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

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

November 26, 2024

Abstract

We develop a prior for VAR autoregressive coefficients that allows for the introduction of flexible non-dogmatic prior beliefs on the shape of the structural impulse responses. We achieve this with a particular setting of the moments of a Normal distribution. Posterior computations are no more demanding than existing prior specifications; yet the methodology provides Bayesian shrinkage directly on impulse responses. Introducing the prior belief that monetary policy shocks generate temporary but persistent effects leads to a hump-shaped response of GDP. The trough occurs with twelve and eighteen months after the shock, depending on how much a-priori persistence we introduce.

JEL classification: C32, E52.

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

[*]We are thankful to Marta Bańbura, Luca Gambetti, Paolo Gelain, Michele Lenza, Christian Matthes, Giovanni Ricco, Giuseppe Ragusa, Raffaella Giacomini, Giorgio Primiceri, Domenico Giannone, Barbara Rossi, Herman van Dijk and Alessio Volpicella for helpful comments and suggestions. We also thank the participants of several seminars and conferences for comments and suggestions.

[†]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 the dynamics induced by structural shocks. 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, Caldara and Iacoviello, 2022), among many others.

Researchers typically hold strong views on what dynamic responses ought to be considered reasonable. Beliefs could be held on the timing and on the persistence of the responses, or on their shape. For example, it is generally believed that a contractionary monetary policy shock should persistently decrease output, 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 leading examples of beliefs researchers have about the likely path of certain variables in responses to structural shocks (Ramey, 2016 Gourinchas and Tornell, 2004).

Unfortunately, macroeconomists are severely constrained in their ability to support the estimation of impulse response functions with meaningful prior beliefs on their timing, their persistence, or their shape. The computational convenience of working with Vector Autoregressive models has been widely acknowledged in the literature (Kilian and Lütkepohl, 2017). Yet, the existing practice of specifying a Minnesota-like prior for vector autoregressive (VAR) parameter does not allow a researcher to introduce, even indirectly, non-dogmatic constraints on the impulse response functions. In fact, in such a setup, prior beliefs about the shape and the timing of impulse responses can only be imposed dogmatically, by rejecting all posterior draws failing to imply the required pattern of responses (Canova and Pappa, 2011).

Beliefs on the shape of the impulse responses can be naturally introduced when estimating Moving Average (MA) models (Plagborg-Møller, 2019). However, MA models are computationally demanding to work with, and may require approximation techniques to obtain estimates of their parameters (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 a 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. In this paper we develop an approach to sharpen inference about IRFs using non-dogmatic beliefs on the timing, the persistence and/or the shape of the impulse responses in structural VAR models. Rather than imposing a full prior distribution directly on the impulse responses, we work with the popular Normal prior on the reduced form VAR parameters, and [FC SUGGESTS leave the prior on the contemporaneous impulse responses unrestricted I AM WORRIED PEOPLE THINK WE CAN WORK WITHOUT IDENTIFYING RESTRICTIONS]. Our prior consists of a new specification of the first moment of the Normal VAR prior that differs from those of the Minnesota specification. The prior specification we propose achieves two goals. First, the first moments of the prior are selected so that the implied distribution of the IRFs is approximately centered around the IRF dynamics a researcher wants to introduce. Second, because the prior on the reduced form autoregressive VAR parameters has a Normal format, it retains the computational convenience of using highly tractable posterior samplings. Hence, one can introduce non-dogmatic beliefs about features of the IRFs by simply replacing the Minnesota-like specification with our proposed specification, which is obtained analytically from the recursive computation of the IRFs. Our approach nests, as a special case, the Minnesota-like prior, as well as the long run prior developed by Giannone et al. (2019).

Identification is a crucial ingredient of impulse response analysis. Our prior for the VAR autoregressive coefficients can be combined with existing contemporaneous identification strategies, for example, zero restrictions, sign restrictions or external instrument restrictions, and it is compatible with both the approaches of Rubio-Ramirez et al. (2010) and of Baumeister and Hamilton (2015) We stress that our contribution is not to provide a new approach to achieve the identification of the structural shocks, but to develop a method to achieve shrinkage via prior beliefs on the dynamics of the impulse responses.

There are three main advantages of our settings. It has been acknowledged that Bayesian shrinkage is more naturally introduced on endogenous functions of the parameters of the model (see, for instance, Van Dijk and Kloek, 1980, Harvey et al., 2007). In the SVAR literature, this type of priors is 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 shrinkage of dynamic responses, which are undoubtedly a key endogenous function of SVAR models (Kilian, 2022). A second advantage is that, contrary to existing approaches, by suggesting the data to produce particular shapes, it can sharpen inference without requiring additional identifying restrictions (Kilian and Murphy, 2012, Amir-Ahmadi and Drautzburg, 2021). The third advantage is that the specification is flexible enough to allow for a combination of tighter beliefs on some impulse responses and looser beliefs on others. Put differently, our approach does not require formulating prior beliefs on the shape of all structural impulse responses.

We illustrate the properties of our proposed prior specification using data simulated from a conventional three-variable New Keynesian model. The model features a very persistent response of the output gap to a government spending shock. As expected, in a large sample the prior is irrelevant, and all prior specifications we consider lead to the same impulse responses. However, in a sample of a realistic size, both a flat and a Minnesota prior lead to posterior IRFs that largely underestimate the degree of persistence of the output gap response. By contrast, our prior makes it possible to introduce the belief that, in the mean, the effect of a government spending shock to the output gap is relatively persistent, while allowing for considerable uncertainty around the mean. As a result, our prior leads to posterior IRFs that mimic the correct half-life of the true responses, provided the prior variance of the autoregressive parameters is not too large.

We study the classical question of how US output responds to a monetary policy surprise. This issue of at what horizon the maximum response occurs has received considerable attention over the last twenty-five years (Christiano et al., 1999, Uhlig, 2005, Antolín-Díaz and Rubio-Ramírez, 2018), but both the shape and the timing of the output response dynamics still remain unsettled. We take a standard six variable VAR model and identify the monetary policy disturbances using impact sign restrictions. A flat and a Minnesota prior for the VAR coefficients produce ~~exponentially declining posterior output responses and no hump~~ a response of output that displays the strongest effect on impact, and no hump-shape. We then introduce the belief that monetary shocks generate persistent mean output effects, a belief which is in line with a wide class of current New Keynesian macroeconomic models. We find that the posterior distribution of output responses displays hump-shaped dynamics, regardless of degree of persistence the prior displays. Importantly, the hump occurs even though our prior does not require the mean responses to be hump-shaped. Interestingly, our results suggest that, depending on how persistent are a-priori assumed to be the responses, it takes between one and one and a half years for the monetary shock to generate its largest output effect. Finally, a one standard deviation shock that increases the federal funds rate by 20 basis points on impact leads to a maximum decrease in real GDP of 0.20%, approximately. Thus, our prior supports the widely-held view that a central bank may be able to affect real economic activity; but this occurs with long and variable lags, and the magnitude of the effect is generally small (Buda et al., 2023).

The relationship with the literature. There is a considerable amount of literature dealing with 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, 2024). Relative to Baumeister and Hamilton (2015, 2024), we focus on impulse responses rather than on structural elasticities. Relative to Baumeister and Hamilton (2018), we consider dynamic responses rather than the impact effect of the shocks. The paper is also related to the work of Barnichon and Matthes (2018), who impose prior shape restrictions on MA models, and to an earlier contribution of Kociecki (2010), who works with recursive identification and considers a jointly Normal prior distributions on the impulse responses. Our approach has the same flavor 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 provide priors for impulse responses.

Our approach is complementary to Plagborg-Møller (2019), who estimates impulse responses directly from MA models. Our approach implicitly assumes that the data has an invertible representation, and ~~restricts the covariance structure of the prior~~ [FABIO, I AM WORRIED THAT THEY WILL SAY WE DO NOT REALLY ADD RESTRICTIONS ON THE VARIANCE OF THE IRF. CAN YOU CHECK IF THE FOLLOWING FLOWS] and is less flexible with respect to the prior variance on the impulse responses. However, it is computationally less demanding and is compatible with a conjugate VAR prior, which has been extensively employed in the literature.

Most of the current Bayesian SVAR literature dealing with sign identification restrictions nowadays discusses the pros and cons of following a two-step approach to the estimation and the identification of the VAR (Baumeister and Hamilton, 2015), whether the failure to update contemporaneous prior beliefs is a problem or not (Inoue and Kilian, 2020), and whether contemporaneous beliefs implied by a standard two-step algorithm are informative or not (Arias et al., 2024). Our paper does not take a stand on these issues. Our approach is consistent with the methodology of Arias et al. (2018), who provide tools for combining sign and zero restrictions, and with any approach imposing prior restrictions on contemporaneous elasticities, for instance Baumeister and Hamilton (2015). Finally, while we do not follow the method of Giacomini and Kitagawa (2021), such an approach can also be used in conjunction with our prior specification to robustify inference.

The outline of the paper. The rest of the paper is organized as follows. Section 2 discusses in details the prior on autoregressive parameters we propose. Section 3 illustrates the properties of our specification using data simulated from a standard DSGE model. Section 4 studies how monetary policy shocks are transmitted to real output. Section 5 concludes. An Online Appendix contains the derivations and the computational details mentioned in the paper.

2 The empirical methodology

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

2.1 The model

We write a 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 serially independent 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. The mapping between model (1) and model (2) is discussed at length, for instance, in Arias et al. (2018). Below, we use notation:

$$\Pi = [\Pi_1, .., \Pi_p], \tag{3}$$

$$\boldsymbol{\pi} = \text{vec}(\Pi), \tag{4}$$

$$\tilde{\boldsymbol{\pi}} = (\boldsymbol{\pi}', \boldsymbol{c}')'. \tag{5}$$

For the rest of this section we assume, without loss of generality, that the data is demeaned so that $\boldsymbol{c}' = \boldsymbol{0}$.

Let Ψ_h denote the impulse response function (IRF) h periods after the shocks, and let M be the maximum horizon of interest for which impulse responses are computed. Ψ_h is of dimensions $k \times k$, with entry $i, j = 1, .., k$ capturing how variable i responds to shock j after h horizons from the shock. 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):

$$\Psi_0 = B, \tag{6a}$$

$$\Psi_1 = \Pi_1 \Psi_0, \tag{6b}$$

$$\Psi_2 = \Pi_1 \Psi_1 + \Pi_2 \Psi_0, \tag{6c}$$

$$\Psi_3 = \Pi_1 \Psi_2 + \Pi_2 \Psi_1 + \Pi_3 \Psi_0, \tag{6d}$$

$$\dots$$

$$\Psi_p = \Pi_1 \Psi_{p-1} + \Pi_2 \Psi_{p-2} + \cdots + \Pi_p \Psi_0, \tag{6e}$$

$$\Psi_{p+1} = \Pi_1 \Psi_p + \Pi_2 \Psi_{p-1} + \cdots + \Pi_p \Psi_1, \tag{6f}$$

$$\dots$$

$$\Psi_M = \Pi_1 \Psi_{M-1} + \Pi_2 \Psi_{M-2} + \cdots + \Pi_p \Psi_{M-p}. \tag{6g}$$

When B is non-singular, (6) provides a one-to-one mapping between the SVAR parameters and the IRF parameters for any $M \geq p$. Thus, any prior beliefs on the SVAR coefficients imply prior beliefs on the elements of the IRFs via the system of equations (6).

2.2 Our approach

Our approach is general along several dimensions, including the identification of the shocks, whether beliefs on the shape of IRFs are held dogmatically or not, and the specification of the covariance matrix of the prior. To provide intuition, we first illustrate the key features of our approach in a simplified setting. We then generalize the approach and we relate it to the existing literature.

2.2.1 Illustration in a simplified environment

For the time being, suppose that:

a) the identification of the economic shocks is achieved via a recursive identification approach;

b) an inverse-Wishart prior is used for Σ, and B is set equal to the Cholesky decomposition of Σ;

c) all k structural shocks are identified;

d) the prior for $\tilde{\boldsymbol{\pi}}$ is Normal independent of Σ,

$$\tilde{\boldsymbol{\pi}} \sim N(\boldsymbol{\mu}, V). \tag{7}$$

As is well known, the joint posterior $p(\tilde{\boldsymbol{\pi}}, \Sigma|Y)$ in this case can be conveniently explored with the Gibbs sampler (Koop and Korobilis, 2010).

We are interested in whether the selection of the hyperparameters $(\boldsymbol{\mu}, V)$ can grant the researcher some flexibility over the implied prior for the IRFs, given that, conditional on B, a prior $p(\tilde{\boldsymbol{\pi}})$ implies a prior for the IRFs via the system (6). It is standard in the literature to set $(\boldsymbol{\mu}, V)$ according to:

$$E((\Pi_h)_{ij}) = \begin{cases} \delta_i, & j = i, h = 1 \\ 0 & \text{otherwise} \end{cases}, \qquad V((\Pi_h)_{ij}) = \begin{cases} \frac{\lambda^2}{h^2}, & j = i \\ \eta \frac{\lambda^2}{h^2} \frac{\sigma_i^2}{\sigma_j^2} & \text{otherwise} \end{cases}, \tag{8}$$

which is typically referred to as Minnesota prior. Note that a flat prior is obtained by letting λ to be large, while the random walk and the white noise specifications can be 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 $\lambda, \eta, \sigma_i, \sigma_j$).[1] While the forecasting properties of a VAR endowed with such prior restrictions are well-documented, the prior allows for no flexibility in designing IRFs shapes. One could constrain the implied IRFs by adding an accept/reject step into the algorithm to ensure that the posterior draws do produce the required shapes (for example, that the response of economic activity to a monetary shock is hump-shaped). However, such a way of proceeding introduces restrictions *dogmatically*, which might not be the intention of the researcher. In addition, the computations may be very inefficient if most of the posterior draws fail to satisfy the candidate restrictions.

Our approach endogenously selects (μ, V) in such a way that the mean of the IRFs has certain a-priori features. Thus, we view our beliefs as a tool to implement posterior Bayesian shrinkage on impulse responses. Let $\bar{\Psi} = (\bar{\Psi}_0, .., \bar{\Psi}_H)$ be an array reflecting the researcher's prior mean on the impulse responses, where H is the maximum horizon up to where beliefs about the IRFs are formulated. Generally, $H \leq M$, as one need not have prior beliefs stretching as far as the horizon of interest. $\bar{\Psi}$ is a high dimensional object, as it includes $k^2(H+1)$ entries. Assume, for the moment, that $H = p$ (the

[1] 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 Σ.

number of lags of the SVAR), and set $\bar{\Psi}_0$ to the expected value of B implied by the prior on Σ and the identification approach used (here, a lower triangular matrix with positive diagonal entries). The researcher sets the remaining entries $(\bar{\Psi}_1, .., \bar{\Psi}_H)$ to capture his beliefs about the dynamics of the IRFs to the structural shocks. Last, substitute $\bar{\Psi}$ into (6), eliminate the first equation and invert the next p equations of system (6) to obtain:

$$\bar{\Pi}_1 = \bar{\Psi}_1 \bar{\Psi}_0^{-1}, \tag{9a}$$

$$\bar{\Pi}_2 = [\bar{\Psi}_2 - \bar{\Pi}_1 \bar{\Psi}_1] \bar{\Psi}_0^{-1}, \tag{9b}$$

$$\bar{\Pi}_3 = [\bar{\Psi}_3 - \bar{\Pi}_1 \bar{\Psi}_2 - \bar{\Pi}_2 \bar{\Psi}_1] \bar{\Psi}_0^{-1}, \tag{9c}$$

$$\dots$$

$$\bar{\Pi}_p = [\bar{\Psi}_p - \bar{\Pi}_1 \bar{\Psi}_{p-1} - \bar{\Pi}_2 \bar{\Psi}_{p-2} - \cdots - \bar{\Pi}_{p-1} \bar{\Psi}_1] \bar{\Psi}_0^{-1}. \tag{9d}$$

$\bar{\Pi} = [\bar{\Pi}_1, .., \bar{\Pi}_p]$ are the values of the VAR autoregressive coefficients associated with the selected $\bar{\Psi}$.

In this simplified setting, replacing the specification of $\boldsymbol{\mu}$ from (8) with

$$E((\Pi_h)_{ij}) = (\bar{\Pi}_h)_{ij}, \qquad h = 1, .., p, \tag{10}$$

is enough to gain control on the implied prior on the impulse responses. In fact, the specified prior $p(\tilde{\boldsymbol{\pi}}, \Sigma)$, the Cholesky identification of B and the equality $E(B) = \bar{\Psi}_0$ jointly imply, via system (6), that $p(\Psi_0, \Psi_1, ...)$ satisfies

$$E(\Psi_h) = \bar{\Psi}_h, \qquad h = 0, 1, \tag{11a}$$

$$\lim_{V \to 0} E(\Psi_h) = \bar{\Psi}_h, \qquad h = 2, .., H, \tag{11b}$$

where (11a) holds for $h = 0$ by assumption, while the remaining conclusions are derived in the Online Appendix. While (11b) is satisfied only in the limit as $V \to 0$, Section 2.4, Section 3 and Section 4 document that the approximation error $E(\Psi_h) - \bar{\Psi}_h$ is negligible for values of V used in the literature, for instance see Bańbura et al. (2010). Put differently [I CHECKED WITH MY WIFE: IT'S EITHER "TO PUT IT DIFFERENTLY", OR "PUT DIFFERENTLY", BUT NOT "PUT IT DIFFERENTLY"], since $\bar{\Psi}$ is selected by the researcher, replacing a Minnesota choice of $\boldsymbol{\mu}$ with an alternative choice which depends on $\bar{\Psi}$, gives the researcher control on the prior expectation of the impulse response, , without imposing that $p(\Psi_0, \Psi_1, ...)$ is joint Normal. [FABIO, YOU SUGGEST SUBSTITUTING THIS WITH "the resulting prior on Ψ is not Normal". WE HAD THIS BEFORE BUT ANDRZEJ ASKED TO REMOVE IT SINCE WE DO NOT REALLY PROVE THAT IT IS NOT NORMAL. WE JUST CANNOT DERIVE IT (unless I missinterpreted his point)] Given that our prior on the VAR coefficients is Normal, standard posterior samplers can be used. .

2.2.2 Generalization for $H > p$

The discussion in subsubsection 2.2.1 focuses on the case in which the researcher formulates beliefs about the IRFs up to p horizons, where p is the number of lags in the SVAR. We will refer to this as *Case a*. It is useful to generalize the approach to the case in which beliefs are expressed up to horizon $H > p$. Since p is typically selected by the researcher, one could in principle only work with *Case a* and increase p up to the intended horizon H. However, in practice, higher values of p make model dimensionality an issue, suggesting that it is better to derive the mapping directly for $H > p$.

Define $\bar{\Pi}$ like in subsubsection 2.2.1 as the value of the SVAR coefficients associated with the first $p+1$ blocks of $\bar{\Psi}$. It can happen, in principle, that when substituting $\bar{\Psi} = (\bar{\Psi}_0, .., \bar{\Psi}_H)$ and $\bar{\Pi} = [\bar{\Pi}_1, .., \bar{\Pi}_p]$ into system (6), all equalities hold despite the fact that $H > p$. In this case $\bar{\Psi} = (\bar{\Psi}_0, .., \bar{\Psi}_H)$ are functionally constrained in a way that a SVAR model with $p < H$ lags can still generate $\bar{\Psi}$ via equation (6). We will refer to this as *Case b*. As discussed in the Online Appendix, the results (11) hold for *Case b*.

When $H > p$ and $\bar{\Psi}$ are not functionally constrained, which we refer to as *Case c*, one could in principle still select $\boldsymbol{\mu}$ as in (10), effectively using only the first $p+1$ entries of $\bar{\Psi}$. This produces smaller approximation errors $E(\Psi_h) - \bar{\Psi}_h$ at shorter horizons at the cost of potentially introducing larger errors at longer horizons. Rather than using such an inefficient approach we consider all horizons H and select $\boldsymbol{\mu}$ as discussed next.

Define $\boldsymbol{b} = \text{vec}(B)$, $\Psi = [\Psi_0, \Psi_F]$, $\Psi_F = [\Psi_1, .., \Psi_H]$, $\boldsymbol{\psi} = \text{vec}(\Psi)$, $\boldsymbol{\psi}_F = \text{vec}(\Psi_F)$ and $\boldsymbol{\psi}_h = \text{vec}(\Psi_h)$, for $h = 0, 1, .., H$. Vectorizing the first $H+1$ equations of the system (6) one obtains:

$$\boldsymbol{\psi}_0 = \boldsymbol{b}, \tag{12}$$

$$\boldsymbol{\psi}_F = R\boldsymbol{\pi}, \tag{13}$$

where

$$R = R_H \otimes I_k \equiv \underbrace{\begin{bmatrix} \Psi_0' & 0 & 0 & \dots & 0 \\ \Psi_1' & \Psi_0' & 0 & .. & 0 \\ \Psi_2' & \Psi_1' & \Psi_0' & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Psi_{p-1}' & \Psi_{p-2}' & \Psi_{p-3}' & \dots & \Psi_0' \\ \Psi_p' & \Psi_{p-1}' & \Psi_{p-2}' & \dots & \Psi_1' \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Psi_{H-1}' & \Psi_{H-2}' & \Psi_{H-3}' & \dots & \Psi_{H-p}' \end{bmatrix}}_{R_H} \otimes I_k. \tag{14}$$

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. Let $\bar{\boldsymbol{\psi}}_h = \text{vec}(\bar{\Psi}_h)$, $\bar{\boldsymbol{\psi}} = (\bar{\boldsymbol{\psi}}_0', \bar{\boldsymbol{\psi}}_F')' = \text{vec}(\bar{\Psi})$, $\bar{\Psi} = [\bar{\Psi}_0, \bar{\Psi}_F]$, $\bar{\Psi}_F = [\bar{\Psi}_1, .., \bar{\Psi}_H]$. Define the artificial random variables W_h, $h = 1, .., H$ as

$$W_1 = \Pi_1 \bar{\Psi}_0 - \bar{\Psi}_1, \tag{15a}$$

$$W_2 = \Pi_1 \bar{\Psi}_1 + \Pi_2 \bar{\Psi}_0 - \bar{\Psi}_2, \tag{15b}$$

$$W_3 = \Pi_1 \bar{\Psi}_2 + \Pi_2 \bar{\Psi}_1 + \Pi_3 \bar{\Psi}_0 - \bar{\Psi}_3, \tag{15c}$$

$$\vdots$$

$$W_p = \Pi_1 \bar{\Psi}_{p-1} + \Pi_2 \bar{\Psi}_{p-2} + \dots + \Pi_p \bar{\Psi}_0 - \bar{\Psi}_p, \tag{15d}$$

$$W_{p+1} = \Pi_1 \bar{\Psi}_p \ + \ \Pi_2 \bar{\Psi}_{p-1} + \dots + \Pi_p \bar{\Psi}_1 - \bar{\Psi}_{p+1}, \tag{15e}$$

$$\vdots$$

$$W_H = \Pi_1 \bar{\Psi}_{H-1} \ + \ \Pi_2 \bar{\Psi}_{H-2} + \dots + \Pi_p \bar{\Psi}_{H-p} - \bar{\Psi}_H. \tag{15f}$$

The system of equations (15) is closely related to the system (6) except that it drops the equation at horizon 0, replaces the matrices Ψ_h with the matrices of hyperparameters $\bar{\Psi}_h$, and drops the last $M - H$ equations. The system (15) can be vectorized as

$$\boldsymbol{w} = \bar{R}\boldsymbol{\pi} - \bar{\boldsymbol{\psi}}_F, \tag{16}$$

where $\boldsymbol{w} = \text{vec}([W_1, .., W_H])$ and $\bar{R}$ is defined in equation (14) after replacing Ψ with $\bar{\Psi}$. If $\bar{\Psi}_0$ is non-singular, the matrix $\bar{R}$ has full column rank. Premultiplying both sides of (16) by $\bar{R}'$ and rearranging the terms gives:

$$\boldsymbol{\pi} = \left(\bar{R}'\bar{R}\right)^{-1}\bar{R}'\boldsymbol{w} + \left(\bar{R}'\bar{R}\right)^{-1}\bar{R}'\bar{\boldsymbol{\psi}}_F. \tag{17}$$

Note that in the system of equations (16) $\boldsymbol{\pi}$ is of lower dimension than $\boldsymbol{w}$ for $H > p$. In other words, given $(\bar{R}, \bar{\boldsymbol{\psi}}_F)$, whether a solution for $\boldsymbol{\pi}$ exists depends on $\boldsymbol{w}$.

From (17), we select $\boldsymbol{\mu}$ as

$$\boldsymbol{\mu} = \left(\bar{R}'\bar{R}\right)^{-1}\bar{R}'\bar{\boldsymbol{\psi}}_F. \tag{18}$$

The three cases of interest can be summarized as follows:

1) When $H = p$ (*Case a*), a SVAR with p lags replicates exactly the pattern of $\bar{\Psi}$, making the system (16) consistent for $\boldsymbol{w} = \mathbf{0}$. This means that (18) is the unique solution at $\boldsymbol{w} = \mathbf{0}$, $\bar{R}$ is square and invertible, and $E(\boldsymbol{\pi})$ simplifies to $\boldsymbol{\mu} = \bar{R}^{-1}\bar{\boldsymbol{\psi}}_F$, which coincides with vec$(\bar{\Pi})$ defined in equation (9).

2) 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 (16) is still consistent for $\boldsymbol{w} = \mathbf{0}$, (18) is still its unique solution, and $\boldsymbol{\mu}$ from (18) still coincides with vec$(\bar{\Pi})$ defined in equation (9) despite $\bar{R}$ not being square and invertible.

3) 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 (16) inconsistent. Yet, $E(\boldsymbol{\pi})$ from (18) is still the unique solution if $\boldsymbol{w} = [\bar{R}\left(\bar{R}'\bar{R}\right)^{-1}\bar{R}' - I]\bar{\boldsymbol{\psi}}_F$. In the Online Appendix we show that (11) do not necessarily hold under *Case c*. However, the approximation error $E(\Psi_h) - \bar{\Psi}_h$ can still be small, and its size can be checked numerically. Indeed, in all the exercises we run corresponding to *Case c*, the approximation error was small, making the methodology suitable even for $H > p$.

2.3 Discussion

The previous section has derived the prior for the VAR coefficients under the assumption that identification is achieved via the recursive approach. This assumption is not needed for our methodology to work. All that is needed is that $E(B) = \bar{\Psi}_0$, an assumption that is not restrictive, given that $\bar{\Psi}_0$ is a hyperparameter in our setup.

Sign restrictions can be used, formulating prior beliefs on $A = B^{-1}$ (as advocated by Baumeister and Hamilton, 2015, Baumeister and Hamilton, 2024), on B (as suggested by Bruns and Piffer, 2023) or on a combination of A and B (see Baumeister and Hamilton, 2018). One can also combine sign with zero restrictions (Binning, 2013, Arias et al., 2018). Identification via external instruments can also be used by adding the instrument to the list of variables in a block recursive SVAR, as in Plagborg-Møller and Wolf (2021), and the specification can be extended to the VARX setting used by Paul (2020). While formulation of the prior for the relevant instantaneous objects is an active field of research, our method does not take a stand on the issue.[2]

In applied work researchers might want to introduce identifying restrictions imposing sign or shape restrictions at horizons greater than the contemporaneous one (for alternative methods to sharpen inference see, for instance, Kilian and Murphy, 2012 and Amir-Ahmadi and Drautzburg, 2021 FABIO, I DON'T UNDERSTAND YOUR COMMENT "Kilian and Murphy do not impose restrictions at horizons greater than 0", WE SAY THAT THAT'S FOR ALTERNATIVE METHODS. I HAVE SWAPPED THE ORDER, JUST TO BE SAFE, LET ME KNOW IF IT DOES NOT FLOW[]). This, however, can be computationally costly if the joint posterior distribution has a large mass on the part of the parameter space that violates the restrictions. Our approach can help, because it can induce a shape on the impulse response that is in line with the intended restrictions, hence reducing the number of draws required to obtain a desired number of draws satisfying the additional restrictions. As we show in Section 4 in an application with sign restrictions on future horizons of the impulse response, our approach can significantly reduce the computational burden to draw from the restricted posterior distributions.

Our prior for $\boldsymbol{\pi}$ is Normal. For this reason, posterior sampling is not in any way

[2] In principle, our approach can introduce non-dogmatic beliefs on the shape of the impulse responses even when no identifying restriction is introduced. For instance, assume that $p(B)$ is centered at $E(B)$ which is freely specified by the researcher. Since our method only requires specifying $\bar{\Psi}$ and ensuring $\bar{\Psi}_0 = E(B)$, it works irrespectively of whether $p(B)$ features sign restrictions or other identifying restrictions. This is the case, for instance, in the illustration in Figure 1 and Figure 2 shown below. Of course, without identifying restrictions the implied impulse responses would carry no economic interpretation and the shocks would not be structural shocks. [FABIO, YOU WERE ASKING ABOUT THE MEANING OF THIS FOOTNOTE. IT IS SUPPOSED TO ADDRESS A POINT OF THE REFEREE, WHO SAYS THAT OUR APPROACH CONDITIONS ON B AND HENCE IS INHERENTLY LINKED TO IDENTIFICATION. HERE I AM SAYING THAT YOU CAN ALSO INTRODUCE NO IDENTIFYING RESTRICTION AND LEAVE THE SHOCKS UNLABELLED, AND STILL INTRODUCE A NON DOGMATIC SHAPE TO THEIR RESPONSE. THIS IS WHAT IS DONE IN NOW FIGURE 1 AND 2, WHERE NO SIGN RESTRICTIONS ARE INTRODUCED]

more challenging than with existing methodologies. If an inverse-Wishart-Uniform prior on (Σ, Q) is employed, the joint posterior can be explored with standard methods: a Gibbs sampler when we specify the prior as in (7), or direct sampling, if the prior (7) is replaced with

$$\tilde{\boldsymbol{\pi}}|\Sigma \sim N(\boldsymbol{\mu}, V_s \otimes \Sigma). \tag{19}$$

Our method also works under this latter specification, and the results from equation (11b) now hold if, in the limit, $V_s \to 0$. When an inverse-Wishart-Uniform prior for (Σ, Q) is not assumed, posterior sampling requires a more involved method for computing the marginal posterior $p(B|Y)$, for example the Metropolis-Hastings algorithm as in Baumeister and Hamilton (2015), the two-step algorithm as in Bruns and Piffer (2023), or the methods discussed in Canova and Pérez Forero (2015) and Waggoner et al. (2016).

It is common in the literature to study the dynamics induced only by a subset of the structural shocks. The prior on $\boldsymbol{\pi}$ affects the joint posterior of the impulse responses of all shocks, hence identification of a subset of shocks is hard to achieve simply changing the prior on $\boldsymbol{\pi}$. FABIO, YOU ARE RIGHT THAT IT DOES NOT WORK WITH CHOLESKY. HERE IS A NEW ATTEMP One solution is to adjust the specification of $p(B)$. For instance, if only a subset of the shocks are identified and identification is achieved using sign restrictions on the impact effect of the shocks, then one can specify $p(B)$ to feature a wide variance for the columns associated with the unidentified shock, and a smaller variance for the columns associated with the identified shocks. While we illustrate this strategy graphically in Section 2.4, the key idea is as follows. The IRF matrix at horizon i is linear in B i.e. $\Psi_i = f(\Pi_1, .., \Pi_i)B$. Let us denote the $j-$th column of Ψ_i as $\boldsymbol{\psi}_i$ and the $j-$th column of B as $\boldsymbol{b}_j$. Then, the IRFs to the $j-$th shock are the sequence $\boldsymbol{\psi}_1 = f(\Pi_1)\boldsymbol{b}_j$, $\boldsymbol{\psi}_2 = f(\Pi_1, \Pi_2)\boldsymbol{b}_j$, $\boldsymbol{\psi}_3 = f(\Pi_1, \Pi_2, \Pi_3)\boldsymbol{b}_j$, ... Notice that $E(\boldsymbol{\psi}_i)$ depends on V (since Π_i's enter in cross and/or multiple products) but not on the prior variance on $\boldsymbol{b}_j$ (since each $\boldsymbol{\psi}_i$ is linear in $\boldsymbol{b}_j$ and $E(B) = \bar{\Psi}_0$). It follows that changing prior variance of $\boldsymbol{b}_j$ has no effect on $E(\boldsymbol{\psi}_i)$. Without loss of generality let us focus on $E(\boldsymbol{\psi}_i\boldsymbol{\psi}_i')$. $E(f(\Pi_1, .., \Pi_i)\boldsymbol{b}_j\boldsymbol{b}_j' f(\Pi_1, .., \Pi_i)')$ must depend on $\boldsymbol{b}_j$ through $E(\boldsymbol{b}_j\boldsymbol{b}_j')$. Large $var(\boldsymbol{b}_j)$ implies large $E(\boldsymbol{b}_j\boldsymbol{b}_j')$, hence large $var(\boldsymbol{\psi}_i)$, for each i. Hence, a large variance $var(\boldsymbol{b}_j)$ allows us to control for prior variance of all IRFs (i.e. at all horizons) with respect to the $j-$th shock. As a result, given a value of V, specifying a large prior variance for some columns of B and a low prior variance for the remaining columns offers a way of making the selection of $\bar{\Psi}$ less relevant for the subset of shocks that are not identified, making the prior diffuse in the relevant dimensions. The prior distribution for B proposed by Bruns and Piffer (2023) is suitable in situations when a researcher is only interested in partial identification.

A key advantage of our approach is that it does not require deriving and integrating the joint prior distribution of Ψ, a procedure that entails complex computation techniques. 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 contributions that work with the transformation between the SVAR and the impulse response parametrization (Kociecki, 2010, Arias et al., 2018).

Because of the way our prior is designed, a Minnesota-like prior selection of $\boldsymbol{\mu}$ emerges as a special case. In fact, the random walk prior, corresponding to $\delta_i = 1$ in equation (8), can be obtained by setting $\bar{\Psi}_h = \bar{\Psi}_0$, $h = 1, .., H$, $H = p$. 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 V is sufficiently small, the Minnesota-like selection given in (8) 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 also be formulated.

Our method enjoys the strengths and the limitations of all Bayesian analyses. It gives a new dimension along which to introduce shrinkage, namely the impulse responses, which are key in structural analyses. It goes without saying that if the researcher introduces tight beliefs that are inconsistent with the data, then the posterior may be driven away from the region where the most likely IRFs are present. As we stressed, our calibrations of V is in line with other works in the literature, including Bańbura et al. (2010). Hence our approach is not introducing prior beliefs that are necessarily more tight than those in standard applications. In addition, prior sensitivity can help to assess if the results are driven by the prior. For instance, one can evaluate whether the posterior variance is driven by posterior uncertainty on B or on $\tilde{\boldsymbol{\pi}}$, see Section 4.

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 of the former shock and capturing well the dynamics it induces. But a prior that is tailored to that purpose will not 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 a good forecasting performance, and this should not be considered a defect of the specification.

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 two 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 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.

FROM FABIO: "Add a paragraph here saying that although the prior on VAR coefficients is normal the prior on IRFs may be non-normal, but as the horizon increase the nonnormalities disapper - which i believe is a nice feature to have ", SEE COMMENT ABOVE. I AM NT SURE WHY WE CAN SAY THAT THE NONNORMALITY DISAPPEARS IN THE LONG RUN, LET'S DISCUSS WITH ANDRZEJ

It is also useful to draw a short comparison with the approaches of Villani (2009), Baumeister and Hamilton (2024) 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, given a prior for the constant, a prior specification for the VAR coefficients. Our approach works the other way around: we formulate prior beliefs on the VAR coefficients that represents certain prior beliefs on the IRFs. Baumeister and Hamilton (2024) also impose priors on a number of functions of the SVAR coefficients but they do this working directly with a SVAR. 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 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 out that it is usually hard to come up with genuine prior for VAR parameters, and standard 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 subjective prior beliefs for observables into a prior for VAR coefficients. Similarly, our approach starts from the premise that researchers are more comfortable with specifying prior beliefs on IRFs rather than VAR coefficients. Hence, our method can be seen as complementary to theirs and useful in different contexts.

Recently Giannone et al. (2019) have suggested a prior that effectively describes beliefs on the long run properties of the data. It turns out that our setup can recover their prior specification. To see this consider equation (17). The prior for the long run is simply a prior on the sum of the VAR coefficients. Thus, by appropriately choosing the elements of $\bar{\boldsymbol{\psi}}_F$ one can either a-priori impose stationarity or unit roots in the data. Note that while the prior of Giannone et al. (2019) is silent about the shape or the persistence of the implied IRFs, our version of the long run prior has built-in particular structures in the IRFs.

2.4 The specification of $(\bar{\Psi}, V)$ and an illustration of our prior

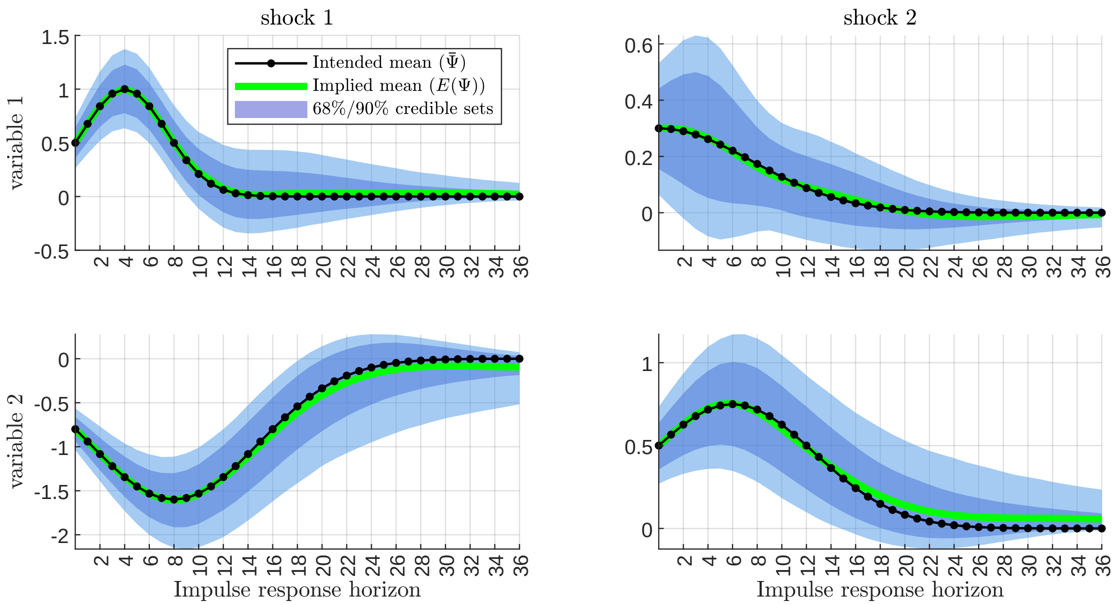
To make our approach operational, one needs to specify $\bar{\Psi}$. This can be a daunting task, due to the dimensionality of the matrix. To reduce the complexity of the problem, we found it convenient to model $\bar{\Psi}$ using the Gaussian basis functions:

$$\bar{\psi}_{ij,h} = a_{ij} \cdot e^{-\left(\frac{(h-b_{ij})^2}{c_{ij}^2}\right)+\frac{b_{ij}^2}{c_{ij}^2}}. \tag{20}$$

For each variable i and each shock j, this specification allows us to span $H+1$ dynamic responses with as few as three scalar parameters (a_{ij}, b_{ij}, c_{ij}). Here a_{ij} captures the impact effect of shock j on variable i, and thus regulates the (i, j) entry of $\bar{\Psi}_0$; b_{ij} is an integer, which pins down the horizon at which the peak effect is reached, and equals 0 if no hump-shaped response is desired. When $b_{ij} > 0$, c_{ij} can be chosen to control the (absolute value) of the size of the peak response relative to the size of the impact effect.

On the other hand, when $b_{ij} = 0$, c_{ij} can be chosen to control the horizon at which the impulse response reaches the half-life relative to the impact effect. Equation (20) re-parametrizes the function used in Barnichon and Matthes (2018) to ensure that the impact effect of shocks is a free parameter, as this is needed to match $\bar{\Psi}_0$ with $E(B)$.[3]

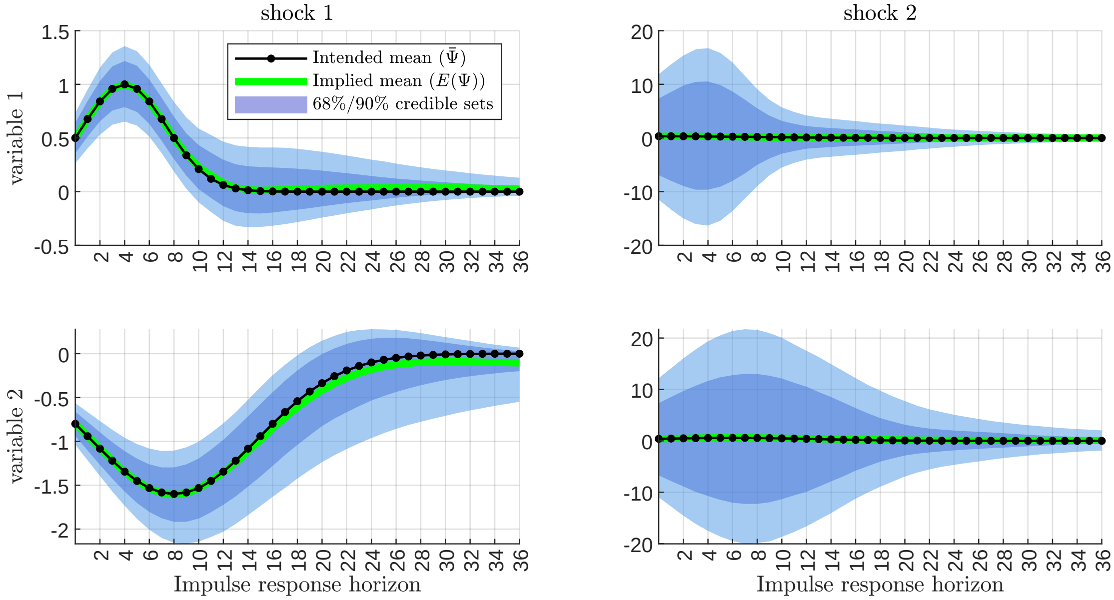
Figure 1: Illustration of our prior (1)



As explained, our approach provides an alternative specification for the prior mean of the VAR coefficients consistent with a-priori beliefs on the mean of the IRFs. It is, however, silent about how to set the covariance matrix of the VAR coefficients V (and thus on the prior uncertainty around the selected IRFs means). We found it convenient to specify V directly, FABIO, I GAVE IT A BETTER SHOCK ~~rather than deriving it as a function of $\bar{\Psi}$ as well as a prior variance on the impulse responses, see the Online Appendix~~ rather than derive V in order to imply a desired covariance structure specified on the IRFs. Selecting V directly ensures that we can achieve an amount of shrinkage comparable to the Minnesota-like prior, benefiting from

[3] If a prior is specified directly on B, $\bar{\Psi}_0$ can be treated as a free parameter and a_{ij} can be selected by the researcher. If a prior is specified on A or (Σ, Q), $\bar{\Psi}_0$ should be set equal to $E(B)$, which can be evaluated numerically. In this latter case, the researcher does not have a direct control on a_{ij}. As for c_{ij}, if $b_{ij} = 0$, we compute $c_{ij} = h_{hl,ij}/\sqrt{-\ln(2)}$ with $h_{hl,ij}$ being the horizon at which the IRF reaches its half-life relative to the impact effect, i.e. $a_{ij}/2$. If $b_{ij} > 0$, we compute $c_{ij} = b_{ij}/(\sqrt{\ln(\tilde{a}_{ij}/a_{ij})})$, with $\tilde{a}_{ij}$ capturing the maximum value of the response at horizon $h = b_{ij}$, with $\text{sign}(a_{ij}) = \text{sign}(\tilde{a}_{ij})$.

Figure 2: Illustration of our prior (2)



the same dimensionality reduction with a handful of hyperparameters, see equation (8). Section 3 in the Online Appendix discusses one way of selecting λ adaptively. Alternatively, with equation (19) at hand, one can also use hierarchical or Empirical Bayes approaches for setting the elements of V_s, as in Giannone et al. (2015). We do not follow a hierarchical approach here, because updating is generally driven by the forecasting performance of the model, which we view as conceptually separate from computing structural impulse analysis.

We conclude this section with an illustration of our prior. ~~We use a bivariate model ($k = 2$) and consider an IRF specified up to 36 horizons after the shock ($H = 36$)~~ THE FOLLOWING MIGHT BE BETTER We consider an IRF specified for 2 variables and 2 shocks up to 36 horizons after the shock ($k = 2$, $H = 36$). $\bar{\Psi}$ contains $k^2(H+1) = 148$ entries. With Gaussian basis functions the selection of 148 separate parameters is replaced with the selection of $3k^2 = 12$ hyperparameters. The black dotted line in Figure 1 gives an illustration of $\bar{\Psi}$ selected via the Gaussian Basis Function as of equation (20). As an example, the top left plot shows the case of a hump-shaped response that takes value 0.5 on impact. It then reaches the peak effect 4 horizons after the shock at a value that is twice as high as the impact effect, before progressively declining to zero. The top-right plot, instead, shows the case of a response that equals 0.5 on impact, and which features no hump-shaped response, with half of the impact effect reached 9 horizons after the shock. FABIO, YOU WERE ASKING ABOUT CASE a OR b. IN THIS PARAGRAPH IT'S JUST THE SELECTION OF $\bar{\Psi}$. WHETHER WE ARE IN CASE a b OR c ONLY MATTERS IN THE NEXT PARAGRAPH, WHERE WE NEED TO SPECIFY THE LENGTH OF THE VAR USED TO MODEL $\bar{\Psi}$. I ALSO REWROTE THE NEXT PARAGRAPH A BIT

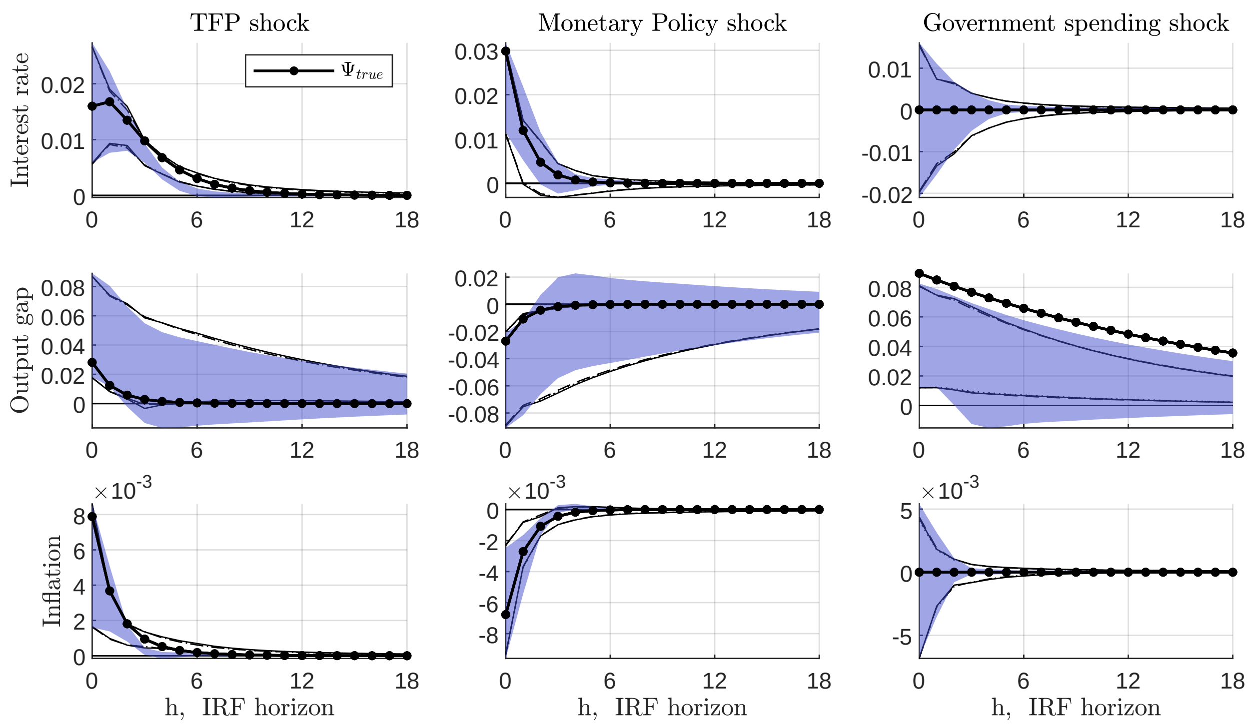
The rest of Figure 1 shows how our prior can capture $\bar{\Psi}$ in a SVAR model with p lags. We add no constant to the SVAR and set $p = 12$, which is only one third of the number of horizons of $\bar{\Psi}$ (*Case b*). We use prior $p(B, \boldsymbol{\pi}) = p(B) \cdot p(\boldsymbol{\pi})$, $\text{vec}(B) \sim N(\text{vec}(\bar{\Psi}_0), 0.02 \cdot I_{k^2})$ and $\boldsymbol{\pi} \sim N(\boldsymbol{\mu}, V)$, with $\boldsymbol{\mu}$ set as (18) given $\bar{\Psi}$, and $V = (0.01)^2 \cdot \text{diag}(1^{-2}, 2^{-2}, .., p^{-2}) \otimes I_{k^2}$. We draw 5,000 times from the prior on $(B, \boldsymbol{\pi})$ and compute the corresponding impulse responses. The figure reports the implied prior on the IRFs by showing the pointwise mean, as well as the 68% and 90% credible sets. As can be seen, the implied expected value of the IRF (shown in the green solid lines) tracks $\bar{\Psi}$ very well (black dotted line). This confirms that our prior can introduce shrinkage on the timing and shape of the impulse responses. Figure 2 complements the analysis by showing that our prior can be used on a subset of shocks and impulse responses. It modifies the specification behind Figure 1 by increasing the prior variance of the second column on B to 50. The figure shows that our prior can accommodate partial IRF beliefs.

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 of An and Schorfheide (2007). The model has three endogenous variables: the output gap, inflation, and the nominal interest rate, and their dynamics are driven by three structural shocks: a TFP shock, a monetary policy shock and a government spending shock. We fix the DSGE parameters using the posterior mean estimates obtained in An and Schorfheide (2007). The shocks are stationary and the model is solved in log deviation from the steady state. Thus, simulations are started from the steady state. We generate two datasets with 250 and 1,150 data points, we discard the first 100 observations, and use the next 50 as a training sample. This means that we have samples of T=1,000 or T=100 for inference.

We estimate four SVAR models, which are identical except for the prior specification for $\boldsymbol{\pi}$. We identify the shocks of the model using sign restrictions on the impact effect of the impulse responses. We restrict the impact effect of the impulse responses to feature the same sign as in the DGP, hence no misspecification in the identification of shocks is present. If the true impulse responses features a zero impact effect of the shock (as is the case for two variables in response to the government spending shock), we leave the corresponding entry of B unrestricted, rather than introduce a zero restriction. We use an inverse-Wishart-Uniform prior for (Σ, Q). All models feature a constant with flat prior centered at zero and include 4 lags. While adding $p > 1$ lags introduces misspecification relative to the data generating process, it helps the visual illustration of our prior, without affecting the results we present.

Figure 3: Posterior distribution, T=1,000

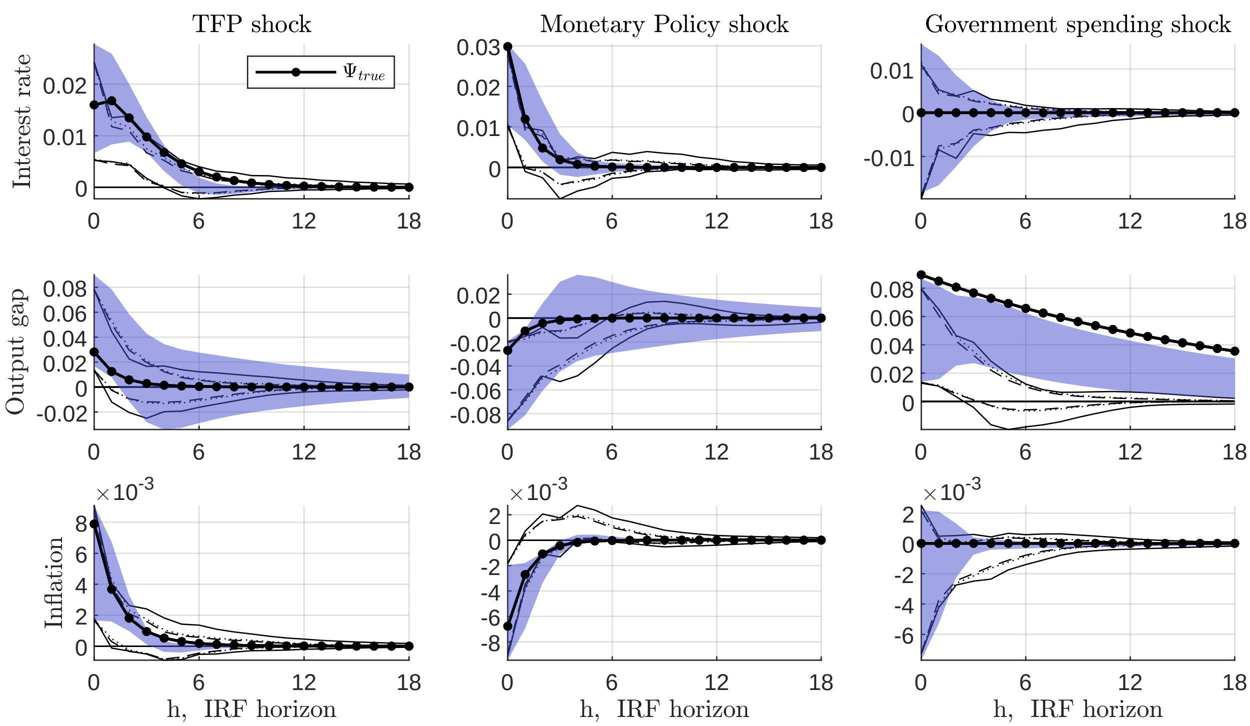


Note: Pointwise 68% credible sets corresponding to the flat prior (solid lines), to the white noise Minnesota prior (dashed lines), to the random walk Minnesota prior (dotted lines), and to our prior (blue shaded area). The dataset includes $T = 1,000$ observations. The responses correspond to a one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

The four models differ in the specification of $(\boldsymbol{\pi}, V)$. The first model uses a flat prior, setting $V^{-1} = 0$. The second and third models use a Minnesota-like prior for $\boldsymbol{\pi}$, setting δ to produce a white noise or a random walk. In both cases the covariance matrix V is set according to equation (8), choosing the hyperparameters as in Canova

(2007), which implies a relatively uninformative prior. In the fourth model we use our prior. We consider responses up to $M = 18$ horizons, and specify the prior mean $\bar{\Psi}$ up to $H = 4$ horizons. Because $H = p$, the setup coincides with *Case a* discussed in Section 2.2. We choose $\bar{\Psi}$ using the Gaussian basis functions as in equation (20), and set the parameters $\{(a_{ij}, b_{ij}, c_{ij})\}_{i=1,2,3;\ j=1,2,3}$ to ensure that $\bar{\Psi}$ approximates the true shape of the impulse responses in the DGP. We specify the variance V as from equation (8), and set λ adaptively, building on the work by Bruns and Piffer (2023). We refer to Section 3 in the Online Appendix for a detailed discussion of the specification.

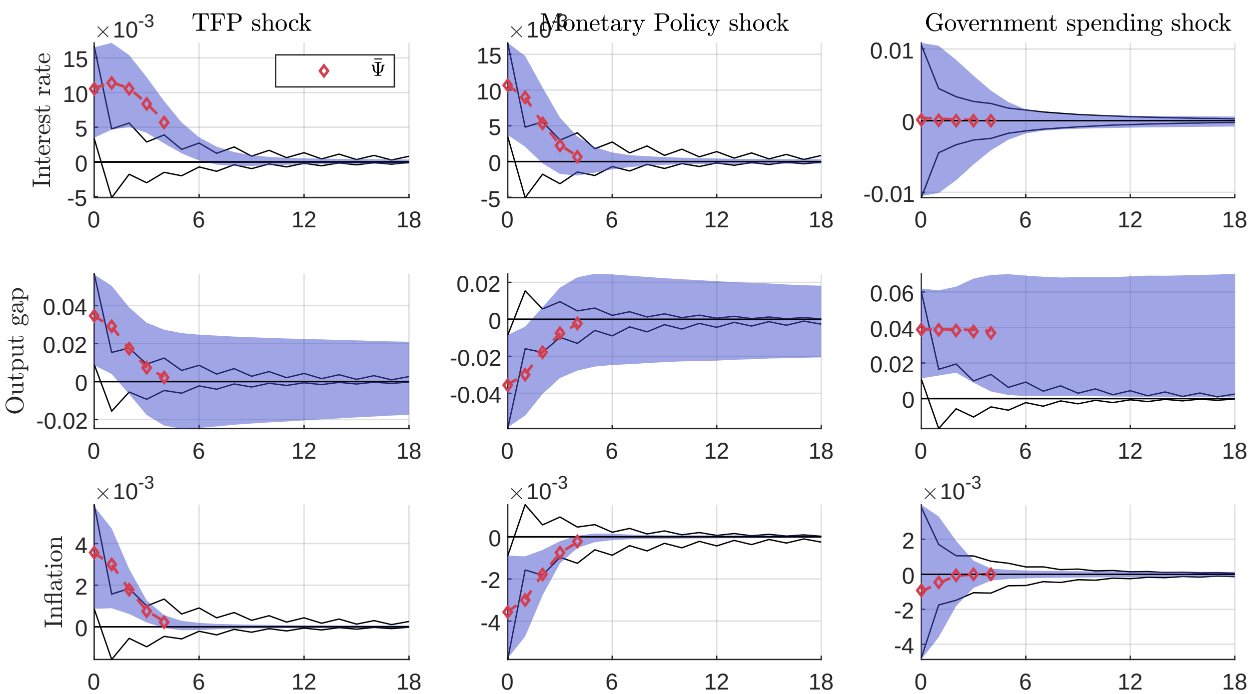
Figure 4: Posterior distribution, T=100



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 3 reports the pointwise 68% credible sets of the posterior IRFs associated with the four priors when T=1,000. The dotted line displays the true impulse responses. We focus the discussion on the response of the output gap to a government spending shock (i.e. the entry (2,3) of the figure), which is particularly persistent in the model. The posterior distribution $p(\Psi_0|Y)$ is similar across prior specifications, and the true value of the instantaneous response of the output gap to the government spending shock lies in the right tail of the distribution. Thus, all four prior selections lead to posterior distributions that tend to underestimate the impact effect of the government spending shock to the output gap. At longer horizons, no material differences emerge, and all four specifications correctly capture the strongly persistent nature of the true response. This result is driven by the fact that, in large samples, $\boldsymbol{\pi}$ is identified from the data. Hence, differences in prior beliefs will vanish asymptotically.

Figure 5: 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, setting $H = p = 4$.

The results differ when a smaller number of observations are used. Figure 4 shows that when $T = 100$, there are differences in the posterior IRFs associated with the flat prior and the 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. In particular, the mean half-life of the response associated with these three prior specifications is about 2-3 horizons while the true half life is 15 horizons. By contrast, because the selection of $\bar{\Psi}$ is informed by the data generating process, our prior leads to posterior IRFs that are more persistent and mimic the true ones.

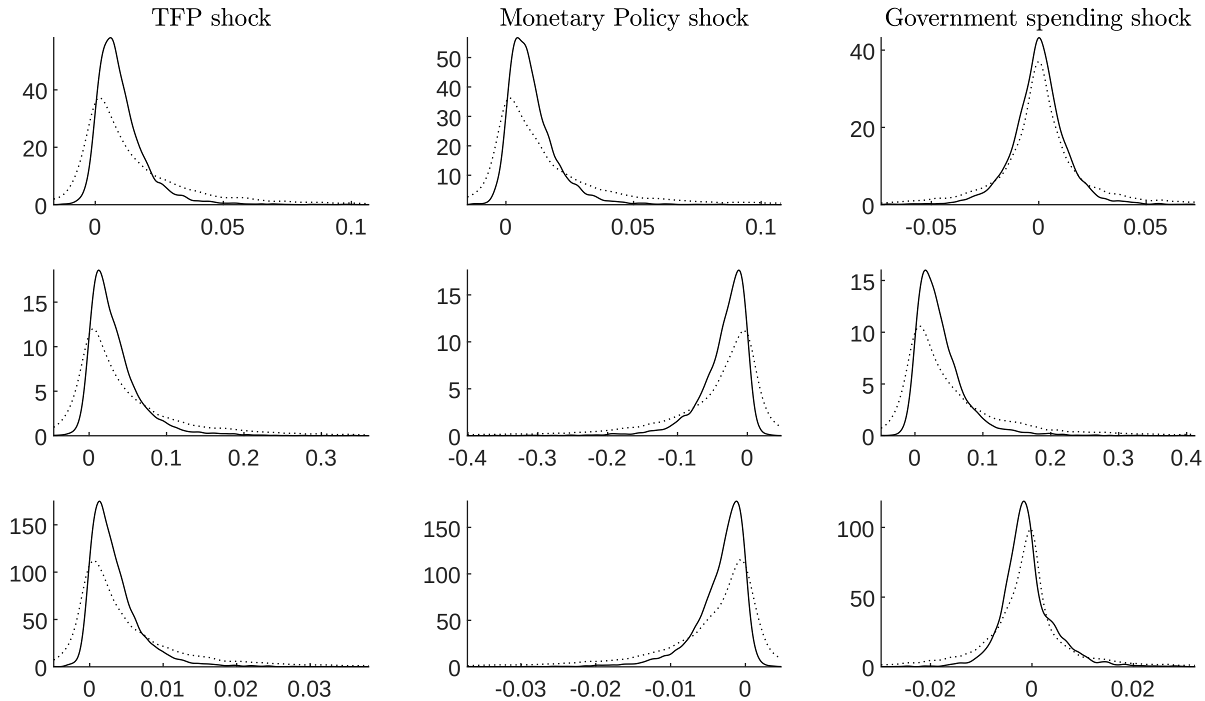
Figure 5 provides further information by exploring the prior distributions underlying Figure 4. The red diamonds show the values of $\bar{\Psi}$, which were specified up to $H = 4$ periods. The credible intervals shown in the shaded area are obtained by drawing from our joint prior distribution $p(\boldsymbol{\pi}, B)$ and computing the associated impulse responses up to horizon 18. The red dashed line shows the expected value of the impulse response for the horizons in which $\bar{\Psi}$ was specified. The figure confirms that it is possible to work with a Normal prior for $\boldsymbol{\pi}$ and to select its moments to imply a prior on the impulse responses centered around the desired trajectory. Note that the red line and red diamonds effectively coincide. Thus, the approximation error due to $V > 0$, equation (11b), is negligible. The grey dashed lines in Figure 5 show that the white noise Minnesota prior indirectly encourages no persistence in the responses. Assessing what the flat prior introduces on impulse responses is not possible under $V^{-1} = 0$. The results associated with the random walk Minnesota prior are not reported because the random walk prior implies nonstationarity, which in turn leads to off-scale credible bands.

The similarity documented by Figure 4 in the posterior distributions associated with the flat prior and two specifications of the Minnesota prior indicate that these priors are relatively uninformative, leading to the posterior distributions being strongly dependent on the sample estimate. Since the latter underestimates the true persistence of the output gap to the government spending shock by roughly 30%, mistakes emerge in the measurement of the true persistence of the response. Importantly, this is true regardless of whether a unit root or a white noise Minnesota prior are used as long as the priors are left sufficiently non-informative.

Figure 6 reports the marginal distribution of the prior on the impulse responses 2 and 4 periods after the shock implied by our selection. As clear from the figure, the marginal prior distribution on the impulse responses is not Normal, as it can be strongly skewed. As discussed in Section 2.3, our method provides control on the first moment of the impulse response without constraining the prior distribution on the impulse responses to be a Normal prior.FC WROTE "I am not sure why you included this here. it is tangential if the prior is normal or not. You can mention it and report it in the appendix, but i really do not see it here"

It is typically suggested that a relatively wide prior variance on the objects of interest should be used "to let the data speak". We find that this is not necessarily

Figure 6: Prior beliefs, marginal distributions

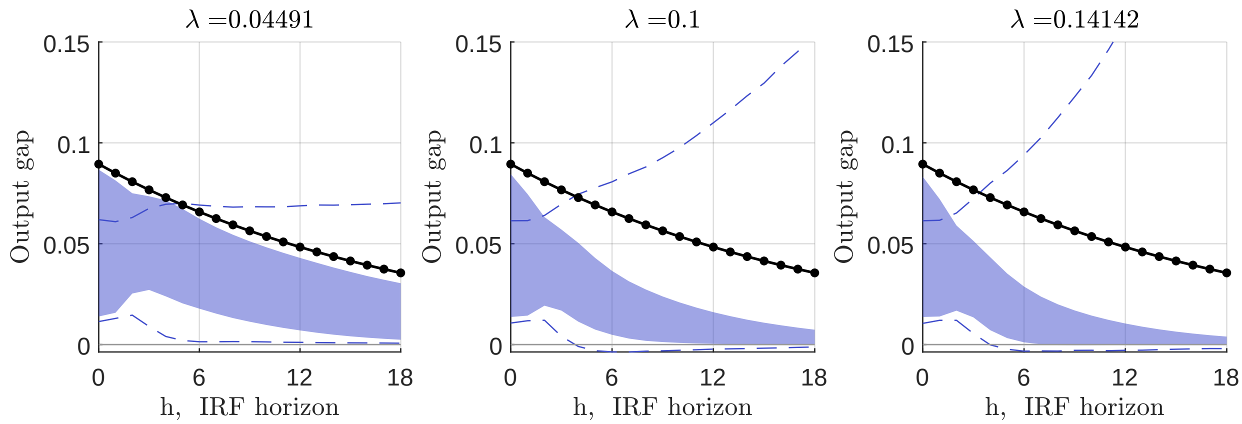


Note: Marginal distribution of the prior on the impulse responses 2 and 4 periods after the shock (continuous and dashed lines, respectively).

the case when performing structural analyses. As λ increases, the prior becomes less informative, but it also loses its informational shape content - approaching the other three priors. Thus, it becomes less useful when the persistence of a response is of interest. Figure 7 provides an illustration. When λ increases, the posterior IRFs obtained with our prior approach gets closer to the I ADDED A VERB, WHICH WAS MISSING posterior associated with the Minnesota and the flat priors. As a consequence, the posterior now fails to capture the true persistence effect when T=100.

So far the analysis has used a single dataset. We replicate the analysis over 100 datasets, holding the data generating process fixed and exploring the posterior distribution under our prior and under the white noise Minnesota prior. As shown in Section 5 of the Online Appendix, our approach consistently produces a response of the output gap to the spending shock that is more persistent than with the Minnesota prior, thanks to the explicit information modelled in the prior distribution. I REORGANIZED THIS PARAGRAPH, DELETED FIGURE 8 AND LEFT THE MORE

Figure 7: 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.

COMPREHENSIVE FIGURES IN THE APPENDIX

4 The output effects of monetary policy shocks

There is an extensive literature quantifying the effects of monetary policy shocks on the real economy (see Christiano et al., 1999, Antolín-Díaz and Rubio-Ramírez, 2018 and Miranda-Agrippino and Ricco, 2021 among many others). One key question often discussed in the literature is whether monetary policy surprises generate their strongest effects on impact, or whether long and variable lags imply that the largest response is delayed (Buda et al., 2023). The question is relevant in the policy debate, given that it directly informs central bankers about their ability to quickly stimulate/contract the real economy if needed.

We use our prior to study how long it takes for a US monetary policy surprise 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 list of variables is common and is consistent with the work of, e.g., Bernanke and

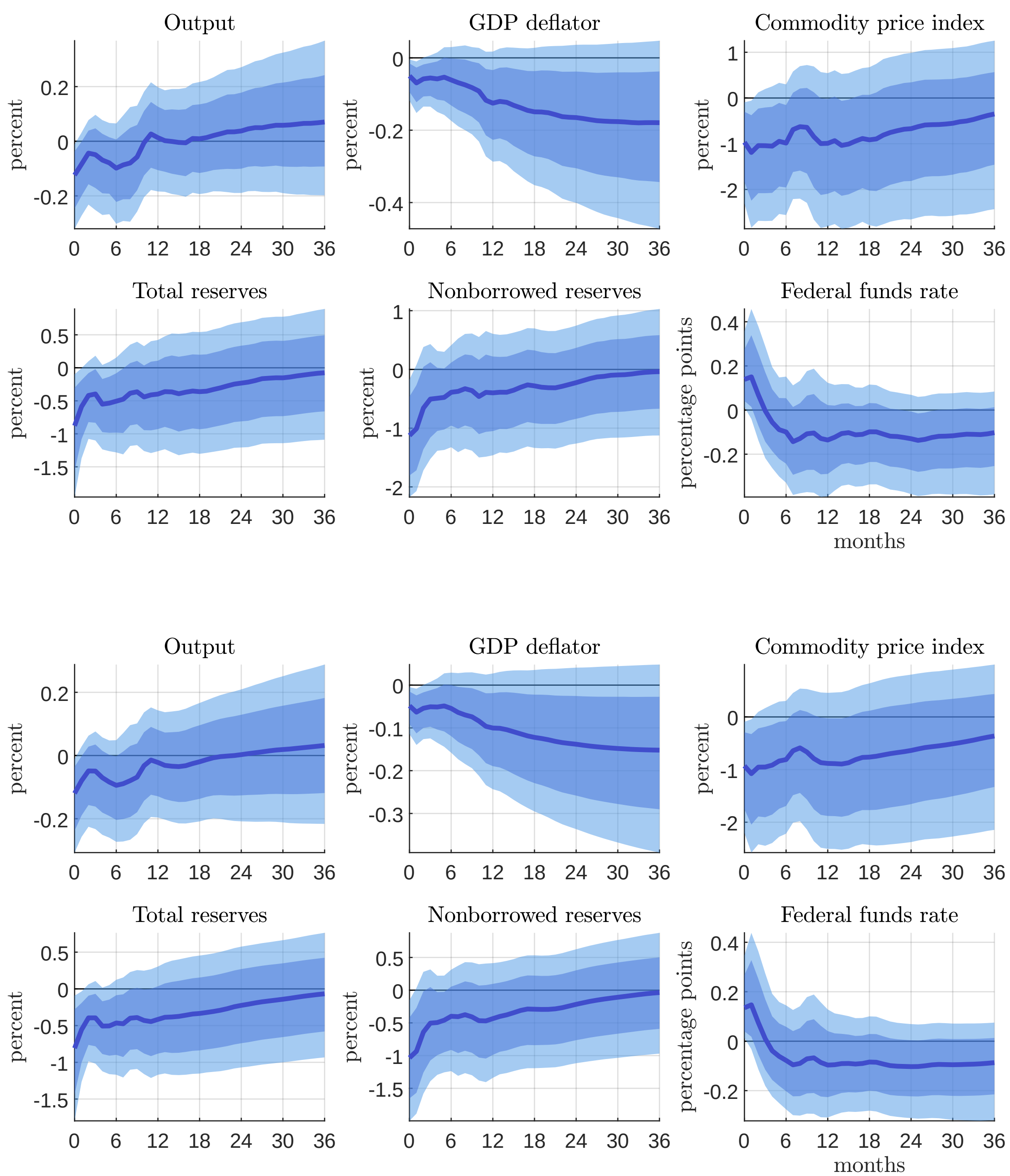
Mihov (1998), Uhlig (2005) and Arias et al. (2019). Following Arias et al. (2019), we use the sample 1965M1 through 2007M6 in the analysis, and estimate a VAR with 12 lags and a constant.

We identify the monetary policy shock via sign restrictions. We assume that an increase in the policy rate decreases all the other variables on impact. Thus, in particular, real GDP falls 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. Note that because of the way we identify the monetary policy shocks, the point raised by Wolf (2020) does not apply. We use an inverse-Wishart-Uniform distribution on (Σ, Q), estimating its parameters in the training sample (the first 20% of the observations), as suggested by Kadiyala and Karlsson (1997). Given that the contribution of our paper is the prior specification of $\boldsymbol{\pi}$ rather than the one of (Σ, Q) or B, we do not further explore alternative prior distributions to introduce prior beliefs on the impact effect of the shocks (see Inoue and Kilian, 2020 for a discussion on the importance of the prior distributions on these objects in structural analyses).

Figure 8 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 $\delta = 1$ and setting the variance V_s as in Kadiyala and Karlsson (1997). We report the pointwise median response together with the 68% and the 90% credible sets for illustration - the responses obtained using the posterior draw closest to the median response are very similar. 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. Furthermore, the IRFs revert back to zero within less than a year and display no hump.

Do the conclusions change if one introduces prior beliefs on the impulse responses directly? We express the prior view that monetary policy shocks generate persistent but temporary effects, in the sense that responses are expected to revert to zero in the medium term. This view is consistent with a large amount of empirical evidence. We stress two important facts: first the beliefs are not imposed dogmatically, in the sense that they can be updated by the data if it wants to do so. Second, we do not assume that the response of any of the variables is humped-shaped to avoid, even indirectly, to lead the conclusions in this direction. Finally, as with the other two specifications, we use a flat prior for the constant, centered at zero.

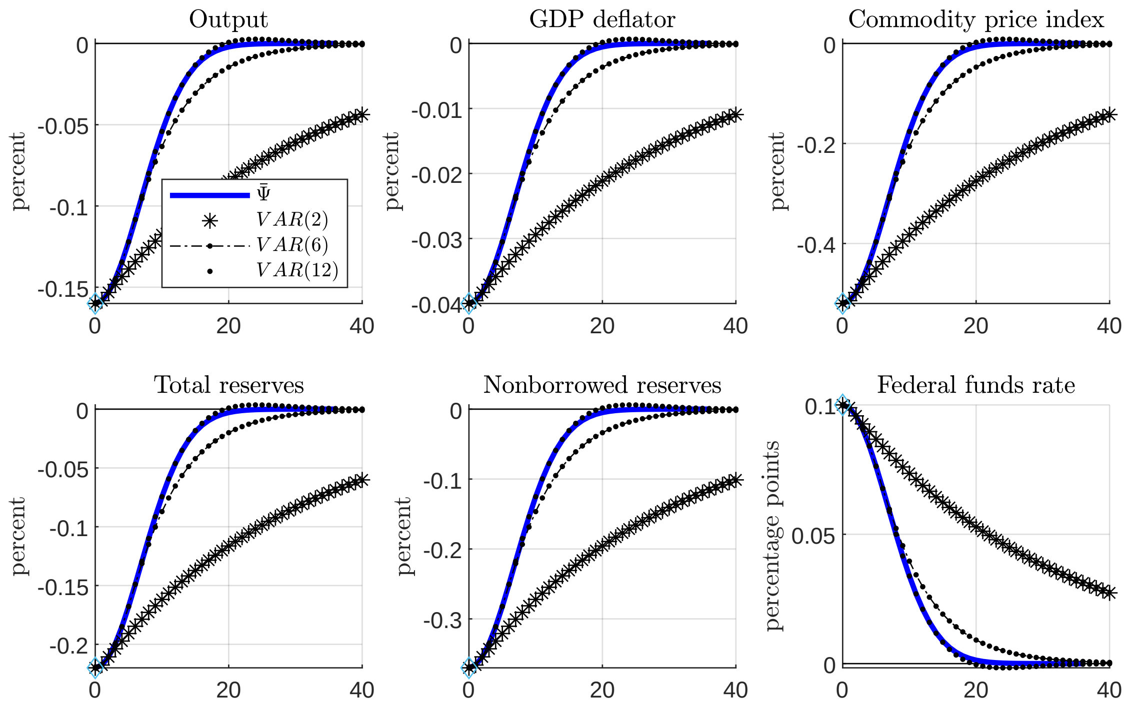
Figure 8: Posterior IRFs for the Flat (top) and Minnesota (bottom) priors



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

Our prior beliefs are made operational as follows. For all variables, we specify $\bar{\Psi}$ using (20), setting $H = 36$. We choose $b_{ij} = 0$, $\forall\ i, j$. In the baseline specification we

Figure 9: Our approach: $\bar{\Psi}$

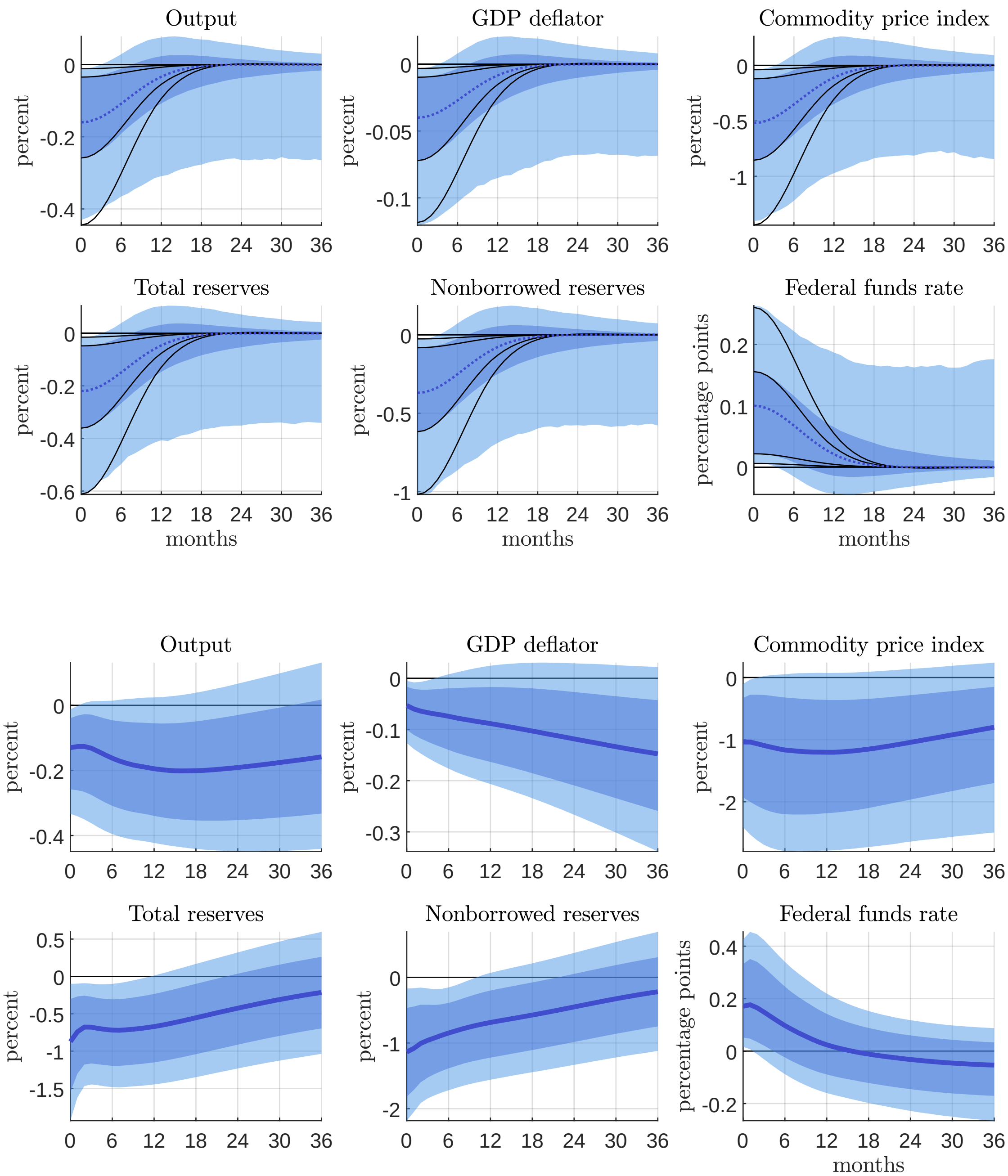


Note: The thick blue line reports $\bar{\Psi}$. The black dashed and dotted lines show the value of the impulse responses associated with a VAR model parametrized as implied from equation (9) using up to $l = 2, 6, 12$ lags of $\bar{\Psi}$.

set c_{ij} to imply that it takes 8 months for the effects of the monetary shock to reach half of the impact response. For convenience, this degree of persistence is introduced equally across all variables and in response to all shocks, although the methodology does not require such a restriction. As for the covariance matrix, we choose V as in equation (19). V_s is set as with the Minnesota prior, and we select λ adaptively. We start the algorithm at $\lambda = 0.005$ and increase it until between 5% and 10% of the joint prior distribution associated with the monetary policy shock is outside of $\pm 2 \cdot \sqrt{\hat{\Sigma}_{ii}}$ for the first $H = 36$ horizons, where $\sqrt{\hat{\Sigma}_{ii}}$ is estimated in the training sample (see the discussion in Section 3 in the Online Appendix). The search produces a value of $\lambda = 0.0331$, which is comparable to values used for the Minnesota prior (Bańbura et al., 2010).

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 (11b) do not hold exactly (this is *Case c* in Section 2.2). Figure 9 helps to assess to what extent a VAR

Figure 10: Our approach: prior (top panel) and posterior (bottom panel) IRFs



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.

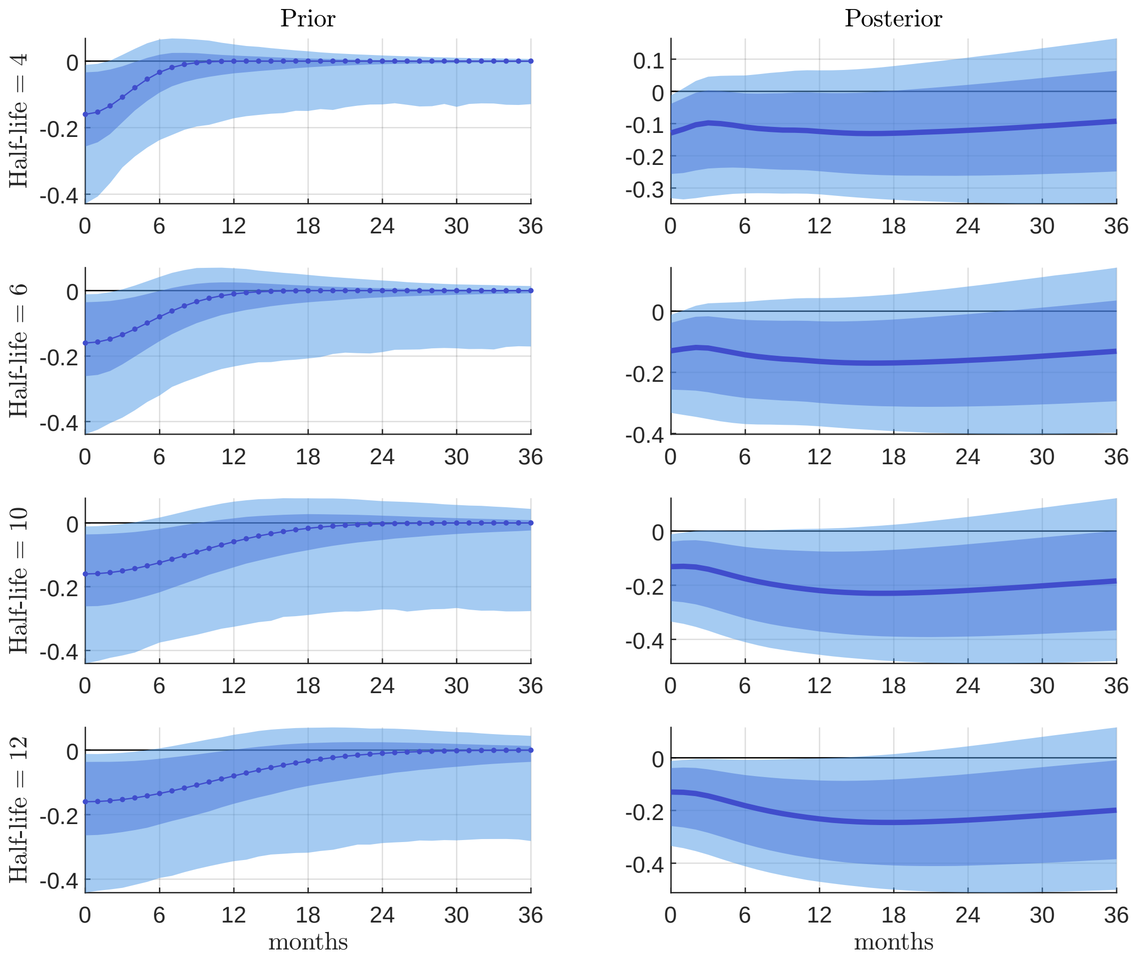
with $p < 36$ lags can match a set of prior impulse responses specified up to horizon 36. For $p = 2, 6, 12$, it shows the impulse responses associated with $\bar{\Pi} = [\Pi_1, .., \Pi_p]$, which are computed via equation (9) using up to entry $\bar{\Psi}_p$ of $\bar{\Psi}$. Clearly, a VAR with 6 lags does a good job in approximating $\bar{\Psi}$, and $p = 12$ implies negligible approximation errors.

The top panel of Figure 10 shows the prior impulse responses to a one standard deviation shock associated with our prior specification. The dotted lines is $\bar{\Psi}$. The shaded areas report the 68% and the 90% credible sets associated with the prior for $\boldsymbol{\pi}$. Indeed the prior distribution is well centered around $\bar{\Psi}$ and does not induce any hump-shaped dynamics. The solid black lines report the 68% and 90% credible sets obtained when setting $\lambda = 0$. This dogmatic specification of the prior on $\boldsymbol{\pi}$ helps to assess how informative is our calibration of λ, and to assess the relative role of prior uncertainty on $(B, \boldsymbol{\pi})$ in determining prior uncertainty on the impulse responses. As expected, prior uncertainty at short horizons is largely driven by prior uncertainty from $p(B)$, while the prior variance of $\boldsymbol{\pi}$ becomes important at longer horizons.

The bottom panel of Figure 10 reports the posterior IRFs. A 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 of about 1 percent and an impact decrease in output by 0.12 percent. Over time, 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 about 0.2 percent, which materializes one year after the shock. Thus, contrary to what we obtained in Figure 8, real GDP does not display the strongest response on impact. We nevertheless stress that posterior uncertainty remains wide, in a way that is quantitatively similar to the one obtained when using the Minnesota or the flat priors.

Next, we assess the sensitivity of the results to alternative prior specifications on the timing of the impulse responses. The baseline specification of our prior set the half-life of the responses equal to 8 months. Figure 11 shows that the results are robust to alternative selections of $\bar{\Psi}$ that imply a half-life of the effects equal to 4, 6, 10 or 12 months. All prior beliefs lead to the conclusion 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 domestic output.

Figure 11: Response of real GDP (robustness)



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

One may wonder why the posterior median responses have wiggles with the Minnesota and the flat prior while this is not the case with our prior. It turns out that while the former two specifications allow for complex posterior AR roots, this is less of a case with our prior. Thus, the hump more clearly emerges.

In the context of this application to monetary policy shocks, we have also explored whether our prior has implications on the computational cost of introducing identifying restrictions not only on the impact effect of the shocks (as in the baseline analysis

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 are progressively introduced up to horizon $h = 0, .., 4$ and the computation times. All codes are run on Matlab on a computer with an Intel i7-7700K 4.2GHz Quad Core processor and 64 GB RAM.

shown above), but also at future horizons. Table 1 reports how many posterior draws are needed to store 100,000 draws that satisfied sign restrictions progressively introduced at horizon 0 (baseline analysis), at horizons 0 and 1, and so on up to horizon 4. The sign restrictions introduced at future horizons are the same as the ones introduced on impact. When identifying sign restrictions are imposed only on impact, the computational cost is approximately the same with the Minnesota and with our priors. When sign restrictions are also introduced at higher horizons it becomes computationally more demanding to impose IRFs restrictions with the Minnesota prior. By contrast, the computational cost under our prior barely changes. Thus, our prior choice can significantly reduce the computational burden of imposing meaningful IRFs restrictions.

Last, we have used this application to explore what our prior distribution implies for the stationarity of the model, relative to the random walk Minnesota prior. We found that 70% of the posterior draws associated with the random walk Minnesota prior lead to a companion form with the maximum eigenvalue (in absolute value) above 1. This number decreases to between 30% and 44% with our prior, depending on the half-life value used. Hence, our approach can help introduce the prior belief that the model is stationary.

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 VAR coefficients that achieves two goals. First, it allows for an explicit introduction of prior beliefs on the shape or the persistence of the impulse responses. Second, it does so by working with a Normal prior distribution which ensures tractable posterior sampling.

We illustrate the properties of the methodology using simulated data from a small scale DSGE model and 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 feature 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 depending on the prior persistence we assume.

References

Amir-Ahmadi, P. and Drautzburg, T. (2021), 'Identification and inference with ranking restrictions', *Quantitative Economics* **12**(1), 1–39.

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).

Antolín-Díaz, J. and Rubio-Ramírez, J. F. (2018), 'Narrative sign restrictions for SVARs', *American Economic Review* **108**(10), 2802–29.

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.

Arias, J. E., Rubio-Ramírez, J. F. and Waggoner, D. F. (2024), 'Uniform priors for impulse responses', *CEPR working paper 18836* .

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. (2024), 'Advances in using vector autoregressions to estimate structural magnitudes', *Econometric Theory* **40**(3), 472–510.

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 Iacoviello, M. (2022), ‘Measuring geopolitical risk’, *American Economic Review* **112**(4), 1194–1225.

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 Pappa, E. (2011), ‘Fiscal policy, pricing frictions and monetary accommodation’, *Economic Policy* **26**(68), 555–598.

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.

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.

Giannone, D., Lenza, M. and Primiceri, G. E. (2019), 'Priors for the long run', *Journal of the American Statistical Association* **114**(526), 565–580.

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.

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.

Wolf, C. K. (2020), 'Svar (mis) identification and the real effects of monetary policy shocks', *American Economic Journal: Macroeconomics* **12**(4), 1–32.