Federal Reserve
Bank of Dallas

The Role of the Prior in Estimating VAR Models with Sign Restrictions

Atsushi Inoue and Lutz Kilian

Working Paper 2030 **December 2020**

Research Department

https://doi.org/10.24149/wp2030

The Role of the Prior in Estimating VAR Models with Sign Restrictions*

Atsushi Inoue* and Lutz Kilian†

First Draft: May 24, 2020
This version: November 29, 2020

Abstract

Several recent studies have expressed concern that the Haar prior typically imposed in estimating sign-identified VAR models may be unintentionally informative about the implied prior for the structural impulse responses. This question is indeed important, but we show that the tools that have been used in the literature to illustrate this potential problem are invalid. Specifically, we show that it does not make sense from a Bayesian point of view to characterize the impulse response prior based on the distribution of the impulse responses conditional on the maximum likelihood estimator of the reduced-form parameters, since the prior does not, in general, depend on the data. We illustrate that this approach tends to produce highly misleading estimates of the impulse response priors. We formally derive the correct impulse response prior distribution and show that there is no evidence that typical sign-identified VAR models estimated using conventional priors tend to imply unintentionally informative priors for the impulse response vector or that the corresponding posterior is dominated by the prior. Our evidence suggests that concerns about the Haar prior for the rotation matrix have been greatly overstated and that alternative estimation methods are not required in typical applications. Finally, we demonstrate that the alternative Bayesian approach to estimating sign-identified VAR models proposed by Baumeister and Hamilton (2015) suffers from exactly the same conceptual shortcoming as the conventional approach. We illustrate that this alternative approach may imply highly economically implausible impulse response priors.

JEL code: C22, C32, C52, E31, Q43.

Keywords: Prior, posterior, impulse response, loss function, joint inference, absolute loss, median.

* Atsushi Inoue, Vanderbilt University, Department of Economics, Nashville, TN 37235-1819. E-mail: atsushi.inoue@vanderbilt.edu.

† Lutz Kilian, Federal Reserve Bank of Dallas, Research Department, 2200 N. Pearl St., Dallas, TX 75201, USA. E-mail: lkilian2019@gmail.com (corresponding author).

1 Introduction

The conventional approach to estimating sign-identified VAR models, as discussed in Uhlig (2005), Rubio-Ramirez, Waggoner and Zha (2010), Arias, Rubio-Ramirez and Waggoner (2018) and Antolin-Diaz and Rubio-Ramirez (2018), involves specifying a Haar prior for the orthogonal rotation matrix Q and a Gaussian-inverse Wishart prior for the parameters A and Σ of the reduced-form VAR model, where A denotes the slope parameters and Σ is the error covariance matrix. The prior for the impulse response vector $\theta = g(A, \Sigma, Q)$ is defined implicitly. A number of recent studies have questioned the extent to which the impulse response estimates from these models are driven by the choice of the prior for Q, given that Q does not enter the likelihood and its prior cannot be overruled by the data (see, e.g., Baumeister and Hamilton 2018, 2019, 2020; Giacomini and Kitagawa 2020; Plagborg-Møller 2019; Watson 2020). Several of these studies have argued for ignoring empirical evidence obtained using the conventional approach because they see no reason for the posterior impulse response estimates and the credible sets reported in applied work to be more plausible than the other responses in the identified set.

This view is based on analysis in Baumeister and Hamilton (2015) who claimed that the Haar prior typically imposed for Q is unintentionally informative about the implied prior for the structural impulse responses. The question raised by Baumeister and Hamilton is indeed an important question for applied work, but we show that the tools they (and several subsequent studies) used to illustrate this problem are invalid. Specifically, Baumeister and Hamilton (2015) proposed characterizing the impulse response prior based on the distribution of the impulse responses conditional on the maximum likelihood estimator (MLE) of the reduced-form parameters. We show that this approach does not make sense from a Bayesian point of view. This point is self-evident because the prior in Bayesian analysis does not, in general, depend on the data. Since the impulse response distribution conditional on the MLE depends on the data by construction, it cannot be the prior.

We formally show that the distribution derived in Baumeister and Hamilton (2015) is not the impulse response prior distribution implied by conventional priors for the model parameters. The flaw in Baumeister and Hamilton's analysis is that the prior for a given impulse response induced by the prior for the rotation matrix and by the sign restrictions also depends on the priors for the reduced-form parameters, which affects the location and shape of the impulse response prior distribution. We show by example that their approach tends to produce highly misleading estimates of the impulse response priors.

This result invalidates the recent evidence against conventional priors for sign-identified VAR models, but leaves unanswered the original question of how widespread unintentionally informative impulse response priors are in applied work and, more importantly, to what extent these implicit impulse response priors affect posterior inference. In this paper, we develop new tools that allow us to examine these questions. We illustrate the use of these tools using examples drawn from the empirical literature. Our results provide a striking contrast to views prevailing in some of the recent literature.

There are four main results. First, we show that there is no evidence that typical sign-identified VAR models estimated using conventional priors are unintentionally informative for the impulse response vector θ. Instead, in models based on static sign restrictions, standard uniform-Gaussian-inverse Wishart priors tend to imply an uninformative prior for θ (in the sense that the impulse response prior in the absence of sign restrictions is centered near zero and is fairly diffuse). This result is general in that it depends only on the prior and not on the data.

Second, the corresponding impulse response posterior for θ is driven largely by the data. This result is consistent with new evidence that, unlike in the stylized example in Baumeister and Hamilton (2015) which implicitly restricted attention to models with a single sign restriction, the extent of the uncertainty about the impulse responses attributable to Q tends to be much smaller in models with multiple sign restrictions. Hence, the prior uncertainty about θ is dominated by the uncertainty about the reduced-form parameters, which is updated based on the data. Our evidence suggests that concerns about the Haar prior for Q have been greatly overstated and that alternative estimation methods are not required in typical applications.

Third, we observe that in models with both static and dynamic sign restrictions the implied prior for θ is necessarily informative. We illustrate by example that even in the latter class models, however, the prior for θ need not be unintentionally informative. Moreover, we show that in this case as well the posterior is driven largely by the data, not by the prior.

Fourth, we demonstrate that the alternative Bayesian approach to estimating sign-identified VAR models introduced in Baumeister and Hamilton (2015), which was intended to avoid the problem of unintentionally informative impulse response priors, suffers from exactly the same conceptual shortcoming as the conventional approach. Based on the oil market model of Baumeister and Hamilton (2019), we illustrate that this alternative approach, which has been used in several recent studies, may imply highly economically implausible impulse response priors.

The remainder of the paper is organized as follows. In Section 2, we derive the joint prior distribution of the structural impulse responses in sign-identified models under the conventional conjugate prior specification. We show that this distribution depends not only on the prior specification for the rotation matrix Q, but also on the prior for the reduced-form parameters. As a result, the posterior of the structural impulse responses is updated based on the data at all horizons of the response function, even though the prior distribution of Q is not. We formally show that the distribution of the impulse responses conditional on the MLE, as reported by Baumeister and Hamilton (2015), is not the impulse response prior. Thus, the approach of plotting the histogram over the identified set of the impulse responses conditional on the MLE of the reduced-form parameters is not informative about the prior distribution for θ nor does it tell us to what extent the posterior distribution of θ depends on the prior. We illustrate this point in the context of a typical empirical application.

We also refute the prevailing view that the uncertainty about Q necessarily results in wide identified sets for the structural responses, even controlling for estimation uncertainty in the reduced-form parameters. This view has been based on a stylized example discussed in Baumeister and Hamilton (2015). Their results, however, are driven by the implicit assumption that there is only one sign restriction in the VAR model. In the presence of multiple sign restrictions, as is typical in applied work, the extent of the uncertainty about the impulse responses attributable to uncertainty about Q tends to be much smaller, mitigating concerns about the Bayes estimators of θ being sensitive to the assumptions about the prior for Q and largely obviating the need for alternative econometric approaches and explaining why the posterior of θ is dominated by the data.

In section 3, we address the question of how to answer the questions raised in Baumeister and Hamilton (2015). We provide simple diagnostic tools that help make explicit the prior on the joint distribution of the structural impulse responses implied by the conventional uniform-Gaussian-inverse Wishart prior (see, e.g., Uhlig 2005; Waggoner, Rubio-Ramirez and Zha 2010; Arias, Rubio-Ramirez and Waggoner 2018, Antolin-Diaz and Rubio-Ramirez 2018). These tools may be used to choose between economically plausible and implausible impulse response prior specifications. We also illustrate how applied users can assess the extent to which the joint posterior distribution of θ is driven by the data as opposed to the prior specification for θ. Our proposal is to compare the joint distribution of the structural impulse responses obtained by drawing from the prior distribution of the model parameters to the joint posterior distribution of the structural impulse responses. We provide metrics that help quantify the extent to which the prior location and dispersion of the joint impulse response distribution is updated by the data.

In section 4, we illustrate the use of these tools in a typical sign-identfied VAR model and show that, given a prior mean of zero for the slope parameters, a conventional uniform-Gaussian-inverse Wishart prior for the VAR model parameters implies an uninformative prior for θ, defined as a prior for the impulse response vector that in the absence of sign restrictions is centered near zero and is fairly diffuse. This result is general in that the prior of the impulse response does not depend on the data. We also show that the posterior of θ is largely driven by the data rather than the prior for Q, contradicting the unsubstantiated claim in Baumeister and Hamilton (2015) that inference based on the impulse response posterior is spurious.

Whereas section 4 focuses on structural VAR models based on static sign restrictions, section 5 examines an empirical example that involves both static and dynamic sign restrictions, which renders the prior of θ necessarily informative. We show that even in the latter case, the conventional uniform-Gaussian-inverse Wishart prior need not imply unintentionally informative impulse response priors, further illustrating the use of the tools developed in section 3. We also confirm that, as in section 4, the posterior is dominated by the data rather than the prior.

In section 6, we show that the alternative Bayesian approach proposed by Baumeister and Hamilton (2015) as a solution to the problem of unintentionally informative impulse response priors suffers from exactly the same conceptual shortcoming as the conventional approach. Their central idea is that we can define priors for the parameters of the structural VAR representation, avoiding the use of rotation matrices and of the Haar prior. The structural impulse responses of interest are then defined as a nonlinear transformation of the parameters of the structural VAR representation. The joint prior distribution of these impulse responses, however, is never derived. We observe that, as in the conventional approach, nothing prevents the implied prior for θ from being unintentionally informative. Our approach to deriving the joint prior for θ accommodates the framework of Baumeister and Hamilton (2015) and may be used to evaluate the prior for θ implied by their alternative prior specification. We illustrate this point based on the global oil market model of Baumeister and Hamilton (2019). We show that, unlike in the conventional approach, there is strong evidence that the implied prior for θ is economically implausible. The concluding remarks are in Section 7.

2 The impulse response prior

In line with many other studies, our premise is that the primary object of interest in structural VAR analysis is the structural impulse response vector θ obtained by stacking the structural responses of interest. Consider the n-dimensional structural VAR(p) model

$$B_0 y_t = B_1 y_{t-1} + ... + B_p y_{t-p} + w_t,$$

where the intercept has been suppressed for expositional purposes, B_i, $i = 0, ..., p$, are $n \times n$ coefficient matrices, and the structural errors, w_t, is mean zero mutually uncorrelated Gaussian white noise. Without loss of generality, we impose the normalization that w_t has the covariance matrix I_n. The corresponding reduced-form VAR(p) representation is

$$y_t = A_1 y_{t-1} + ... + A_p y_{t-p} + u_t,$$

where $A_i = B_0^{-1} B_i$, $i = 1, ..., p$, and $u_t = B_0^{-1} w_t$. The variance-covariance matrix of the reduced-form error, u_t, is $\Sigma = B_0^{-1} B_0^{-1\prime}$. Identification is typically achieved by imposing restrictions on B_0^{-1}. In models identified based on exclusion restrictions, the vector of structural impulse responses can be written as $\theta = g(A, \Sigma)$, where $A = [A_1, ..., A_p]$ and $g(\cdot)$ denotes a nonlinear function. In the case of sign-identified models, $B_0^{-1} = PQ$, where P is the lower triangular Cholesky decomposition of Σ with positive elements on the diagonal and Q denotes an $n \times n$ orthogonal matrix such $QQ' = I_n$. Hence, $\theta = g(A, \Sigma, Q)$, where $g(\cdot)$ denotes a nonlinear function. For details of the specification, estimation and identification of this class of models the reader is referred to Kilian and Lütkepohl (2017).

2.1 The conventional uniform-Gaussian inverse Wishart prior

The conventional approach to estimating sign-identified VAR models is to postulate an inverse Wishart prior for Σ and a Gaussian prior for $[\nu, A_1, ..., A_p]$ conditional on Σ, combined with an independent Haar prior for Q, which may be viewed as a uniform prior in the space of possible rotation matrices. Many applications of sign-identified models rely on a uniform-diffuse Gaussian-inverse Wishart prior (see, e.g., Uhlig 2005). Although such diffuse priors imply a uniform-Gaussian-inverse Wishart posterior, they are not proper priors, making it impossible to compute the summary statistics proposed in Section 3. For expository purposes, we therefore work with a version of the popular conjugate reduced-form Gaussian-inverse Wishart prior proposed by Karlsson (2013), as specified by Antolin-Diaz and Rubio-Ramirez (2018). This so-called Minnesota prior involves a linear decay of the prior standard deviation at higher lags, sets the relative tightness of other variables in a given equation to 1, and sets the prior standard deviation of the first own lag in each equation to 0.2. The prior variance is scaled based on estimates of the innovation variance obtained from fitting univariate AR(1) models to each variable.

Given the Haar prior for Q, as originally proposed by Uhlig (2005) and Rubio-Ramirez, Waggoner and Zha (2010), these assumptions imply a uniform-Gaussian-inverse Wishart prior for structural VAR models identified by sign restrictions. It should be noted that our approach could be adapted to other proper priors, as illustrated in Section 6. While one could entertain other prior specifications, our objective for now is to follow as closely as possible current practice, as discussed in Uhlig (2005), Rubio-Ramirez et al. (2010), Arias et al. (2018) and Antolin-Diaz and Rubio-Ramirez (2018).

2.2 The implied prior for θ in sign-identified models

The conventional approach is for the user to specify a joint prior for the parameters Q, A, and Σ, without examining the implied prior for the vector of the structural impulse responses, $\theta = g(A, \Sigma, Q)$. Next, we derive this implied prior distribution by the change-of-variable method, building on Inoue and Kilian (2019). For expository purposes we abstract from narrative inequality restrictions, as discussed in Antolin-Diaz and Rubio-Ramirez (2018), and from combinations of sign and zero restrictions, as discussed in Arias et al. (2018). Given the rotation matrix Q, define the $n \times n$ skew-symmetric matrix S by

$$S \;=\; I_n - 2(I_n + Q)^{-1}, \tag{1}$$

and let s be the $\frac{n(n-1)}{2} \times 1$ vector that consists of the below-diagonal elements of S. When Q is uniformly distributed over the space of $n \times n$ real matrices such that $Q'Q = I_n$ and $|Q| = 1$, the density of s is given by

$$f(s) \;=\; \left(\Pi_{i=2}^{n} \frac{\Gamma\left(\frac{i}{2}\right)}{\pi^{\frac{i}{2}}}\right) \frac{2^{\frac{(n-1)(n-2)}{2}}}{|I_n + Q|^{n-1}}, \tag{2}$$

(see equation 4 in León et al., 2006, p. 415).

Let D_n denote the $n^2 \times \frac{(n(n+1)}{2}$ duplication matrix of zeros and ones such that $\text{vec}(M) = D_n \text{vech}(M)$ for any $n \times n$ symmetric matrix M (see Definition 4.1 in Magnus (1988), p. 55). Let D_n^+ denote the Moore-Penrose inverse of D_n so that $\text{vech}(M) = D_n^+ \text{vec}(M)$ (see Definition 6.1 in Magnus (1988), p.94). K_n is the $n^2 \times n^2$ commutation matrix such that $\text{vec}(M') = K_n \text{vec}(M)$ for any $n \times n$ matrix M (see Magnus and Neudecker (1999), pp. 46–47). L_n is the $\frac{n(n+1)}{2} \times n^2$ elimination matrix of zeros and ones such that $vec(M) = L_n' vech(M)$ for any lower triangular matrix M (see Definition 5.1 of Magnus (1988), p. 76),

Proposition 1 (Joint prior density of θ): Let $\theta = g(A, \Sigma, Q)$ denote the $n^2(H+1) \times 1$ vector of structural impulse responses, where H is the maximum impulse response horizon. Then the joint prior density of θ is given by

$$f(\theta) \ = \ 2^{\frac{n(n+1)}{2}} |\Sigma|^{-\frac{(np-1)}{2}} |I_n + S|^{-(n-1)} \left| \frac{\partial \text{vec}(\Phi)'}{\partial \text{vec}(A)} \frac{\partial \text{vec}(\Phi)}{\partial \text{vec}(A)'} \right|^{\frac{1}{2}} f_1(A|\Sigma) f_2(\Sigma) f(s), \tag{3}$$

where $f(\theta)$, $f_1(A|\Sigma)$, $f_2(\Sigma)$ are the prior densities of θ, A conditional on Σ and Σ, respectively, S satisfies equation (1), and s denotes the below-diagonal elements of S.

The proof of Proposition 1 can be found in the appendix.

2.3 The distribution of θ conditional on the MLE

As emphasized in Baumeister and Hamilton (2015) and Watson (2020), in the conventional Bayesian approach to estimation and inference, abstracting from estimation uncertainty, summary statistics about the impulse responses in the identified set depend only on the prior for Q. The following propositions makes this point more formally.

Proposition 2 (Joint prior density of θ conditional on A and Σ): Let $\theta = g(A, \Sigma, Q)$ denote the $n^2(H+1) \times 1$ vector of structural impulse responses, where H is the maximum impulse response horizon. Then the joint prior density of θ, given A and Σ, is given

$$h(\theta|A, \Sigma) \ = \ \left| J_U'(I_n \otimes (P'P + P'\Phi'\Phi P)) J_U \right|^{\frac{1}{2}} f(s). \tag{4}$$

The proof is in the appendix. Unlike in Proposition 1, the joint prior density for θ no longer depends on the priors for A and Σ.

Baumeister and Hamilton (2015) propose to assess the implications of the uniform Gaussian-inverse Wishart prior for the prior of individual elements of θ based on the distribution $h(\theta|\widetilde{A}, \widetilde{\Sigma})$, where $\widetilde{A}$ and $\widetilde{\Sigma}$ denote the MLE of A and Σ, respectively. It is immediately clear upon inspection that $h(\theta|\widetilde{A}, \widetilde{\Sigma})$ differs from the prior distribution we derived in (3). Hence, $h(\theta|\widetilde{A}, \widetilde{\Sigma})$ is not the prior, invalidating any conclusions about how informative the uniform-Gaussian inverse Wishart prior is for the structural impulse responses based on $h(\theta|\widetilde{A}, \widetilde{\Sigma})$. Another, simpler, way of making this point is to observe that $h(\theta|\widetilde{A}, \widetilde{\Sigma})$ is a function of MLE and hence depends on the data. Obviously, in the standard Bayesian framework employed in estimating sign-identfied VAR models, the prior does not depend on the data. As a result, the stylized example provided in Baumeister and Hamilton (2015) to illustrate how the conventional Bayesian approach to estimating sign-identified VAR models implies unintentionally informative prior distributions for the impact responses is missing the point. Effectively, Baumeister and Hamilton criticize the priors employed in the literature based on diagnostics that depend on the data, which is at odds with the Bayesian approach to impulse response inference.

Essentially the same approach as in Baumeister and Hamilton (2015, 2018, 2019, 2020) has also been used by Watson (2020, p. 189) who refers to the element-wise distribution of θ over $h(\theta|\widetilde{A}, \widetilde{\Sigma})$ as the impulse response prior for a given element of θ induced by the prior for Q and the sign restrictions. This language is misleading since the actual impulse response priors induced by the prior distribution for Q also depend on priors for the reduced-form parameters, as we showed earlier. Moreover, the approach employed by Baumeister and Hamilton (2015) and Watson (2020), among others, is pointless from a Bayesian point view as a characterization of the prior as well as as a characterization of the posterior, since the estimation uncertainty about A and Σ does not vanish in finite samples. As the next subsection illustrates, their approach tends to be highly misleading in practice.

2.4 Empirical illustration

We illustrate the concerns with the approach of Baumeister and Hamilton (2015) and Watson (2020) using a simple, but realistic example drawn from the empirical literature. Consider a structural VAR model identified by sign and exclusion restrictions that examines the determination of household inflation expectations. The model includes 12 lags and is estimated with an intercept. The variables include the log real price of U.S. motor gasoline ($rpgas$), CPI headline inflation (π) and the mean one-year inflation expectation in the Michigan Survey of Consumers (π^{exp}). The model explains variation in these three variables in terms of three mutually uncorrelated structural shocks: (1) A nominal gasoline price shocks; (2) a shock to the core CPI (defined as the CPI excluding gasoline); and an idiosyncratic shock to household inflation expectations that is not reflected in current prices. We postulate that

$$\begin{pmatrix} u_t^{rpgas} \\ u_t^{\pi} \\ u_t^{\pi^{\text{exp}}} \end{pmatrix} = \begin{bmatrix} + & - & 0 \\ + & + & 0 \\ + & + & + \end{bmatrix} \begin{pmatrix} w_t^{\text{nominal gasoline price}} \\ w_t^{\text{core CPI}} \\ w_t^{\text{idiosyncratic inflation expectation}} \end{pmatrix} \tag{5}$$

Further discussion of the rationale of these restrictions and of the VAR model specification can be found in Kilian and Zhou (2020a) whose analysis in turn is motivated by Coibion and Gorodnichenko (2015). The model is estimated based on a conventional uniform-Gaussian-inverse Wishart prior with the prior means of all slope parameters set to zero. Since identification is achieved by a combination of sign and exclusion restrictions on the impact responses, estimation requires the use of the importance sampler proposed in Arias et al. (2018). Our analysis focuses on individual elements of θ to facilitate comparisons with the discussion in Baumeister and Hamilton (2015).

Baumeister and Hamilton's first proposal is to evaluate the support of $h(\theta|\widetilde{A}, \widetilde{\Sigma})$, where $\widetilde{A}$ and $\widetilde{\Sigma}$ denote the MLE of A and Σ, respectively. The upper panel of Figure 1 shows estimates of this type of identified set for selected responses in the empirical example discussed in the preceding section. In all cases, the identified sets form a narrow region, indicating a small degree of uncertainty associated with Q. What we learn from this exercise is that in samples large enough for the MLE to be estimated precisely, there remains residual uncertainty about the value of the structural impulse responses associated with Q, but that uncertainty in our empirical application is small.

This exercise is intended to illustrate the frequentist properties of the conventional Bayesian approach, but conditioning on the data in constructing the MLE, while still drawing from the Haar prior for Q, renders this exercise meaningless from a Bayesian point of view. In particular, plotting the identified set conditional on the MLE of the reduced-form parameters does not tell us anything about the identified set for θ under the uniform-Gaussian-inverse Wishart prior for the model parameters. The lower panel of Figure 1 shows the identified sets for the impulse responses under this prior, with the boundaries truncated to match the dimensions of the upper panel, for illustrative purposes. Panel (b) illustrates that this identified set covers nearly the entire parameter region, whereas the one conditional on the MLE does not. Thus, these two identified sets are fundamentally different objects.

Moreover, plotting the identified set, as in Figure 1, tells us nothing about the distribution of the responses over the identified set. This is why Baumeister and Hamilton (2015), in addition, plot the histogram over the identified set obtained conditional on the MLE for selected individual impulse responses. This may be viewed as an approximation to the marginal density for each element of θ implied by $h(\theta|\widetilde{A}, \widetilde{\Sigma})$. The upper panel of Figure 2 shows such histograms for the impact response of the real price of gasoline to a positive shock to the nominal retail price of gasoline. This impact response is restricted to be positive. It also shows the corresponding histograms constructed from the identified sets under the prior and under the corresponding posterior. The lower panel in Figure 2 shows the corresponding results for the same response function at horizon 2.

Focusing on the impact response, the histogram conditional on the MLE in panel (a) is tightly concentrated, whereas the histogram over the identified set under the prior in panel (b) is not. Not only is the shape of these histograms rather different, but so is the location and the dispersion, underscoring our point that Baumeister and Hamilton's and Watson's approach does not capture the impulse response prior, which makes it impossible to ascertain whether the impulse response prior is reasonable or not.

Given that our tools allow us to construct the prior for the impact response, how plausible is the prior in panel (b)? One may have expected the prior distribution to decay smoothly in the positive range, as the value of the response increases. However, even though the impulse response distribution is truncated from below at zero, the mode of the prior distribution is strictly positive. This result follows from the fact that the distribution of $\sqrt{\Sigma_{11}}$ has no mass at zero and is strictly positive with probability 1, and that Q_{12} has strictly positive support after imposing the sign restrictions, as implied by the histogram in the first column of Figure 2. While the prior in panel (b) does not look unreasonable in that it embodies the maintained sign restriction and is fairly diffuse, it is hard to know how economically plausible this prior for the impact response is.

The latter question does not matter much, however, given strong evidence that, whatever feature of this prior a user may disagree with, is overruled by the data. Panel (c) in Figure 2 shows the corresponding histogram over the identified set under the posterior. Two facts stand out. First, the median of this distribution has shifted to the right, compared to the prior in panel (b), and the dispersion has been greatly reduced, providing compelling evidence that the posterior is driven by the data rather than the prior. Thus, there is no support for the conjecture in Baumeister and Hamilton (2015) and Watson (2020) that the impulse response posterior tends to be dominated by the prior. Second, the median of the posterior value in panel (c) is distinctly larger than the median conditional on the MLE in panel (a).

The lower panel of Figure 2 shows the corresponding results for the response of the real gasoline price after two months. Unlike in the upper panel, there is no sign restriction on this response. As a result, the prior distribution is centered on a value close to zero with considerable probability mass on values to the left and to the right of zero. This prior distribution is consistent with the view that we do not know a priori whether this response is positive or negative, which makes a central tendency near zero a reasonable representation of our uncertainty about this response. In fact, it is more reasonable than a truncated uniform distribution centered on zero would have been, since very large and very small values of the response are less likely. Much the same reasoning has been applied to the prior for the slope coefficients in VAR modeling for many years. Moreover, the support is wide enough to cover any value of this response that would be reasonable a priori. Thus, not only is this impulse response prior absolutely reasonable, but the posterior distribution in panel (c) indicates that the posterior is quite different from the prior. The posterior median is far from the center of the prior distribution and the posterior distribution is highly concentrated around the posterior median. As in the previous example, the distribution conditional on the MLE in panel (a) tells us nothing about the prior nor does it match the posterior. We conclude that there is no substitute for a proper Bayesian analysis of this problem.

The evidence in the upper panels of Figures 1 and 2 is instructive for four reasons. First, it shows that the identified set conditional on the MLE is very narrow, indicating that whatever assumptions we make about the prior for Q will not have much of an effect on the impulse response posterior. This finding is in sharp contrast to the illustrative examples discussed in Baumeister and Hamilton (2015), which imply much wider identified sets, conditional on the reduced-form MLE. It is the perception that sign-identified imply wide identified sets that has prompted Plagborg-Møller (2019) and Giacomini and Kitagawa (2020), among others, to develop alternative methods of inference that avoid specifying a prior for Q. In fact, these studies explicitly cite Baumeister and Hamilton

(2015) as the motivation for their alternative approaches. It is fair to say that the literature may not have evolved in this direction, were it not for the fact that Baumeister and Hamilton (2015) overstated the importance of the prior for Q for the structural impulse responses.

Our evidence suggests that the identified set conditional on the MLE in more realistic settings tends to be so narrow that few applied users would even bother to think about the role of the prior for Q. Why do we reach a different conclusion? The key difference is that Baumeister and Hamilton (2015) in making their case against the Haar prior for Q implicitly assumed that there is only one sign restriction in the structural VAR model, regardless of n. This assumption is not representative for applied work. Even for $n = 2$, applied users would typically use four sign restrictions and in larger models the number of sign restrictions tends to be much larger because, in practice, many static sign restrictions are required to identify the structural shocks in sign-identified models. Since each of these sign restrictions tends to truncate the identified set for the impulse responses in the model, the resulting identified sets tend to be much tighter than suggested by the stylized examples in Baumeister and Hamilton (2015). This result, of course, relies on the structural VAR model being fully identified, as in our empirical example. When only one shock is identified in the structural model, in contrast, there are far fewer sign restrictions and hence the identified set is much larger.

Second, the presence of additional static sign restrictions elsewhere in the model changes the location and the shape of the prior distribution over the identified set of the responses and invalidates the analytical derivations of these prior distributions conditional on the MLE in Baumeister and Hamilton (2015). For example, for VAR models with $n = 3$, as in our empirical example, Baumeister and Hamilton (2015) derive the result that the density of the impact response is a truncated uniform distribution conditional on $\widetilde{\Sigma}$. The histogram in panel (a) of Figure 2, shows that the distribution of the impact response is far from uniform in general. It shows increasingly higher probability mass, as values approach the upper bound of the identified set. This result again follows from the imposition of additional sign restrictions.

Third, the fact that the uncertainty about the responses attributable to Q tends to be small implies that the impulse response prior will be dominated by the priors for Σ (and possibly A), which are updated based on the data. This fact helps explain why the posterior is dominated by the data in models based on static sign restrictions.

Fourth, an immediate implications of this finding is that the choice of the prior over the support of Q in this application is largely inconsequential because any draw in the identified set for θ would imply substantively identical impulse responses. Whether this prior is assumed to be uniform or not has little effect on the Bayes estimator of the responses.

Finally, it should be noted that our central point about the invalidity of the approach proposed by Baumeister and Hamilton (2015) and Watson (2020) does not depend on the choice of the loss function and the construction of the Bayes estimator. In the next section, we develop an alternative approach to evaluating whether conventional priors for sign-identified VAR models are unintentionally informative about the impulse response prior and whether they distort posterior inference about the impulse responses.

3 An alternative approach to assessing impulse response priors

Our objective in this section is to examine the implications of the conventional uniform-Gaussian-inverse Wishart prior for sign-identified VAR models for the joint prior distribution of the vector θ. We also examine the implications of that prior for the joint posterior of θ, which has received little or no attention in the literature so far, beyond unsubstantiated claims that this posterior is driven by the implicit impulse response prior. One key difference from earlier studies is that we focus on inference about vectors of impulse responses rather than selected responses one at a time. The latter approach has been shown to be inappropriate for characterizing the shape and comovement of vectors of impulse responses as well as the uncertainty surrounding Bayes estimates of θ.[1]

For expository purposes, we conduct our analysis under absolute loss. The assumption of absolute loss was also maintained by Baumeister and Hamilton (2015) and Watson (2020), among many other studies, making it a natural starting point. Our analysis could be adapted to other loss functions including quadratic loss and Dirac delta without affecting the points we are making.

3.1 Characterizing the joint prior under absolute loss

As before, let θ denote the vector of unknown impulse responses obtained by appropriately stacking all n_{irf} impulse responses of interest into a vector. Since there are n^2 impulse response functions for horizon $h = 0, 1, 2, ..., H$, in the n-dimensional autoregressive model, if all responses are included, $n_{irf} = n^2(H + 1)$. Let $\theta \in \Theta$, where Θ is the set of all structural impulse response functions that satisfy the identifying restrictions imposed on the structural VAR model. We abstract from the details of the construction of θ, which may differ from one structural VAR model to another, because our analysis in this section does not depend on these details. The space Θ may be approximated by drawing from the distribution of the model parameters and simulating the distribution of the structural impulse responses that are consistent with the identifying assumptions.

Given that the joint posterior of θ is typically summarized based on the Bayes estimator of the impulse response vector and the corresponding joint credible set, it is useful to compute similar summary statistics for the joint prior distribution, building on results in Inoue and Kilian (2020) who derived the Bayes estimator of θ under absolute loss and the corresponding joint posterior credible set. Let $L(\theta, \overline{\theta})$ denote a loss function that maps from $\Theta \times \Theta$ to $\Re$ and define the impulse

[1] The importance of joint inference about impulse response vectors has been made both in the frequentist literature (see Lütkepohl, Staszewska-Bystrova and Winker 2015a,b, 2018; Inoue and Kilian 2016; Kilian and Lütkepohl 2017; Bruder and Wolf 2018; Montiel Olea and Plagborg-Møller 2019) and in the Bayesian literature (see Sims and Zha 1999; Inoue and Kilian 2013, 2019, 2020).

response estimator as

$$\widehat{\overline{\theta}} \;=\; \text{argmin}_{\overline{\theta}\in\Theta} E_\theta[L(\theta,\overline{\theta})], \tag{6}$$

where $\overline{\theta} \in \Theta$ denotes an action and the expectation is with respect to the joint prior distribution of $\theta \in \Theta$, derived in Section 2.2. Under absolute loss,

$$L(\theta,\overline{\theta}) \;= \sum_{j=1}^{n_{irf}} |\theta_j - \overline{\theta}_j|, \tag{7}$$

where θ_j and $\overline{\theta}_j$ are the jth elements of θ and $\overline{\theta}$, respectively. In practice, the impulse response estimator under absolute loss may be approximated by

$$\widehat{\overline{\theta}}_M = \text{argmin}_{\overline{\theta}\in\widehat{\Theta}_M} \frac{1}{M}\sum_{i=1}^{M}\sum_{j=1}^{n_{irf}} |\theta_j^{(i)} - \overline{\theta}_j|, \tag{8}$$

where $\theta^{(i)}$ is the ith prior draw and $\widehat{\mathit{\Theta}}_M$ consists of M prior draws.

Similarly, we define the $(1-\alpha)100\%$ joint lowest prior risk region as

$$\Theta_{1-\alpha,L} \;=\; \{\overline{\theta}\in\Theta : E_\theta(L(\theta,\overline{\theta})) \leq c_{1-\alpha,L}\}, \tag{9}$$

where $c_{1-\alpha,L}$ is the smallest number such that the prior probability of $\Theta_{1-\alpha,L}$ is $1-\alpha$ and L refers to the loss function. Under absolute loss, this joint prior credible set is estimated as

$$\widehat{\Theta}_{1-\alpha,M} = \left\{ \overline{\theta}\in\widehat{\Theta}_M : \frac{1}{M}\sum_{i=1}^{M}\sum_{j=1}^{n_{irf}} |\theta_j^{(i)} - \overline{\theta}_j| \leq c_{1-\alpha} \right\} \tag{10}$$

where $\widehat{\overline{\theta}}_M$ denotes the impulse response estimator and $c_{1-\alpha,AL}$ is the smallest values such that $\widehat{\Theta}_{1-\alpha,M}$ has prior probability $1-\alpha$ in the limit, as the number of prior draws approaches infinity. In practice, these joint credible set may be constructed by sorting$\frac{1}{M}\sum_{i=1}^{M} L(\theta^{(i)},\theta^{(1)})$,

$\frac{1}{M}\sum_{i=1}^{M} L(\theta^{(i)}, \theta^{(2)})$,... $\frac{1}{M}\sum_{i=1}^{M} L(\theta^{(i)}, \theta^{(M)})$, in ascending order and retaining the first $(1-\alpha)100\%$ draws, starting with the draw with the lowest value.

3.2 Summary statistics for the difference between the joint prior and the joint posterior distribution of the structural responses

It is straightforward to visually assess the joint prior distribution of θ based on the Bayes estimate of the responses and the responses in the joint credible set. Likewise, it is straightforward to visually compare the posterior of θ to the prior distribution or to compare posteriors derived under alternative priors. It nevertheless can be useful to quantify these differences based on summary statistics.

By construction, the joint posterior density of θ takes the same form as $f(\theta)$ in Proposition 1 with $f_1(A|\Sigma)$ and $f_2(\Sigma)$ denoting the posterior densities of A conditional on Σ and Σ, respectively. This allows us to apply the same diagnostics to draws from the posterior distribution of θ, as for the draws from its prior distribution. Since θ depends on Σ at all horizons and, in addition, on A at all horizons but the impact period, it follows immediately that the posterior distribution of θ at all horizons depends on the data. This means that, in practice, Bayes inference about θ conditional on the data is never completely determined by the prior for Q.

Building on the notation in Section 3.1, we propose four summary statistics for measuring the extent to which the joint posterior distribution of the impulse responses differs from the corresponding prior distribution. When these statistics depend on the loss function, we focus on the example of absolute loss, for expository purposes. Our discussion for now focuses on comparing the estimates of the responses and the joint credible sets under the prior and under the posterior, which is a natural starting point, since these are the estimators an applied user presumably would focus on.

First, we consider the Hausdorff distance metric

$$\begin{aligned} d_H(\Theta^{prior}, \Theta^{posterior}) &= \max\Bigg\{ \sup_{\theta^{prior} \in \Theta^{prior}} \inf_{\theta^{posterior} \in \Theta^{posterior}} d(\theta^{prior}, \theta^{posterior}), \\ &\qquad \sup_{\theta^{posterior} \in \Theta^{posterior}} \inf_{\theta^{prior} \in \Theta^{prior}} d(\theta^{prior}, \theta^{posterior}) \Bigg\}, \end{aligned} \tag{11}$$

which refers to the greatest of all the distances from a point in one set to the closest point in the other set. Here Θ^{prior} denotes the set of draws in the joint prior credible set defined in section 3.1 and $\Theta^{posterior}$ the set of draws in the corresponding joint posterior credible set, defined in Inoue and Kilian (2020). The subscripts have been dropped for notational convenience. If the two credible sets are identical, this distance measure will be zero, which would indicate that the posterior draws do not depend on the data at all. Implicit in this approach is the presumption that all impulse responses are measured in the same unit (say, percent). If not, this measure may be applied to subsets of responses measured in the same units or an alternative loss function has to be used (see Inoue and Kilian 2020).[2]

Second, we can summarize the maximum change and the average change in the central tendency of the distribution, as measured by the location of the estimate of the responses, in one number each, as long as all impulse responses are measured in the same units. For example, the average change in the location of the impulse response estimator based on the posterior distribution relative to that based on the prior distribution is

$$d_L \equiv \sum_{i=1}^{n_{irf}} \left| \widehat{\theta}_{M,i}^{posterior} - \widehat{\theta}_{M,i}^{prior} \right| \Big/ n_{irf}. \tag{12}$$

where $|\cdot|$ denotes the absolute value, $\widehat{\theta}_M^{prior}$ is the impulse response estimator defined in section 3.1 and $\widehat{\theta}_M^{posterior}$ denotes the corresponding Bayes estimator defined in Inoue and Kilian (2020). To the

[2] For related applications of the Hausdorff distance metric see, e.g., Chernozhukov, Hong and Tamer (2007).

extent that impulse responses are measured in different units, the same approach may be applied to the responses of individual variables.

Third, we investigate by how much the dispersion of the impulse response draws about the measure of central tendency implied by the loss function is reduced, as the prior and the likelihood are combined to form the posterior.

Step 1: For all members of the credible set $\Theta^{posterior}$ except for the element $\widehat{\theta}_M^{posterior}$, define the dispersion about $\widehat{\theta}_M^{posterior}$ as

$$C^{posterior} = \frac{\sum_{j=1}^{M(1-\alpha)} \sum_{i=1}^{n_{irf}} \left| \theta_i^{(j),posterior} - \widehat{\theta}_{M,i}^{posterior} \right|}{M(1-\alpha)n_{irf}},$$

where $|\cdot|$ denotes the absolute value.

Step 2: Compute the analogous dispersion measure based on the impulse responses drawn from the prior distribution, C^{prior}.

Step 3: Then the percent change in the dispersion of the impulse response estimates, as the prior is updated based on the likelihood, is given by

$$d_C \equiv 100 * \left(C^{posterior} - C^{prior}\right) / C^{prior}. \tag{13}$$

This metric is labeled d_c because it tells us about changes in the concentration of the probability mass about the impulse response estimate.

Fourth, to capture changes in the signs of the impulse responses, as the prior information is updated based on the data, we transform the impulse response vector to a binary sequence of positive and negative signs (excluding responses that are restricted to zero). With some abuse of notation, define the Jaccard-like metric $1 - |\Theta^{prior} \cap \Theta^{posterior}| / |\Theta^{prior} \cup \Theta^{posterior}|$, where $|\Theta^{prior} \cup$

$\Theta^{posterior}|$ is the number of uniquely distinct sign patterns of the impulse responses under the prior distribution and the posterior distribution combined and $|\Theta^{prior} \cap \Theta^{posterior}|$ is the number of shared sign patterns of the impulse responses (see Jaccard 1908). Then the number of pairs of prior and posterior impulse responses that have exactly the same pattern of signs is given by

$$|\Theta^{prior} \cap \Theta^{posterior}| \;=\; \sum_{i=1}^{|\Theta^{prior}|} \sum_{j=1}^{|\Theta^{posterior}|} I(\theta_k^{(i),prior}\theta_k^{(j),posterior} > 0,\; \forall\, k = 1, 2, ..., n^2(H+1))$$

where $|\Theta^{prior}|$ and $\left|\Theta^{posterior}\right|$ are the cardinalities of Θ^{prior} and $\Theta^{posterior}$, respectively, and $I(\cdot)$ denotes the indicator function. $|\Theta^{prior} \cap \Theta^{posterior}|$ measures the number of uniquely distinct sign patterns that are shared by the impulse responses under the prior and under the posterior distribution. Furthermore, let $|\Theta^{prior} \cup \Theta^{posterior}|$ denote the cardinality of $\Theta^{prior} \cup \Theta^{posterior}$. Then $|\Theta^{prior} \cup \Theta^{posterior}|$ measures the number of uniquely distinct sign patterns encountered in the prior and the posterior combined, and

$$d_J(\Theta^{prior}, \Theta^{posterior}) \;=\; 1 - \frac{|\Theta^{prior} \cap \Theta^{posterior}|}{|\Theta^{prior} \cup \Theta^{posterior}|} \in [0, 1].$$

A value of 1 indicates that there is no common sign pattern, whereas a value of 0 indicates that all sign patterns are the same, corresponding to no updating. This distance metric is most useful when applied to one response function at a time, since the odds that there is no common sign pattern approach 1 when considering all elements of θ.[3]

Our discussion so far has focused on the impulse response estimator and on the draws contained in the corresponding credible sets under the prior and under the posterior. It should be noted

[3] When restricting attention to the estimated joint $(1-\alpha)100\%$ credible set, the cardinality of $|\Theta^{prior}|$ is $M_1(1-\alpha)$, where M_1 is the number of admissible prior draws; the cardinality of $|\Theta^{posterior}|$ is $M_2(1-\alpha)$, where M_2 is the number of admissible posterior draws; and the cardinality of $\Theta^{prior} \cup \Theta^{posterior}$ is $(M_1+M_2)(1-\alpha)$. In practice, the dimension of the impulse response vector is adjusted to exclude responses that are restricted to zero, since these responses do not change by construction.

that it would be equally valid to compare all admissible prior draws and posterior draws. The distance and dispersion metrics discussed above could be easily adapted for that purpose. Another natural measure of the distance between the prior and the posterior distribution of θ would be their Kullback-Leibler divergence:

$$d_{KL} = \int_{\Theta} f(\theta|y_1, ..., y_T) \log \left(\frac{f(\theta|y_1, ..., y_T)}{f(\theta)} \right) d\theta \tag{14}$$

where $f(\theta|y_1, ..., y_T)$ and $f(\theta)$ are the posterior and prior densities of θ, respectively.

In the next two sections, we illustrate by example the role of the prior for impulse response inference in the context of several applications drawn from the empirical literature. We first examine how informative conventional priors on A, Σ, and Q are for θ. We then investigate to what extent these impulse response priors are overruled by the data. It should be noted that lack of evidence that a prior is being overruled by the data does not necessarily imply that the prior is unduly informative for the posterior. It is also possible that this prior simply lines up well with the data. In the latter case, evidence that the posterior is robust to alternative prior specifications adds credibility to the conclusions. The summary statistics discussed above are useful not only for characterizing the extent to which impulse response priors are updated by the data. The same type of summary statistics may also be used to quantify the sensitivity of the joint posterior distribution to alternative prior specifications. The only difference is that in the latter case we compare the posteriors under alternative prior specifications to the posterior under an arbitrarily determined benchmark prior.

4 Sign-identified VAR models based on static sign restrictions

In examining sign-identified VAR models, it is useful to differentiate between models identified by static sign restrictions, possibly combined with exclusion restrictions, as discussed in this section, before turning to models with additional dynamic sign restrictions, as discussed in section 5. For expository purposes, in this section, we examine the same structural VAR model example already discussed in section 2.4. The baseline prior sets the prior mean of the slope parameters to zero (Prior 1), which is a reasonable starting point in models with stationary variables, and, in our examples, produces posterior distributions that closely match those obtained using a uniform-diffuse Gaussian-inverse Wishart prior. An alternative prior specification raises the prior mean of the first own lag in each equation to 0.9, implying high persistence in the responses (Prior 2). Finally, we experiment with replacing the estimated innovation variances by fixed values of 0.01, which changes the prior mean of Σ as well as the prior variance of the slope parameters (Priors 3 and 4). These choices are summarized in Table 1.

If we are interested in the implications of the prior for the shapes and comovement of the structural response functions, as applied users typically are, it is essential to report statistics summarizing the joint prior distribution of θ rather than the marginal distributions. Figure 3a examines the question of whether Prior 1 is informative for the joint impulse response prior. The prior for impact responses that are restricted to zero are degenerate and may be dropped from consideration. Except when a sign restriction is imposed on impact, the remaining response estimates are close to zero, consistent with the impulse response prior being largely uninformative. The estimates of the responses that are restricted to be positive or negative tend to be at the lower end of the range of responses in the joint prior credible set. Thus, there is no evidence that the impulse response prior for this model favors large responses.

In practice, what matters most is not how informative the impulse response prior is, but how much this prior affects the impulse response posterior. Because this posterior depends on aspects of the prior that are revised based on the data and aspects that are not, with these components being nonlinearly combined, it is not clear a priori how sensitive the impulse response posterior is to the impulse response prior. Our tools allow us to address this question.

Figure 3b shows that in this model, as in the earlier macroeconomic model based on exclusion restrictions, the impulse response posterior is dominated by the data rather than by the prior. For example, the large and persistent response of the real price of gasoline to a nominal gasoline price shock is driven by the data, as are the responses of inflation expectations to the expectations shock and the nominal gasoline price shock. Nor does the prior necessarily pin down the magnitude of the impact response. For example, the impact responses of headline inflation to a nominal gasoline price shock and to a core CPI shock are much larger in magnitude than under the prior. In addition, the uncertainty about the impulse responses may be smaller than under the prior in some cases and may be substantially larger in other cases. In short, there is no evidence to support the conclusion that impulse response estimates derived from the framework of Waggoner, Rubio-Ramirez and Zha (2010), Arias, Rubio-Ramirez and Waggoner (2018), and Antolin-Diaz and Rubio-Ramirez (2018) are necessarily driven by their prior.

The results in Figure 3 are derived under Prior 1, but this conclusion holds more generally. Table 2a provides evidence that all four of the impulse response priors we considered are overturned by the data to varying degrees. The location shift of the impulse response estimate may be as high as 0.12 percentage points in absolute terms with an average revision of 0.01 percentage points. The largest revision relates to the headline inflation response. At an annualized rate, this translates to a shift by 1.44 percentage points of inflation, which is economically significant. The dispersion of the impulse responses may fall by as much as 98%. Not a single draw for θ matches the sign patterns under the prior. The Hausdorff distance measure also suggests substantial revisions to the joint density of the impulse responses.

Table 2b shows that the impulse response posterior is not sensitive to changes in the prior specification, including some specifications with more informative impulse response priors. The location of the impulse response estimate never changes by more than 0.02 percentage points in absolute terms. The average change is 0.00 percentage points. The changes in the Hausdorff dispersion measure and in the concentration about the impulse response estimate is one or two orders of magnitude smaller than in Table 2a. There are still changes in the sign patterns of θ, but not as much as in Table 2a. Thus, the posterior estimates appear robust to the prior.

Although this is only one example, we obtained similar results for other models that are fully identified based on static sign restrictions (possibly in combination with zero restrictions). This is not an accident. Our central result that priors for θ tend to be uninformative does not depend on the data. It depends only on the specification of the uniform-Gaussian-inverse Wishart prior and the sign restrictions and hence applies more generally to other VAR models. The sensitivity of the posterior to the prior for Q, in contrast, depends on how many sign restrictions are imposed in estimating the VAR model and how binding these restrictions are. The example we considered is typical for applied work in that it involves a fully identified model. When sign restrictions are used only sparsely, it is possible for the identified set to be larger and for the posterior to be less responsive to the data. The tools proposed in section 3 can be used to assses this question.

5 Sign-identified VAR models based on both static and dynamic sign restrictions

We now turn to an example of a structural VAR model identified not only by static sign restrictions, as in the preceding example, but also by dynamic sign restrictions and narrative restrictions. This difference matters because dynamic sign restrictions render the implied impulse response prior informative in at least some dimensions. This may be true even for impulse responses that are not directly restricted. As in section 4, we ask whether these priors are perhaps unintentionally informative.

Our illustrative example is the widely used global oil market VAR model of Kilian and Murphy (2014). This model has been reexamined and extended in a range of subsequent studies including Kilian and Lee (2014), Kilian (2017), Herrera and Rangaraju (2020), Zhou (2020), Cross, Nguyen and Tran (2020), Rausser and Stürmer (2020), and Kilian and Zhou (2020a,b). The data are monthly and the model includes 24 lags and an intercept. The model variables are the percent change in global crude oil production ($\Delta prod$), an appropriate measure of the global business cycle (rea), the log real price of oil ($rpoil$), and the change in global crude oil inventories (Δinv). The structural shocks include a flow demand shock, a flow supply shock, a storage demand shock and a residual demand shock (defined as the complement to all other shocks) that captures, for example, changes in the intensity of oil use, as users substitute toward more or less oil-intensive technologies. The structural shocks are identified based on static and dynamic sign restrictions, complemented by elasticity bounds and additional narrative inequality restrictions, which require the use of importance sampling, as discussed in Antolin-Diaz and Rubio-Ramirez.(2018). The model heavily relies on static sign restrictions.

$$\begin{pmatrix} u_t^{\Delta prod} \\ u_t^{rea} \\ u_t^{rpoil} \\ u_t^{\Delta inv} \end{pmatrix} = \begin{bmatrix} - & + & + & \\ - & + & - & \\ + & + & + & \\ & & + & \end{bmatrix} \begin{pmatrix} w_t^{\text{flow supply}} \\ w_t^{\text{flow demand}} \\ w_t^{\text{storage demand}} \\ w_t^{\text{residual demand}} \end{pmatrix} \tag{15}$$

The sign restrictions associated with the flow supply shock are applied not only on impact, but for the first 12 months. The estimation period is 1973.2-2018.6. Further details of the construction of the data and of the elasticity bounds and narrative restrictions can be found in Zhou (2020). The original specification involved a diffuse prior for the reduced-form parameters. Here we employ the prior specification corresponding to Prior 1 in Table 1. This change has no effect on the posterior estimates.

The central conclusion from this class of models is that flow demand shocks have large and persistent effects on global real activity and on the real price of oil. Storage demand shocks also have large and persistent effects on the real price of oil. By comparison, the effects of flow supply shocks on global real activity and the real price of oil are more modest. It has been suggested in the literature, but never demonstrated, that this conclusion is an artifact of the implicit prior for θ. The concern is that conventional priors on A, Σ, and Q may be unintentionally informative about θ because $\theta = g(A, \Sigma, Q)$. We address this concern by first deriving the implied prior for θ and then examining the extent to which the posterior estimates of θ are driven by the data rather than the prior. For expository purposes, we restrict attention to the responses of global real activity and the real price of oil to flow supply, flow demand and storage demand shocks.

5.1 Empirical results based on the Kilian and Murphy (2014) model

Figure 4a confirms that in the presence of dynamic sign restrictions and narrative sign restrictions, the prior distribution of θ is more informative than in the preceding empirical illustration. Consider first the estimate of the responses based on the prior for θ. Figure 4a shows a sharp fall in global real activity on impact in response to a negative flow supply shock, followed by a slightly negative response for the next 12 months. The prior favors a comparatively large impact response of the real price of oil to a negative flow supply shock that only gradually dies out, mirroring the corresponding response of global real activity. In contrast, the impact response of global real activity to a positive demand shock is positive by construction, but the remaining responses are effectively zero. The same goes for the response of the real price of oil to a positive flow demand shock, making these priors mutually consistent. Finally, the real price of oil jumps on impact and then declines in response to a positive storage demand shock, consistent with economic theory (see Alquist and Kilian 2010). This price increase is associated with a modest decline in real activity.

This joint prior is consistent with the traditional view that (1) unexpected oil supply disruptions have large negative effects on real activity in the short run and cause substantial oil price increases, that (2) positive flow demand shocks cause only a short-lived expansion of global real activity without a sustained effect on the real price of oil, and (3) that positive storage demand shocks cause the real price of oil to jump on impact, followed by a gradual decline, as predicted by economic theory, while its effects on real activity are only modestly negative. It is fair to say that this prior is stacking the deck against the modern view that global real activity is largely driven by flow demand shocks, but this was explicitly intended when Kilian and Murphy (2014) imposed their dynamic sign restrictions on the effects of flow supply shocks. The 68% joint credible set for the prior allows for considerable uncertainty about the responses.

Not only is the prior specification in Kilian and Murphy (2014) economically defensible, but Figure 4b demonstrates that this prior matters little for the impulse response posterior. The data overrule the prior to a large extent. For example, the response of the real price of oil to an unexpected flow supply disruption is more modest than in the prior, although the joint credible set recognizes the uncertainty about the magnitude of this response. The posterior also shows a much more modest response of real activity to flow supply shocks than the prior. There is no evidence of a sharp contraction on impact. At the same time, the responses to a positive flow demand shock are much larger and more persistent than suggested by the prior. The response of real activity to a positive flow demand shock is much larger than under the prior and hump-shaped. This economic expansion is associated with a corresponding large and sustained increase in the real price of oil.

Finally, the response of the real price to a storage demand shock is more sustained, while preserving its overall shape, while the response of real activity to the storage demand shock is shrunk toward zero.

In short, the posterior distribution of θ is largely driven by the data rather than the prior. This conclusion is reinforced by the distance metrics reported in Table 3. Not only do the data increase the concentration of the distribution about the impulse response estimate, but there are large shifts in the location of the responses in the credible set. Thus, there is no evidence that the conventional prior underlying this model is economically unappealing, unintentionally informative, or driving the posterior inference.

6 The alternative approach of Baumeister and Hamilton (2015)

There is a common perception in the literature that the alternative set of priors for sign-identified VAR models introduced by Baumeister and Hamilton (2015, 2018, 2019, 2020) addresses the conceptual shortcomings of conventional priors for sign-identified models discussed in this paper. Indeed, that is the explicit argument made by these authors. In this section, we show that their proposal for postulating priors on the structural model parameters implies an impulse response prior of unknown form, necessitating the use of the same tools we discussed in the context of the conventional approach to estimating sign-identified models. The reason is simple. Imposing explicit prior distributions on the parameters of $B_0, ..., B_p$ in the structural VAR representation

$$B_0 y_t = B_1 y_{t-1} + ... + B_p y_{t-p} + w_t$$

as proposed by Baumeister and Hamilton (2015), is not equivalent to specifying an explicit prior on the vector of structural impulse responses, which is defined by the nonlinear transformation

$\theta = \widetilde{g}(B_0, B_1, ..., B_p)$. Even if a prior on $B_0, ..., B_p$ may be defended on economic grounds, after applying the change-of-variable method, the prior on θ may be unintentionally informative. This conclusion remains true, if one is specifying a prior on one or more elements of B_0^{-1}. Since users of this approach have no idea what their implied prior for θ is, the proposal of Baumeister and Hamilton (2015) does nothing to address the concerns that motivated our analysis.

This conclusion should not be surprising since, in general, one cannot control the prior of the model in two dimensions at the same time. Thus, there is nothing to choose between their approach and the conventional approach on a priori grounds. Contrary to some assertions in the literature, the critique that sign-identified models should not be used without analyzing the implied prior for θ applies equally to the approach of Baumeister and Hamilton (2015).[4]

The tools developed in this paper are designed not only for the conventional Bayesian approach to estimating sign-identified VAR models, but they also accommodate the alternative approach proposed by Baumeister and Hamilton, allowing us to evaluate how informative Baumeister and Hamilton's impulse response priors are and how sensitive the joint posterior of the structural responses is to that prior.[5]

6.1 Empirical results based on the Baumeister and Hamilton (2020) model

We illustrate this point based on the global oil market model proposed in Baumeister and Hamilton (2019). As in the last section, in the interest of space, we focus on the responses of global real activity and of the real price of oil. The four structural shocks of interest are an oil supply shock, an oil-market specific demand shock (referred to as the consumption shock), an oil inventory demand shock

[4]A different way of interpreting Baumeister and Hamilton's approach is that they are specifying an implicit nonuniform prior for Q that is invariant to the data. There is nothing wrong necessarily with this approach, but this interpretation makes it clear why such a prior is conceptually no different from the conventional approach and may be unintentionally informative for the prior of the impulse response.

[5]Baumeister and Hamilton in their own work only report prior median response functions, or, as in the empirical example below, they do not report priors for the impulse response functions at all.

and a shock to global economic activity. Our analysis is based on Baumeister and Hamilton's data and preferred VAR(12) model specification, which includes a number of nonstandard features such as additive classical measurement error in oil inventories. In general, the data, prior restrictions, and model specification differs from the global oil market model in the previous section, so estimates should not be compared directly. For details the reader is referred to the original source.

Figure 5a show the priors for θ implied by the baseline model in Baumeister and Hamilton (2019). The prior distributions are generated using the same type of Metropolis-Hastings algorithm as used in the original study, except that the proposal density is centered on the prior mode. For now, we focus on the impulse response estimate, abstracting from the tremendous degree of prior uncertainty about the impulse responses. It is striking that the only shock in Figure 5a that is a priori expected to substantially raise the real price of oil is the oil supply shock. In contrast, the response of the real price of oil to the demand shocks in the model is either quite small, near zero, or even negative.

It is useful to consider each shock in isolation. First, an unexpected oil supply disruption causes a persistent increase in the real price of oil by 2.5% on impact and by 6.6% after 12 months. The corresponding decline in global real activity gradually approaches -1.8% after 12 months and reaches -3.1% after 18 months. This prior view is economically reasonable, if one is a firm believer in the importance of oil supply shocks.

Second, an unexpected increase in global real economic activity is associated with a sustained increase in the level of real activity over time that reaches 3.8% after 18 months. An obvious question is how to reconcile the persistently positive and increasing response of economic activity over 18 months with the negligible or even negative response of the real price of oil. This pattern is clearly at odds with conventional views about the relationship between economic expansions and the real price of oil. Even if one believed economic activity shocks to be unimportant for the real price of oil, one would not postulate a negative price response to a positive shock to economic activity.

Third, we know a priori that an oil-market specific shock to the demand for oil that raises the real price of oil, must lower global real activity. The response of the real price of oil to such an oil-market specific demand shock in Figure 5a is modestly positive on impact and is essentially zero after a few months. This is not an unreasonable prior view, except that it raises the question of how such a consumption shock at the same time raises global real activity by 0.7% after 18 months.

Fourth, it is not clear how to reconcile the flat response of the real price of oil to a storage demand shock with standard economic theory for storable commodities, which implies a jump in the price on impact followed by a gradual decline.

These observations highlight that the economic plausibility of impulse response priors often depends on the comovement across several response functions and on the shape of these response functions, making it essential to evaluate the joint prior of the impulse responses rather than their marginal prior distribution. Our analysis shows that the joint impulse response prior implied by Baumeister and Hamilton's prior on the structural VAR model parameters is economically implausible in several dimensions. The nature of this prior was neither intended by the authors nor has it been discussed in the literature.

There are also questions about the extent of the uncertainty embodied in the 68% joint impulse response credible set. For example, the prior puts considerable probability mass on persistent declines in the real price of oil in response to an unexpected oil supply disruption. Even on impact the 68% joint credible set includes oil price responses as high as 97% and as low as -1101%. It is not clear why anyone would consider a large negative response a priori plausible, not to mention one as large as -1101%. After all, an oil supply disruption is expected to raise the real price of oil. One possible remedy would be to impose additional dynamic sign restrictions on the responses of the real price of oil. Nor is it clear why anyone would consider a 97% increase in the real price of oil on impact a likely outcome. In fact, after 18 months, the range of the 68% credible set has grown to -2.97e+16%to 3.41e+16%. This suggests that there is lots of room to tighten the prior.

Similar problems also apply to the other responses in the joint credible set in Figure 5a. For example, most observers would agree that a positive shock to global real activity is not likely to cause a decline of 1.37e+15% in global real activity after 18 months or, for that matter, an increase of 1.21e+15%. These extreme realizations are likely to be an artifact of expressing the model variables in growth rate. After imposing relatively diffuse priors on the model coefficients, the cumulative impulse responses often become explosive, which helps explains the shape of the joint credible set for the level responses.[6] This problem is not restricted to the analysis in Baumeister and Hamilton (2019), but is particularly visible in this example. One possible remedy would be the imposition of bounds on the dominant root in the prior.

Overall, this impulse response prior would hardly have been the starting point of a researcher thinking about the prior specification for the impulse responses in a global oil market model. Does it matter that Baumeister and Hamilton (2019) rely on an unintentionally informative impulse response prior? To some extent it does and to some extent it does not. Figure 5b shows substantial revisions to the Bayes estimate based on the joint posterior of the impulse responses. For example, the response of global real activity to an unexpected oil supply disruption is revised up from -3.1% after 18 months to -0.1%, while the corresponding response of the real price of oil shrinks from 6.7% to 2.9%. There is little evidence of a strong recessionary impact on the global economy and the price response is modest at best. The data correct the strong views embodied in the prior for the impulse responses. Likewise, the responses to an economic activity shock now look more reasonable than in the prior. The response of the real price of oil peaks after 6 months at 5.1% and remains

[6] It should be noted that, as a result, for this prior the difference between inference based on marginal impulse response distributions and based on the joint distribution is even greater than for the posterior.

persistently high at longer horizons. The response of global real activity after 18 months is 1.8% rather than 3.8%. However, the responses to the consumption demand shock remain economically implausible in that a modest, but persistent positive effect on the real price of oil of up to 2.2% is accompanied by a temporary increase by as much as 0.1% in global real activity. Similarly, the economically implausible response of oil inventories to the inventory demand shocks remains.

In addition, the data tend to substantially narrow the uncertainty about the impulse responses in all cases, which is not surprising since the prior did not impose tight restrictions on the dynamics of the response functions. Table 4 confirms this impression. There is a strong increase in the concentration of the distribution about the impulse response estimate of θ. On average, the data shift the location of the response of the real price of oil much more than the response of real activity.

6.2 Limitations of the proposed methodology

The preceding example illustrates that deriving the implied prior for θ can be helpful in detecting whether priors on the structural model parameters have questionable implications for the prior of θ. Unlike in the earlier empirical illustrations, in the Baumeister and Hamilton (2019) example, there is reason to question the implied prior for θ. One way of addressing these concerns would be to impose additional dynamic sign restrictions and to bound the dominant autoregressive root.[7] Even if there were no evidence that the impulse response prior is distorting the posterior distribution of θ, however, it would be a mistake to blindly view estimates of the posterior distribution of θ as summaries of what we learn from the data. Since posterior inference is always conditional on the modeling assumptions and restrictions on the support of the impulse responses distribution, posterior estimates are only as credible as the underlying economic structure.

[7] An alternative approach was proposed by Plagborg-Møller (2019) who imposes priors directly on the parameters of the structural moving average representation. Even in the latter case, however, the diagnostic tools we proposed remain useful for assessing how informative such priors are and how much the joint posterior of θ depends on these priors.

In practice, it may not always be possible to detect flawed modeling assumptions simply by studying the central tendency of the prior of θ. Nor does the fact that a prior for θ may look fairly agnostic necessarily mean that this prior is innocuous. Seemingly reasonable prior or posterior draws for the impulse responses may be unrealistic, for example, because they violate elasticity bounds that are not imposed in estimation or because they fail to satisfy narrative sign restrictions that are not imposed in estimation. For example, in the Baumeister and Hamilton (2019) model, the oil price spike of 1990, following Iraq's invasion of Kuwait, is in substantial part attributed to an unexpected idiosyncratic increase in consumers' demand for oil in the second half of 1990, which does not seem economically plausible. Likewise, restrictions on the dynamics of the VAR process or the use of inappropriate data transformations may distort the posterior impulse responses. The more agnostic the prior distribution of θ seems, the more likely this problem is to arise. This point has been forcefully made by Kilian and Murphy (2012) and Kilian (2019). Thus, the credibility of the posterior estimates of θ ultimately rests on how carefully the underlying structural model has been designed and parameterized.

7 Concluding remarks

Several recent studies have voiced concerns about the priors typically used in estimating sign-identified VAR models, as described in Uhlig (2005), Rubio-Ramirez et al. (2010), Arias et al. (2018) and Antolin-Diaz and Rubio-Ramirez (2018). There is a consensus in the literature that, in the words of Watson's (2020), only "good" impulse response priors lead to "good" inference in sign-identified VAR models, where a "good" impulse response prior is generally understood to mean a prior that is economically plausible or that is uninformative.

Perhaps surprisingly, adequate tools for assessing the impulse response priors implied by conventional prior specifications do not exist. We first showed that approximating the impulse response prior based on the distribution of the impulse responses estimates conditional on the MLE of the reduced-form parameters, as proposed by Baumeister and Hamilton (2015, 2018, 2019, 2020) and Watson (2020), does not make sense from a Bayesian point of view, because the prior does not depend on the data. The flaw in this approach is that the prior for a given impulse response induced by the prior for the rotation matrix and by the sign restrictions also depends on the priors for the reduced-form parameters, which affects the location and shape of the impulse response prior distribution. We showed by example that this approach tends to produce highly misleading estimates of the impulse response priors. Our analysis invalidates the evidence provided in the recent literature against the conventional approach to estimating sign-identified VAR models.

We then derived appropriate tools to differentiate good impulse response priors from bad ones and illustrated their use in a range of representative empirical applications. Our evidence suggests that unduly informative impulse response priors are the exception rather than the rule. Our evidence does not support the view that conventional priors for sign-identified VAR models tend to be unintentionally informative about the vector of structural responses, θ. Nor does it support the view that the substantive conclusions of users of conventional prior specifications follow from their implicit prior on θ. Thus, there is no basis for the argument that this widely used approach must be abandoned and replaced by alternative prior specifications. Nor is there a basis for disregarding the empirical evidence compiled in the literature based on this approach. In fact, our evidence calls into question the common view that much of the uncertainty about the impulse response can be attributed to uncertainty about the rotation matrix Q. We showed that in typical models with many static sign restrictions the identified set is quite narrow, controlling for estimation uncertainty in the reduced-form parameters, which helps explain why the posterior of θ in practice is largely determined by the data. Our results contrast sharply with prevailing views in some of the recent literature on sign-identified VAR models.

Posterior inference about θ is econometrically valid, given any prior for Q (see Uhlig 2017). This does not necessarily mean that constructing the posterior for θ will also be economically sensible. This will only be the case under under any one of three conditions that can be verified in practice. First, the conventional approach will be sensible when the implied prior for θ, can be shown to be uninformative, except for the imposition of sign restrictions. An illustrative example showed that under conventional prior specifications this condition is met in typical VAR models based on static sign restrictions. This example is general because the prior for θ does not depend on the data. It depends entirely on the uniform-Gaussian-inverse Wishart prior for the VAR model parameters and hence applies more generally. Second, the conventional approach will be sensible if the implied prior for θ can be given an economic interpretation. We showed by example that even in models that combine static and dynamic sign restrictions, which renders the prior for θ necessarily informative, the impulse response prior need not be unintentionally informative. Third, even if the prior for θ turns out to be unintentionally informative, relative to the views of the researcher, the conventional approach will remain sensible, as long as the data effectively overturn the prior for θ. The evidence in our paper suggests that, in all empirical examples we considered, the posterior for θ is largely determined by the data rather than the prior for θ. We noted that this is more likely to be the case in models based on many sign restrictions than in sparsely identified models.

We furthermore demonstrated that the alternative Bayesian approach of Baumeister and Hamilton (2015), which was explicitly intended to circumvent the problem of unintentionally informative priors for θ, suffers from the same conceptual drawback as the conventional Bayesian approach. We illustrated this point by characterizing the implicit joint prior for θ employed by Baumeister and Hamilton (2019) in modeling the global oil market. Our analysis shows that some aspects of this prior are difficult to reconcile with standard economic reasoning. Our analysis calls into question suggestions in the literature that Baumeister and Hamilton's approach is inherently superior to the conventional approach. Our evidence suggests that the conventional approach is more likely to be robust to the prior for the model parameters than this alternative approach.

The issues discussed in this paper are distinct from the point that Bayesian inference in sign-identified VAR models does not asymptotically coincide with frequentist inference when using the conventional approach in the literature or when using the alternative approach of Baumeister and Hamilton (2015). Since the frequentist approach and the Bayesian approach ask different questions, it is not surprising that they may arrive at different answers (see Uhlig 2017). For VAR users concerned with the frequentist properties of Bayesian methods in large samples, a natural approach would be to report sets of posterior impulse response means, as proposed by Giacomini and Kitagawa (2020). As we showed, however, in many applications based on multiple sign restrictions the identified sets are largely unaffected by the prior uncertainty about Q, so the choice of method is likely to be less important than often thought, as long as one accounts for the joint dependence across the impulse responses.

Acknowledgments

The views expressed in the paper are those of the authors and do not necessarily represent the views of the Federal Reserve Bank of Dallas or the Federal Reserve System. We thank participants at the 2020 IAAE webinar and the Federal Reserve Bank of Atlanta research seminar for helpful comments. We have also benefited from discussions with Juan Rubio-Ramirez, Juan Antolin-Diaz, and Benjamin Wong. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix: Proofs

Proof of Proposition 1: Let P denote the $n \times n$ lower-triangular Cholesky decomposition of Σ with positive elements on the diagonal and let $\Phi = [\Phi_1' \; \Phi_2' \; \cdots \; \Phi_H']'$ where Φ_i is the ith reduced-form vector moving average coefficient matrix. Then it follows from Theorem 2 in Arias et al. (2019) and equation (3) in Inoue and Kilian (2019) that

$$\begin{aligned} f(\theta) &= \left| \frac{\partial[\mathrm{vec}(A)' \; \mathrm{vec}(P)' \; s']'}{\partial \theta} \frac{\partial[\mathrm{vec}(A)' \; \mathrm{vec}(P)' \; s']}{\partial \theta'} \right| f(B, \mathrm{vech}(P), s) \\ &= \left| \frac{\partial[\mathrm{vec}(A)' \; \mathrm{vec}(P)' \; s']'}{\partial \theta} \frac{\partial[\mathrm{vec}(A)' \; \mathrm{vec}(P)' \; s']}{\partial \theta'} \right| f(B|\Sigma) f(\Sigma) f(s) \\ &= \left| \frac{\partial[\mathrm{vec}(\Phi)' \; \mathrm{vec}(P)' \; s']'}{\partial \theta} \right| \left| \frac{\partial \mathrm{vec}(\Phi)'}{\partial \mathrm{vec}(A)} \frac{\partial \mathrm{vec}(\Phi)}{\partial \mathrm{vec}(A)'} \right|^{\frac{1}{2}} f(B|\Sigma) f(\Sigma) f(s). \end{aligned} \tag{A.1}$$

It follows from equations (13) and (17) in Inoue and Kilian (2019) that the first Jacobian is given by

$$\left| \frac{\partial[\mathrm{vec}(\Phi)' \; \mathrm{vec}(P)' \; s']'}{\partial \theta} \right| = |(U' \otimes I_n) L_n' \; (I_n \otimes P) J_U| |P|^{nH} = 2^{-\frac{n(n-1)}{2}} |I_n + Q|^{n-1} \Pi_{i=1}^{n-1} a_{ii}^{n-i}, \tag{A.2}$$

where

$$J_U = 2[(I_n - S'^{-1} \otimes (I_n - S)^{-1}] \tilde{D}_n. \tag{A.3}$$

Equations (16) and (18) in Inoue and Kilian (2019) imply that

$$\left| \frac{\mathrm{vech}(\Sigma)}{\mathrm{vech}(P)'} \right| = \left| D_n^+ [(P \otimes I_n) + (I_n \otimes P) K_n] L_n' \right| = 2^n \Pi_{i=1}^n p_{ii}^{n-i+1}, \tag{A.4}$$

where p_{ii} is the (i, i)th element of P.

Proof of Proposition 2: It follows from equation 12 in Inoue and Kilian (2019) that

$$\frac{\partial \theta}{\partial s'} = \begin{bmatrix} (I_n \otimes P)J_U \\ (I_n \otimes \Phi P)J_U \end{bmatrix} \tag{A.5}$$

Thus Proposition 2 follows from Theorem 2 in Arias et al. (2018).

References

1. Alquist, R., Kilian, L. 2010. What do we learn from the price of crude oil futures? Journal of Applied Econometrics 25, 539-573. https://doi.org/10.1002/jae.1159

2. Antolin-Diaz, J., Rubio-Ramirez, J.F., 2018. Narrative sign restrictions for SVARs. American Economic Review 108, 2802-2839. https://doi.org/10.1257/aer.20161852

3. Arias, J., Rubio-Ramirez, J.F., Waggoner, D.F., 2018. Inference based on SVARs identified with sign and zero restrictions: Theory and applications. Econometrica 86, 685-720. https://doi.org/10.3982/ECTA14468

4. Baumeister, C., Hamilton, J.D., 2015. Sign restrictions, vector autoregressions, and useful prior information. Econometrica 83, 1963-1999. https://doi.org/10.3982/ECTA12356

5. Baumeister, C., 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. https://doi.org/10.1016/j.jmoneco.2018.0

6. Baumeister, C., Hamilton, J.D., 2019. Structural interpretation of vector autoregressions with incomplete identification: Revisiting the role of oil supply and demand shocks. American

Economic Review 109, 1873-1910. https://doi.org/10.1257/aer.20151569

7. Baumeister, C., Hamilton, J.D., 2020. Drawing conclusions from structural vector autoregressions identified on the basis of sign restrictions. Journal of International Money and Finance, forthcoming.

8. Bruder, S., Wolf, M., 2018. Balanced bootstrap joint confidence bands for structural impulse response functions. Journal of Time Series Analysis 39, 641-664. https://doi.org/10.1111/jtsa.12289

9. Chernozhukov, V., H. Hong, H., Tamer, E., 2007. Estimation and confidence regions for parameter sets in econometric models. Econometrica 75, 1243-1283. https://doi.org/10.1111/j.1468-0262.2007.00794.x

10. Coibion, O., Gorodnichenko, Y., 2015. Is the Phillips curve alive and well after all? Inflation expectations and the missing disinflation. American Economic Journal: Macroeconomics 7, 197-232. https://doi.org/10.1257/mac.20130306

11. Cross, J.L., Nguyen, B.H., Tran, T.D., 2020. The role of precautionary and speculative demand in the global market for crude oil. Manuscript, Norwegian School of Business.

12. Giacomini, R., Kitagawa, T. 2020. Robust Bayesian inference for set-identified models. Econometrica, forthcoming.

13. Herrera, A.M., Rangaraju, S.K., 2020. The effects of oil supply shocks on U.S. economic activity: What have we learned? Journal of Applied Econometrics 35, 141-159. https://doi.org/10.1002/jae.2735

14. Inoue, A., Kilian, L., 2013. Inference on impulse response functions in structural VAR models. Journal of Econometrics, 177, 1-13. https://doi.org/10.1016/j.jeconom.2013.02.009

15. Inoue, A., Kilian, L., 2016. Joint confidence sets for structural impulse responses. Journal of Econometrics 192, 421-432. https://doi.org/10.1016/j.jeconom.2016.02.008

16. Inoue, A., Kilian, L. 2019. Corrigendum to 'Inference on impulse response functions in structural VAR models' [J. Econometrics 177 (2013), 1-13]. Journal of Econometrics. 209, 139-143. https://doi.org/10.1016/j.jeconom.2017.08.020

17. Inoue, A., Kilian, L., 2020. Joint Bayesian inference about impulse responses in VAR models. Manuscript, Federal Reserve Bank of Dallas.

18. Jaccard, P., 1908. Nouvelles recherches sur la distribution florale. Bulletin de la Société Vaudoise des Sciences Naturelles 44, 223-270.

19. Karlsson, S., 2013. Forecasting with Bayesian vector autoregressions. Handbook of Economic Forecasting, vol 2B, ed. by G. Elliott and T. Timmermann. North Holland, Elsevier, 791-897. https://doi.org/10.1016/B978-0-444-62731-5.00015-4

20. Kilian, L., 2017. The impact of the fracking boom on Arab oil producers. Energy Journal 38, 137-160. https://doi.org/10.5547/01956574.38.6.lkil

21. Kilian, L. 2019. Facts and fiction in oil market modeling. Manuscript, Federal Reserve Bank of Dallas.

22. Kilian, L., Lee, T.K. 2014. Quantifying the speculative component in the real price of oil: The role of global oil inventories. Journal of International Money and Finance 42, 71-87. https://doi.org/10.1016/j.jimonfin.2013.08.005

23. Kilian, L., Lütkepohl, H., 2017. Structural vector autoregressive analysis. Cambridge University Press, New York. https://doi.org/10.1017/9781108164818

24. Kilian, L., 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, 1166-1188. https://doi.org/10.1111/j.1542-4774.2012.01080.x

25. Kilian, L., Murphy, D.P. 2014. The role of inventories and speculative trading in the global market for crude oil. Journal of Applied Econometrics 29, 454-478. https://doi.org/10.1002/jae.2322

26. Kilian, L., Zhou, X. 2020a. Oil Prices, Gasoline Prices and Inflation Expectations: A New Model and New Facts. Manuscript, Federal Reserve Bank of Dallas.

27. Kilian, L., Zhou, X. 2020b. Does drawing down the U.S. strategic petroleum reserve help stabilize oil prices? Journal of Applied Econometrics, forthcoming. https://doi.org/10.1002/jae.2798

28. León, C.A., Massé, J.-C., Rivest, L.P., 2006. A statistical model for random rotations. Journal of Multivariate Analysis 976, 412-430. https://doi.org/10.1016/j.jmva.2005.03.009

29. Lütkepohl, H., Staszewska-Bystrova, A., Winker, P., 2015a. Confidence bands for impulse responses: Bonferroni vs. Wald. Oxford Bulletin of Economics and Statistics 77, 800-821. https://doi.org/10.1111/obes.12114

30. Lütkepohl, H., Staszewska-Bystrova, A., Winker, P., 2015b. Comparison of methods for constructing joint confidence bands for impulse response functions. International Journal of Forecasting 31, 782-798. https://doi.org/10.1016/j.ijforecast.2013.08.003

31. Lütkepohl, H., Staszewska-Bystrova, A., Winker, P., 2018. Calculating joint confidence bands for impulse response functions using highest density regions. Empirical Economics 55, 1389-1411. https://doi.org/10.1007/s00181-017-1325-3

32. Magnus, J.R., 1988. Linear Structures. Oxford University Press, Oxford, UK.

33. Magnus, J.R., Neudecker, H., 1999. Matrix differential calculus with applications in statistics and econometrics, 2nd ed.Wiley, Chichester, UK.

34. Montiel Olea, J.L., Plagborg-Møller, M., 2019. Simultaneous confidence bands: Theory,

implementation, and an application to SVARs. Journal of Applied Econometrics 34, 1-17. https://doi.org/10.1002/jae.2656

35. Plagborg-Møller, M., 2019. Bayesian inference on structural impulse response functions. Quantitative Economics 10, 145-184. https://doi.org/10.3982/QE926

36. Rausser, G., and Stūrmer, M., 2020. A dynamic analysis of collusive action: The case of the world copper market, 1882-2016. Manuscript, Federal Reserve Bank of Dallas.

37. Rubio-Ramirez, J.F., Waggoner, D., Zha, Z., 2010. Structural vector autoregressions: Theory of identification and algorithms for inference. Review of Economic Studies 77, 665-696. https://doi.org/10.1111/j.1467-937X.2009.00578.x

38. Sims, C.A., Zha, T., 1999. Error bands for impulse responses. *Econometrica* 67, 1113-1156. https://doi.org/10.1111/1468-0262.00071

39. Uhlig, H., 2005. What are the effects of monetary policy on output? Results from an agnostic identification procedure. Journal of Monetary Economics 52, 381-419. https://doi.org/10.1016/j.jmoneco.200

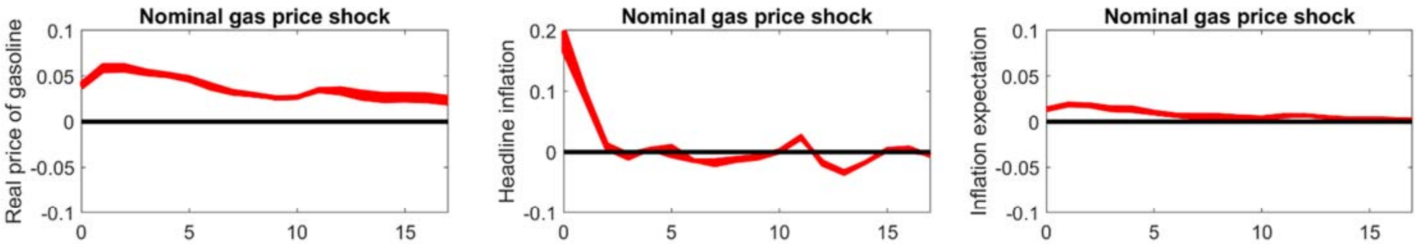
40. Uhlig, H., 2017. Shocks, sign restrictions, and identification. Advances in Economics and Econometrics: Eleventh World Congress, edited by Honoré, B., Pakes, A., Piazzesi, M., Samuelson, L., Cambridge University Press, pp. 95-127. https://doi.org/10.1017/9781108227223.004

41. Watson, M.W., 2020. Comment on 'On the empirical (ir)relevance of the zero lower bound' by D. Debortoli, J. Gali, and L. Gambetti. NBER Macroeconomics Annual 2019 34, 182-193. https://doi.org/10.1086/707179

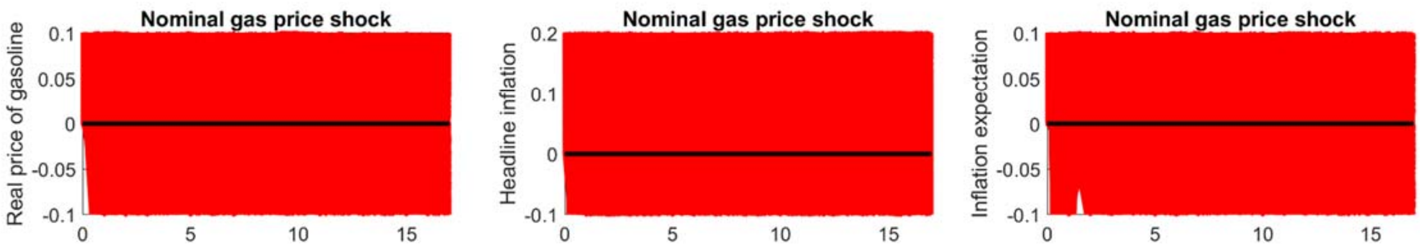
42. Zhou, X., 2020. Refining the workhorse oil market model, Journal of Applied Econometrics 35, 130-140. https://doi.org/10.1002/jae.2743

Figure 1: Identified set under prior versus identified set conditional on MLE

(a) Estimate of identified set conditional on MLE

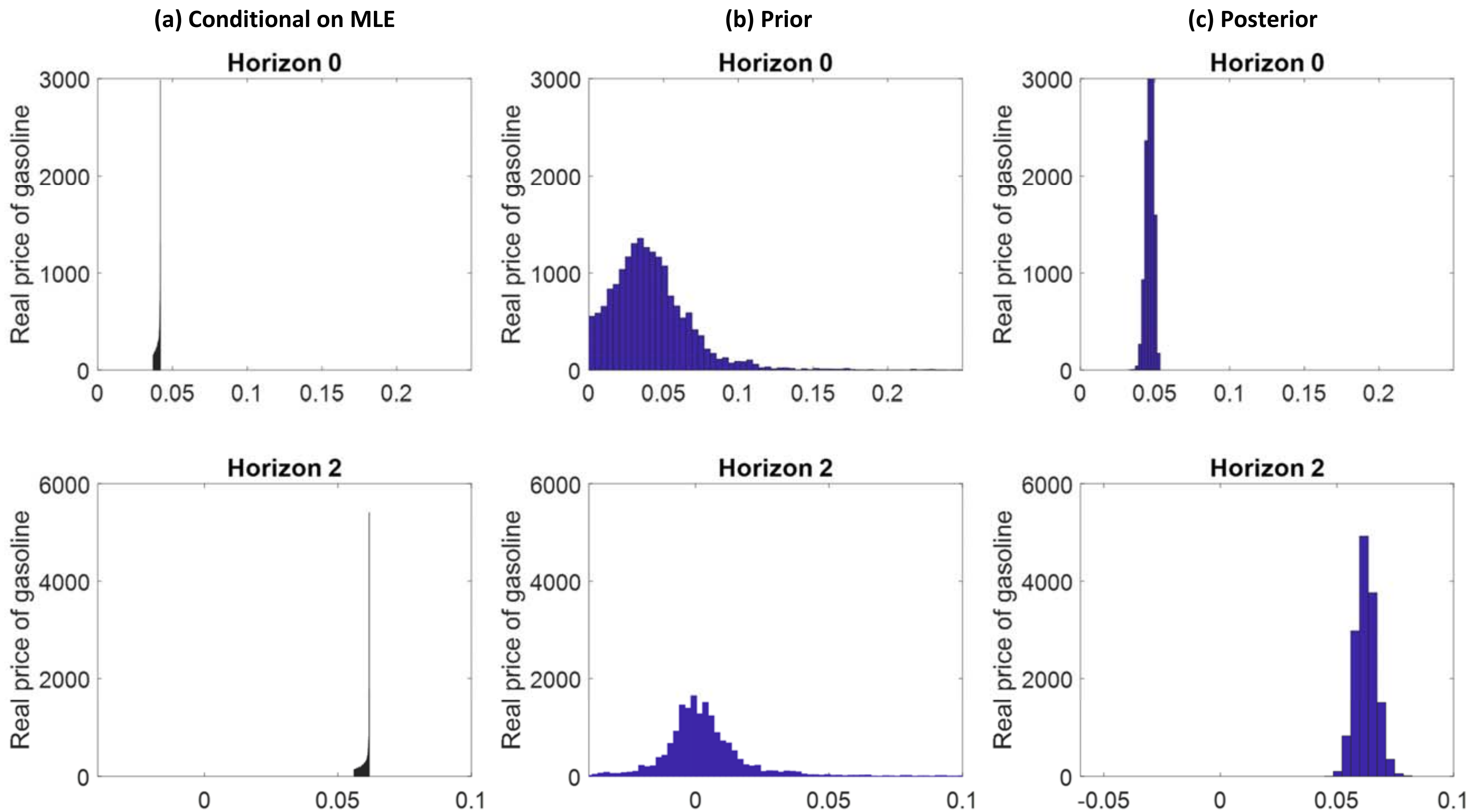


(b) Estimate of identified set under prior



NOTES: The axes in panel (b) have been adjusted to make the plots comparable with panel (a).

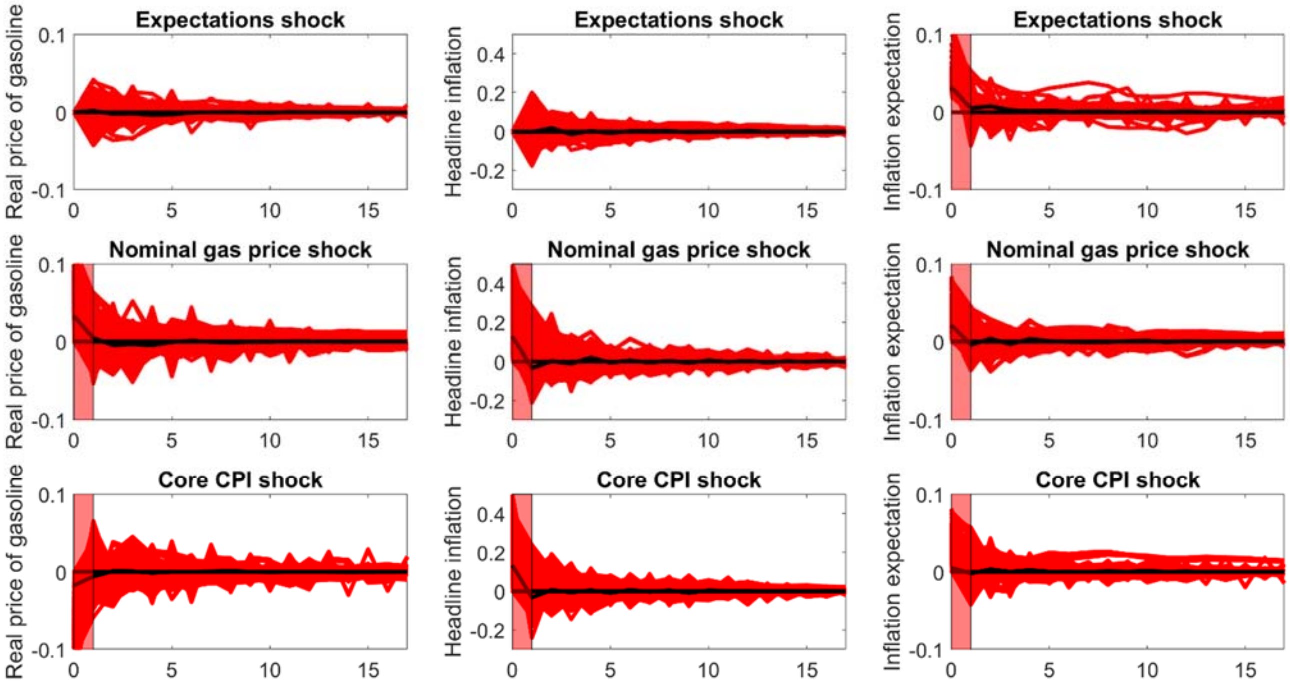
Figure 2: Histograms over identified sets of selected responses to the nominal gasoline price shock



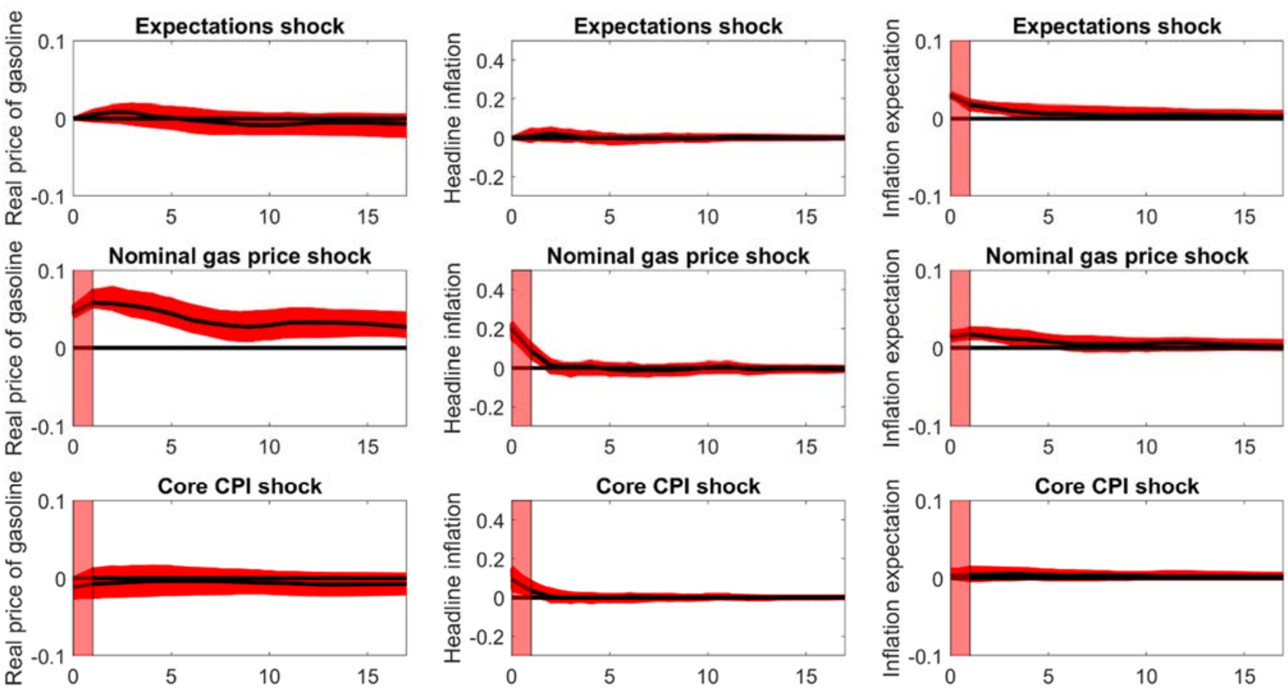
NOTES: The horizontal axis has been adjusted to make the plots comparable.

Figure 3: Impulse response inference in inflation expectations model under prior 1

(a) Impulse response prior



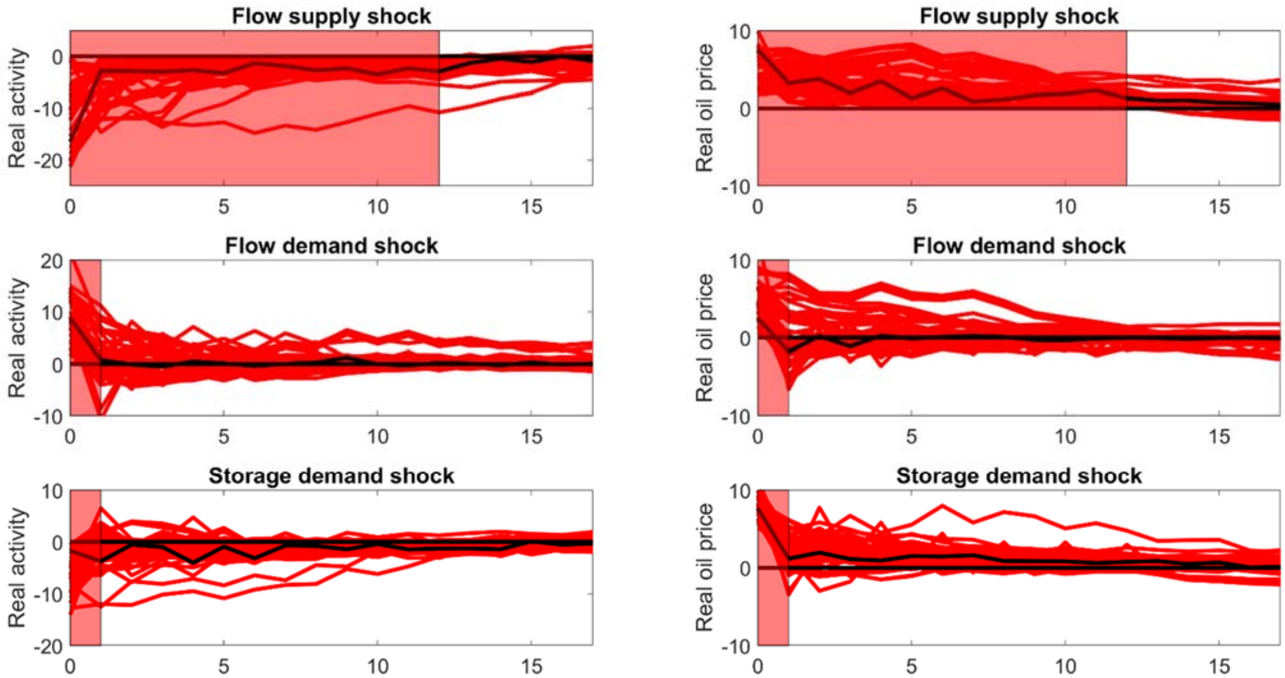
(b) Impulse response posterior



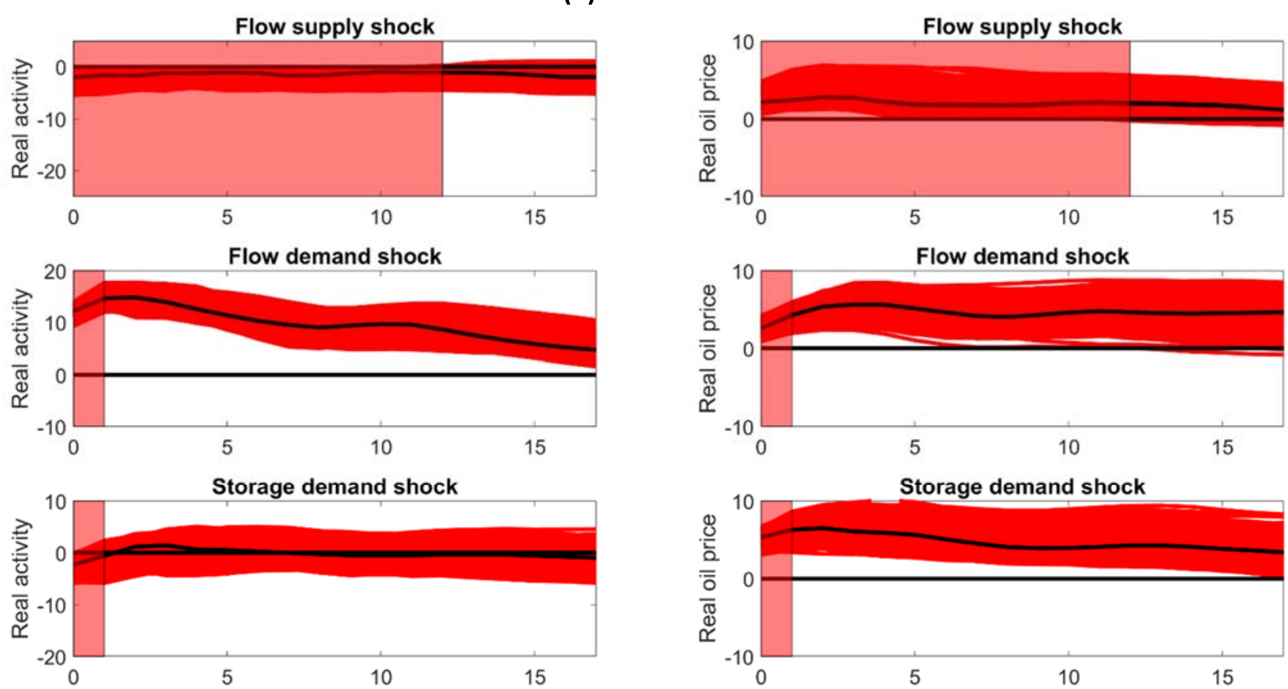
NOTES: Horizons at which sign restrictions are imposed are shown as shaded areas. Bayes estimates under absolute loss are shown in black and 68% joint credible sets in red.

Figure 4: Impulse response inference in Kilian and Murphy (2014) global oil market model

(a) Prior



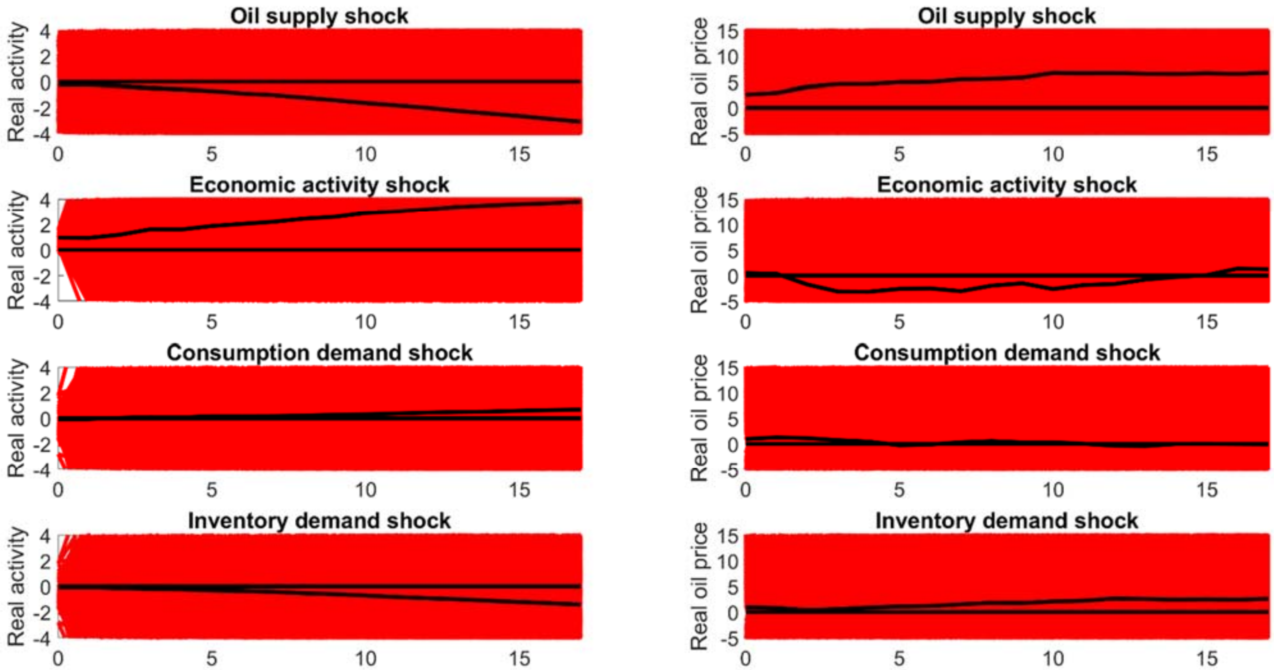
(b) Posterior



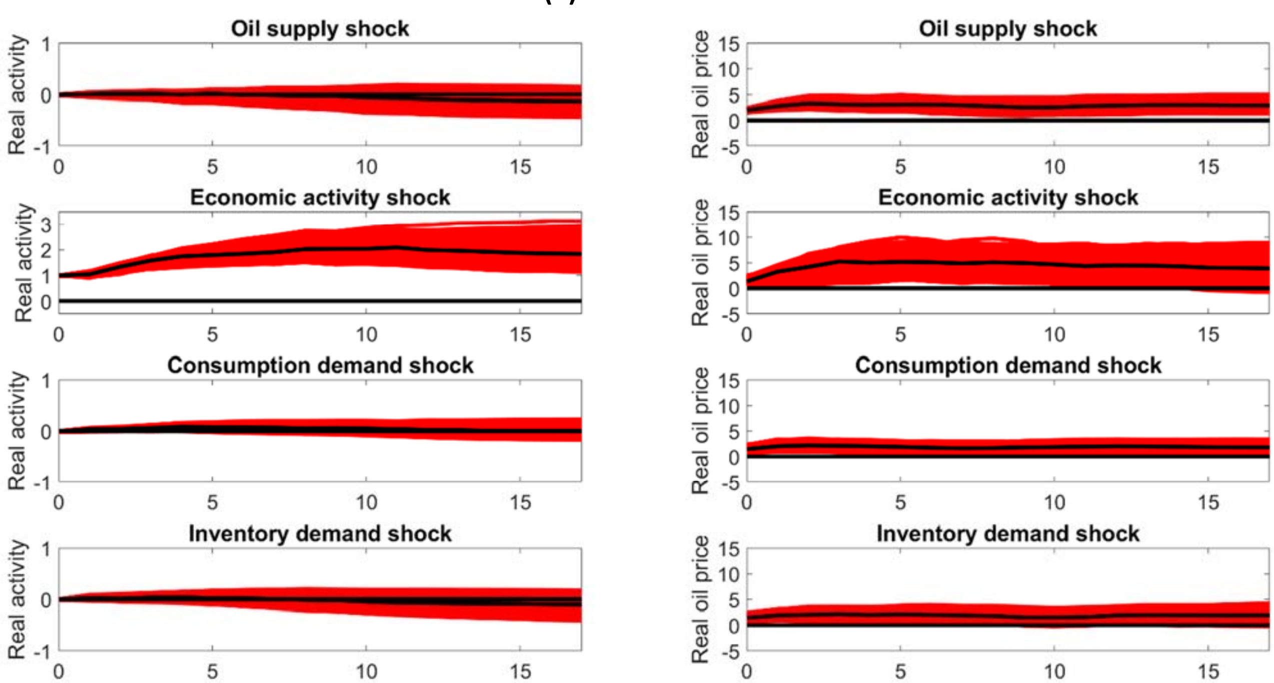
NOTES: Horizons at which sign restrictions are imposed are shown as shaded areas. Bayes estimates under absolute loss are shown in black and 68% joint credible sets in red.

Figure 5: Impulse response inference in Baumeister and Hamilton (2019) global oil market model

(a) Prior



(b) Posterior



NOTES: Bayes estimates under absolute loss are shown in black and 68% joint credible sets in red.

Table 1: Alternative Reduced-Form Minnesota Prior Specifications

	Prior mean of first own lag	Prior Innovation variances
Prior 1	0	AR(1) innovation variance
Prior 2	0.9	AR(1) innovation variance
Prior 3	0	0.01
Prior 4	0.9	0.01

Table 2a: The Extent to Which the Impulse Response Priors are Overturned Based on the Data
Inflation Expectations Model

	Hausdorff distance metric (Percentage points)	Absolute location shift in estimator of responses (Percentage points)		Percent change in dispersion about estimator	Jaccard sign distance metric
	d_H	d_L		d_C	d_J
		Max	Mean		
Prior 1	2.9	0.12	0.01	-51.6	1
Prior 2	172.8	0.11	0.03	-98.2	1
Prior 3	2.9	0.12	0.01	-47.1	1
Prior 4	170.0	0.13	0.02	-98.1	1

Table 2b: Sensitivity of the Impulse Response Posterior to Alternative Priors
Inflation Expectations Model

	All Results Normalized Relative to Impulse Response Posterior Based on Prior 1				
	Hausdorff distance metric (Percentage points)	Absolute location shift in estimator of responses (Percentage points)		Percent change in dispersion about estimator	Jaccard sign distance metric
	d_H	d_L		d_C	d_J
		Max	Mean		
Prior 2	0.5	0.02	0.00	0.2	0.97
Prior 3	0.6	0.02	0.00	5.9	1.00
Prior 4	0.6	0.01	0.00	8.2	1.00

Table 3: The Extent to Which the Impulse Response Priors are Overturned Based on the Data
Kilian and Murphy (2014) Global Oil Market Model

	Hausdorff distance metric (Percentage points)	Absolute location shift in estimator of responses (Percentage points)		Percent change in dispersion about estimator	Jaccard sign distance metric
	d_H	d_L		d_C	d_J
		Max	Mean		
Real activity	248.5	14.94	4.16	-36.5	1.00
Real oil price	198.3	6.73	3.09	-18.6	0.98

Table 4: The Extent to Which the Impulse Response Priors are Overturned Based on the Data
Baumeister and Hamilton (2014) Global Oil Market Model

	Hausdorff distance metric (Percentage points)	Absolute location shift in estimator of responses (Percentage points)		Percent change in dispersion about estimator	Jaccard sign distance metric
	d_H	d_L		d_C	d_J
		Max	Mean		
Real activity	244.6	2.95	0.75	-93.2	1.00
Real oil price	370.0	8.34	2.63	-74.7	1.00