-1} + \Pi_2\bar{\Psi}_{p-2} + \cdots + \Pi_p\bar{\Psi}_0 - \bar{\Psi}_p, \tag{8d}$$
$$W_{p+1} = \Pi_1\bar{\Psi}_p \ + \ \Pi_2\bar{\Psi}_{p-1} + \cdots + \Pi_p\bar{\Psi}_1 - \bar{\Psi}_{p+1}, \tag{8e}$$
$$\vdots$$
$$W_H = \Pi_1\bar{\Psi}_{H-1} \ + \ \Pi_2\bar{\Psi}_{H-2} + \cdots + \Pi_p\bar{\Psi}_{H-p} - \bar{\Psi}_H. \tag{8f}$$

The system of equations (8) is closely related to the system (3) except that it drops the equation at horizon 0, replaces the matrices Ψ_h with the matrices of hyperparameters $\bar{\Psi}_h$, $h = 1, .., H$, and drops the last $M - H$ equations. The system (8) can be vectorized as

$$\boldsymbol{w} = \bar{R}\boldsymbol{\pi} - \bar{\boldsymbol{\psi}}_F, \tag{9}$$

where $\boldsymbol{w} = \text{vec}([W_1, .., W_H])$ and $\bar{R}$ is defined in equation (6) after replacing Ψ with $\bar{\Psi}$. If $\bar{\Psi}_0$ is non-singular (which we assume), the matrix $\bar{R}$ has full column rank.

Σ.

Premultiplying both sides of (9) by $\bar{R}'$ and rearranging the terms gives:

$$\boldsymbol{\pi} = \left(\bar{R}'\bar{R}\right)^{-1}\bar{R}'\boldsymbol{w} + \left(\bar{R}'\bar{R}\right)^{-1}\bar{R}'\bar{\boldsymbol{\psi}}_F. \tag{10}$$

Since $\boldsymbol{\pi}$ is of lower dimension than $\boldsymbol{w}$ for $H > p$, the system of equations (9) is not necessarily consistent. Put it differently, given $(\bar{R}, \bar{\boldsymbol{\psi}}_F)$, whether a solution for $\boldsymbol{\pi}$ exists depends on the value of $\boldsymbol{w}$.

Our prior on $\boldsymbol{\pi}$ is a Normal and it is specified as follows

$$\boldsymbol{\pi}|B \sim N\big(\boldsymbol{\mu}, V(B)\big), \tag{11}$$

$$\boldsymbol{\mu} = \left(\bar{R}'\bar{R}\right)^{-1}\bar{R}'\bar{\boldsymbol{\psi}}_F. \tag{12}$$

The rationale behind (12) comes from the system of equations (9). When $H = p$ (which we refer to in the text and the Online Appendix as *Case a*), then a VAR with p lags can replicate exactly the pattern in $\bar{\Psi}$, making the system (9) consistent for $\boldsymbol{w} = \mathbf{0}$. This means that (12) is the unique solution, and $E(\boldsymbol{\pi})$ simplifies to $\boldsymbol{\mu} = \bar{R}^{-1}\bar{\boldsymbol{\psi}}_F$, since $\bar{R}$ is squared and invertible. When $H > p$ but $\bar{\Psi}$ is selected such that it can be replicated by a VAR with p lags (*Case b*), then the system (9) is still consistent for $\boldsymbol{w} = \mathbf{0}$ and (12) is still its unique solution. Lastly, when $H > p$ but no parametrization of a VAR with p lags exists that replicates $\bar{\Psi}$ (*Case c*), then $\boldsymbol{w} = \mathbf{0}$ makes the system (9) inconsistent. Yet, $E(\boldsymbol{\pi})$ from (12) is still the unique solution if $\boldsymbol{w} = [\bar{R}\left(\bar{R}'\bar{R}\right)^{-1}\bar{R}' - I]\bar{\boldsymbol{\psi}}_F$.

The covariance matrix of our prior is, for the moment, left unspecified. We remain flexible about $V(B)$ to allow a researcher to freely select it in a way that potentially depends on B. For example, for computational purposes, one can impose a Kronecker structure and set

$$V(B) = V_H \otimes BB'. \tag{13}$$

where V_H mimics the structure of a Minnesota prior covariance matrix.

It is immediate to note that our prior selection replaces the arbitrary Minnesota choice of $\boldsymbol{\mu}$ with an endogenous selection that depends on $\bar{\Psi}$. It is harder to see, but not difficult to prove, that our prior choice implies convenient properties on the prior distribution of Ψ. Assume, without loss of generality, that $p(B)$ implies $E(B) = \bar{\Psi}_0$. Section 1 in the Online Appendix proves that under *Case a* and *Case b* the following results hold:

$$E(\Psi_h) = \bar{\Psi}_h, \qquad h = 0, 1, \tag{14a}$$

$$\lim_{V(B)\to 0} E(\Psi_h) = \bar{\Psi}_h, \qquad h = 2, .., H. \tag{14b}$$

In words this means that it is possible to specify a Normal prior distribution on Π that implies a prior on the IRFs approximately centered around $\bar{\Psi}$. An appropriately chosen $V(B)$ gives researchers control on the width of the credible intervals on Ψ, while guaranteeing that $E(\Psi_h) \approx \bar{\Psi}_h$, for $h = 2, .., H$. Recall that *Case a* requires $H = p$, a condition that can be easily satisfied given that p is a free parameter in our setting. We stress also that under *Case c* the approximation error $E(\Psi_h) - \bar{\Psi}_h$ can still be small, and its size can be checked numerically depending on the application and on the values of $\bar{\Psi}$ used. While equation (14b) is based on the limiting behavior of our prior, all of our simulations and applications suggest that the approximation is very satisfactory, making the methodology suitable for applied work.

2.3 Discussion

A key advantage of our approach is that it does not require deriving and integrating the joint prior distribution of Ψ implied by our prior, a procedure that entails complex computation techniques. Instead, our analysis only requires working with the expectation operator of a multivariate system of equations, see Section 1 of the Online Appendix. This is a considerable advantage relative to the existing contributions that work with the transformation between the SVAR and the impulse response parametrization (Kociecki, 2010, Arias et al., 2018).

Because our approach leaves $p(B)$ unrestricted, it can be combined with a number of identification strategies. One can specify the model as in equations (2a)-(2e) and work with an inverse-Wishart-Uniform prior distribution on (Σ, Q), possibly restricted to imply a combination of zero and sign restrictions on B or on $A = B^{-1}$ (Rubio-Ramirez et al. (2010) , Binning, 2013, Arias et al., 2018). One can also form a prior directly on A (as advocated by Baumeister and Hamilton, 2015, Baumeister and Hamilton, forthcoming), on B (as suggested by Bruns and Piffer, 2023) or on a combination of A and B (see Baumeister and Hamilton, 2018). In all cases, setting $\bar{\Psi}_0 = E(B)$ ensures that $\bar{\Psi}$ is compatible with $p(B)$. Identification via external instruments can also be used by adding the instrument as the first variable in a recursive

SVAR as in Plagborg-Møller and Wolf (2021) and the specification can be extended to the VARX setting used by Paul (2020). Longer horizon identification restrictions can be introduced using an accept-reject algorithm, as in Uhlig (2005). Section 4 shows that doing so with our approach is computationally less demanding than with a flat or Minnesota-like priors, because our prior distribution already encourages the selected signs over multiple horizons. This significantly reduces the number of draws that are rejected by the algorithm, facilitating posterior sampling and computations in general.

It is common to estimate SVAR models and study only the dynamics induced by a subset of structural shocks. Our approach is compatible with this strategy. One way to achieve this is to work with a prior distribution $p(B)$ that features a wide variance for the shocks that are not identified and a tighter variance on the shocks that are identified. In this case, equations (3) imply very wide prior beliefs on the impulse responses associated with the non-identified shocks, making the selection of $\bar{\Psi}$ for such shocks irrelevant. The prior distribution proposed by Bruns and Piffer (2023) is suitable for this purpose, as it treats each entry of B as independent Normal distributions, potentially truncated to be positive or negative.

Our prior on $\boldsymbol{\pi}$ is Normal. For this reason, posterior sampling is not in any way more challenging than with existing methodologies, as the conditional posterior $p(\boldsymbol{\pi}|Y, B)$ remains Normal. An inverse-Wishart-Uniform prior on (Σ, Q) leads to a posterior that can be explored using standard methods: a Gibbs sampler or direct sampling, depending on whether $V(B)$ is restricted to have a Kronecker structure as in equation (13) or not. If a more general prior $p(B)$ is used, posterior sampling requires a more involving method for computing the marginal posterior $p(B|Y)$, for example a Metropolis-Hastings algorithm as in Baumeister and Hamilton (2015), the two-step algorithm as in Bruns and Piffer (2023), or the methods in Canova and Pérez Forero (2015) and Waggoner et al. (2016). All in all, our setup provides flexibility on the specification of the prior distribution on the impulse responses at no additional computational cost.

It is common in the literature to assess the "goodness" of a prior specification using the forecasting performance of the posterior. We warn against using such an exercise when the scope is structural analysis for at least two important reasons. First, forecasting and structural performance are not two sides of the same coin. For example, suppose there are two shocks in the data, one that explains 90 percent and one that explains 10 percent of the variance of inflation. Suppose that one is interested in the dynamics induced by the latter shock. Good forecasting performance for inflation requires proper identification the former shock and capturing well the dynamics it induces. But a prior that is tailored to that purpose will not be able to tell us much about the dynamics in response to the second shock. To put this result differently, a good forecasting performance is neither a necessary nor a sufficient condition for good structural inference. By the same token, a prior that flexibly accommodates prior beliefs on certain impulse responses need not have good forecasting performance, but this should not be considered a defect.

Second, SVARs often suffer from deformation problems, see Canova and Ferroni (2022). Because systems tend to be small, structural shocks may be confounded, making structural analysis typically biased. Still, deformed systems may have good forecasting performance as long as enough lags are used. Thus, one may be able to produce decent forecasts even when the structural model is misspecified and the dynamics in response to the shocks distorted. For these reasons, we find it inappropriate to judge a prior specification, which is specifically designed for structural objects, using the forecasting performance of the implied posterior model. If anything, introducing prior beliefs consistent with the true response of a structural shocks in a deformed system increases the ability of the posterior distribution to reflect some true features of the shock, despite the misspecification present in the model.

It is also useful to draw a short comparison with the approaches of Villani (2009) and Andrle and Benes (2013). Villani (2009) writes the VAR system in deviation from the steady states and designs priors for the steady states, which are endogenous functions of the VAR coefficients. These priors imply, in turn, a prior specification for the VAR coefficients. Our approach works the other way around: we impose prior restrictions on the VAR coefficients that represents certain prior beliefs on the IRFs. Andrle and Benes (2013) provide priors for endogenous objects of a structural (DSGE) model, such as the sacrifice ratio. These priors imply, in turn, priors on the structural parameters of the model that enter the functions of interest. The main difference here is that a DSGE model rather than a VAR model is used in the exercise.

Our prior is also related to Jarociński and Marcet (2019) who propose to formulate a prior directly on observable variables instead of the parameters of the VAR. They rightly point that it is usually hard to come up with genuine prior for VAR parameters. In addition, priors for the VAR parameters may imply priors for observables that are hard to defend, for example, huge future yearly output growth. To address this problem, they develop a framework for translating the prior for observables into a prior for VAR parameters which is in line with truly subjective prior beliefs. Similarly, our approach starts from the premise that researchers are more comfortable specifying their prior beliefs on impulse response functions rather than VAR parameters. Hence, our method can be seen as complementary to the one by Jarociński and Marcet (2019) and useful in different contexts.

2.4 The specification of $(\bar{\Psi}, V(B))$

To make our approach operational we need to select $(\bar{\Psi}, V(B))$. We found it convenient to set the entries of $\bar{\Psi}$ using the Gaussian basis functions

$$\bar{\psi}_{ij,h} = a_{ij} \cdot e^{-\left(\frac{(h-b_{ij})^2}{c_{ij}}\right)+\frac{b_{ij}^2}{c_{ij}^2}}, \tag{15}$$

where a_{ij} captures the impact effect of shock j on variable i, and is set equal to the i, j entry of $E(B)$. b_{ij} is an integer scalar, pinning down the horizon at which the peak effect is reached, and equals 0 when no hump-shaped response is desired. When $b_{ij} > 0$, c_{ij} can be freely set to control for how much higher the peak effect is (in absolute value) relative to the impact effect, while for $b_{ij} = 0$, c_{ij} controls the horizon at which the impulse response reaches the half-life relative to the impact effect. Equation (15) re-parametrizes the function used in Barnichon and Matthes (2018) to ensure that the impact effect of the impulse response is a free parameter, as this is needed to make $\bar{\Psi}_0$ coincide with $E(B)$.

$V(B)$ can be set freely to indirectly control for the uncertainty in the prior beliefs on the impulse responses. We found it convenient to select $V(B)$ as with the Minnesota-like prior, hence achieving the same dimensionality reduction to a handful of hyperparameters, see equation (7). We propose to select λ adaptively, as discussed in detail in Section 3. Alternatively, with equation (13) in hand, one can also use a hierarchical or an Empirical Bayes approach for setting the elements of $V(B)$, as in Giannone et al. (2015). We do not follow this approach, because the update in the hierarchical method by Giannone et al. (2015) is driven by the forecasting properties of the model, which we view as conceptually separate from structural analysis. We stress that our methodology is compatible with alternative selections of $(\bar{\Psi}, V(B))$.

Our setup nests a Minnesota-like prior selection of $(\boldsymbol{\mu}, V)$ as a special case. As mentioned $V(B)$ can be set as with the Minnesota-like prior. As for $\boldsymbol{\mu}$, the random walk prior, corresponding to $\delta_i = 1$ in equation (7), can be obtained by setting $\bar{\Psi}_h = \bar{\Psi}_0$, $h = 1, .., H$, $H = p$ and the white noise prior, i.e. $\delta_i = 0$, can be obtained by setting

$\bar{\Psi}_h = 0$, $h = 1, .., H$, $H = p$. Thus, when λ is sufficiently small, the Minnesota-like selection given in (7) is consistent with the belief that the responses to the structural shocks are either very persistent (the random walk prior) or not persistent at all (the white noise prior). In contrast, our prior offers more flexibility as intermediate persistence cases are possible and shape beliefs of any form can be introduced non-dogmatically.

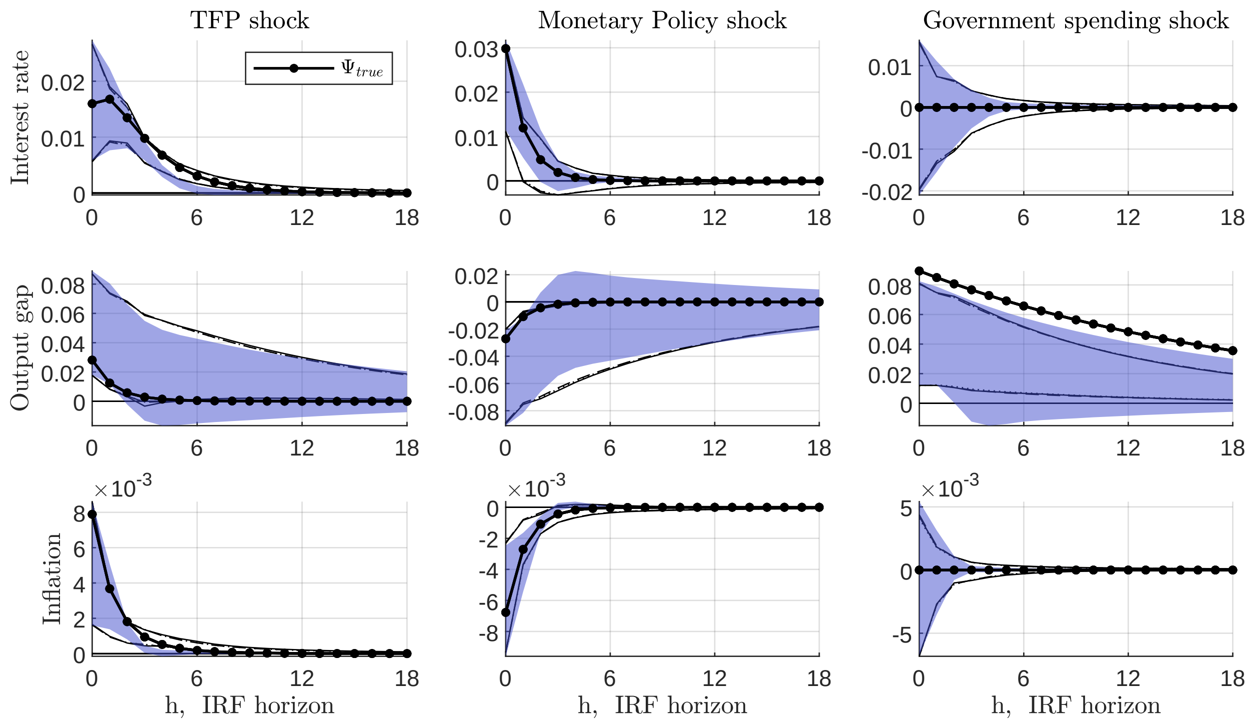
3 The features of our prior with simulated data

We illustrate the properties of our prior specification using data simulated from a stylized small-scale DSGE model. We build on the three-variable New Keynesian model by An and Schorfheide (2007). The model has three variables: the output gap, inflation, and the interest rate, and their dynamics are driven by three structural shocks: a TFP shock, a monetary policy shock and a government spending shock. We calibrate the parameters of the DSGE model as in An and Schorfheide (2007). As the model is stationary, simulations are started from the unconditional mean, and we discard the first 100 observations. We then generate two datasets with 100 and 1000 observations, plus 50 observations used as training sample.

We estimate four SVAR models, which are identical except for the specification for $\boldsymbol{\pi}$. All models include 4 lags and a constant, on which we use a flat prior centered at zero. While adding $p > 1$ lags introduces a misspecification to the model relative to the data generating process, it helps the visual illustration of our prior, without affecting the results we present (see discussion below). For simplicity, we use an inverse-Wishart-uniform prior for (Σ, Q) and identify the three shocks by sign restrictions on impact. The restrictions are consistent with the true impact responses embedded in the model; thus no misspecification in the identification of shocks is present. The discussion below focuses on the responses to a government spending shock. In the model, this disturbance induces no response in the interest rate and in inflation. Hence, we introduce no restrictions for the corresponding entries of B.

We use prior independence between $\boldsymbol{\pi}$ and Σ, hence no Kronecker structure on $V(B)$ is introduced. The first model uses a flat prior on $\boldsymbol{\pi}$. The second and third model use a Minnesota prior on $\boldsymbol{\pi}$, setting the prior mean to reflect either a white noise or a random walk. We set the covariance matrix of the Minnesota prior as discussed in Canova (2007), which implies a relatively uninformative prior choice. For our prior we select $\bar{\Psi}$ so as to reproduce as close as possible the shape of the true

Figure 1: Posterior distribution, T=1000



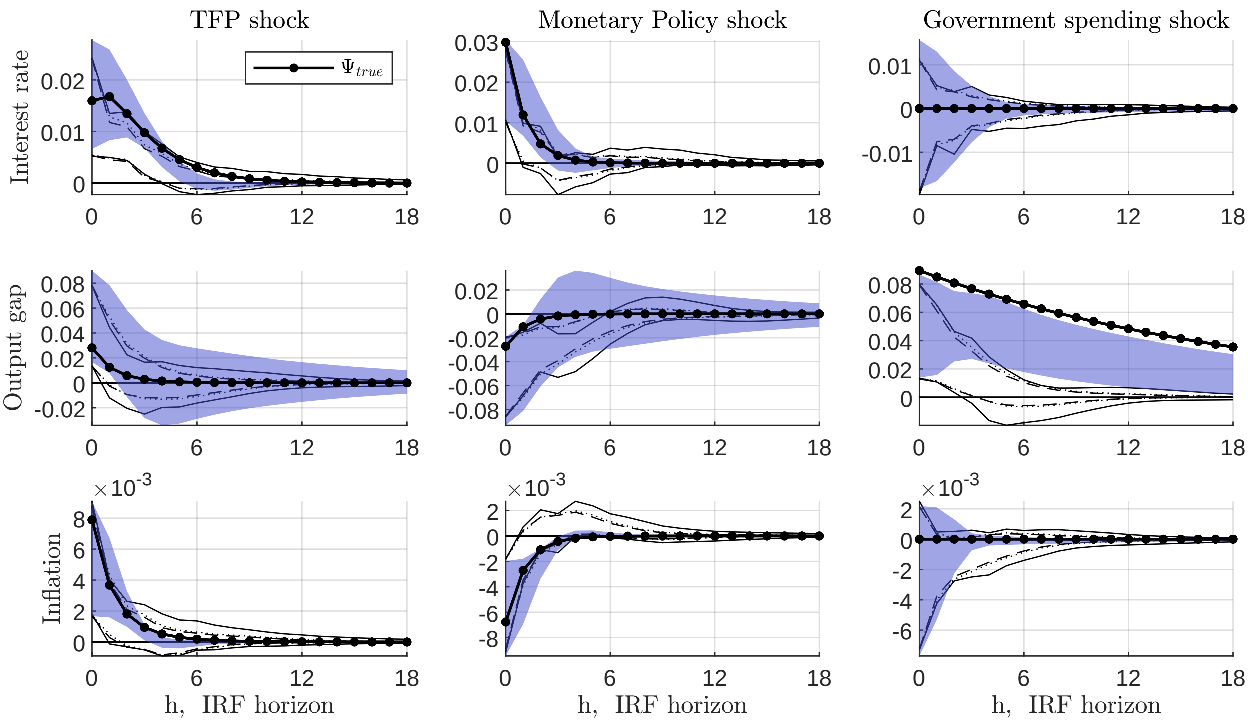
Note: Pointwise 68% credible sets corresponding to the flat prior (solid lines), to the white noise Minnesota prior (dashed lines), to the random walk Minnesota prior (dotted lines), and to our prior (blue shaded area). The dataset includes $T = 1,000$ observations. The responses correspond to a one standard deviation shock.

impulse responses. We study impulse responses up to 18 horizons, and specify the prior mean on the impulse responses up to $H = 4$ horizons. Because $H = p$, the setup coincides with *Case a* discussed in Section 2.2. We specify $\bar{\Psi}$ as in equation (15). We set a_{ij} equal to the ij entry of $E(B)$, as discussed in Section 2.3. We set $b_{ij} = 1$ for $i = 1,\ j = 1$ and zero otherwise, to capture the true timing of the peak effects of the responses. Last, we set c_{ij} to capture the true half-life of the responses. As for the variance, we set $V(B)$ as under the Minnesota prior.

For our prior, λ is chosen adaptively. As discussed in Bruns and Piffer (2023), equation (2d) implies the constraint $-\sqrt{\Sigma_{ii}} < B_{ij} < \sqrt{\Sigma_{ii}}$. We estimate $\sqrt{\hat{\Sigma}_{ii}}$ using, for each dataset, the training sample with the first 50 observations. We take $\gamma_i = [-2\cdot\sqrt{\hat{\Sigma}_{ii}},\ +2\cdot\sqrt{\hat{\Sigma}_{ii}}]$ to be a reasonable bound for the dynamics of the vector of impulse responses over the first 18 horizons, which are the horizons of interest for the impulse response. We start from $\lambda = 0.01$ and progressively increase it until between 5% and

10% of the joint prior probability lies outside of the chosen interval, irrespectively of the variable, the shock, and the horizon. This ensures an appropriately informative yet non-dogmatic prior. The search algorithm delivers a baseline $\lambda = 0.04491$, which is in line with the selection by Bańbura et al. (2010) for a medium/large VAR model. Section 2 in the Online Appendix provides the details on the iterative procedure used for selecting λ.

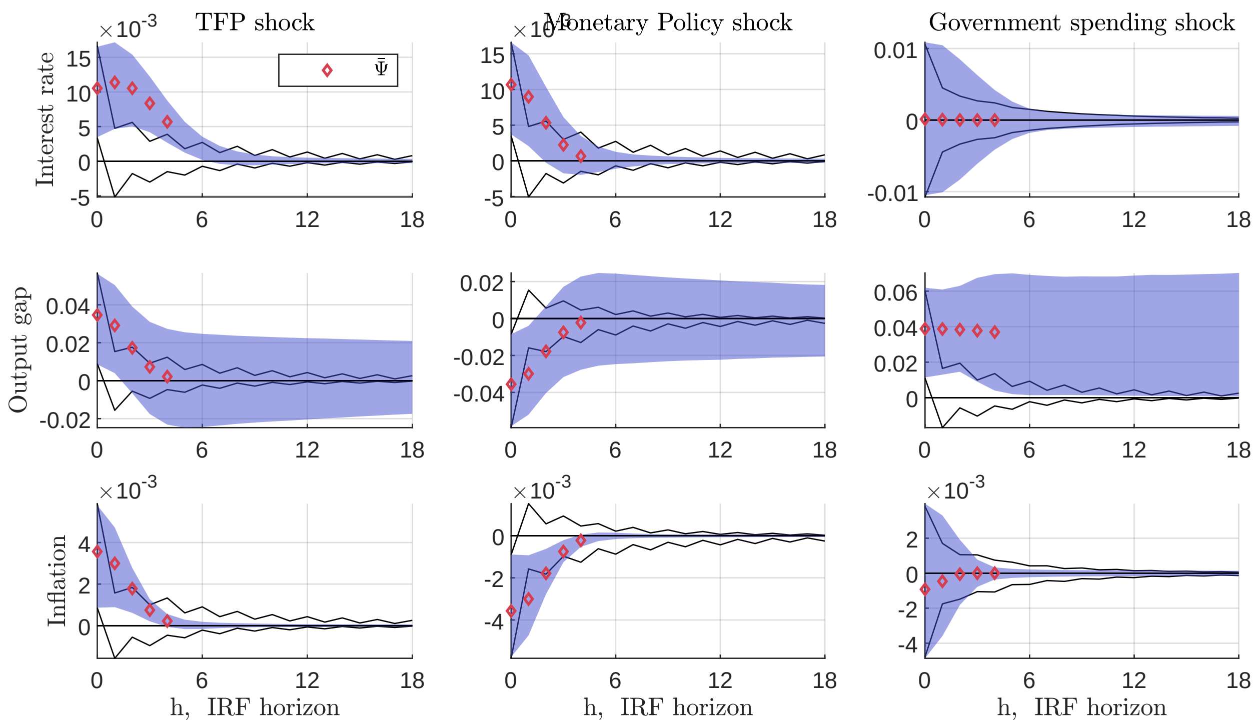
Figure 2: Posterior distribution, T=100



Note: Pointwise 68% credible sets corresponding to the flat prior (solid lines), to the white noise Minnesota prior (dashed lines), to the random walk Minnesota prior (dotted lines), and to our prior (blue shaded area). The dataset includes $T = 100$ observations. The responses correspond to a one standard deviation shock.

Figure 1 reports the pointwise 68% credible sets of the posterior IRFs associated with the four priors when T=1000. The dotted line displays the true impulse responses. The posterior distribution $p(B|Y)$ is similar across prior specifications, and the true value of the instantaneous response of output gap to the government spending shock lies in the right tail. The figure shows that all four prior selections underestimate the impact effect of the government spending shock to the output gap, although we stress that the reported credible sets are 68% bands. Our analysis centers on the ability to correctly infer the true persistence of the responses, rather than the correct impact effect. At longer horizons, no material difference emerges, as all four specifications correctly capture the strongly persistent nature of the true response. This result is driven by the fact that, in large samples, Π is identified from the data. Hence, differences in prior beliefs asymptotically vanish.

Figure 3: Prior beliefs



Note: Pointwise 68% credible sets associated with the white noise Minnesota prior (continuous lines) and our prior (shaded area). The red diamonds show $\bar{\Psi}$ used as our prior, which was specified setting $H = p = 4$.

The results differ in a small sample. Figure 2 shows that when $T = 100$, an important difference emerges in the posterior distribution of the IRFs associated with either the flat and the random walk and white noise Minnesota priors on the one hand, and our prior on the other. The first three specifications strongly underestimate the persistence of the output gap responses to the government spending shock. They find that the half-life of the response is about 2-3 quarters while the true half life is 18 quarters. Our prior, instead, leads to posterior IRFs that approach the true IRFs. This holds not only for the response of output gap to the government spending shock, but also, for example, for the response of inflation to the TFP shock and to the monetary policy shock. Thus, as mentioned earlier, shape restrictions on one impulse response may be beneficial for other impulse responses.

Figure 3 interprets the results by plotting the prior distributions underlying Figure 2. The dotted line shows the values of $\bar{\Psi}$, which was specified up to horizon $H = p = 4$. The corresponding expected value of Π computed using $\boldsymbol{\mu} = \bar{R}^{-1}\bar{\boldsymbol{\psi}}_F$ is

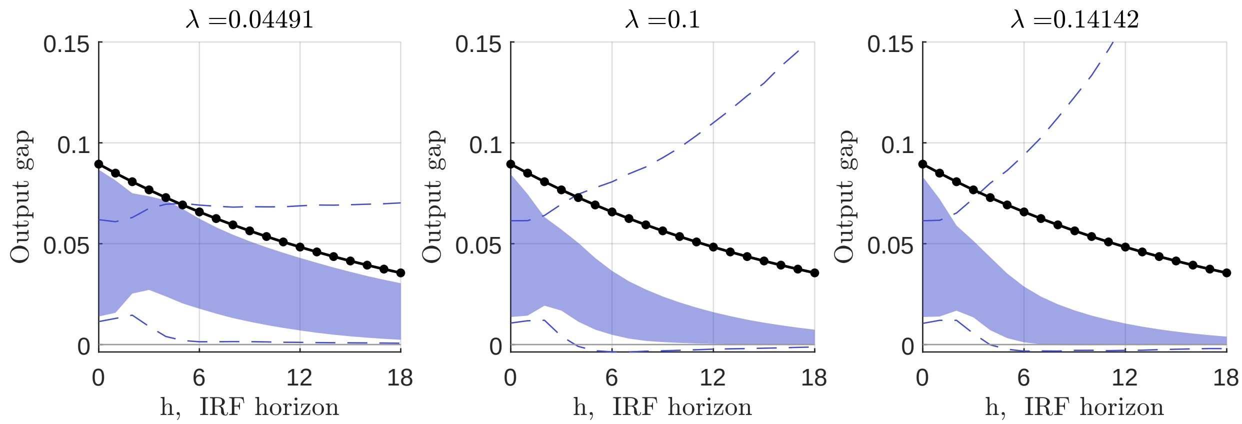
$$E(\Pi_1) = \begin{pmatrix} 0.960 & 0.006 & 0.299 \\ 0.005 & 0.967 & -1.244 \\ 0.000 & 0.007 & 0.776 \end{pmatrix}, \qquad E(\Pi_2) = \begin{pmatrix} -0.173 & 0.000 & 0.106 \\ 0.007 & -0.032 & -1.721 \\ 0.000 & -0.003 & -0.177 \end{pmatrix},$$
$$E(\Pi_3) = \begin{pmatrix} -0.053 & -0.004 & -0.013 \\ 0.001 & 0.015 & -0.508 \\ 0.000 & -0.004 & 0.006 \end{pmatrix}, \qquad E(\Pi_4) = \begin{pmatrix} -0.001 & -0.002 & -0.038 \\ -0.002 & 0.019 & -0.002 \\ 0.000 & 0.001 & 0.010 \end{pmatrix}.$$

The credible intervals shown in the shaded area in Figure 3 are obtained by drawing from the joint prior distribution $p(\boldsymbol{\pi}, B)$ and computing the associated impulse responses up to horizon 18.

The figure confirms that it is possible to work with a Normal prior for $\boldsymbol{\pi}$ and calibrate its moments to imply a prior on the vectorized impulse responses centered around the desired trajectory. Our prior introduces beliefs that are, by construction, in line with the persistence of the response. The results hold also when $H = 1$. However, such a choice implies that $\bar{\Psi}$ in Figure 3 is very short, making the illustration less informative. The grey dashed lines in Figure 3 show that the white noise Minnesota prior indirectly encourages no persistence in the impulse responses. Assessing what the flat prior introduces on impulse responses is not possible, while results available on request indicate that the random walk Minnesota prior introduces very unstable and large credible bands. The similarity documented by Figure 2 in the posterior distributions associated with the flat prior and both specifications of the Minnesota prior indicate that these priors are relatively uninformative, leading to the posterior distribution being strongly affected by the sample estimate. Since the latter underestimates the true persistence of the output gap to the government spending shock in the sample by roughly 30%, one fails to measure the true persistence of the response.

It is typically suggested that a relatively wide prior variance on the object of interest should be used “to let the data speak”. We find that this is not necessarily the case when performing structural analyses. Increasing λ makes the prior less informative. However, as the prior loses its informational shape content, it becomes less useful

Figure 4: Increase in prior variance



Note: The figure reports the effect of a government spending shock on the output gap, together with pointwise 68% credible sets associated with our prior as λ increases. The dotted line reports the true impulse response. The prior is represented by the dashed line, the posterior by the shaded area.

when one wants to answer questions such as what is the persistence of a response. As shown in Figure 4, when λ increases the posterior IRFs obtained with our prior approaches the posterior IRFs associated with the Minnesota and the flat priors. As a consequence, the posterior now fails to capture the true persistence effect in a small sample.

4 The output effects of monetary policy shocks

There is a large literature quantifying the effects of monetary policy shocks on the real economy (see Christiano et al., 1999 and Miranda-Agrippino and Ricco, 2021 among many others). One key question often discussed in the literature is whether the monetary policy intervention generates its strongest effects on impact, or whether long and variable lags imply that the largest response occurs with a delay (Buda et al., 2023). This question is relevant in the policy debate, given that it directly informs central bankers about their ability to quickly stimulate the real economy when needed.

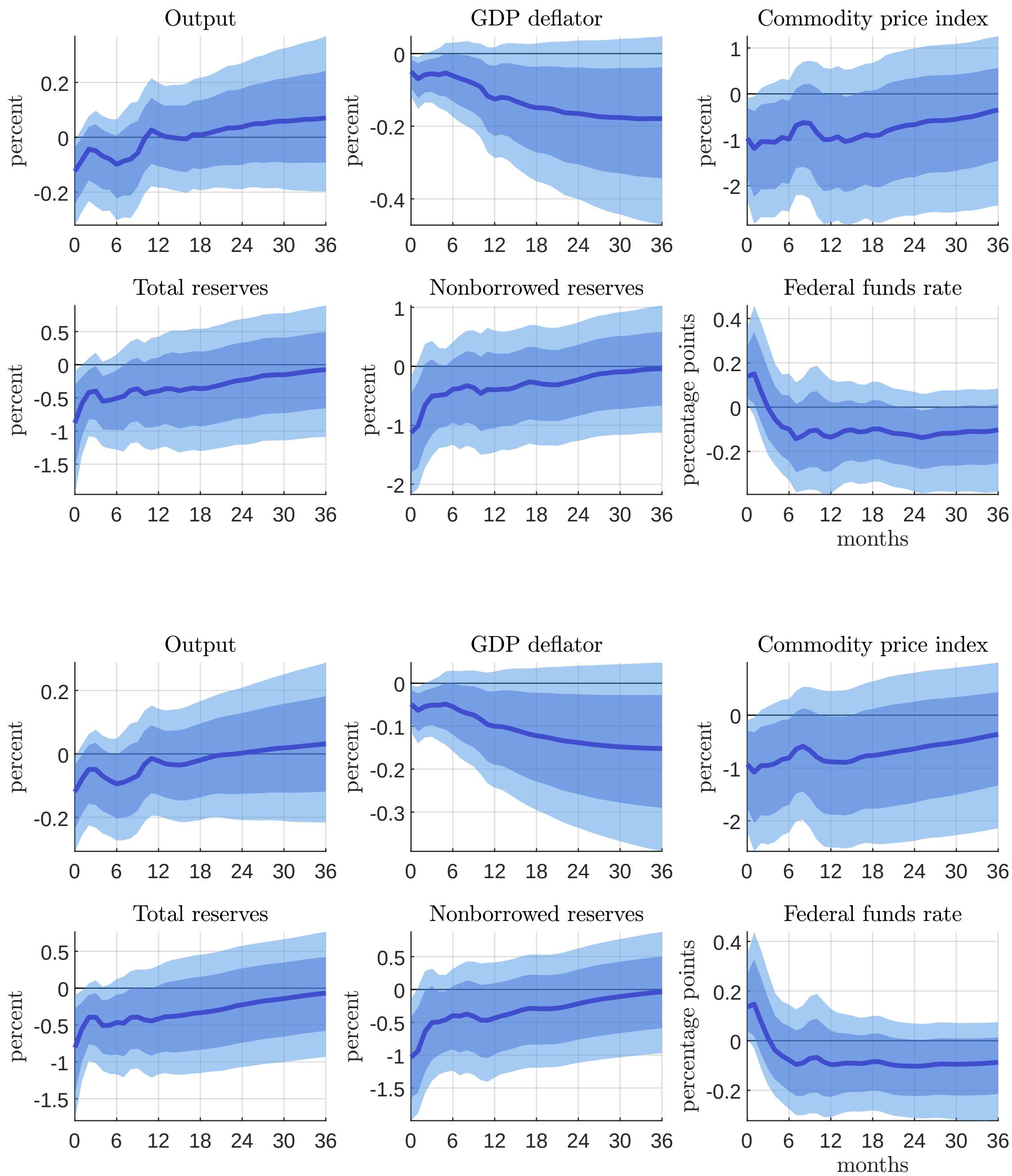
We use our prior methodology to study how long it takes for a US surprise monetary intervention to affect US real economic activity. We use a SVAR model with six variables: real GDP, the GDP deflator, the commodity price index, total reserves, nonborrowed reserves, and the federal funds rate. All variables enter in log except for the federal funds rate. The data is monthly, and real GDP and the GDP deflator are interpolated using either industrial production data or the consumer price and the producer price indexes. The set of variables used is common in the literature and is consistent with the work of, e.g., Bernanke and Mihov (1998), Uhlig (2005) and Arias et al. (2019). Following Arias et al. (2019), we use the data in the period 1965M1 through 2007M6, in a VAR with 12 lags and a constant. As in Arias et al. (2019) and others in the literature, our sample stops before the Great Financial Crisis.

We identify the monetary policy shock via sign restrictions: we assume that an increase in the policy rate decreases on impact all variables. Thus by construction, real GDP decreases on impact in response to a contractionary monetary policy surprise. The key question is whether the effect of the monetary policy disturbance is largest on impact or whether real GDP further decreases, displaying a hump. For the remaining shocks no restrictions are imposed, except the normalization that the diagonal entries of B are positive. We use an inverse-Wishart-uniform distribution on (Σ, Q), estimating in the training sample (the first 20% of the observations) the hyperparameters of the inverse Wishart, as suggested by Kadiyala and Karlsson (1997). While we agree with Baumeister and Hamilton (2015, 2018) that explicit prior distributions can be expressed directly on structural parameters of interest, we employ the indirect inverse-Wishart-Uniform prior for computational convenience. We stress that the contribution of our paper concerns $\boldsymbol{\pi}$ rather than (Σ, Q) or B, making the actual prior on these parameters less crucial, see also Inoue and Kilian (2020) for a discussion on the importance of prior distributions in structural analysis.

Figure 5 shows the posterior impulse responses to a one standard deviation shock for two alternative prior specifications of $\boldsymbol{\pi}$: the flat and the Minnesota-like priors. We use a conjugate specification for the latter, setting the mean of $\boldsymbol{\pi}$ to reflect a random walk for all variables and setting the variance of $\boldsymbol{\pi}$ as in Kadiyala and Karlsson (1997). We report the pointwise median response together with the 68% and the 90% credible sets. The figure shows that when prior beliefs are represented by these two priors, the largest (in absolute value) posterior real GDP response is on impact. The IRFs revert back to zero within less than a year and display no hump.

Would the conclusions change if one introduces prior beliefs on the shape of the impulse responses? We do not assume humped-shaped beliefs on response of any of the variables to avoid leading the results in this direction, even indirectly. Instead, we assume that monetary policy shocks generate persistent but temporary effects. This view is consistent with a large amount of evidence. We stress that this belief is not

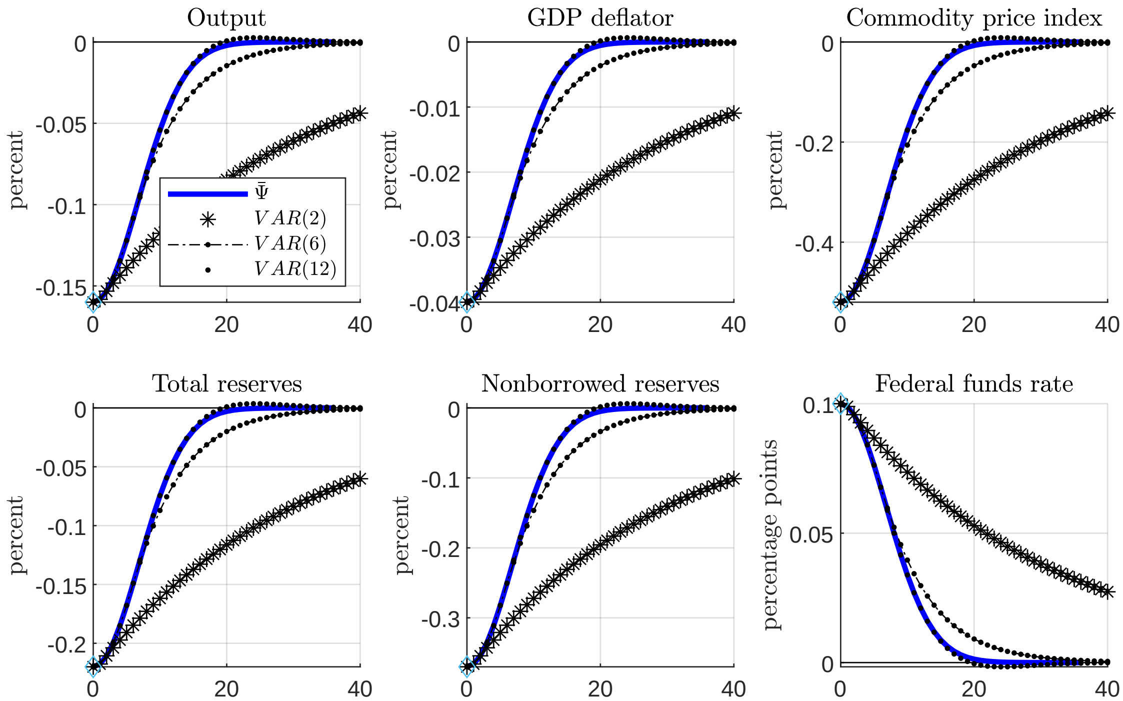
Figure 5: Posterior IRFs: the Flat (top) and Minnesota (bottom) priors



Note: Pointwise median as well as 68% and 90% credible sets. The responses correspond to a one standard deviation shock.

imposed dogmatically, and can be updated by the data if it wants to do so. As with the other two cases, we use a flat prior for the constant, centered at zero.

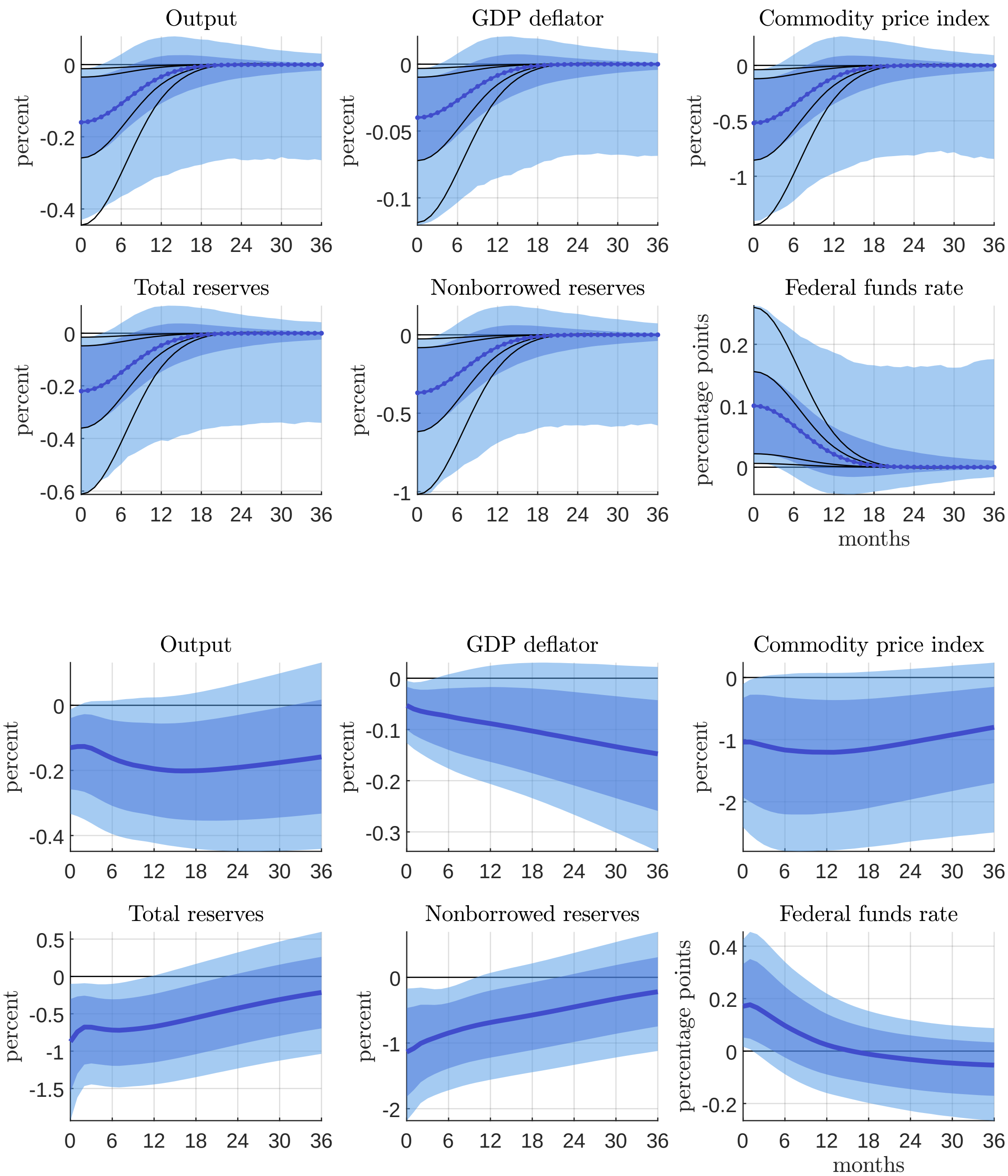
Figure 6: Our approach: $\bar{\Psi}$



Note: The thick blue line reports $\bar{\Psi}$. The black dashed and dotted lines show the value of the impulse responses associated with a VAR model with l lags parametrized as implied by $\{\bar{\Psi}_h\}_{h=0}^{l}$, with $l = 2, 6, 12$.

We make our prior beliefs operational as follows. For all variables, we specify $\bar{\Psi}$ using (15), setting $H = 36$. We set $b_{ij} = 0$, $\forall\ i, j$, and c_{ij} to allow for different degrees of persistence of the response to the shock. The baseline case assumes that it takes 8 months for the effects of the monetary shock to reach half of the impact response. We introduce the same beliefs across all variables and in response to all shocks for convenience, although the methodology does not require such a restriction. Since we do not introduce functional restrictions on $\bar{\Psi}$, a VAR with $p = 12$ lags need not replicate the dynamics of $\bar{\Psi}$ up to $H = 36$ horizons, and results (14b) do not hold exactly (this is *Case c* in Section 2.2). Figure 6 reports the IRFs associated with a VAR(l) with $l = 2, 6, 12$ parametrized as implied by $\{\bar{\Psi}_h\}_{h=0}^{l}$, when $l = 2, 6, 12$ (see equations (10) in the Online Appendix). Clearly, a VAR with 6 lags does a great job in approximating $\bar{\Psi}$, and $p = 12$ implies negligible approximation errors. As for the covariance matrix, we specify $V(B)$ as in equation (13). V_H is set as with the Minnesota prior, and we select λ adaptively. We start the algorithm at $\lambda = 0.005$ and

Figure 7: Our approach: prior (top panel) and posterior (bottom panel)



Note: The top panel reports $\bar{\Psi}$ (dotted line), the pointwise 68% and 90% prior credible set (shaded areas), and the same set when only prior uncertainty in B is present (solid lines). The bottom panel shows the pointwise median as well as the pointwise 68% and 90% credible sets. The responses correspond to a one standard deviation shock.

increase λ until the joint prior distribution associated with the monetary policy shock leaves between 5% and 10% outside of $\pm 2 \cdot \sqrt{\hat{\Sigma}_{ii}}$ for the first $H = 36$ horizons, with $\sqrt{\hat{\Sigma}_{ii}}$ estimated in the training sample (see the discussion in Section 3). The search produces a value of $\lambda = 0.0331$.

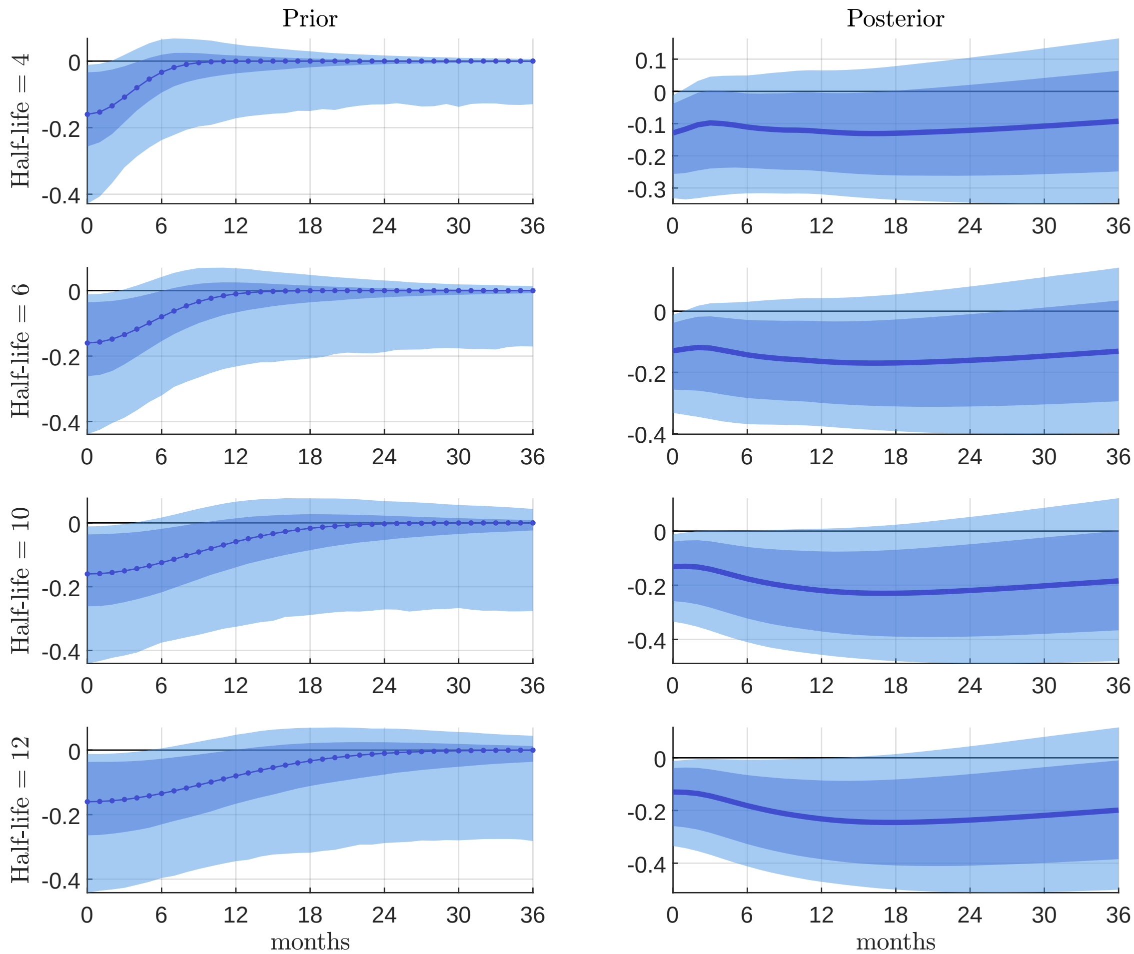
Figure 7 shows the prior (top panel) and the posterior (bottom panel) impulse responses to a one standard deviation shock associated with our setup. The dotted line in the top panel is $\bar{\Psi}$, while the shaded areas report the 68% and the 90% credible sets associated with our prior. The solid lines report the 68% and 90% credible sets obtained with only prior uncertainty in B, setting $\lambda = 0$. This helps to understand the implications of a relatively tight prior on $\boldsymbol{\pi}$. As should be expected, prior uncertainty is largely driven by $p(B)$ at shorter horizons, while the prior variance of $\boldsymbol{\pi}$ becomes relevant at longer horizons. The top panel shows that indeed the prior distribution is well centered around $\bar{\Psi}$, as should be expected. It also shows that the prior does not induce any hump-shaped dynamics.

The bottom panel of Figure 7 reports the posteriors. The one standard deviation monetary contraction leads to an impact median increase in the federal funds rate close to 20 basis points. The monetary contraction generates an impact decrease in the commodity price index by around 1 percent and an impact decrease in output by 0.12 percent. As for the dynamics, the interest rate reverts back within a year, while real GDP and the commodity price index further decrease displaying a hump-shaped response. The monetary contraction leads to a strongest effect on real GDP of 0.2 percent, which materializes one year after the shock. Contrary to the case from Figure 5, real GDP does not display the strongest response on impact. We stress that posterior uncertainty remains wide, in a way that is quantitatively similar to when using the Minnesota or the flat prior.

Next, we assess the sensitivity of the results to alternative prior specifications on the timing of the impulse responses. Figure 8 shows that the results are robust to alternative selections of $\bar{\Psi}$ that imply faster or slower responses to the monetary policy shocks (we set the half-life of the effects on all variables to 4, 6, 10 or 12 months). All prior beliefs lead to conclude that real GDP responds with a hump. The maximum effect occurs between a year and a year and a half, but never before one year or after one year and a half. This result thus supports the widely-held view that it takes time for the central bank to affect the real economy and that long and variable lags constrain its ability to affect output.

Our prior also has implications for the stationarity of the VAR model. We found

Figure 8: Response of real GDP (robustness)



Note: The dotted line in the left column represents $\bar{\Psi}$, the solid lines in the right column represents the pointwise median. Shaded areas represent pointwise 68% and 90% credible set for the prior and the posterior.

that 70% of the posterior draws associated with the Minnesota prior lead to a companion form with the maximum eigenvalue above 1. This number decreases to between 30% and 44% with our prior, depending on the half-life value used. Complex eigenvalues did not appear in either case. Table 1 reports how many posterior draws were needed to store 100,000 draws that satisfied sign restrictions at horizon 0 (baseline analysis), at horizons 0 and 1, and so on up to horizon 4. When sign restrictions are introduced only on impact, the computational cost is approximately the same with

Table 1: Computational burden as restrictions are introduced on higher horizons

up to horizon	Minnesota		Our prior	
0	2,030,062	21m03s	1,868,909	22m18s
1	3,173,015	32m45s	2,171,074	26m53s
2	6,638,708	1h6m51s	2,438,543	30m44s
3	12,969,568	2h14m33s	2,686,903	37m15s
4	26,533,676	4h34m33s	2,912,228	40m30s

Note: The table reports how many posterior draws were needed to store 100,000 that satisfy the sign restrictions, which were introduced up to horizon $h = 0, .., 4$. It also reports computations times. All codes are run on Matlab on a computer with an Intel i7-7700K 4.2GHz Quad Core processor and 64 GB RAM.

the Minnesota prior and with our prior. When sign restrictions are also introduced at higher horizons it becomes computationally more demanding to work with the Minnesota prior. By contrast, the computational cost under our prior barely changes. Thus, our prior choice can significantly reduce the computational burden when sign restrictions are introduced at longer horizons.

5 Conclusions

Bayesian VAR models are frequently used to estimate impulse response functions to structural shocks of interest. This paper develops a tractable prior distribution for SVAR coefficients that achieves two goals. First, it allows for an explicit introduction of prior beliefs on the shape of the impulse responses. Second, it does so by working with a Normal prior distribution which, in turn, ensures tractable posterior sampling.

We develop the methodology and illustrate its properties using simulated data from a small scale DSGE model. We use this simulation exercise to show how the key hyperparameters of the prior can be specified. We then use the prior we suggest to investigate how long it takes for a monetary policy shock to generate its strongest effect on real GDP. We show that the popular flat and Minnesota priors lead to posterior IRFs that features no hump-shaped response of real GDP. By contrast, our prior, which is set to mimic the belief that monetary policy shocks generate persistent yet temporary effects on the economy, leads to a posterior that features a hump-shaped response of real GDP. We also estimate that it takes between one year and a year and a half to obtain the maximum effect on output.

References

An, S. and Schorfheide, F. (2007), 'Bayesian analysis of DSGE models', *Econometric reviews* **26**(2-4), 113–172.

Andrle, M. and Benes, J. (2013), 'System priors: Formulating priors about DSGE models' properties', *IMF working paper* (13/257).

Arias, J. E., Caldara, D. and Rubio-Ramirez, J. F. (2019), 'The systematic component of monetary policy in SVARs: an agnostic identification procedure', *Journal of Monetary Economics* **101**, 1–13.

Arias, J. E., Rubio-Ramírez, J. F. and Waggoner, D. F. (2018), 'Inference based on Structural Vector Autoregressions identified with sign and zero restrictions: Theory and applications', *Econometrica* **86**(2), 685–720.

Bańbura, M., Giannone, D. and Reichlin, L. (2010), 'Large Bayesian vector auto regressions', *Journal of Applied Econometrics* **25**(1), 71–92.

Barnichon, R. and Matthes, C. (2018), 'Functional approximation of impulse responses', *Journal of Monetary Economics* **99**, 41–55.

Baumeister, C. and Hamilton, J. D. (2015), 'Sign restrictions, structural vector autoregressions, and useful prior information', *Econometrica* **83**(5), 1963–1999.

Baumeister, C. and Hamilton, J. D. (forthcoming), 'Advances in using vector autoregressions to estimate structural magnitudes', *Econometric Theory* .

Baumeister, C. J. and Hamilton, J. D. (2018), 'Inference in structural vector autoregressions when the identifying assumptions are not fully believed: Re-evaluating the role of monetary policy in economic fluctuations', *Journal of Monetary Economics* **100**, 48–65.

Bernanke, B. S. and Mihov, I. (1998), 'Measuring monetary policy', *The Quarterly Journal of Economics* **113**(3), 869–902.

Binning, A. (2013), 'Underidentified SVAR models: A framework for combining short and long-run restrictions with sign restrictions'.

Bruns, M. and Piffer, M. (2023), 'A new posterior sampler for Bayesian structural vector autoregressive models', *Quantitative Economics* **4**(14), 1221–1250.

Buda, G., Carvalho, V. M., Corsetti, G., Duarte, J. B., Hansen, S., Ortiz, Á., Rodrigo, T. and Rodríguez Mora, J. V. (2023), 'Short and variable lags', *Robert Schuman Centre for Advanced Studies Research Paper* (22).

Caldara, D. and Kamps, C. (2017), 'The analytics of SVARs: a unified framework to measure fiscal multipliers', *The Review of Economic Studies* **84**(3), 1015–1040.

Canova, F. (2007), *Methods for applied macroeconomic research*, Vol. 13, Princeton University Press.

Canova, F. and Ferroni, F. (2022), 'Mind the gap! Stylized dynamic facts and structural models', *American Economic Journal: Macroeconomics* **14**(4), 104–135.

Canova, F. and Pérez Forero, F. J. (2015), 'Estimating overidentified, non-recursive, time varying coefficients structural VARs', *Quantitative Economics* **6**(2), 309–358.

Christiano, L. J., Eichenbaum, M. and Evans, C. L. (1999), 'Monetary policy shocks: What have we learned and to what end?', *Handbook of Macroeconomics* **1**, 65–148.

Doan, T., Litterman, R. and Sims, C. (1984), 'Forecasting and conditional projection using realistic prior distributions', *Econometric reviews* **3**(1), 1–100.

Dwyer, M. (1998), 'Impulse response priors for discriminating structural vector autoregressions'.

Ferreira, L. N., Miranda-Agrippino, S. and Ricco, G. (2023), 'Bayesian local projections', *The Review of Economics and Statistics* pp. 1–45.

Giacomini, R. and Kitagawa, T. (2021), 'Robust Bayesian inference for set-identified models', *Econometrica* **89**(4), 1519–1556.

Giannone, D., Lenza, M. and Primiceri, G. E. (2015), 'Prior selection for vector autoregressions', *Review of Economics and Statistics* **97**(2), 436–451.

Gilchrist, S. and Zakrajšek, E. (2012), 'Credit spreads and business cycle fluctuations', *American Economic Review* **102**(4), 1692–1720.

Gordon, S. and Boccanfuso, D. (2001), 'Learning from structural vector autoregression models', *Manuscript, Universite Laval, Quebec City* .

Gourinchas, P.-O. and Tornell, A. (2004), 'Exchange rate puzzles and distorted beliefs', *Journal of International Economics* **64**(2), 303–333.

Harvey, A. C., Trimbur, T. M. and Van Dijk, H. K. (2007), 'Trends and cycles in economic time series: A Bayesian approach', *Journal of Econometrics* **140**(2), 618–649.

Inoue, A. and Kilian, L. (2020), 'The role of the prior in estimating VAR models with sign restrictions'.

Jarociński, M. and Marcet, A. (2019), 'Priors about observables in vector autoregressions', *Journal of Econometrics* **209**(2), 238–255.

Kadiyala, K. R. and Karlsson, S. (1997), 'Numerical methods for estimation and inference in bayesian var models', *Journal of Applied Econometrics* **12**(2), 99–132.

Kilian, L. (2022), 'Comment on Giacomini, Kitagawa, and Read's "narrative restrictions and proxies"', *Journal of Business & Economic Statistics* **40**(4), 1429–1433.

Kilian, L. and Lütkepohl, H. (2017), *Structural vector autoregressive analysis*, Cambridge University Press.

Kilian, L. and Murphy, D. P. (2012), 'Why agnostic sign restrictions are not enough: understanding the dynamics of oil market VAR models', *Journal of the European Economic Association* **10**(5), 1166–1188.

Kociecki, A. (2010), 'A prior for impulse responses in Bayesian structural VAR models', *Journal of Business & Economic Statistics* **28**(1), 115–127.

Koop, G. and Korobilis, D. (2010), 'Bayesian multivariate time series methods for empirical macroeconomics', *Foundations and Trends in Econometrics* **3**(4), 267–358.

Litterman, R. B. (1986), 'Forecasting with Bayesian vector autoregressions five years of experience', *Journal of Business & Economic Statistics* **4**(1), 25–38.

Miranda-Agrippino, S. and Ricco, G. (2019), Bayesian vector autoregressions: Estimation, *in* 'Oxford Research Encyclopedia of Economics and Finance'.

Miranda-Agrippino, S. and Ricco, G. (2021), 'The transmission of monetary policy shocks', *American Economic Journal: Macroeconomics* **13**(3), 74–107.

Paul, P. (2020), 'The time-varying effect of monetary policy on asset prices', *Review of Economics and Statistics* **102**(4), 690–704.

Piffer, M. and Podstawski, M. (2018), 'Identifying uncertainty shocks using the price of gold', *The Economic Journal* **128**(616), 3266–3284.

Plagborg-Møller, M. (2019), 'Bayesian inference on structural impulse response functions', *Quantitative Economics* **10**(1), 145–184.

Plagborg-Møller, M. and Wolf, C. K. (2021), 'Local Projections and VARs Estimate the Same Impulse Responses', *Econometrica* **89**(2), 955–980.

Ramey, V. A. (2016), 'Macroeconomic shocks and their propagation', *Handbook of macroeconomics* **2**, 71–162.

Rubio-Ramirez, J. F., Waggoner, D. F. and Zha, T. (2010), 'Structural vector autoregressions: Theory of identification and algorithms for inference', *The Review of Economic Studies* **77**(2), 665–696.

Uhlig, H. (2005), 'What are the effects of monetary policy on output? Results from an agnostic identification procedure', *Journal of Monetary Economics* **52**(2), 381–419.

Van Dijk, H. K. and Kloek, T. (1980), 'Further experience in Bayesian analysis using monte carlo integration', *Journal of Econometrics* **14**(3), 307–328.

Villani, M. (2009), 'Steady-state priors for vector autoregressions', *Journal of Applied Econometrics* **24**(4), 630–650.

Waggoner, D. F., Wu, H. and Zha, T. (2016), 'Striated Metropolis–Hastings sampler for high-dimensional models', *Journal of Econometrics* **192**(2), 406–420.