Identification through Heterogeneity

Pooyan Amir-Ahmadi
Thorsten Drautzburg

CESIFO WORKING PAPER NO. 6359
CATEGORY 6: FISCAL POLICY, MACROECONOMICS AND GROWTH
FEBRUARY 2017

ISSN 2364-1428

CESifo *Working Paper No. 6359*

Identification through Heterogeneity

Abstract

Set identification in Bayesian vector autoregression (VARs) is becoming increasingly popular while facing recent criticism about potentially unwanted prior dominance and underrepresented bounds of the identified set. This can lead to biased inference even in large samples. Common estimation strategies in high dimensions or with tight restrictions can prove to be highly inefficient or even practically infeasible. We propose to include micro data on heterogeneous entities for the estimation and identification of vector autoregressions to achieve sharper inference. First, we provide conditions when imposing a simple ranking of impulse responses will sharpen inference in bivariate and trivariate VARS. Importantly, we show that this sharpening also applies to variables not subject to ranking restrictions. Second, we develop two types of inference to address recent criticism: (i) A prior-robust posterior over the bounds of the identified set and (ii) a fully Bayesian sampling algorithm that allows us to efficiently include an agnostic prior over the non-identifiable parameters. Third, we apply our methodology to US data to identify productivity news and defense spending shocks. We find that under both algorithms the bounds of the identified sets shrink substantially under heterogeneity restrictions relative to standard sign restrictions.

JEL-Codes: C320, E320, E620.

Keywords: Bayesian VAR, sign restrictions, set identification, micro data, news shocks, defense spending.

Pooyan Amir-Ahmadi
Department of Economics
University of Illinois
at Urbana-Champaign
Urbana / Illinois / USA
pooyan@illinois.edu

Thorsten Drautzburg
Federal Reserve Bank of Philadelphia
Philadelphia / PA / USA
tdrautzburg@gmail.com

February 14, 2017
We would like to thank Jonás Arias, Mark Bognanni, Jesús Fernández-Villaverde, Frank Schorfheide, Mark Watson, Jonathan Wright, and audiences at the FRB Philadelphia, MEG 2016, NBER-NSF 2016 SBIES, U Penn, and SEA 2016 for comments and Nick Zarra for research assistance. All remaining errors are our own and the views expressed herein are our own views only. They do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia, the Federal Reserve System, or its Board of Governors.

1 Introduction

Since the seminal paper by Sims (1980), the structural vector autoregressive (SVAR) model remains the workhorse for assessing the dynamic causal effects of economic shocks that drive macroeconomic variables. The model by itself is under-identified and requires additional prior identifying assumptions from outside the model which has been the key challenge in the SVAR literature since. Typical inference methods impose zero restrictions to achieve just- or over-identification, such as short-run schemes (Sims, 1980; Christiano et al., 1999; Sims and Zha, 1998) or long-run schemes (Blanchard and Quah, 1989). But even though the resultant point identification is desirable, the underlying assumptions are often too strong and restrictive to defend confidently and rarely backed by theory according to Canova (2007). Alternatively, Uhlig (2005), Faust (1998), and Canova and De Nicolo (2002) have pioneered the "*agnostic identification*" approach, proposing to identify shocks by restricting the sign of the responses of selected variables consistent with conventional wisdom. This approach relaxes strong zero restrictions and has become increasingly popular in the recent literature. It belongs to the econometric class of set identified models.

We make three contributions to the class of set identified dynamic time series models: (1) We show how to use micro data for the sharper identification of macroeconomic shocks by imposing a ranking of responses. (2) We characterize analytically when ranking restrictions on one set of variables, including, but not limited to micro data, can sharpen the inference for another variable, such as a macro time series. (3) We propose two inference methods: As a *conservative* inference method, we devise a sampling algorithm to transparently characterize the posterior bounds of set identified shocks. This algorithm is immune to any shape of the prior distribution. We also devise a new efficient sampling algorithm that characterize the posterior distribution over the identified set for an agnostic prior. While most common algorithms within the Bayesian framework lead to a large fraction of draws to be inadmissible empty sets, our approach finds for all posterior reduced form draws the corresponding identified set and its distribution. This is particularly important when restrictions are tight or the application is complex, for example, because of its high dimensionality.

Heterogeneity restrictions can stem from different a priori shock elasticities of different industries, regions or households. A-priori, heterogeneity could be supported or derived either from theory or data. For example, consider identifying a fiscal policy shock as an unexpected increase in government defense spending. Manufacturing industry A might have, a priori, a (much) higher exposure to those shocks relative to sector B, if the government is a key client in the former industry and a negligible client in the latter industry. Hence, we may reasonably expect industry A to expand more than B following a defense spending shock. We label such restrictions *heterogeneity restrictions*. In this paper we show both analytically and in two quantitative applications that our approach sharpens inference substantially, even by the challenging standards of the recent set-identified VAR literature.

Recent contributions on set identified SVARs have pointed to limitations related to (i) the correct and transparent measures of uncertainty about the bounds on the identified impulse response functions (Moon and Schorfheide, 2012; Moon et al., 2013; Giacomini and Kitagawa, 2014) and (ii) the unwanted dominance of seemingly uninformative priors (Baumeister and Hamilton, 2015; Arias et al., 2014). Baumeister and Hamilton (2015) explore the role of the typical uniform prior and its one-to-one dominance in the posterior. They explore prior elicitation directly on the structural representation of the SVAR by truncating its direction to be consistent with the imposed sign restrictions derived from theory defining a shock of interest to be identifed. Giacomini and Kitagawa (2014) propose a method to conduct inference that is robust to the choice of prior on the non-identifiable part, i.e. the rotation matrix mapping reduced form to structural impulse vectors. Their approach delivers multiple posteriors, allows to summarize the corresponding robustified credible regions and bounds on the posterior mean. Similar to them, we also pursue prior-robust inference, in the spirit of not imposing distributions one may be uncomfortable with. However, we focus on characterizing bounds on the impulse-responses, following Moon and Schorfheide (2012) and Moon et al. (2013). To address unwanted prior dominance Moon and Schorfheide (2012) suggest to construct and report credible sets for the identified set by providing a "*transparent parametrization*" of the set. They discuss how any prior, no matter how uninformative, can lead to overly informative inference. Bayesian credible sets thus lie strictly within the frequentist identified set. Building on Moon and Schorfheide (2012), Moon et al. (2013) provide a valid frequentist approach to inference in sign restricted VARs. We report prior robust posteriors using the quasi-analytical characterization of the identified set of impulse-responses in the spirit of Moon et al. (2013).

We combine this literature with a separate advance in empirical macroeconomics: The use of the growing microeconomic datasets to answer macroeconomic questions. Including disaggregate data by itself holds three promises. First, richer, *heterogeneous* data tends to exhibit more variation which helps for identification, as in Nekarda and Ramey (2011). Second, it provides a more detailed analysis of the transmission mechanism and the heterogeneous effects of shocks themselves. For example, Anderson et al. (forthcoming) analyze the transmission of monetary policy shocks to groups of households and Perotti (2008) looks how fiscal shocks transmit to different industries. Third, the richer data may better reflect the relevant information set of economic agents which can be important in the presence of news shocks (Beaudry and Portier, 2014).

Specifically, we use the extra dimension of *heterogeneous* data both for the estimation and importantly the identification of SVARs. We show that it is particularly helpful for the identification: In bi- and trivariate VARs we provide analytical conditions under which imposing heterogeneity restrictions leads to strictly smaller identified sets for the responses to a macroeconomic shock while avoiding unwanted prior information by following the approach of Moon and Schorfheide (2012) and Moon et al. (2013). As we highlight in our analysis of the trivariate VAR, heterogeneity restrictions can also lead to set reduction for responses of traditional macro variables that are not directly affected by heterogeneity restrictions. This occurs under simple conditions on the reduced form correlation between forecast errors. We verify in two applications that this set reduction for variables not directly affected by heterogeneity restrictions is also relevant in practice.

Where do heterogeneity restrictions come from and how can we motivate them? Heterogeneity restrictions could come from theory or could be based on estimates from separate sets of micro data. We provide two applications and take the latter route to motivate heterogeneity restrictions. In our first application we identify the dynamic effects of productivity news shocks while in our second application we identify fiscal policy shocks. Beaudry and Portier (2006) have argued that productivity news shocks are important drivers of business cycles. Our heterogeneity assumption is that productivity news move the stock returns of R&D intensive industries more. We measure the R&D intensity using Compustat data for either five or ten Fama and French (1997) industries. We find that the restriction that more innovative sectors responds more sharpens our inference substantially: Prior-robust bounds on the impulse-responses shrink by up to 45% for consumer confidence and by more than one third for output and employment. The fully Bayesian responses have a much more pronounced shape with heterogeneity restrictions and imply a slow increase in TFP relative to its trend that is consistent with technology diffusion. We also find that our conclusions change little when we impose a soft zero restriction on initial TFP, lending further support to our identification scheme.

In our second application we identify a defense spending shock financed through higher taxes. In the spirit of Nekarda and Ramey (2011) we characterize the macroeconomic effects with the help of the differential effects on manufacturing industries. Our heterogeneity assumption is that shipments of all manufacturing industries rise, but more so in industries with a higher share of sales to the government, as measured by the input-output linkages computed by Nekarda and Ramey (2011). One of the variables included in this application is real federal debt and its response is left unrestricted. With heterogeneity restrictions, but not with sign restrictions, we find evidence that despite the tax increase, federal debt rises in response to spending shocks. Intriguingly, this finding already applies to the identified set at the posterior mean, highlighting the power of heterogeneity restrictions at the micro level for macro variables.

While our focus is on identifying impulse-response functions, the identification scheme also has implications for other moments derived from the identified shocks in VARs. For example, Arias et al. (2015) examine the implied policy rules of a set identified SVAR, while Uhlig (2003) directly targets the maximization of the forecast error variance decomposition (FEVD) of selected variables to identify specific shocks. Focusing on FEVD, we show how to compute the identified set of the explained variance that is consistent with sign- or heterogeneity restrictions, building on the work by Faust (1998). Typically, reported FEVD in set identified SVARs tend to be very wide, hence equally consistent with theories that either render driving shock being negligible or the key driver. See Uhlig (2005) for a comparison and discussion of FEVD under agnostic and recursive identification in the case of contractionary monetary policy shocks. In our applications we show that the heterogeneity restrictions reduce the upper bound of the forecast error variance significantly.

Our approach and our contributions naturally carry over to factor-augmented VARs and dynamic factor models in general as in Amir-Ahmadi and Uhlig (2015), panel VARs as in De Graeve and Karas (2014),[1] or time-varying parameter VARs with stochastic volatility (henceforth TVP-VAR) popularized by Primiceri (2005) and Cogley and Sargent (2005).Our proposed identification scheme is independent of the specific statistical model and our analytical results regarding its efficiency are static and thus independent of how the dynamics are modeled. In addition, the algorithms we develop can also easily be applied to more general VAR-style models. Given that inference in these models is already more demanding, having an efficient algorithm for sampling over the identified set becomes even more important than in our VAR application.

This paper is structured as follows. First, we set up the general statistical model and identification problem. In this general framework, we discuss prior-robust inference and provide analytical characterizations of the identified sets in bivariate and trivariate models. Second, we discuss a simple estimation algorithm that recovers the full identified set. Third, we provide two empirical examples of large VARs applied to data from different US industries and US metro areas.

2 Model

Here we set up the standard Bayesian VAR framework that we work with throughout the paper before we define sign and heterogeneity restrictions. We discuss that heterogeneity restrictions weakly narrow the identified set and provide sufficient conditions for identified sets to have positive measure. To illustrate the concept of heterogeneity restrictions, we provide examples of possible applications. Last, we provide conditions when heterogeneity restrictions, compared to pure sign restrictions, lead to strict reduction and no set reduction in bivariate and trivariate VARs.

2.1 Setup

We work with a Gaussian VAR with a conjugate prior over the identifiable reduced form parameters. Specifically, the $p \times 1$ vector of observables Y_t depends on k lags and has *iid* normally distributed forecast errors e_t.

$$Y_t = \mu + \sum_{\ell=1}^{k} B_\ell Y_{t-\ell} + e_t, \qquad e_t \overset{iid}{\sim} \mathcal{N}(0, \Sigma). \tag{2.1}$$

Structural VARs are underidentified and require a number of additional restrictions to provide a one-one mapping of the reduced form innovations e_t to structural shocks ϵ_t by factoring the variance-covariance matrix Σ. This can be summarized by the following relation:

$$e_t = A\epsilon_t, \qquad \epsilon_t \overset{iid}{\sim} \mathcal{N}(0, I_p), \qquad \Sigma = AA'. \tag{2.2}$$

[1] De Graeve and Karas (2014) have used what we call identification through heterogeneity to identify shocks from banking panics in a data set of Russian banks. While their application is intriguing, our focus here is different. First, we focus on formally characterizing asymptotically identifiable sets. Second, we propose to generalize their specific application to exploit heterogeneity in a broad range of applications and dynamic econometric models.

In addition to this generic VAR restriction, we want to impose restrictions on the signs of impulse-response functions. We now lay out the notation needed to formalize these restrictions. We define impulse vectors following Uhlig (2005):

Definition 1. The vector $a \in \mathbb{R}^p$ is called an *impulse vector*, iff there is some matrix A, so that $AA' = \Sigma$ and so that a is a column of A.

Trivially, the columns of the lower Cholesky matrix $\tilde{A}$ are impulse vectors – but generally not structural impulses. In addition, for any rotation matrix $Q = [q_1, \dots, q_p]$, the columns of $\tilde{A}Q$ are impulse vectors. Thus, without loss of generality we can express impulse vectors as:

$$a = \tilde{A}q, \qquad ||q|| = 1. \tag{2.3}$$

We use $||\cdot||$ to denote the Euclidean norm throughout.

In general, we can then write our full model as:

$$p(Y^T, B, \Sigma, Q) = \ell(B, \Sigma|Y^T)\pi_{rf}(B, \Sigma)\pi_Q(Q|B, \Sigma), \tag{2.4}$$

where ℓ is the likelihood function, π_{rf} denotes the prior over the reduced form parameters, and π_Q denotes the prior over Q that could be conditional on the reduced form parameters. We assume standard a conjugate prior over (B, Σ) and take these parameters as given for now. We will discuss estimation in the next section. Now we focus on what we can learn from beliefs about impulse responses and the reduced form parameters about Q. We proceed assuming that π_Q has full support over the identified set.

2.2 Sign and heterogeneity restrictions

To learn about Q and identify structural impulse vectors, we impose qualitative restrictions on the impulse-responses Q induces. To define these restrictions, we need extra notation. We use the companion form $X_t = B_X X_{t-1} + A\epsilon_t$ of the VAR (2.1) to express impulse-responses after the initial impact. The response at horizon h is then given by

$$r_a(h) = \begin{bmatrix} I_p & 0_{p,p\times(k-1)} \end{bmatrix} (B_X)^h \begin{bmatrix} a \\ 0_{p\times(k-1),1} \end{bmatrix} \tag{2.5}$$

We are now equipped to define sign restrictions, following Amir-Ahmadi and Uhlig (2015). Imposing sign restrictions is equivalent to picking a list $\mathbb{L}_{SR} \subseteq \{(s, n)|s \in \{-1, 1\}, n \in \{1, \dots, p\}\}$ of variables n and signs s as well as a restriction horizon $H \geq 0$[2].

Definition 2. The impulse vector a satisfies the *sign restrictions* $(\mathbb{L}_{SR}, H)$ iff $s \times r_a(h)_n \geq 0$ for all $(s, n) \in \mathbb{L}_{SR}$ and $h \in \{0, \dots, H\}$.

[2]Note, that extending the list to have have potentially different binding horizons for each pair of inequality restrictions would be straightforward. For ease of notation we refrain from this exposition here.

We define heterogeneity restrictions similarly, except that they are defined for a list of pairs of variables (n, m) and an associated strength $\lambda \in \mathbb{R}_+$. Define $\mathbb{L}_{HR} \subseteq \{(s, n, m, \lambda) | s \in \{-1, 1\}, (n, m) \in \{1, \ldots, p\}^2, \lambda(n - m) \neq 0, \lambda \geq 0\}$.

Definition 3. The impulse vector a satisfies the *heterogeneity restrictions* $(\mathbb{L}_{HR}, H)$ iff $s \times r_a(h)_n \geq \lambda s \times r_a(h)_m$ for all $(s, n, m, \lambda) \in \mathbb{L}_{HR}$ and $h \in \{0, \ldots, H\}$.

Note that with a specific prior over Q and for given percentiles of the posterior distribution, the heterogeneity restrictions can produce more dispersed posteriors percentiles: The tighter restrictions can shift mass away from the center of the prior towards the tails of the distribution, as we will see in our applications. For the (distribution-free) identified set, however, there is a clear sense in which heterogeneity restrictions are tighter than sign restrictions: Heterogeneity restrictions can nest the standard sign restrictions. If they do so, the identified set is weakly smaller.

Lemma 1. Write $\mathbb{L}_{SR} = \{(s^{(j)}, n^{(j)}) | j = 1, \ldots, J\}$ for the full set of sign restrictions and write $\mathbb{L}_{HR} = \{(s^{(j)}, n^{(j)}, m^{(j)}, \lambda^{(j)}) | j = 1, \ldots, J\}$ for the analogous set of heterogeneity restrictions. If for all $j = 1, \ldots, J$ $n_{SR}^{(j)} = n_{HR}^{(j)}$ and $\lambda^{(j)} \geq 0$, then identified the set for a induced by $\mathbb{L}_{HR}$ is weakly smaller than the set for a induced by $\mathbb{L}_{SR}$.

Proof. (Sketch.) Note that for $\lambda^{(j)} = 0$, the restrictions in $\mathbb{L}_{HR}^{(j)}$ imply the restrictions in $\mathbb{L}_{SR}^{(j)}$ given that $n_{SR}^{(j)} = n_{HR}^{(j)}$. □

Below we provide conditions under which the identified sets are strictly smaller than with pure sign restrictions in the cases of $p = 2$ and $p = 3$.

Heterogeneity restrictions may also apply when no sign restrictions are available because we can only sign the difference in the responses. For example, we might know that lump-sum fiscal transfers raise the expenditure of highly leveraged households more than those with low leverage. Depending on how the transfers are financed, some household might actually cut expenditures, for example if they pay most taxes. In that case we might want to impose only heterogeneity restrictions that do not nest the standard sign restrictions.[3]

Can heterogeneity restrictions be too tight and result in empty identified sets? We now provide sufficient conditions to guarantee a non-empty identified set. While the focus on impact restrictions is more restrictive than our empirical specifications, the same intuition applies when we can rule out overshooting responses or the restricted horizon is short enough.

Formally, if heterogeneity restrictions are imposed on impact only and satisfy the order condition $J \leq p$ and a rank condition, there is always a set of impulse-vectors a that are consistent with the heterogeneity restrictions.

Lemma 2. Assume $H = 0$, $J \leq p$, all $n^{(j)}$ are distinct. Let Λ be a $J \times p$ matrix of zeros, except for $\lambda^{(j)}$s in the $(j, m^{(j)})$ positions, $j = 1, \ldots, J$. Let E be a $J \times p$ matrix of zeros, except for ones

[3] In a sense, also in that case the heterogeneity restrictions are stronger as the sign restrictions could leave an unrestricted set of $a = \tilde{A}q$ subject only to $||q|| = 1$, while the heterogeneity restrictions imply restrictions.

in the $(j, n^{(j)})$ positions, $j = 1, \dots, J$. If $M \equiv E - \Lambda$ is of rank J, then the set for a induced by $\mathbb{L}_{HR}$ has positive (Lebesgue) measure.

Proof. Let S be a $J \times J$ diagonal matrix with the direction of the restrictions on its diagonal $S_{j,j} = s^{(j)}$. Then the heterogeneity restriction in $\mathbb{L}_{HR}$ are equivalent to

$$S \underbrace{(E - \Lambda)}_{=M} \tilde{A} q \geq 0.$$

Since M is of rank J by assumption, we can re-write $M = U[D, 0_{J\times(p-J)}]V$, where U, V are orthogonal matrices of dimension J and p, respectively. D is a $J \times J$ diagonal matrix with non-zero entries along its main diagonal. Define $\tilde{M} = V'[D^{-1}; 0_{(p-J)\times J}]U'$. Note that $M\tilde{M} = I$.

Now define

$$\tilde{q} \equiv \tilde{A}^{-1}\tilde{M}S^{-1}1_{J\times 1}.$$

Note that $\tilde{q}$ is non-zero. To see this, assume by contradiction that $\tilde{q} = 0_{p\times 1}$. Equivalently, after left-multiplying by $\tilde{A}$ and then by M, $M\tilde{M}S^{-1}1_{J\times 1} = M0_{p\times 1} = 0_{J\times 1}$. But $M\tilde{M} = I$ and since S is invertible, $1_{J\times 1} = 0_{J\times 1}$, a contradiction. Thus $||\tilde{q}|| > 0$.

Let $q \equiv \frac{\tilde{q}}{||\tilde{q}||}$. Then:

$$SM\tilde{A}q = ||\tilde{q}||^{-1}SM\tilde{A}\tilde{A}^{-1}\tilde{M}S^{-1}1_{J\times 1} = ||\tilde{q}||^{-1}1_{J\times 1} > 0,$$

where the inequality is taken elementwise. Since the inequality is strict, by continuity there exists a $\delta > 0$ such that all $\hat{q}$ with $||\tilde{q} - \hat{q}|| < \delta$ small enough can be rescaled so that $\tilde{A}\frac{\hat{q}}{||q||}$ satisfies $\mathbb{L}_{HR}$. Thus, the set of admissible a has positive (Lebesque) measure. □

This Lemma is also useful to guide the design of heterogeneity restrictions. To see this note that if the rank of M equals $R < J$ only a degenerate solution with zero Lebesgue measure may exist. Consider the case that $J = 2$ and $M = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$. In that case only $q \propto \left[1, \frac{\tilde{A}_{11} - \tilde{A}_{21}}{\tilde{A}_{22}}\right]$, scaled to have unit norm, is a possible solution. Thus if we want to increase the odds that a non-degenerate solution exists we have to rule out cycles: This is natural on economic grounds, but we need to formalize this notion. Restricting ourselves to no more restrictions than variables and focusing on chains of restrictions is sufficient for the rank condition in Lemma 2.

In our application, we always impose heterogeneity restrictions for groups of variables. While this restriction is by no means necessary, we now show that this type of restriction is sufficient for the rank condition in the previous Lemma 2.

Corollary 1. Assume $H = 0$, $J \leq p$, all $n^{(j)}$ are distinct, and there is at most one restriction $\mathbb{L}_{HR}^{(j)}$ with $m^{(j)} = n$ and $\lambda^{(j)} > 0$ for each variable $n = 1, \dots, p$. Furthermore, heterogeneity restrictions come in non-overlapping groups $\mathbb{G} = \{j1, j2, \dots, \bar{j}\}$ with $s^{(j)} = s^{(\ell)} = s^{\mathbb{G}}$ for all $j, \ell \in \mathbb{G}$ with one $\lambda^{(j1)} = 0$, i.e.:

$$
\begin{aligned}
&0 \leq s^{\mathbb{G}} r(a)_{n^{(j1)}} \\
&s^{\mathbb{G}} \lambda^{(j2)} r(a)_{n^{(j1)}} = s^{\mathbb{G}} \lambda^{(j2)} r(a)_{m^{(j2)}} \leq r(a)_{n^{(j2)}} \qquad \text{using } n^{(j1)} = m^{(j2)} \\
&\dots \\
&s^{\mathbb{G}} \lambda^{(\bar{j})} r(a)_{m^{(\bar{j})}} \leq r(a)_{n^{(\bar{j})}}
\end{aligned}
$$

Then the set for a induced by $\mathbb{L}_{HR}$ has positive (Lebesgue) measure.

Proof. To keep the notation simpler, assume that the variables with restrictions are ordered first in the VAR, such that $n^{(j)} = j$ for $j = 1, \dots, J$. Otherwise the proof below holds after multiplication with appropriate permutation matrices.

Since groups are non-overlapping, we have that the rows of E, Λ involving any variables $j \in \mathbb{G}$ do not involve any variables $j' \in \mathbb{G}', \mathbb{G}' \neq \mathbb{G}$. Note that E, Λ are zero except for: (1) Positions $\{(j1, j1), \dots, (\bar{j}, \bar{j})\}$ in E, which are unity, and (2) positions $\{(j1, m^{(j1)}), \dots, (\bar{j}, m^{(\bar{j})})\}$ in Λ, which equal $\lambda^{(j1)}, \dots, \lambda^{(\bar{j})}$, respectively. Proceed by Gaussian Elimination.

Note that $\lambda^{(j1)} = 0$ by assumption. Then, multiplying row $j1$ by $-\lambda^{(j2)}$ and adding it to row $j2$ ensure that $M_{j2,\circ} - \lambda^{(j2)} M_{j1,\circ} = E_{j2,\circ} - \Lambda_{j2,\circ} - \lambda^{(j2)} E_{j2,\circ} = E_{j2,\circ} = e_{j2}$ – a zero row except for one entry equal to unity.

Now assume that $M_{jn,\circ} = E_{jn,\circ} = e_{jn}$. Multiplying multiplying row jn by $-\lambda^{(jn+1)}$ and adding it to row jn ensure that $M_{jn+1,\circ} - \lambda^{(jn+1)} M_{jn,\circ} = E_{jn+1,\circ} = e_{jn+1}$. Continue until $jn + 1 = \bar{j}$. Thus, we can rewrite $M_{j1,\circ}, \dots, M_{\bar{j},\circ}$ as a linear combination of the independent basis vectors $E_j = e_j, j \in \mathbb{G}$. Thus, their rank equals the cardinality of $\mathbb{G}$.

Since groups are non-overlapping, the total rank is the cardinality of all groups, which equals J. Thus, Lemma 2 applies. ☐

Note that the logic underlying our existence results does not generally hold when $H \geq 1$ because dynamic restrictions involve interaction terms between restrictions of potentially different sign or reversal to the mean that is not monotone. Heterogeneity restrictions and simple sign restrictions alike can lead us to reject reduced form draws in these cases.

2.3 Equivalence to change of variables

Note that in simple settings there is an equivalence between heterogeneity restrictions and sign restrictions with an appropriate change of variables. Let $[1, 0]$ and $[\lambda, -1]$ be the rows of M encoding the heterogeneity restrictions on $Y_t = [Y_{1,t}, Y_{2,t}]'$. Then this heterogeneity restriction is equivalent to two standard univariate sign restrictions $[1, 0]$ and $[0, -1]$ in a VAR of $\tilde{Y}_t = [Y_{1,t}, \lambda Y_{1,t} - Y_{2,t}]$ with associated Cholesky factor:

$$\tilde{\tilde{A}} = \begin{bmatrix} 1 & 0 \\ -\lambda & 1 \end{bmatrix} \begin{bmatrix} \tilde{A}_{1,1} & 0 \\ \tilde{A}_{2,1} & \tilde{A}_{2,2} \end{bmatrix} = \begin{bmatrix} \tilde{A}_{1,1} & 0 \\ \lambda \tilde{A}_{1,1} - \tilde{A}_{2,1} & -\tilde{A}_{2,2} \end{bmatrix}.$$

For example, take $Y_{1,t}$ to be the nominal interest rate and $Y_{2,t}$ to be the inflation rate. Then the first restriction identifies an increase in the nominal interest rate and the second restriction requires the ex post real rate to rise. Equivalently, we can represent these restrictions as sign restrictions in a bivariate VAR with the nominal and the ex post real interest rate.

More generally, if there are $J = p$ full rank heterogeneity restrictions in a VAR of $\{Y_t\}$ these are equivalent to standard sign restrictions in a VAR of $\{\tilde{Y}_t\} = \{MY_t\}$ with covariance matrix $\tilde{\Sigma} \equiv M\Sigma M'$. Here, $M = E - \Lambda$. Our argument can, thus, alternatively be viewed as a theory of the VAR observables. Our setup is, however, more general because we do not require the order condition $J = p$ but can allow for more restrictions than observables.

2.4 Strength of the restrictions

How do we choose the strength of the heterogeneity restrictions? Mathematically, what are reasonable values for λ? If we have a notion that we want to rank the responses of different sectors qualitatively, the case of $\lambda = 1$ might be the most natural. However, in this case can also think of λ as expressing our degree of confidence in the measured heterogeneity. For example, setting $\lambda = \frac{1}{2}$ expresses a weaker ranking. However, also qualitative beliefs about macroeconomic relationships can generate bounds. For example, we might want to assume that the (ex post) real interest rate rises in response to contractionary monetary policy shock rises, implying

$$\lambda \times IRF_{\text{FFR},0} \geq IRF_{\text{inflation},0} \quad \text{with} \quad \lambda = 1$$

The following example generates a $\lambda \neq 1$: If we believed in a Phillips Curve relationship between inflation and unemployment whose slope below a certain threshold, we might specify that in response to demand shocks the following restriction holds

$$IRF_{\text{inflation},0} \leq -\lambda \times IRF_{\text{unemployment},0} \quad \text{with} \quad \lambda = 2$$

The choice of $\lambda = 2$ is a conservative value. For example, Blanchard et al. (2015) report that the 75th percentile of the point estimates of the slopes across all countries and time periods is smaller than 1.75.

2.5 Different variations of heterogeneity restrictions

While we focus on short-run heterogeneity restriction in the remainder of the paper, we note that the approach is versatile and applies also to the following variations:

Soft zero restrictions: We can also use varying ranking intensities λ to impose approximate zero restrictions, i.e., soft zero restrictions. For example, Christiano et al. (1999) identify monetary policy shocks via zero short-run restrictions, imposing, among other things, that real output cannot respond contemporaneously to monetary policy shocks. Here, we could also impose an analogous, but less dogmatic, soft zero restriction by imposing for a small value of λ for the following restrictions

$$-\lambda \times IRF_{\text{FFR},0} \leq IRF_{\text{GDP},0} \leq \lambda \times IRF_{\text{FFR},0} \quad \text{with} \quad \lambda = 0.01$$

Long-run zero restrictions: It is straightforward to extend our suggested heterogeneity restrictions to the case of set identified long-run restrictions. One could either implement long-run neutrality based on soft zero restrictions as detailed above. For example, to impose approximate monetary neutrality impose:

$$-\lambda \times IRF_{\text{FFR},0} \leq IRF_{\text{GDP},\infty} \leq \lambda \times IRF_{\text{FFR},0} \quad \text{with} \quad \lambda = 0.01.$$

Here, $IRF_{\text{GDP},\infty}$ is the respective long-run impulse response of real GDP to a contractionary monetary policy shock.

Long-run heterogeneity restrictions: Implementing a long-run identification scheme under heterogeneity restrictions is also straightforward. Consider the case of productivity news shocks and of two industries, A and B. A is more R&D intensive than B. To impose that the long-run impulse response of productivity industry A be stronger than productivity in industry B impose:

$$IRF_{\text{Productivity in A},\infty} \geq \lambda IRF_{\text{Productivity in B},\infty} \quad \text{with} \quad \lambda = 0.01.$$

Fully Bayesian analysis of models with extreme values of λ can be challenging. To that end, we develop an efficient algorithm in Section 3 that works well even with soft zero restrictions.[4]

2.6 Characterizing the identified set analytically

When can we expect heterogeneity restrictions to be most useful? Here we first follow Moon et al. (2013) to characterize the identified set analytically in a bivariate VAR with impact restrictions only. We show that for the common restrictions, associated with $\lambda_{HR}^{(j)} = 0$, the identified set for $a_{n^{(j)}}$ can be strictly or weakly smaller, depending on reduced form parameters. For $a_{n^{(j')}}$ with $\lambda_{HR}^{(j')} > 0$, however, we find that the identified set is strictly smaller, except for degenerate cases. Either type of restriction has the more bite the more negative the reduced form correlation of forecast errors. Trivially, heterogeneity restrictions have the more bite, the stronger the known degree of

[4] Note that existing Bayesian analyses of sign and zero restrictions often inadvertently impose non-stated beliefs in the identification. See Arias et al. (2014), who explain how to combine sign and hard zero restrictions in a fully Bayesian fashion.

heterogeneity and the higher the ratio of conditional standard deviations. We also show that the results generalize to trivariate VARs. Trivariate VARs allow us to distinguish between variables directly affected by heterogeneity restrictions and other variables subject only to sign restrictions.

2.6.1 Bivariate VAR with impact restrictions

We impose two restrictions to identify the first shock. In a bivariate VAR we can use (2.3) to express these restrictions as:[5]

Standard sign-restrictions	Heterogeneity restrictions	
$q_1\tilde{A}_{1,1} \geq 0$	$q_1\tilde{A}_{1,1} \geq 0$	(2.6a)
$q_1\tilde{A}_{2,1} + q_2\tilde{A}_{2,2} \geq 0$	$(q_1\tilde{A}_{2,1} + q_2\tilde{A}_{2,2}) - \lambda q_1\tilde{A}_{1,1} \geq 0$	(2.6b)

Since the heterogeneity restriction nests the standard sign restriction for $\lambda = 0$, we now focus on this more general case.

To understand the implied restrictions, it is useful to write the elements of the Cholesky factor $\tilde{A}$ in terms of the correlation and variances of the reduced form errors.[6] We can then rewrite (2.6) as:

$$q_1 \geq 0 \tag{2.7a}$$

$$q_2 \geq \left(\lambda\frac{\tilde{A}_{1,1}}{\tilde{A}_{2,2}} - \frac{\tilde{A}_{2,1}}{\tilde{A}_{2,2}}\right)q_1 = \Big(\lambda\underbrace{\frac{\tilde{A}_{1,1}}{\tilde{A}_{2,2}}}_{>0} - \frac{\rho}{\sqrt{1-\rho^2}}\Big)q_1. \tag{2.7b}$$

In (q_1, q_2) space, q_2 has to lie in the plane above the ray through the origin with slope $-\frac{\rho}{\sqrt{1-\rho^2}}$ with pure sign restrictions. he slope depends on correlation between the reduced form forecast errors. Heterogeneity restrictions can always flip the slope for λ high enough.

Intersecting the set described by (2.7) with the unit circle yields Figure 2.1, following Moon et al. (2013): First, q_1 is positive. Second, q_2 lies above the straight line through the origin, that may have positive or negative slope. The slope is increasing in λ. Last, (q_1, q_2) are confined to the unit circle since $||q|| = 1$. Given $\lambda > 0$, three cases can arise: (a) the reduced form correlation is positive and dominates the positive contribution of the heterogeneity restriction, (b) the reduced form correlation is positive, but the contribution from the heterogeneity restriction dominates, (c)

[5] As written, we impose one sign and one heterogeneity restriction. An example is identifying a cost shock in a competitive industry for which we observe prices and quantities. In the presence of decreasing returns to scale and with elastic demand, we know that minus the quantities fall more strongly than the prices within that industry. For example, let $Q_t = \epsilon_t^a L_t^{1-\phi}$ be the production function and $Q_t = \epsilon_t^b P_t^{-\omega}$, $\phi \in (0,1)$ and $\omega > 1$. Then prices equal $\ln p_t = \ln\bar{w} + \frac{\phi}{1-\phi}\ln q_t - \frac{1}{1-\phi}\ln\epsilon_t^a$ and $\ln q_t = \ln\epsilon_t^b - \omega\ln p_t$. In equilibrium, $\frac{\partial\ln p_t}{\partial\ln\epsilon_t^a} = \frac{(1-\phi)^2}{1+\phi(\omega-1)}$ and $-\frac{\partial\ln q_t}{\partial\ln\epsilon_t^a} = \frac{\omega}{1-\phi}\frac{\partial\ln p_t}{\partial\ln\epsilon_t^a} > \frac{\partial\ln p_t}{\partial\ln\epsilon_t^a}$.

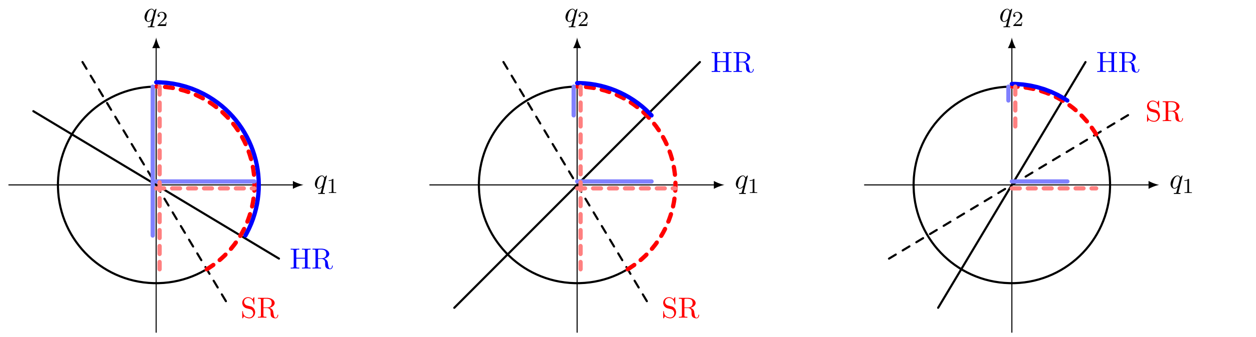
[6] Note that the Cholesky decomposition $\Sigma = \tilde{A}\tilde{A}'$ can be written as: $\tilde{A}_{1,1} = \sqrt{\Sigma_{11}}$, $\tilde{A}_{2,1} = \frac{\Sigma_{21}}{\tilde{A}_{1,1}} = \tilde{A}_{2,2}\frac{\rho}{\sqrt{1-\rho^2}}$, and $\tilde{A}_{2,2} = \sqrt{\Sigma_{22} - (\tilde{A}_{2,1})^2} = |\tilde{A}_{2,1}|\sqrt{1/\rho^2 - 1}$. Σ is the covariance matrix of the forecast errors and ρ is the reduced form correlation between the forecast errors.

the correlation is negative so that both contributions are positive. In cases (b) and (c), the marginal full set for q_1 is strictly smaller with heterogeneity restrictions. In case (a), the marginal set for q_1 is $[0, 1]$ in both cases, but the full identified set for q_2 is strictly smaller with heterogeneity restrictions.

(a) $\lambda\tilde{A}_{1,1} - \tilde{A}_{2,1} \leq 0, \tilde{A}_{2,1} \geq 0$ Weak heterogeneity restriction Positive correlation

(b) $\lambda\tilde{A}_{1,1} - \tilde{A}_{2,1} \geq 0, \tilde{A}_{2,1} \geq 0$ Strong heterogeneity restriction Positive correlation

(c) $\lambda\tilde{A}_{1,1} - \tilde{A}_{2,1} \geq 0, \tilde{A}_{2,1} \leq 0$ Strong heterogeneity restriction Negative correlation



The full identified set is given by the intersection of the unit circle with the $q_1 \geq 0$ plane and the plane above the HR and SR lines, respectively. The resultant joint set on the unit circle as well as the marginal sets on the axes are marked in red (and solid) lines for the case of heterogeneity restrictions (HR) and in dashed and blue lines for the traditional sign restriction (SR). The HR set is strictly smaller on the unit circle – this always translates into a tighter set for q_2 and, in cases (b) and (c) also in smaller sets for q_1. We show in the text that this also translates to tighter sets for a_1 and a_2 in the HR case.

Figure 2.1: Graphical representation of the identified set for the two types of restrictions

However, we are not interested in the set of admissible q per se, but in the induced set for $a = \tilde{A}q$. Since $a_1 = \tilde{A}_{11}q_1$, we can simply read off the results from Figure 2.1. Appendix A.1 summarizes the identified sets for both a_1 and a_2. Proposition 1 uses this characterization to summarize when we have a strict set reduction for the responses. Since $\tilde{A}_{11} = \sqrt{\Sigma_{11}}$ and $\tilde{A}_{21} = \frac{\Sigma_{21}}{\sqrt{\Sigma_{11}}}$, these restrictions depend only on the reduced form variances and covariances.

Proposition 1. The identified set for the structural impulse a_1 from (2.6) is strictly smaller under heterogeneity restrictions than under sign restrictions iff $\lambda\tilde{A}_{11} - \tilde{A}_{21} > 0$. The identified for a_2 is strictly smaller unless $\lambda\tilde{A}_{11} = \tilde{A}_{21}$.

Proof. This follows directly from comparing the sets listed in Appendix A.1 for $\lambda = 0$ and $\lambda > 0$. □

Note that independent of the presence of heterogeneity restrictions or sign restrictions, a negative reduced form correlation leads to a smaller identified set of q_1 and, consequently, of a_1. These sets are, in turn, smaller with heterogeneity restrictions. The differences in the sets are most pronounced when the correlation is positive, but the heterogeneity restriction is strong – compared to the reduced form standard deviations of the second variable relative to the first.

Intuitively, we find set reductions with sign restrictions if the reduced form correlation between the variables is of the opposite sign than the one attributed to the identified shock: In this case, the identified shock cannot account for the entire impact response – else the VAR could not generate the observed reduced form correlation. This intuition not only applies to the case of heterogeneity restrictions, but also for the reduced form correlation between variable 1, $[1, 0]Y_t$ and the linear combination $[-\lambda, 1]Y_t$.

When are the two sets equal? Note that when there is only one common shock to variables 1 and 2 whereas also a second shock affects variable 2, the identified sets in response for the common shock are necessarily equal. The second shock is an idiosyncratic shock to variable 2, such as an industry-specific shock: $A = [a_{11}, 0; a_{21}, a_{22}]$. In this case, $\tilde{A} = A$. Assume a positive covariance. Then, $\tilde{A}_{21} = \kappa \tilde{A}_{11}$ for $\kappa = \frac{a_{21}}{a_{11}}$. If $\lambda = \kappa$, the two sets are equal.[7]

Proposition 1 implies that for λ large enough, identified sets for both responses a_1, a_2 are strictly smaller. A different way to understand our results is through Proposition 4 in Amir-Ahmadi and Uhlig (2015). They show that in a bivariate VAR, all possible sign restrictions are spanned by two sign restrictions with maximal 180° angle. Standard sign restrictions as defined above imply an angle of 90°, whereas heterogeneity restrictions imply an angle of more than 90°.[8] Here, as $\lambda \nearrow \infty$, the angle spanned by the heterogeneity restriction approaches 180°. In this case, our identified sets for a_2 converge to a point mass at $\tilde{A}_{22}$. Note that this case arises when we impose a soft zero restriction: For large λ, we are constraining the response of variable one, i.e., $IRF_{0,1}$ to lie in $IRF_{0,1} \in [0, \lambda^{-1} IRF_{0,2}]$. Given that $IRF_{0,2} \geq 0$, the limit of $\lambda \nearrow \infty$ is point identification. Large but finite λ correspond to "soft" point identification.

Note that the idea of ranking the responses of two different variables to one shock carries over to ranking the response of a single variable to two different shocks: The response of the first variable to the two shocks can be written as $a_{1,1}(Q) = \tilde{A}_{11} \begin{bmatrix} q_{1,1} & q_{1,2} \end{bmatrix}$ subject to $|| \begin{bmatrix} q_{1,1} & q_{1,2} \end{bmatrix} || = 1$. Assuming positive responses, the heterogeneity restriction then takes the form of $q_{1,1} \geq 0$ and $\lambda q_{1,1} \geq q_{1,2} = \sqrt{1 - q_{1,1}^2} \geq 0$ so that $q_{11} \geq \frac{1}{\sqrt{1+\lambda^2}} > 0$. Because $q_{12} = \sqrt{1 - q_{11}^2} = \frac{|\lambda|}{\sqrt{1+\lambda^2}} > 0$, we have a strict set reduction.

2.6.2 Trivariate VAR with impact restrictions

Proposition 1 shows that impulse response of the variable on the right-hand side of the heterogeneity restriction belongs to a strictly smaller identified set with heterogeneity restriction compared to sign restrictions under conditions on the reduced form conditional covariance. Higher dimensional cases are more complicated. However, in the trivariate case, there is a set of sufficient conditions that parallel the necessary and sufficient conditions of the bivariate case. These sufficient conditions also imply either equal sized sets or a strict set reductions for the variable that is not involved in the heterogeneity restrictions.

We begin by stating the heterogeneity restriction for the trivariate case – to obtain the sign

[7] With a negative covariance, both the sign and the heterogeneity restrictions are violated. In this case, the heterogeneity restrictions will, mistakenly, lead us to be more confident about the wrong restrictions.

[8] Since $[1, 0][-\lambda, 1]' < 0$ but $[1, 0][0, 1] = 0$ the angle implied by sign restrictions is wider.

restrictions, set $\lambda = 0$.

$$q_1 \tilde{A}_{11} \geq 0 \tag{2.8a}$$

$$q_1 \tilde{A}_{21} + q_2 \tilde{A}_{22} \geq 0 \tag{2.8b}$$

$$q_1 \tilde{A}_{31} + q_2 \tilde{A}_{32} + q_3 \tilde{A}_{33} \geq \lambda(q_1 \tilde{A}_{21} + q_2 \tilde{A}_{22}) \tag{2.8c}$$

Proposition 2. The identified set for the structural impulse a_1 from (2.8) is strictly smaller under heterogeneity restrictions than under sign restrictions if $\lambda \tilde{A}_{21} > \tilde{A}_{31}$ and $\tilde{A}_{31} > 0$. The identified set for a_1 is equal under heterogeneity and sign restrictions if $\lambda \tilde{A}_{21} \leq \tilde{A}_{31}$ and $\tilde{A}_{21} \geq 0$.

Proof. See Appendix A.2.1. □

The intuition from Proposition 1 also explains Proposition 2: Consider a case where shock identification calls for positive comovements between the variables. The sufficient condition applies to the case where the reduced form correlations are the same as the correlations conditional on the shock. The heterogeneity restriction strictly sharpens inference if in the space of transformed variables the conditional correlation has the opposite sign from the reduced form correlation.

Proposition 2 implies that heterogeneity restrictions can sharpen the inference also on standard macro variables, say variable 1, even if the heterogeneity restrictions only involve micro variables 2 and 3. Again, since $\tilde{A}_{1i} = \frac{\Sigma_{1i}}{\sqrt{\Sigma_{11}}}$, these conditions involve only the reduced form covariances between the forecast errors.[9]

In the Appendix, we provide three examples that show that the sufficient condition in Proposition 2 has bite in real world applications: (1) We analyze the workhorse New Keynesian model of the nominal interest rate, a measure of real activity, and the rate of inflation. (2) We look at fiscal policy in a VAR of GDP, spending, and taxes, motivated by Blanchard and Perotti (2002). (3) We also look at a news shock, in a VAR with GDP, TFP, and a stock index. In these examples, we consider a range of values for λ that implement a "soft" zero restriction on, respectively, real activity, government spending, and current TFP, motivated by Beaudry and Portier (2006). In the New Keynesian application, the sufficient condition for equal sets applies and we verify that for any λ, the identified sets for the macro variable (the interest rate) is unchanged. In the fiscal policy application, we find that the sufficient condition for set reduction applies for modest λ. The set reduction builds up to about 10-15% of the impact response of GDP. The results are similar for the third application, with a set reduction of up to 7.5% for the output response.

What happens if there are only two aggregate shocks and the responses of variables 2 and variables 3 to both shocks satisfy the heterogeneity restriction in population? We show in Appendix A.2.3 that in this case the heterogeneity restriction simply becomes redundant and we are left with two simple sign restrictions, (2.8a) and (2.8b), to identify the shock of interest. Thus, when

[9] Note that the same logic would apply in the somewhat peculiar case of a p dimensional VAR in which $\Sigma_{1i} \geq 0$ for $i = 1, \ldots, p$ with $p - 3$ positivity restrictions appended to (2.8) – or no restrictions on the extra $p - 3$ restrictions.

responses to all aggregate shocks satisfy the heterogeneity restrictions, heterogeneity restrictions have no bite.[10]

3 Estimation

The estimation uncertainty of the identified impulse response functions stems from two sources – the size of the identified set and the uncertainty about the reduced form parameters. We consider two types of inference: First, we consider prior-robust inference (Algorithm 1) about the identified set. Second, we also consider an efficient fully Bayesian inference (Algorithm 2). We provide numerical algorithms for both schemes and begin by summarizing inference over reduced form parameters.

3.1 Reduced form parameter uncertainty

We quantify the uncertainty about the reduced form parameters using a Bayesian approach. This approach is also perfectly valid from a frequentist perspective. The posterior distribution is standard for our Gaussian Bayesian VAR.

Specfically, stacking all the coefficients in a vector β and denoting the forecast error variance by Σ we have the following conjugate prior distribution over the reduced form parameters:

$$\beta \sim \mathcal{N}(\bar{\beta}_0, N_0^{-1} \otimes \Sigma) \tag{3.1}$$

$$\Sigma \sim \mathcal{W}_p(\nu_0(\bar{\Sigma}_0)^{-1}, \nu_0). \tag{3.2}$$

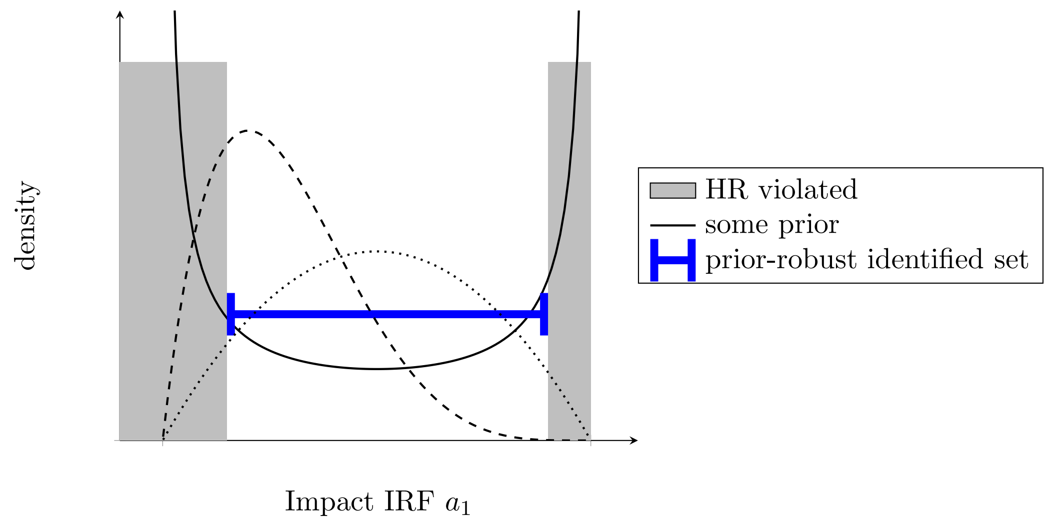
The marginal posterior distribution for Σ^{-1} is a Wishart-distribution, from which we draw directly. Given the draw for Σ^{-1}, we can draw from the conditional normal distribution for the coefficients B.

3.2 Prior-robust inference

In a standard BVAR with sign restrictions such as Uhlig (2005), the posterior distribution of impulse-response functions results from integrating out both the rotation matrix Q and the reduced form parameter uncertainty. However, there are many possible prior distributions over Q that may imply different shapes for the resultant IRF, as we illustrate in Figure 3.1. Baumeister and Hamilton (2015) point out that the commonly used prior that Q be uniformly distributed in the space of orthonormal matrices does not translate to a uniform distribution within the identified set. We also find this in our applications below. Additionally, Arias et al. (2014) argue that practitioners have combined sign and zero restrictions in ways that introduced unnoticed prior information.

We argue that one can address the criticism by Baumeister and Hamilton (2015) and Arias et al. (2014) by being conservative and choosing the worst case prior possible over Q. However, when we are conservative about the distribution of Q, we still know how to quantify the posterior

[10] Trivially, the sets identified with heterogeneity restrictions equal those identified with pure sign restrictions.



Parameterizing $a = \tilde{A}q$ means that the impact response of the first variable $a_1 = \tilde{A}_{11}q_1$. The figure shows that different possible priors, here coming from the family of beta-distributions, can generate very different densities over the structural parameters. Here we focus instead on the full identified set, highlighted in blue.

Figure 3.1: Impulse-vector for first variable in VAR for different beta-priors over q_1

distribution over the reduced-form parameters (β, Σ), and we should use this information that transparently reflects the data.

Thus, we follow Moon and Schorfheide (2012) to compute the infimum and supremum over all admissable rotation matrices Q. This set is distribution free, as we compute the infimum and supremum over the set of all prior distributions over admissable rotation matrices. We compute this set conditional on the reduced form parameters (β, Σ). While this set is robust to any full-support prior over rotation matrices, we still care about the parameter uncertainty: Some parameter combinations (β, Σ) have very low posterior probability. These parameter draws may or may not have large bounds for the impulse response functions attached to them, but since the data tells us that these have very lower posterior density, we argue that we should communicate this. We therefore compute a distribution over the $[\inf, \sup]$-bounds that reflects the posterior reduced-form parameter uncertainty.

Formally, define the posterior distribution over the IRF for variable j at horizon h given the prior π over the rotation vectors q as:

$$\tilde{F}^{\pi}_{j,h}(x) = \int \int_q \mathbf{1}_{\{r_{\tilde{A}q}(h;\Sigma,\beta)_j \leq x\}} \mathbf{1}_{\{r_{\tilde{A}q}(s;\Sigma,\beta)_n \geq \lambda r_{\tilde{A}q}(s;\Sigma,\beta)_m \forall (n,m,\lambda) \in \mathbb{L}^{(j)}_{HR} \forall j=1,\dots J\}} \pi(q) dq \times p(\Sigma, \beta; Y^T) d\Sigma d\beta$$

In contrast, we define the prior-robust posterior distribution over the IRFs as:

$$F_{j,h}(x) = \int \sup_{\pi, q | \pi(q) > 0} \mathbf{1}_{\{r_{\tilde{A}q}(h;\Sigma,\beta)_j \leq x\}} \mathbf{1}_{\{r_{\tilde{A}q}(s;\Sigma,\beta)_n \geq \lambda r_{\tilde{A}q}(s;\Sigma,\beta)_m \forall (n,m,\lambda) \in \mathbb{L}^{(j)}_{HR} \forall j=1,\dots J\}} p(\Sigma, \beta; Y^T) d\Sigma d\beta$$

Our prior-robust inference avoids taking a stance on the shape of the prior over the identified set. It is therefore "frequentist-friendly" in the language of DiTraglia and García-Jimeno (2016). It sidesteps both the criticism by Baumeister and Hamilton (2015) that priors can dominate the inference and the criticism by Moon et al. (2013) that traditional sign identified Bayesian VARs misrepresents the identified set. Our approach follows the principle of transparent parameterization detailed in Schorfheide (2016).

In contrast to the simple sampling scheme for the reduced form parameters, characterizing the bounds of the identified set via Monte Carlo integration is hard, particularly in higher dimensions and can become impractical. We therefore rely on the following numerical algorithm to compute the identified sets. It mimics the analytical approach that we use to characterize the identified set in the bivariate and trivariate VAR examples.

Algorithm 1 Prior robust inference

1. Draw $B^{(d)}$ and $\Sigma^{(d)}$ from $p(B, \Sigma|Y)$.
2. Given $B^{(d)}$ and $\tilde{A}^{(d)} = \text{chol}(\Sigma^{(d)})$, compute the following matrix:

$$W \equiv \begin{bmatrix} S(E-\Lambda)\mathcal{B}_0^{(d)}\tilde{A}^{(d)} \\ S(E-\Lambda)\mathcal{B}_1^{(d)}\tilde{A}^{(d)} \\ \cdots \\ S(E-\Lambda)\mathcal{B}_H^{(d)}\tilde{A}^{(d)} \end{bmatrix},$$

   where

$$\mathcal{B}_h^{(d)} = \begin{cases} \sum_{s=0}^{h}(B^{(d)})^s & j \text{ if estimated in growth rates,} \\ (B^{(d)})^h & j \text{ if estimated in levels.} \end{cases}$$

3. For each variable $i = 1, \ldots, p$ and for each horizon $s = 0, \ldots, S$ solve the following problems

$$\min_q \text{ and } \max_q \quad e_j' \mathcal{B}_s^{(d)} \tilde{A}^{(d)} q$$
$$\text{s.t.} \quad Wq \leq 0,$$
$$||q|| = 1$$

   Save the resulting values as upper and lower bounds.

Our Algorithm 1 is related to those of Faust (1998) and Giacomini and Kitagawa (2014). It resembles the algorithm of Giacomini and Kitagawa (2014) without their computation of posterior means. Unlike them, we have no need to approximate the bounds using stochastic integration. Note that the numerical optimization problem in the algorithm has a simple structure: A linear objective and inequality constraint, and an equality constraint with gradient $2q$. We find that Matlab's `fmincon`[11] solves the problem efficiently. For high-dimensional problems we can run the algorithm in parallel, given independent posterior draws for $B^{(d)}$ and $\tilde{A}^{(d)} = \text{chol}(\Sigma^{(d)})$.[12]

[11]We experimented with different algorithms and solvers to ensure robustness of the results.

[12]In the language of Giacomini and Kitagawa (2014) and Kline and Tamer (forthcoming), we find that for the

IRFs are just one implication of the set-identified shock. The identified set also has implications for policy rules (see Arias et al., 2015) and the forecast error variance decomposition. Since Christiano et al. (1999) argue that structural VARs can fail to identify policy rules even when they identify IRFs correctly, we focus on the forecast error variance decomposition. In Appendix B we follow Uhlig (2003) to show that the forecast error variance for variable i from horizon 0 to H associated with the orthonormal vector q can be expressed as:

$$q' S_{i,H} q, \qquad S_{i,H} \equiv \sum_{h=0}^{H} (H + 1 - h)(e_i \mathcal{B}_h^{(d)} \tilde{A})'(e_i \mathcal{B}_h^{(d)} \tilde{A}). \tag{3.3}$$

We can now compute bounds on the forecast error variance contribution of any variable i up to horizon H by replacing the objective $e_j' \mathcal{B}_s^{(d)} \tilde{A}^{(d)} q$ in the previous algorithm with $q' S_{i,H} q$. For given parameters this approach is the algorithm used in Faust (1998) to assess whether the finding that monetary policy shocks only explain a small proportion of output are indeed robust.

3.3 Fully Bayesian inference

If a researcher has beliefs that provide information in addition to the sign restrictions, she might want to impose these beliefs. Here, we provide a framework for conducting inference under the belief that the rotation vector q is distributed uniformly over the unit n-sphere, conditional on lying in the identified set. Because the identified set can be small, we provide an algorithm for drawing from this set that is efficient and leads to a perfect acceptance rate.

Our prior belief that conditional on a given reduced form draw, whose associated identified set is non-empty, the rotation vector q is distributed uniformly on the unit n-sphere corresponds to the following complete Bayesian model:[13]

$$p(Y, B, \Sigma, q; R(\cdot)) = p(Y|B, \Sigma) p(B, \Sigma) p(q|B, \Sigma; R(\cdot)), \tag{3.4a}$$

$$p(q|B, \Sigma; R) = \frac{\mathbf{1}\{R(B, \Sigma) q \leq 0\}}{\int_{\mathbb{Q} \cap \{\tilde{q} | R(B,\Sigma)\tilde{q} \leq 0\}} d\tilde{q}} \tag{3.4b}$$

In practice, we found that is can be extremely difficult to sample from $p(q|B, \Sigma; R(\cdot))$ when R has many restrictions. We therefore device an efficient algorithm for drawing from the posterior (Algorithm 2). To do this, we use the fact that our restricted set is scale free and that a draw from the multivariate normal distribution rescaled to have zero norm is uniformly distributed on

applications reported here, the posterior plausibility of our restrictions is always 100%. We found lower posterior plausibilities only in a priori unreasonable specifications of our heterogeneity restrictions.

[13] Note that our unconditional prior is agnostic in the sense of Arias et al. (2014): Without restrictions, $p(B, \Sigma, q) = p(B, \Sigma, q\tilde{Q})$ for any orthogonal matrix $\tilde{Q}$. The conditional posterior, however, need not be conditionally agnostic in their language, however, because the size of the identified set enters the probability of q via $\int_{\mathbb{Q} \cap \{\tilde{q} | R(B,\Sigma)\tilde{q} \leq 0\}} d\tilde{q}$. Note that this prior implies that the marginal data density is unaffected by the prior over q when the identified set is never empty: $p(Y) = \int\int\int p(Y|B,\Sigma)p(B,\Sigma)p(q|B,\Sigma;R(\cdot))dqdBd\Sigma = \int\int p(Y|B,\Sigma)p(B,\Sigma)\int p(q|B,\Sigma;R(\cdot))dqdBd\Sigma = \int\int p(Y|B,\Sigma)p(B,\Sigma)dBd\Sigma$.

the unit n-sphere.[14] We formally state these facts in Lemma 3.

Lemma 3. If $x \overset{iid}{\sim} \mathcal{N}(0, I_n)$ and $Wx \leq 0$, then $q = \frac{x}{||x||}$ is a uniform draw from the unit n-sphere that satisfies $Wq \leq 0$.

Proof. The fact that $\frac{y}{||y||}$ has uniform measure over the space of unit vectors follows from the rotation invariance of the zero-mean multivariate normal distribution. More formally and generally, from Theorem 3.3 in Stewart (1980) or Theorem 9 in Rubio-Ramírez et al. (2010), if $X = [x_1, \ldots, x_n]$ with $x_i \overset{iid}{\sim} \mathcal{N}(0, I)$, the matrix Q obtain from the QR decomposition of X has, after normalizing R from the QR decomposition to have a positive diagonal, the uniform distribution on the space of orthogonal matrices. The first column of Q is then simply $q = \frac{x_1}{||x_1||}$. Thus, if $y \overset{iid}{\sim} \mathcal{N}(0, I_n)$, then $q = \frac{y}{||y||}$ is distributed uniformly on the unit n-sphere.

Thus, for any Borel set $\mathcal{A}$, $\Pr\left\{\frac{y}{||y||} \in \mathcal{A}\right\} = \Pr\{q \in \mathcal{A}\}$. Now consider the truncated distribution: $q|Wq \leq 0$. Let $\mathcal{S} = \{z \in \mathbb{R}^n | Wz \leq 0\}$. Notice that $\Pr\left\{\frac{y}{||y||} \in \mathcal{S}\right\} = \Pr\{y \in \mathcal{S}\}$ because $y = 0$ has zero probability. It follows that the truncated distributions are equal: $\Pr\left\{\frac{y}{||y||} \in \mathcal{A} \cap \mathcal{S}\right\} \Pr\left\{\frac{y}{||y||} \in \mathcal{S}\right\}^{-1} = \Pr\{q \in \mathcal{A} \cap \mathcal{S}\} \Pr\{q \in \mathcal{S}\}^{-1}$. □

Lemma 3 allows us to draw efficiently from the truncated unit n-sphere efficiently by drawing from the truncated multivariate normal distribution subject to inequality constraints. Practically, we use the Gibbs sampling algorithm in Li and Ghosh (2015). More efficient direct samplers such as Botev (2016), which uses a recursive sampler based on the LQ decomposition of the W matrix of restrictions, are available when the number of restrictions is no larger than the dimension of q.[15][16]

While Algorithm 2 is designed to draw from the posterior associated with the specific prior (3.4b), we could adapt it for other beliefs by introducing a reweighting step. Specifically, the conditional structure of the prior over (B, Σ, q) in (3.4b) makes our prior hierarchical. This prior belief differs from the prior in Uhlig (2005) and Arias et al. (2014) that implies a flat distribution of q independent of the reduced form parameters. Under regularity conditions, Algorithm 2 can be adapted to produce draws from posteriors induced by other prior beliefs. To do this, we would importance-sample by reweighing the draws from the posterior based on (3.4b) by measure of the restricted set $Wq \leq 0$.

[14] Note that if we had a set of restrictions $\{q|Wq \leq b\}$ for $b \neq \mathbf{0}$, then $Wx \leq b$ does not imply that $W\frac{x}{||x||} \leq b$ – for example, if an equality is strict and $||x|| < 1$. This limits our algorithm to scale-free problems.

[15] We simply use the inverse normal CDF in `Matlab` to draw from its truncated distribution, unlike Li and Ghosh (2015). The inverse standard normal CDF transform is accurate up to ±8. Simulating draws from both the Li and Ghosh (2015) method and the inverse normal method showed that the Li and Ghosh (2015) method was no more accurate in the tails and in some occasions less accurate. Also experimenting with an alternative approximation to the inverse normal CDF produced indistinguishable results.

[16] Notice that the thinning step 3(c)ii in Algorithm 2 is not strictly necessary. However, thinning increases the effective sample size and therefore ensures that comparisons between the measure of different sets associated with restrictions R and R' are not driven by differences in the effective sample size.

Algorithm 2 Fully Bayesian inference

1. Draw $B^{(d)}$ and $\Sigma^{(d)}$ from $p(B, \Sigma|Y)$.

2. Given $B^{(d)}$ and $\tilde{A}^{(d)} = \text{chol}(\Sigma^{(d)})$, compute the following matrix:

$$W \equiv \begin{bmatrix} S(E-\Lambda)\mathcal{B}_0^{(d)}\tilde{A}^{(d)} \\ S(E-\Lambda)\mathcal{B}_1^{(d)}\tilde{A}^{(d)} \\ \dots \\ S(E-\Lambda)\mathcal{B}_H^{(d)}\tilde{A}^{(d)} \end{bmatrix},$$

where

$$\mathcal{B}_h^{(d)} = \begin{cases} \sum_{s=0}^{h}(B^{(d)})^s & j \text{ if estimated in growth rates,} \\ (B^{(d)})^h & j \text{ if estimated in levels.} \end{cases}$$

3. Draw from $p(q|B^{(d)}, \Sigma^{(d)}; R)$

   (a) Initialize $x^{(d,0)} = \frac{x^c}{||x^c||}$ where x^c is the Chebychev center of the set $Wx \leq 0$:

   $$x^c = \arg\min_x \max_r \quad r \qquad \text{s.t.} \quad W_{i,\circ}q + r||W_{i,\circ}|| \leq 0 \forall i.$$

   (b) Verify that the identified set is non-empty, i.e., proceed if $\max_i |x_i^c| > 0$. Otherwise, go back to Step 1.

   (c) Draw $\bar{\ell}$ realizations of $q^{(d,\ell)}$ using the following Gibbs sampler:

      i. For $\ell = 1, \dots, \hat{\ell} + f \times \bar{\ell}$:

      - For $m = 1, \dots n$, draw $x_m^{(d,\ell)}$ from the univariate truncated normal distribution truncated to $[l_m^{(d,\ell)}, u_m^{(d,\ell)}]$.
      - The upper bound is: $u_m^{(d,\ell)} = \min\left\{\infty, \min_{\{j:W_{jn}>0\}} -\frac{W_{jn}x_{-m}^{(d-\mathbf{1}\{n>m\},\ell)}}{W_{jn}}\right\}$.
      - The lower bound is: $l_n^{(d,\ell)} = \max\left\{-\infty, \max_{\{j:W_{jn}<0\}} -\frac{W_{jn}x_{-m}^{(d-\mathbf{1}\{n>m\},\ell)}}{W_{jn}}\right\}$.

      ii. Drop the first $\hat{\ell}$ draws and then keep every fth draw.

      iii. For the remaining draws, compute $\mathcal{B}_s^{(d)}\tilde{A}^{(d)}\frac{x^{(d,\ell)}}{||x^{(d,\ell)}||}$.

4 Applications

4.1 News shocks

In our first application, we analyze a productivity news shock, in the spirit of Beaudry and Portier (2006). Beaudry and Portier (2006), Barsky and Sims (2012) and others summarized in Beaudry and Portier (2014) have argued that these shocks can be important. For example, Beaudry and Portier (2014) estimate that at the two year horizon, between 50 and 80% of the variance in consumption, investment, GDP, and hours can be explained by news shocks. These results are obtained using zero restrictions. Beaudry et al. (2011) used sign restrictions to identify news shocks. However, Arias et al. (2014) show their approach uses prior information that is not acknowledged and, when implemented only with the stated prior, inference becomes imprecise. We now show how adding readily available information on industry returns sharpens inference substantially, compared to only using macro time series.

Our added assumption is that productivity news moves the stock returns of the most innovative sectors the most. To keep the estimation simple, we focus on the five-industry classification by Fama and French (1997).[17] For firms within each industry, we compute the distribution of R&D intensities, measured as the ratio of the three-year moving average of R&D expenses relative to a lagged measure of firm size. Figure 4.1 displays the distribution of the R&D intensity, pooled across firm-years, for each of the five industries using either gross operating income or total assets as a measure of size.[18] While we focus on the 5-industry classification for simplicity, we show below that our results hold up using the finer 10-industry data.

We define a news shock to raise real GDP, employment, productivity, and consumer confidence as well as cumulative real stock returns. Based on the R&D intensities in Figure 4.1, we impose the following ranking on industry returns: (1) Health and High Tech returns increase more than those in Manufacturing, (2) Manufacturing returns increase more than those in the Consumer and Other industries, and (3) Stock returns in the Consumer and Other industries increase. We impose these restrictions on impact and in the two subsequent quarters. Below we also report an extension that imposes a (soft) zero restriction on initial TFP, in the spirit of Beaudry and Portier (2006).

Our VAR includes a total of nine variables. We allow for four lags, as a rule of thumb for quarterly data, and estimate the model in levels with a quadratic trend.[19] We work with a flat prior for the coefficients B and Jeffrey's prior for the covariance matrix Σ. Throughout, we take 500 reduced form draws, and 10,000 draws from the Gibbs-sampler over q, keeping every 10th draw.

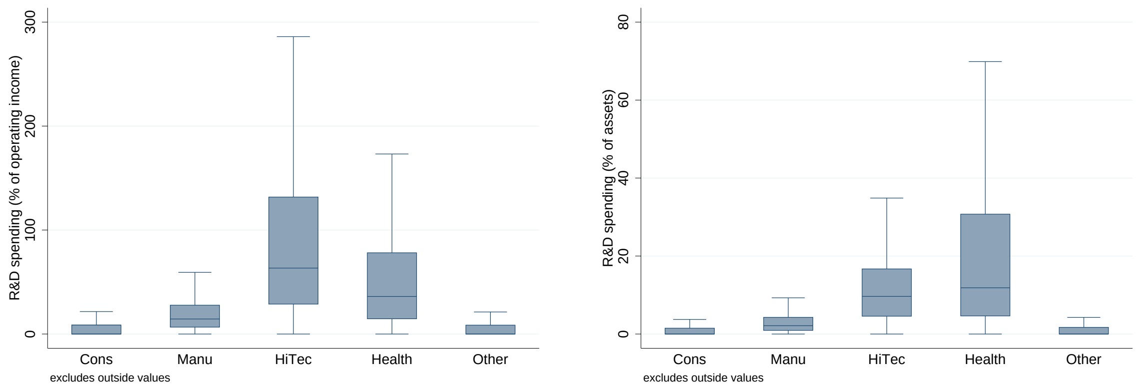
Turning to the results, we discuss the impulse responses first, because this is where we impose the restrictions. We then discuss the forecast error variance decomposition. Then we turn to an analysis which restrictions are the most important, and conclude the discussion with a robustness

[17] The returns are available in Kenneth's French's data library:
`http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

[18] We use Compustat data and drop observations with negative net sales, assets, or employment. Also, we keep only firms that are incorporated in the US and whose records are denominated in US dollars. For our analysis, we winsorize the data at the 1st and 99th percentile year by year.

[19] In unreported robustness checks we estimated the model in first differences with comparable results.

(a) Relative to operating income　　(b) Relative to assets



The boxes show the median along with the interquartile range of the R&D intensity for each of the five industries in the coarsest Fama and French (1997) classification. The upper whiskers end in the values just above to the 75th percentile plus 1.5 times the interquartile range, and analogous for the lower whiskers. We measure firm size either as the lagged three year moving average of operating income or total assets.

Figure 4.1: R&D intensity by industry in the 5-industry Fama and French (1997) classification

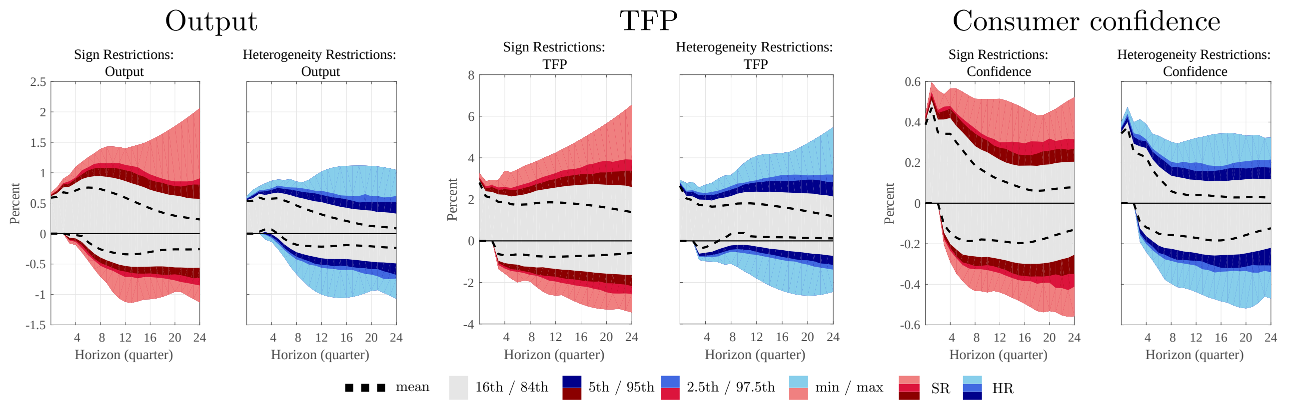
check. In what follows, we focus on a select number of results, but provide the full set of results in Appendix D.1

4.1.1 Impulse response functions

Figure 4.2 and 4.4 show the impulse response functions for the four macro variables in the VAR, computed according to Algorithm 1 and 2, respectively. Figure 4.2 shows the posterior distribution over the bounds of the identified set, i.e., the prior-robust posterior. Throughout, we contrast the results with pure sign restrictions, in red, and the heterogeneity restrictions, in blue. All responses are based on the same reduced form draws, all of which were consistent with the restrictions.

To begin, let us discuss what we learn when we remain robust to the prior over the distribution within the identified set. Figure 4.2 shows these results for the four macro variables and two industry returns: The black dashed lines show the bounds of the identified set at the posterior mean. The shaded areas show the posterior distribution over these bounds – specifically, the 68%, 90% and 95% percentile plus the most extreme bounds. We find that the distributions over the identified sets always include zero after the restrictions are lifted – here in the third quarter. In the short-term, most of the uncertainty reflects the width of the identified set, whereas at longer horizons parameter uncertainty dominates.

Can we still draw substantive conclusions while remaining prior-robust? We see different shapes in the identified sets, with support for a build-up in the output and employment responses, but a smooth decline in consumer confidence. Little can be inferred about the shape of the TFP response when we take parameter uncertainty into account. The exception is TFP: At the posterior mean, the lower bound for TFP excludes zero after two years when we impose heterogeneity restrictions. Similarly, the shown industry returns turn significantly negative at the posterior mean with heterogeneity restrictions. With parameter uncertainty, the analysis mainly serves to bound the responses and the heterogeneity restriction again sharpens the results significantly: For example, for GDP the sign restrictions tell us that the one standard deviation news shock may raise GDP (relative to trend) by almost 1.4%. The heterogeneity restrictions imply that the GDP increases no more than 0.8%. We find comparable set reductions of up to 50% for consumer confidence and almost 40% for employment (for numbers, see Table D.1(a) in the Appendix).

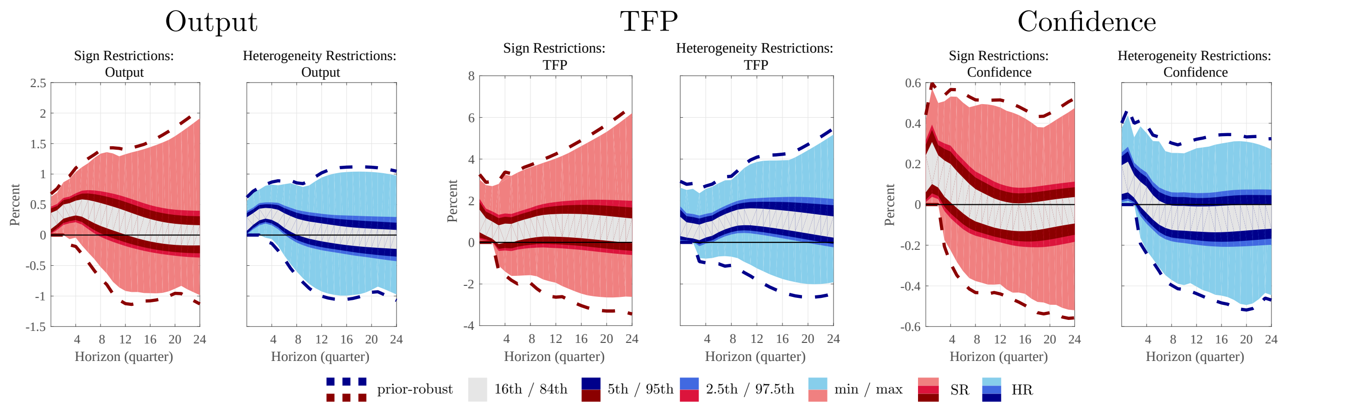


Comparing the identified set at the posterior mean (black dashed lines), we find that heterogeneity restrictions reduced the identified set substantially compared to sign restrictions, but after the restrictions are eased, we can only infer that TFP increases in the medium run. However, parameter uncertainty is pervasive at longer horizons and when taking it into account we can only bound the responses with the restrictions. The bounds with heterogeneity restrictions are a third smaller for output at the 95th percentile and almost fifty percent smaller for consumer confidence.

Figure 4.2: Prior-robust responses of output, TFP, and consumer confidence productivity news shock

We now turn to the fully Bayesian analysis and begin with a technical point: Figure 4.3 shows the fully Bayesian posterior over the three macro responses: The plots show the most extreme realizations from the sampler (the lightest shaded areas) along with the quasi-analytical bounds familiar from Figure 4.2. Throughout, the outermost credible sets of the fully Bayesian posterior track the outermost prior-robust bounds closely. This speaks to the ability of our algorithm to sample the entire parameter space. However, the upper and lower 2.5% of the posterior mass are often as wide or wider than the inner 95% – which is why we "zoom" in for our discussion of the IRFs.

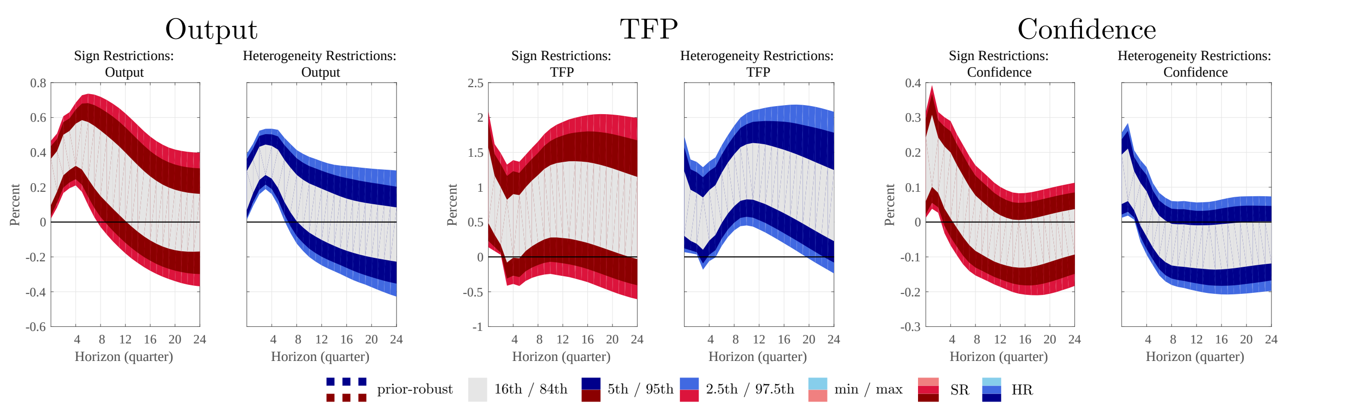
What can researchers with conditionally agnostic priors learn from sign or heterogeneity restrictions? Figure 4.4 shows the belief that q is conditionally uniformly distributed sharpens our inference substantially: The pointwise 95% confidence sets shown all exhibit well-defined shapes. While we could say little about the shape of the TFP response while being robust to any prior, our fully Bayesian posterior implies that TFP increases in a hump-shaped fashion in response to a productivity news shock, plausibly reflecting technology diffusion. Inference about the hump is much sharper with heterogeneity restrictions which place the peak increase in TFP between 0.5 and 2%, about three years after the initial shock. This causes a hump-shaped expansion in output, peaking one year out between 0.2% and 0.5% with 95% confidence, according to the model



Here we compare the coverage of the identified set using Monte-Carlo integration using Algorithm 2 compared to the quasi-analytical characterization using Algorithm 1. We find that the posterior mass is concentrated in the center of the identified sets, but the top and bottom 2.5% density cover the identified set well.

Figure 4.3: Comparison of fully Bayesian coverage and prior-robust bounds

with heterogeneity restrictions or 0.2% to 0.7% with sign restrictions only. Consumer confidence increases 0.05% to 0.2% on impact (0.3% with sign restrictions) and then reverts back to zero after one year. Overall, we see economically sensitive responses that are much sharper with heterogeneity restrictions.



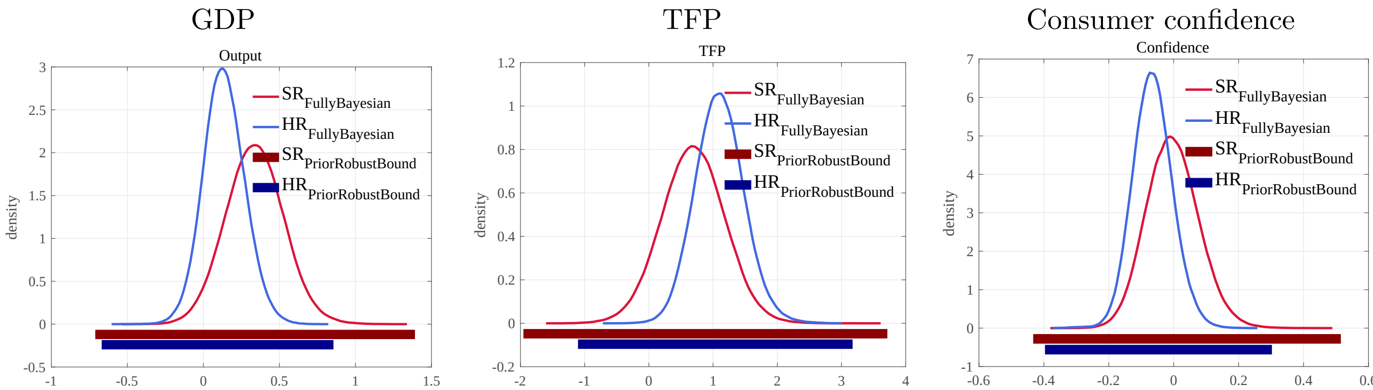
Heterogeneity restrictions sharpen the inference with sign restrictions in economically meaningful ways: TFP is found to increase in a hump-shaped fashion, consistent with slow technology diffusion. We find a smaller increase in output and consumer confidence with heterogeneity restrictions.

Figure 4.4: Fully Bayesian responses of macro variables to productivity news shock.

The confidence sets alone could obscure irregular posterior distributions – but in Figure 4.5 we show that they do not.[20] The posterior densities are unimodal and largely symmetric. The plot also confirms that the densities assign positive measure to almost the extremes of the distribution over identified sets, shown as thick lines underneath the zero line. On the substantive side the densities show that for the three macro variables except TFP, the posterior mass is shifted to the

[20] At short horizons, when the restrictions are still binding, we sometimes observe higher densities around zero, reflecting the truncation.

left using heterogeneity restrictions. For TFP, the mass is shifted, in contrast, to the right. For all four variables the densities are markedly more concentrated with heterogeneity restrictions.



Heterogeneity restrictions lead to both a reduction in the identified set, here integrated over all reduced form parameters, and the dispersion of the fully Bayesian responses, show as density plots two years after impact. Heterogeneity restrictions both lead to less dispersed distributions of responses, but can also shift mass away from zero. For example, the TFP and consumer confidence responses are zero with sign restrictions, but significantly positive for TFP (at the 95% level) and negative (at the 68% level) for consumer confidence.

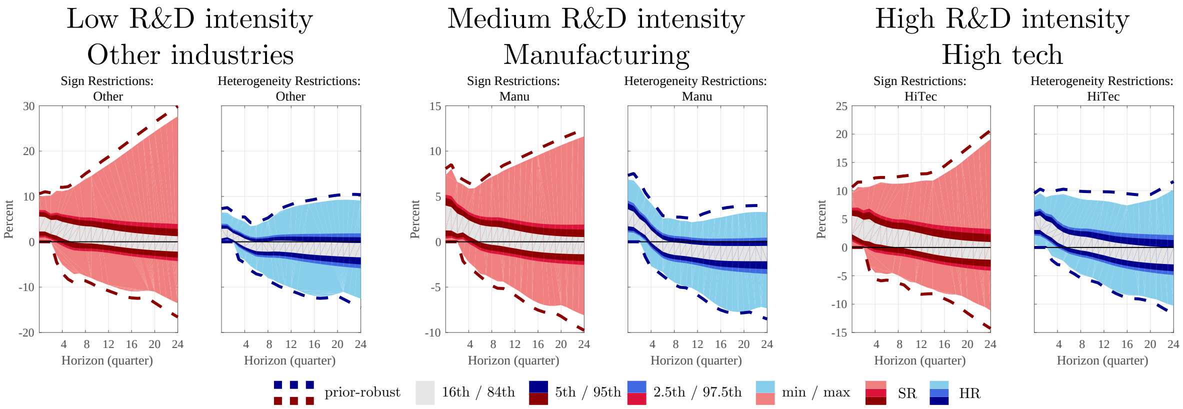
Figure 4.5: Distribution of responses to productivity news shock: Macro variables two years out.

We now turn to the micro-responses of cumulative industry stock returns in Figure 4.6. Heterogeneity restrictions on these responses yields tighter bounds on the macro variables. In addition, we find that the heterogeneity restrictions provide enough structure to rule out identified sets for the responses that simply drift off into positive or negative territory. Take, for example, the return in other industries. Even though we bound the return in this industry with that in manufacturing only up to the second quarter, it rules out expansionary paths over the entire horizon and hints at a smooth reversal of the initial expansion. We find similar patterns for the all five industries and show here one of each category: the low R&D intensive other industries, manufacturing, and high tech. Considering the fully Bayesian posterior reveals very pronounced shapes and reduce the magnitudes of responses up to 45% on impact and 70% one year our (Table D.1(b)).

4.1.2 Forecast error variance decomposition

We now analyze the forecast error variance decomposition. For simplicity, we focus on results at the posterior mean. We normalize the explained variance by the unconstrained optimum without restrictions, i.e., we divide by the maximal variance that could be attained by any single shock, as derived by Uhlig (2003).

Table 4.1 summarizes the variance reduction relative to pure sign restrictions for each variable at horizons of up to six years. Table D.2 in the Appendix shows that with sign restrictions alone the restrictions are uninformative because the forecast error variance contribution typically attain the maximal contribution, particularly at horizons of up to four quarters. Heterogeneity restrictions, in contrast, bring a variance reduction of 8% to 33% at the one year horizon for the four macro



We rank the responses of stock returns of industries from zero to two quarters according to their R&D intensity. Our main purpose was to sharpen inference about macro variables, but the restrictions also help to rule out paths for cumulative returns that drift off. For the prior-robust inference, this just bring a substantial set reduction. For the Fully Bayesian responses, this reveals a swift mean reversion of returns that is fastest for the industry with the lowest R&D intensity.

Figure 4.6: Responses of (cumulative) industry returns to productivity news shock.

variables. These reductions typically persist up to six years. Overall, the variance reductions due to heterogeneity restrictions, also for the industry returns, are substantial. Intuitively, they mirror the reductions in the magnitude of impulse-responses, summarized in Table D.1 in the Appendix.

Reduction in Maximum FEV (% of maximal FEV)

	Horizon H (quarters)									
Variable	0	1	2	3	4	8	12	16	20	24
Output	18.5	18.6	18.9	19.7	20.8	23.0	26.0	28.2	29.3	29.6
TFP	11.6	10.8	9.6	8.8	8.4	7.2	6.3	6.1	6.2	6.5
Confidence	21.0	26.9	29.8	31.6	32.9	35.5	35.1	33.2	30.8	28.6
Employment	12.8	12.6	11.7	13.5	16.5	24.8	29.0	31.0	31.6	31.7
Consumers	40.0	37.7	37.6	38.1	38.7	40.2	41.0	40.8	39.9	38.4
Manu	16.1	17.7	18.9	19.8	20.3	21.1	21.2	21.0	20.7	20.4
HiTec	16.8	15.3	15.1	15.2	15.6	18.5	20.5	21.7	22.2	22.3
Health	23.2	24.4	24.7	25.4	26.6	31.0	34.1	36.3	37.5	37.9
Other	52.3	53.4	55.4	56.9	57.9	60.2	63.0	65.0	65.8	65.8

We find that the forecast error variance shrinks by about 10% to 35% for macro variables other than TFP over all horizons thanks to the heterogeneity restrictions. For the micro variables, the reduction is ranges from roughly 15 to 65%. The variance contribution is expressed in percent of the total forecast error variance up to horizon H. All contributions are computed at the posterior mean.

Table 4.1: Reduction in maximum forecast error variance explained by productivity news due to heterogeneity restrictions

4.1.3 Important restrictions

We impose tighter restrictions to achieve sharper identification. Which of these restrictions matter? Do they reduce the importance of the "standard" restrictions?

We answer this question using the output of the prior-robust Algorithm 1: For the bound on the response of each variable over horizon $h = 0, \ldots, \bar{H}$, we simply count how often a given restriction binds with equality. If we took away such a restriction, the bounds would widen.[21] We repeat this for each reduced form draw. Last, we compute the fraction of binding instances. Table 4.2 reports these statistics both for GDP and averaged across all four macro variables.

Comparing how often a restriction on industry return binds with $\lambda = 0$ relative to how often it binds with actual heterogeneity imposed ($\lambda = 1$) tells us which restriction matters. Consider the Table 4.2(a): The restriction that manufacturing returns increase more than λ times returns in other industries binds 68% of the time for the upper bound of GDP with heterogeneity, compared to 4% of the time as a simple sign restriction. We conclude that this restriction is important for bringing about the reduction in the identified set for GDP. The same is true when averaged across all macro variables, see Table 4.2(b).

Surprisingly, we find that the importance of the return heterogeneity restrictions does not generally diminish how often the sign restrictions on macro variables bind with equality.[22] Rather than being slack more often the sign restriction on consumer confidence even binds much more often in the case of the GDP response (in 14% more of instances). The other three macro sign restrictions bind about as often. Thus, information from heterogeneity restrictions complements the standard sign restrictions.

4.1.4 Robustness

Are our results an artifact of the particular industry classification we chose? They are not. We double the number of industries in our VAR and using the same procedure to order the micro data responses. In Figure D.3 we contrast the fully Bayesian responses along with the min and max of the prior-robust results for the four macro variables and the two datasets. With either dataset we obtain significant set reductions with heterogeneity restrictions, for example, a reduction of the maximum GDP response from 1.2% to about 0.6% at the three year horizon. The fully Bayesian confidence sets also imply pronounced shapes. The one exception is the TFP response. Here we find no build-up, but just a lasting increase in the TFP level when using ten rather than five industries to impose the micro data restrictions.

4.1.5 Soft zero restrictions

[21] Formally, we compute which rows of $Wq \leq 0$ in Algorithm 1 are equal to zero and then map the rows of W back into English. We pool the same restriction when imposed at multiple horizons. In practice, we need to set a numerical threshold for equality. We tried both 10^{-8} and 10^{-4} and obtained almost identical results – unsurprisingly, given our threshold for the optimizer of 10^{-12}.

[22] Our second application shows a different pattern, revealing that this finding is not mechanical.

(a) Binding restrictions (as a share of all cases) – GDP

Restrictions	Lower Bound SR ($\lambda = 0$)	Lower Bound HR ($\lambda = 1$)	Upper Bound SR ($\lambda = 0$)	Upper Bound HR ($\lambda = 1$)
Output>0	0.07	0.07	0.15	0.14
TFP>0	0.05	0.04	0.64	0.63
Confidence>0	0.09	0.11	0.23	0.36
Employment>0	0.26	0.22	0.27	0.27
Consumers>0	0.13	0.16	0.01	0.12
Other>0	0.32	0.49	0.00	0.04
Manu$>\lambda\times$Consumers	0.16	0.41	0.04	0.25
Manu$>\lambda\times$Other	0.16	0.26	0.04	0.65
Health$>\lambda\times$Manu	0.28	0.45	0.05	0.39
HiTec$>\lambda\times$Manu	0.18	0.29	0.11	0.30

(b) Binding restrictions (as a share of all cases) – all macro IRFs

Restrictions	Lower Bound SR ($\lambda = 0$)	Lower Bound HR ($\lambda = 1$)	Upper Bound SR ($\lambda = 0$)	Upper Bound HR ($\lambda = 1$)
Output>0	0.08	0.08	0.17	0.19
TFP>0	0.23	0.18	0.49	0.50
Confidence>0	0.11	0.14	0.25	0.39
Employment>0	0.31	0.26	0.23	0.19
Consumers>0	0.11	0.24	0.03	0.15
Other>0	0.17	0.39	0.03	0.12
Manu$>\lambda\times$Consumers	0.16	0.41	0.10	0.39
Manu$>\lambda\times$Other	0.16	0.36	0.10	0.56
Health$>\lambda\times$Manu	0.28	0.42	0.12	0.38
HiTec$>\lambda\times$Manu	0.16	0.27	0.10	0.26

We quantify the importance of restrictions by computing for which fraction of responses, averaged over horizons zero to 24 quarters and all reduced form draws, any given restriction is binding. Although we could distinguish at which restricted horizon any restriction is binding, we sum them together for horizons zero to three. We find that the restrictions that manufacturing returns exceed those in other industries is particularly important for sharpening both the GDP response and the four macro responses more generally. Introducing heterogeneity restrictions does not diminish the importance of the sign restrictions on macro variables.

Table 4.2: Importance of constraints for identifying macro impulse-responses: Incidence of binding constaints.

Beaudry and Portier (2006) identify the news shock by a Cholesky decomposition that imposes that the impact of the news shock on the level of TFP be zero: TFP news should not, by themselves, raise TFP immediately. Here we incorporate this assumption as a "soft" zero restriction on the initial TFP response.[23] Table 4.3 shows that this extra restriction yields a very powerful additional set reduction: For output, this reduces the maximal FEV explained by the TFP news by roughly half, an additional reduction of 15 to 30% compared to heterogeneity restrictions alone. For employment we also obtain a significantly lower maximal FEV of 15 to 25%, while consumer confidence is largely unaffected. By construction, the FEV for TFP that can be explained drops dramatically at short horizons, but rises with the forecast horizon.

Reduction in Maximum FEV due to heterogeneity restrictions (% of maximal FEV)
plus soft zero restriction on initial TFP

	Horizon H (quarters)									
Variable	0	1	2	3	4	8	12	16	20	24
Output	51.2	50.4	50.4	50.2	50.0	48.1	47.3	46.5	45.8	45.0
TFP	99.2	98.1	94.6	90.7	87.1	76.0	66.1	57.4	50.4	44.9
Confidence	22.9	28.8	31.9	33.8	35.1	37.3	36.7	35.0	33.1	31.4
Employment	34.3	36.5	37.2	39.0	41.4	46.3	46.3	45.5	44.5	43.6
Consumers	61.4	59.2	58.8	58.4	57.7	55.1	53.6	51.9	49.7	47.2
Manu	44.0	44.6	45.1	45.2	45.0	43.0	41.0	39.3	37.8	36.5
HiTec	36.0	32.1	30.1	28.8	28.2	28.0	28.3	28.4	28.2	27.7
Health	24.1	25.0	25.1	25.7	26.8	31.1	34.3	36.5	37.8	38.2
Other	68.4	68.9	70.0	70.3	70.2	69.5	70.6	71.6	71.7	71.2

Compared to the results with heterogeneity restrictions but without the soft zero restriction (Table 4.1), we observe a significant further reduction in the maximal forecast error variance attributable to the fiscal shock. By construction, this is most pronounced for TFP at the short run, but the identified shock becomes more important for TFP in the medium term. Reduction for output and employment are between 20 and 30%. The contribution is expressed in percent of the total forecast error variance up to horizon H. All contributions are computed at the posterior mean.

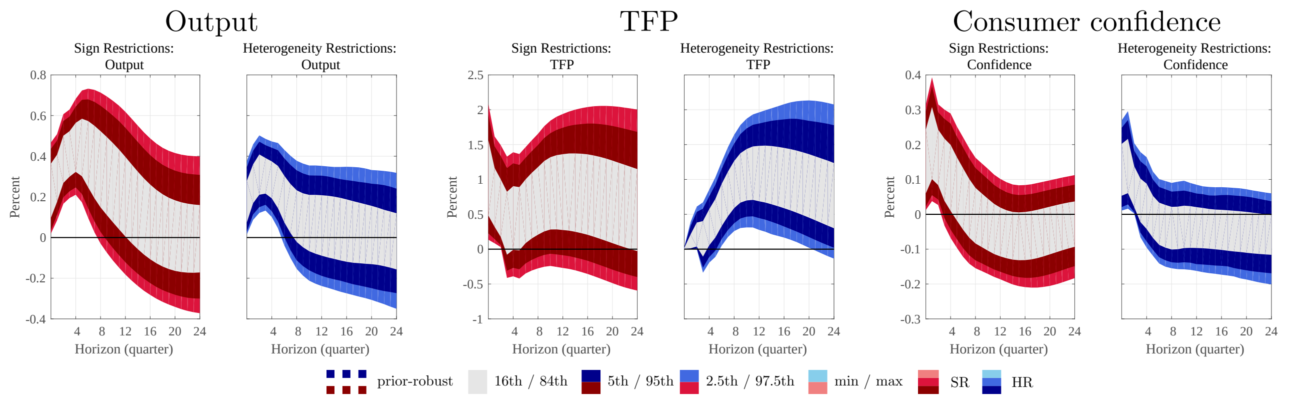
Table 4.3: Reduction in maximum forecast error variance explained by productivity news: Heterogeneity restrictions without and with soft zero restriction

While the explained forecast error variance drops with extra restrictions, we find that the impulse-responses change little. Figure 4.7 shows the corresponding responses. By construction, the initial response of TFP is (almost) zero, but there is still a substantial build-up in TFP in the medium term. The shape of the output and consumer confidence response change little. We conclude that our identification scheme is robust to the added assumption that TFP news have no contemporaneous effect on TFP.

4.1.6 Alternative sampler

The more bite a restriction has, the harder it often is to draw while respecting this restriction with the standard algorithm of drawing q uniformly and keeping only admissible draws. For example, Inoue and Kilian (2013) report that they need 20,000 draws of the rotation vector for numerical

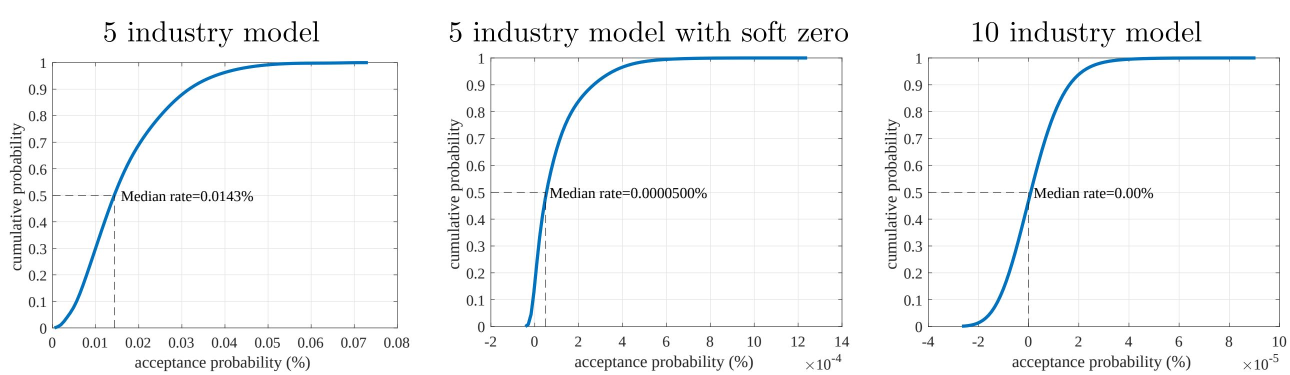
[23] Formally we impose on impact that: Output $> 10 \times$ TFP, in addition to: TFP > 0.



Qualitatively, imposing that TFP increase (hardly) on impact in response to a news shock changes our results very little: The build-up in TFP is more pronounced, but takes away little from the medium-term effects. The response of output, but not confidence is further muted compared to just heterogeneity restrictions in Figure 4.4.

Figure 4.7: Fully Bayesian responses of macro variables to productivity news shock with soft zero restriction on initial TFP.

stability. Our approach delivers, in contrast, a perfect acceptance rate. In many of our applications we would be unable to provide fully Bayesian results without it, as Figure 4.8 illustrates: The median acceptance rate across reduced form draws in our benchmark VAR is just 0.014%. Still, it is strictly positive for all draws based on 5 million attempts. With the slightly harder problem of including a soft zero restriction, the median acceptance rate drops by three orders of magnitude, to 0.00005%. Without the soft zero restriction, but in the higher dimensional version with ten industry returns, the median acceptance rate is even 0. We conclude that Algorithm 2 is important in practice.



Using a simple uniform proposal density, as customary since the seminal paper by Uhlig (2005), becomes impractical with tight sign restrictions. We show the distribution of acceptance probabilities as a function of the reduced form parameter draws. The acceptance probability is based on 5 million draws for each vector of reduced form parameters.

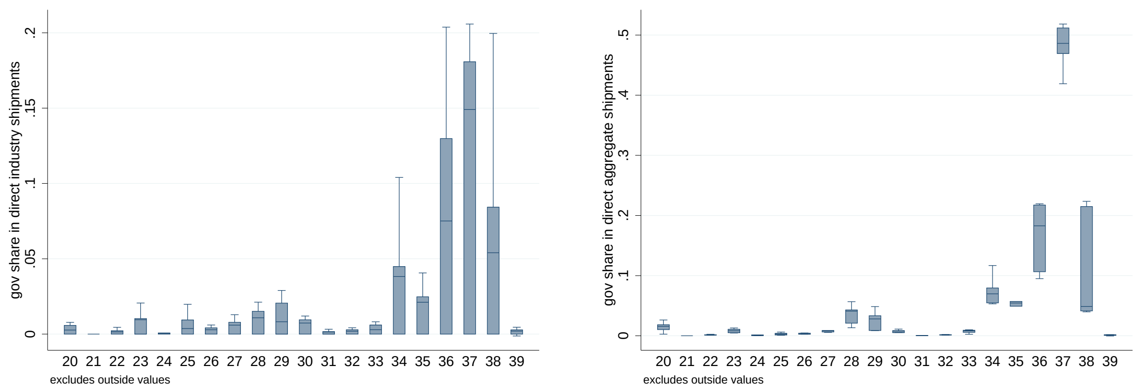
Figure 4.8: Distribution of acceptance probabilities for uniform proposal density over reduced form draws

4.2 Fiscal shocks

In this application, we analyze a defense spending shock financed through higher taxes. We use the NBER manufacturing database and IO-table information from Nekarda and Ramey (2011) as the source of our micro data. For tractability, we aggregate to the 20 SIC2 manufacturing industries. Our heterogeneity restriction is that shipments of all manufacturing industries rise, but more so in the industries of which the government is an important client.

Figure 4.9 measures the importance of the government for each SIC2 industry by showing the median and distribution of the government share in direct shipments (left panel) and overall shipments to the government (right panel).[24] For both measures we can clearly see that the sub-industries 36 (electronics), 37 (transportation), and 38 (sensors etc.) are the most exposed to the government. Our strategy is to pick two industries at the two of the distribution, two in the middle, and two in the bottom of the distribution. We choose the aggregate of 36 and 37 as the most exposed industries, 29 (petrol and refineries) and 35 (equipment) as industries with an intermediate exposure, and 21 (tobacco) and 24 (lumber) as those with the lowest exposure.

(a) Relative to industry shipments (b) Relative to aggregate shipments



The boxes show the median along with the interquartile range of the importance of government shipments each of 2-digit SIC manufacturing industry. The upper whiskers end in the values just above to the 75th percentile plus 1.5 times the interquartile range, and analogous for the lower whiskers. We obtain the data from Nekarda and Ramey (2011). We choose the aggregate of 36 (electronics) and 37 (transportation) as the industries most exposed to the government, 29 (petrol and refineries) and 35 (equipment) as industries with an intermediate exposure, and 21 (tobacco) and 24 (lumber) as those with the lowest exposure.

Figure 4.9: Importance of government shipments by industry

Our VAR includes annual data on defense spending, GDP, the real market value of federal debt, total hours worked (all in logs and per capita terms), and the average marginal tax rate in addition to shipments from the six industries.[25] We estimate an annual VAR with one lag in levels and

[24]Because later we aggregate industries up to the SIC2 level, we focus on direct shipments to avoid double-counting of indirect shipments.

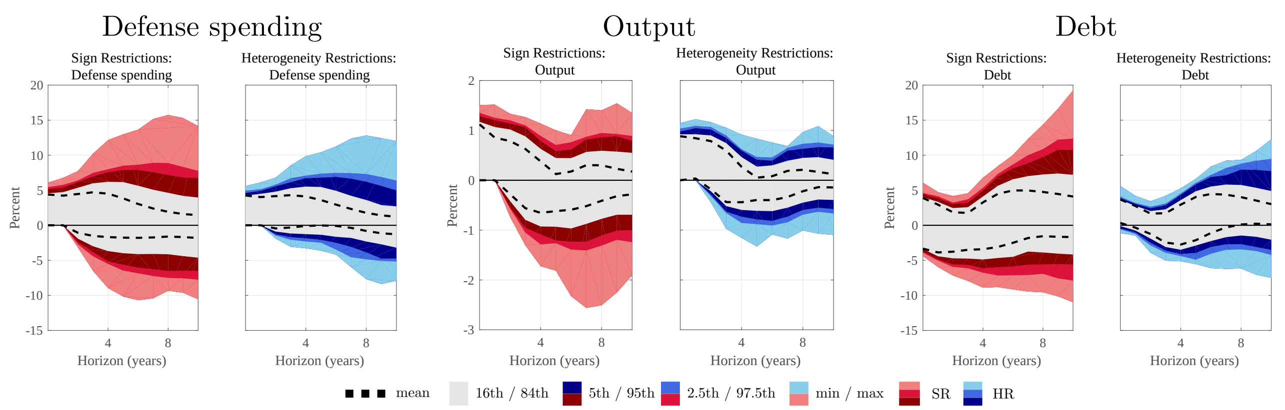
[25]The macro data, except for debt, is taken from Ramey (2011). We use the Dallas Fed data for the nominal market value of federal debt, available at `https://www.dallasfed.org/research/econdata/govdebt`, and deflate it by the CPI.

include a linear-quadratic trend to remove low-frequency movements.[26]

We define a tax-financed defense spending shock as follows: Defense spending, GDP, total hours, and the average marginal tax rate increase for two years. The most exposed industries increase their shipments more than those with more modest exposure. The modestly exposed industries, in turn, increase shipments more than the industries with no exposure. Output in the lumber and tobacco industries weakly increases. Debt is free to respond in any way. Let us now turn to the implied responses.

4.2.1 Impulse response functions

Inference that is robust to the prior distribution is hard – but the heterogeneity restrictions tighten bounds and allow qualitative inference at the posterior mean. Figure 4.10 shows the results for the three macro variables. As in the analysis of the news shocks, at short horizons the uncertainty is modest and mostly due to the width of the identified that would also prevail at the posterior mean. At longer horizons, parameter uncertainty compounds the uncertainty about the identified set at the posterior mean. Heterogeneity restrictions lower the upper bound – by between 8% (hours at the three year horizon) and 48% (tax rates at the 8 year horizon), see Table D.5. In addition, at the posterior mean heterogeneity restrictions permit inference about the shape of the responses. The unrestricted debt response is positive, despite the tax increase. Defense spending and hours remain high persistently, while output quickly reverts back towards zero (Figure D.7 shows all responses).



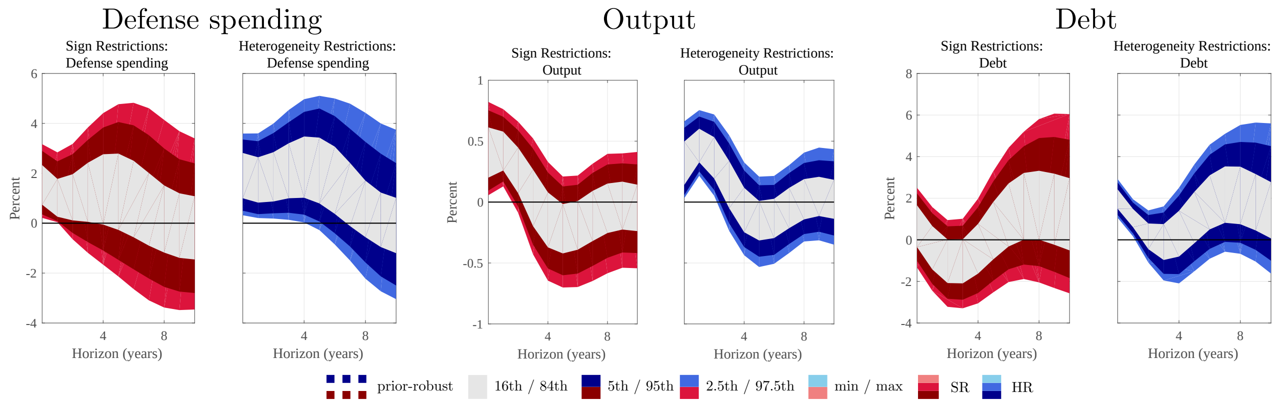
Heterogeneity restrictions yield set reductions that allow to sign responses for both defense spending and debt at the posterior mean, even though the debt response is left unrestricted. While parameter uncertainty blurs these findings, heterogeneity restrictions still lower the 95th percentile of upper bounds by between 10% and 25% for the variables shown across the different horizons.

Figure 4.10: Prior-robust responses to defense spending shock: Macro variables

A Bayesian would find that her beliefs that rotation matrices are uniformly distributed allow her to sharpen inference significantly, because there is little mass in the extremes of the identified

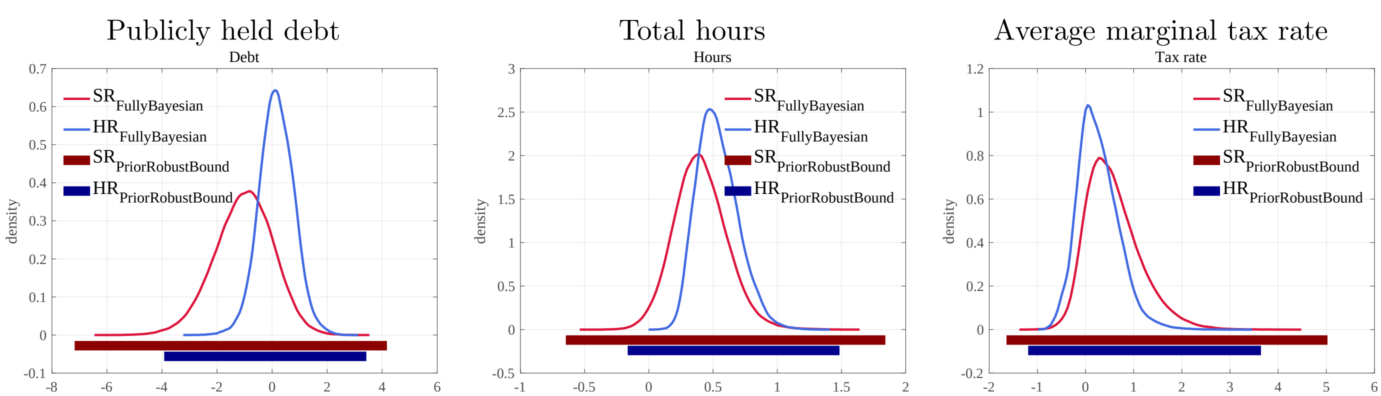
[26] With sign restrictions alone, reduced form draws with explosive eigenvalues often dominated the tails. We therefore decided to reject draws with eigenvalue above 1.03 in absolute value to give sign restrictions a better chance to stand their own against heterogeneity restrictions.

set.[27] Already with sign restrictions, a Bayesian can say that output increases for two years after impact, along with hours worked and tax rates. With heterogeneity restrictions, however, we isolate a more persistent increase in spending up to four years, find a hump-shaped increase also in output, and find clear evidence that debt rises initially and, with some confidence, up to eight years after the shock. The range for a one standard deviation defense spending increase is wide, around 0.25% to 3.5%, with initial increases in GDP of 0.1% to 0.6%.[28]



Both sign and heterogeneity restrictions produce economically sensible responses: Defense spending shocks are persistent and raise output above trends for two years after impact. Heterogeneity restrictions, however, allow sharper inference that reveals a more persistent increase in defense spending, and thus lower discounted multipliers, as well as a pronounced increase in debt even though part of the spending increase is tax financed.

Figure 4.11: Prior-robust responses to defense spending shock: Macro variables



Here we show the identified set and the fully Bayesian posterior densities over the responses of three macro variables three years after the initial shock, integrated over the uncertain parameters. We find that some densities are skewed, but all are unimodal. For total hours and tax rates the sharper inference is particularly clear, but for hours the heterogeneity restrictions also rule out negative responses.

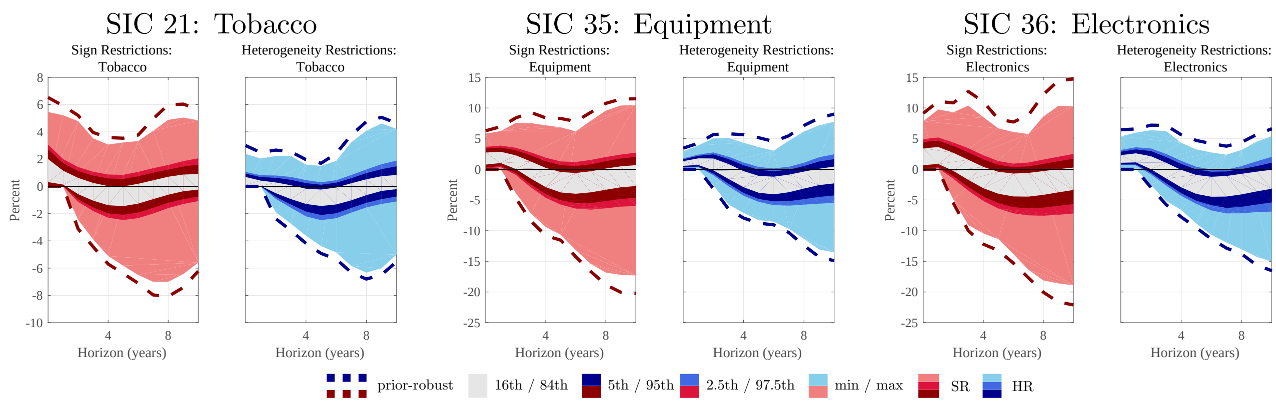
Figure 4.12: Distribution of responses to defense spending shock: Macro variables three years out.

[27] However, the tails of the numerically computed posterior confidence sets do cover almost all of the identified set. See Figure D.7 in the Appendix.

[28] Based on the medians, the impact multiplier is thus roughly three (0.3% over (0.05 $\times$ 2%)), but this is driven by a large increase in the present discounted value of defense spending that generates a big negative wealth effect.

Figure 4.12 shows the Bayesian densities and the prior-robust bounds at the three year horizon. It illustrates the set reductions for debt, hours, and tax rates. They are particularly strong for debt and tax rates: The left tail of the identified set for the debt distribution shrinks by 1.5p.p. and the mode of the fully Bayesian density shifts to the right by a similar amount. On the flipside, for taxes heterogeneity restrictions rule out large tax increases and shifts the distribution to the left. For total hours worked, the shift in the distribution due to heterogeneity rules out very negative responses of total hours worked.

Shipments in the industries used for heterogeneity restrictions exhibit a sensible pattern, see Figure 4.13. As for aggregate output, industry shipment revert back to values around zero after three years. Heterogeneity restriction cut the prior-robust upper bounds by 17% to 63% and almost shrink the width of the inner 95% credible sets – see Table D.5.



We rank the responses of real shipments of industries from zero to one year according to the importance of government purchases to the industry. Our main purpose was to sharpen inference about macro variables, but the restrictions also help to tighten the range of possible responses, by between 35% and 65% on impact. While the identified set is consistent with a wide range of responses, the fully Bayesian posterior indicates with 95% probability that there is mean reversion after the initial increase in shipments.

Figure 4.13: Responses of select industry shipments to defense spending shock.

4.2.2 Forecast error variance decomposition

Already pure sign restrictions help to significantly reduce the forecast error variance that could be attributed to a defense spending shock at horizons of one year and higher: According to Table D.6 in the Appendix, the macro response account could account for about 60-70% of the variance relative to what the single most important shock could account for. But heterogeneity restriction sharpen the inference even further: they imply that only 40% (71.8% minus 32.4%)of the tax rate variation at the one year horizon could be due to the spending shock, compared to 72% under sign restrictions alone. The reduction in the forecast error variance for output is almost as strong, with more modest reductions for the other macro variables.

Reduction in Maximum FEV due to heterogeneity restrictions (% of maximal FEV)

	Horizon H (year)					
Variable	0	1	2	3	4	8	10
Defense spending	5.5	4.3	4.2	4.2	4.2	4.3	4.3
Output	34.9	22.2	13.9	10.4	9.2	6.9	6.3
Debt	8.0	6.3	5.1	4.4	3.9	1.1	2.5
Hours	8.3	6.3	4.5	3.6	3.0	2.6	2.9
Tax rate	29.8	31.3	32.4	33.2	33.6	32.2	31.2
Tobacco	49.4	47.6	45.6	43.1	40.5	33.9	32.1
Lumber	65.0	52.0	44.4	40.1	37.7	30.1	27.7
Petrol	46.5	43.4	41.7	40.0	38.5	34.0	32.0
Equipment	61.5	59.1	50.4	38.0	30.2	18.7	22.2
Electronics	59.0	57.4	53.8	48.8	42.9	20.8	13.9
Transportation	29.4	16.7	10.6	8.0	6.8	4.7	4.3

We find mixed evidence on the reduction in the maximum forecast error variance (FEV) when comparing heterogeneity restrictions to pure sign restrictions. While there is a substantial set reduction of more than 30% for tax rates at all horizons, the set reduction lasts only a few years for output and is below 10% for the other macro variables. The contribution is expressed in percent of the total forecast error variance up to horizon H. All contributions are computed at the posterior mean.

Table 4.4: Maximum forecast error variance explained by defense spending shocks: Results for sign restrictions and heterogeneity restrictions

4.2.3 Important restrictions

Looking at which micro restrictions are important for our results, we find that the restrictions that petrol shipments increase more than tobacco shipments (rather than just zero) and the restriction that transportation shipment increase more than equipment shipments (rather than zero) matter the most: For example, for the lower bound of the macro variables, the restriction on transportation shipments binds 94% of the time but only 27% of the time when taking a pure non-negativity constraint. See Table 4.5.

We also find that heterogeneity restrictions substitute for restrictions on defense spending and tax rates: For example, the non-negativity constraint on defense spending binds for 39% of the lower bounds of macro variables, but only for 20% when we also have heterogeneity restrictions. This stands in contrast to our finding for the news shock, where they complemented sign restrictions on macro variables.

For lower bounds, the most important macro restriction is that taxes rise. It binds half of the time when we compute lower bounds on macro variables and almost 20% of the time for their upper bounds. Overall, the restriction that transportation shipments rise more than equipment shipments is the most important, binding for 92% of the lower and 65% of the upper bounds on macro variables. Second comes the restriction that petrol rise more than lumber, binding in 55-60% of all cases.

Binding restrictions (as a share of all cases) – all macro IRFs

Restrictions	Lower Bound SR ($\lambda = 0$)	Lower Bound HR ($\lambda = 1$)	Upper Bound SR ($\lambda = 0$)	Upper Bound HR ($\lambda = 1$)
Defense spending> 0	0.39	0.19	0.23	0.13
Output> 0	0.04	0.10	0.07	0.05
Hours> 0	0.23	0.17	0.08	0.08
Tax rate> 0	0.58	0.50	0.19	0.18
Tobacco> 0	0.51	0.71	0.56	0.75
Lumber> 0	0.36	0.46	0.26	0.40
Petrol> $\lambda\times$Tobacco	0.38	0.44	0.23	0.38
Petrol> $\lambda\times$Lumber	0.38	0.57	0.23	0.49
Equipment> $\lambda\times$Tobacco	0.10	0.23	0.26	0.35
Equipment> $\lambda\times$Lumber	0.10	0.27	0.26	0.42
Electronics> $\lambda\times$Petrol	0.25	0.27	0.40	0.38
Electronics> $\lambda\times$Equipment	0.25	0.45	0.40	0.43
Transportation> $\lambda\times$Petrol	0.27	0.39	0.29	0.59
Transportation> $\lambda\times$Equipment	0.27	0.94	0.29	0.63

We quantify the importance of restrictions by computing for which fraction of responses, averaged over horizons zero to 10 years and all reduced form draws, any given restriction is binding. Although we could distinguish at which restricted horizon any restriction is binding, we sum them together for horizons zero and one. We find that the restrictions that transportation shipment exceed those in the equipment industry is particularly important for sharpening the five macro responses. Introducing heterogeneity restrictions diminishes the importance of some sign restrictions on macro variables, in particular the requirement that defense spending rise.

Table 4.5: Importance of constraints for identifying macro impulse-responses: Incidence of binding constaints.

4.2.4 Robustness

Here, we ask how specification uncertainty affects our results. Does controlling for low frequency movements via trends matter? The comparison between responses with and without deterministic trends in the VAR in Figure D.9 in the Appendix show that for the macro variables, our findings are robust: While the specification without trends generally implies much wider credible sets, the set reduction with heterogeneity restrictions remains substantial. On impact it varies between 5% for hours worked and 15% for tax rates. While it remains modest for hours, it rises to 32% at the four year horizon and over 50% at the eight year horizon for tax rates. Most other variables experience set reduction of 10-20%. Importantly, the heterogeneity restriction indicate a more prolonged increase in defense spending and a likely increase in debt.

5 Conclusion

While inference in VARs with sign restrictions is popular, recent papers have pointed to three issues: First, an intransparent representation of the identified set. Second, unwanted or unnoticed prior dominance. Third, impractical inference in complex models. Prior-robust and accurately represented bounds can be large. Here we propose (a) algorithms for prior-robust and efficient inference and (b) the use of micro data to develop new sets of restrictions. We rank the response of micro time series to shocks according to the heterogeneous attributes of underlying industries. We develop intuitive conditions for bivariate and trivariate models when these ranking restrictions, or heterogeneity restrictions, lead to sharper inference, both at the macro and micro level.

To implement our approach in quantitative models, we develop algorithms for both prior-robust and efficient fully Bayesian inference. The prior-robust algorithm provides a distribution over the bounds of the identified set of the object of interest – impulse-responses or variance decompositions. The fully Bayesian algorithm is a novel way to draw from sign restrictions with a 100% acceptance rate, by exploiting a connection between the truncated uniform distribution of rotation vectors and a truncated multivariate normal distribution. We find that the algorithm works well in several examples, sampling even in the tails of the identified set and with soft zero restrictions. Both algorithms are of independent interest.

With these tools at hand we turn to two applications: First, we identify productivity news shocks with the help of stock return information on sectors with different R&D intensities. Second, we identify a defense spending shock with the help of information on the importance of the government as a client. We find that heterogeneity restrictions on micro data, but not pure sign restrictions, allow inference about the shape of responses for several macro variables at the posterior mean without imposing any prior over the space of rotation matrices. More generally, we find that heterogeneity restrictions cut the size of the identified set significantly, with reductions of up to 50% for macro variables in both applications.

Heterogeneity restrictions also help to sharpen fully Bayesian inferences. Interestingly, however, the extra restrictions do not simply shrink the response towards zero, but can shift posterior mass away from zero. In the productivity news example, we find that heterogeneity restrictions reveal an intuitive hump-shaped increase in TFP, whereas sign restrictions would also be consistent with a only a short-lived response. In the fiscal policy example, we find with heterogeneity restrictions that debt increases significantly after a tax-financed spending shock, whereas sign restrictions do not allow to sign the response.

Our approach to inference allows us to provide information about the importance of the different restrictions. We find that for many macro variables, the tighter heterogeneity restrictions do not substitute for the pure sign restrictions on macro variables, but represent genuinely new information. It also allows us to isolate which restrictions are most relevant in our analysis.

References

Amir-Ahmadi, Pooyan and Harald Uhlig, "Sign Restrictions in Bayesian FaVARs with an Application to Monetary Policy Shocks," Working Paper 21738, National Bureau of Economic Research November 2015.

Anderson, Emily, Atsushi Inoue, and Barbara Rossi, "Heterogeneous consumers and fiscal policy shocks," *Journal of Money Credit and Banking*, forthcoming.

Arias, Jonas E., Dario Caldara, and Juan Rubio-Ramirez, "The Systematic Component of Monetary Policy in SVARs: An Agnostic Identification Procedure," 2015. unpublished, Duke University.

_ , Juan F. Rubio-Ramirez, and Daniel F. Waggoner, "Inference Based on SVARs Identified with Sign and Zero Restrictions: Theory and Applications," Dynare Working Papers 30, CEPREMAP January 2014.

Barsky, Robert B. and Eric R. Sims, "Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence," *American Economic Review*, June 2012, *102* (4), 1343–77.

Baumeister, Christiane and James D. Hamilton, "Sign Restrictions, Structural Vector Autoregressions, and Useful Prior Information," *Econometrica*, 09 2015, *83* (5), 1963–1999.

Beaudry, Paul and Franck Portier, "Stock Prices, News, and Economic Fluctuations," *American Economic Review*, September 2006, *96* (4), 1293–1307.

_ and _ , "News-Driven Business Cycles: Insights and Challenges," *Journal of Economic Literature*, December 2014, *52* (4), 993–1074.

_ , Deokwoo Nam, and Jian Wang, "Do Mood Swings Drive Business Cycles and is it Rational?," NBER Working Papers 17651, National Bureau of Economic Research December 2011.

Blanchard, Olivier and Roberto Perotti, "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output," *The Quarterly Journal of Economics*, 2002, *117* (4), 1329–1368.

_ , Eugenio Cerutti, and Lawrence Summers, "Inflation and Activity – Two Explorations and their Monetary Policy Implications," NBER Working Papers 21726, National Bureau of Economic Research November 2015.

Blanchard, Olivier Jean and Danny Quah, "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, September 1989, *79* (4), 655–73.

Botev, Z. I., "The normal law under linear restrictions: simulation and estimation via minimax tilting," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 2016, pp. n/a–n/a.

Canova, Fabio, "Methods for Applied Macroeconomic Research," *Princeton University Press*, 2007.

_ and Gianni De Nicolo, "Monetary disturbances matter for business fluctuations in the G-7," *Journal of Monetary Economics*, September 2002, *49* (6), 1131–1159.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans, "Monetary policy shocks: What have we learned and to what end?," in J. B. Taylor and M. Woodford, eds., *Handbook of Macroeconomics*, Vol. 1 of *Handbook of Macroeconomics*, Elsevier, 1999, chapter 2, pp. 65–148.

Cogley, Timothy and Thomas J. Sargent, "Drift and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S.," *Review of Economic Dynamics*, April 2005, *8* (2), 262–302.

De Graeve, Ferre and Alexei Karas, "Evaluating Theories of Bank Runs with Heterogeneity Restrictions," *Journal of the European Economic Association*, 2014, *12* (4), 969–996.

DiTraglia, Francis J. and Camilo García-Jimeno, "A Framework for Eliciting, Incorporating, and Disciplining Identification Beliefs in Linear Model," 2016. unpublished, University of Pennsylvania.

Fama, Eugene F. and Kenneth R. French, "Industry costs of equity," *Journal of Financial Economics*, 1997, *43* (2), 153 – 193.

Faust, Jon, "The robustness of identified VAR conclusions about money," *Carnegie-Rochester Conference Series on Public Policy*, December 1998, *49* (1), 207–244.

Fernald, John G, "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity.," FRBSF Working Paper 2012-19 (updated March 2014)., FRB SF 2014.

Giacomini, Raffaella and Toru Kitagawa, "Inference about Non-Identied SVARs," CeMMAP working papers CWP45/14, Centre for Microdata Methods and Practice, Institute for Fiscal Studies November 2014.

Inoue, Atsushi and Lutz Kilian, "Inference on impulse response functions in structural {VAR} models," *Journal of Econometrics*, 2013, *177* (1), 1 – 13.

Kline, Brendan and Elie Tamer, "Bayesian Inference In A Class Of Partially Identified Models," *Quantitative Economics*, forthcoming.

Li, Yifang and Sujit K. Ghosh, "Efficient Sampling Methods for Truncated Multivariate Normal and Student-t Distributions Subject to Linear Inequality Constraints," *Journal of Statistical Theory and Practice*, 2015, *9* (4), 712–732.

Moon, Hyungsik Roger and Frank Schorfheide, "Bayesian and Frequentist Inference in Partially Identified Models," *Econometrica*, 2012, *80* (2), 755–782.

_ , _ , and Eleonora Granziera, "Inference for VARs Identified with Sign Restrictions," Technical Report, University of Southern California 2013.

Nekarda, Christopher J. and Valerie A. Ramey, "Industry Evidence on the Effects of Government Spending," *American Economic Journal: Macroeconomics*, January 2011, *3* (1), 36–59.

Perotti, Roberto, "In Search of the Transmission Mechanism of Fiscal Policy," in "NBER Macroeconomics Annual 2007, Volume 22" NBER Chapters, National Bureau of Economic Research, February 2008, pp. 169–226.

Primiceri, Giorgio, "Time Varying Structural Vector Autoregressions and Monetary Policy," *Review of Economic Studies*, July 2005, *72* (3), 821–852.

Ramey, Valerie A., "Identifying Government Spending Shocks: It's all in the Timing," *The Quarterly Journal of Economics*, 2011, *126* (1), 1–50.

Rubio-Ramírez, Juan F., Daniel F. Waggoner, and Tao Zha, "Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference," *Review of Economic Studies*, 2010, *77* (2), 665–696.

Schorfheide, Frank, "Macroeconometrics - A Discussion," Discussion 2016.

Sims, Christopher A, "Macroeconomics and Reality," *Econometrica*, January 1980, *48* (1), 1–48.

_ **and Tao Zha**, "Bayesian Methods for Dynamic Multivariate Models," *International Economic Review*, November 1998, *39* (4), 949–68.

Stewart, G. W., "The Efficient Generation of Random Orthogonal Matrices with an Application to Condition Estimators," *SIAM Journal on Numerical Analysis*, 1980, *17* (3), 403–409.

Uhlig, Harald, "What moves GNP?," *draft, Humboldt Universität zu Berlin*, 2003.

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

A Derivations

A.1 Bivariate VAR(0)

Recall the restrictions:

$$a_1 \equiv q_1 \tilde{A}_{11} \geq 0$$
$$a_2 \equiv q_1 \tilde{A}_{21} + q_2 \tilde{A}_{22} \geq \lambda q_1 \tilde{A}_{11}$$

Trivially, the lower bound for a_1 of zero is always within our set: $\underline{a}_1 = 0$.

Note that if the heterogeneity restriction binds with equality, we have that:

$$q_1 = \frac{\tilde{A}_{22}}{\sqrt{\tilde{A}_{22}^2 + (\tilde{A}_{21}^2 - \lambda \tilde{A}_{11})^2}} \qquad q_2 = \pm \frac{|\tilde{A}_{21} - \lambda \tilde{A}_{11}|}{\sqrt{\tilde{A}_{22}^2 + (\tilde{A}_{21}^2 - \lambda \tilde{A}_{11})^2}}$$

Case (a) $\tilde{A}_{21} \leq 0$.

- Upper bound for a_2: Since $q_1 \geq 0$, the upper bound for a_2 is, trivially, $\bar{a}_2 = \tilde{A}_{22}$.
- Lower bound for a_2: Since $\tilde{A}_{22} > 0$, the lower bound is attained by the largest q_1 and the lowest q_2, i.e. with a binding heterogeneity restriction for $q_2 > 0$. Then: the lower bound for a_2 is $\underline{a}_2 = \frac{\lambda \tilde{A}_{11} \tilde{A}_{22}}{\sqrt{\tilde{A}_{22}^2 + (\tilde{A}_{21}^2 - \lambda \tilde{A}_{11})^2}}$.
- Upper bound for a_1: $\bar{a}_1$ is also associated with the binding heterogeneity restriction: $\bar{a}_1 = \frac{\tilde{A}_{22}}{\sqrt{\tilde{A}_{22}^2 + (\tilde{A}_{21}^2 - \lambda \tilde{A}_{11})^2}} \tilde{A}_{11}$.

Case (b) $\lambda \tilde{A}_{11} - \tilde{A}_{21} \leq 0, \tilde{A}_{21} \geq 0$.

- Upper bound for a_2: a_2 is now weakly positive and the heterogeneity constraint is slack. The SOC for the unique interior extremum to be a maximum always holds. At the interior extremum, $q_1 = \frac{\tilde{A}_{21}}{\sqrt{\tilde{A}_{22}^2 + \tilde{A}_{21}^2}}$ and $q_2 = \frac{\tilde{A}_{22}}{\tilde{A}_{21}} q_1$. Thus: $\bar{a}_2 = \sqrt{\tilde{A}_{22}^2 + \tilde{A}_{21}^2}$.
- Lower bound for a_2: A negative q_2 is now possible, but constrained by the heterogeneity constraint, as its RHS is increasing faster in q_1 than its LHS. Thus, the lower bound is associated with a binding heterogeneity constraint and $\underline{a}_2 = \frac{\lambda \tilde{A}_{11} \tilde{A}_{22}}{\sqrt{\tilde{A}_{22}^2 + (\tilde{A}_{21}^2 - \lambda \tilde{A}_{11})^2}}$.
- Upper bound for a_1: Since $q_2 = 0, q_1 = 1$ is possible, the upper bound is simply $\bar{a}_1 = \tilde{A}_{11}$.

Case (c) $\lambda \tilde{A}_{11} - \tilde{A}_{21} \geq 0, \tilde{A}_{21} \geq 0$ or $0 \leq \rho \leq \lambda \frac{\sqrt{\Sigma_{11}}}{\sqrt{\Sigma_{22}}}$.

- Upper bound for a_2: We proceed by brute force, checking whether the heterogeneity constrained is binding at the unconstrained maximum. We find that if $\lambda \leq \frac{\tilde{A}_{22}^2 + \tilde{A}_{21}^2}{\tilde{A}_{11} \tilde{A}_{21}} = \frac{\Sigma_{22}}{\Sigma_{21}}$, the heterogeneity constraint is slack. Thus:

$$\bar{a}_2 = \begin{cases} \sqrt{\tilde{A}_{22}^2 + \tilde{A}_{21}^2} & \lambda \leq \frac{\tilde{A}_{22}^2 + \tilde{A}_{21}^2}{\tilde{A}_{11} \tilde{A}_{21}} = \frac{\Sigma_{22}}{\Sigma_{21}} = \frac{1}{\rho} \frac{\sqrt{\Sigma_{22}}}{\sqrt{\Sigma_{11}}} \\ \frac{\lambda \tilde{A}_{11} \tilde{A}_{22}}{\sqrt{\tilde{A}_{22}^2 + (\tilde{A}_{21}^2 - \lambda \tilde{A}_{11})^2}} & \lambda \geq \frac{\tilde{A}_{22}^2 + \tilde{A}_{21}^2}{\tilde{A}_{11} \tilde{A}_{21}} = \frac{\Sigma_{22}}{\Sigma_{21}} = \frac{1}{\rho} \frac{\sqrt{\Sigma_{22}}}{\sqrt{\Sigma_{11}}} \end{cases}$$

- Lower bound for a_2: Since the interior extremum is always a maximum, we check the corners. Comparing the two corners, we find:

$$\underline{a}_2 = \begin{cases} \tilde{A}_{22} & \lambda \geq \frac{1}{2}\frac{\tilde{A}_{22}^2+\tilde{A}_{21}^2}{\tilde{A}_{11}\tilde{A}_{21}} = \frac{1}{2}\frac{\Sigma_{22}}{\Sigma_{21}} = \frac{1}{2}\frac{1}{\rho}\frac{\sqrt{\Sigma_{22}}}{\sqrt{\Sigma_{11}}} \\ \frac{\lambda\tilde{A}_{11}\tilde{A}_{22}}{\sqrt{\tilde{A}_{22}^2+(\tilde{A}_{21}^2-\lambda\tilde{A}_{11})^2}} & \lambda \leq \frac{1}{2}\frac{\tilde{A}_{22}^2+\tilde{A}_{21}^2}{\tilde{A}_{11}\tilde{A}_{21}} = \frac{1}{2}\frac{\Sigma_{22}}{\Sigma_{21}} = \frac{1}{2}\frac{1}{\rho}\frac{\sqrt{\Sigma_{22}}}{\sqrt{\Sigma_{11}}} \end{cases}$$

- Upper bound for a_1: $\bar{a}_1$ is also associated with the binding heterogeneity restriction: $\bar{a}_1 = \frac{\tilde{A}_{22}}{\sqrt{\tilde{A}_{22}^2+(\tilde{A}_{21}^2-\lambda\tilde{A}_{11})^2}}\tilde{A}_{11}$.

A.2 Trivariate VAR(0)

A.2.1 Identified set

Here we only consider bounds for a_1. We seek a solution to the following problem:

$$\min_q \text{ or } \max_q \tilde{A}_{11}q_1 \tag{A.1a}$$

$$\text{s.t. } ||q|| = 1 \tag{A.1b}$$

$$\tilde{A}_{11}q_1 \geq 0 \tag{A.1c}$$

$$\tilde{A}_{21}q_1 + \tilde{A}_{22}q_2 \geq 0$$

$$\underbrace{(\tilde{A}_{31} - \lambda\tilde{A}_{21})}_{\equiv\tilde{A}_{31}^\lambda} q_1 + \underbrace{(\tilde{A}_{32} - \lambda\tilde{A}_{22})}_{\equiv\tilde{A}_{32}^\lambda} q_2 + \tilde{A}_{33}q_3 \geq 0 \tag{A.1d}$$

Since $\tilde{A}_{ii} > 0 \forall i$, we can write equivalently:

$$\min_q \text{ or } \max_q \sqrt{1-(q_2)^2-(q_3)^3}$$

$$\text{s.t. } \tilde{A}_{21}\sqrt{1-(q_2)^2-(q_3)^3} + \tilde{A}_{22}q_2 \geq 0$$

$$\underbrace{(\tilde{A}_{31} - \lambda\tilde{A}_{21})}_{\equiv\tilde{A}_{31}^\lambda} \sqrt{1-(q_2)^2-(q_3)^3} + \underbrace{(\tilde{A}_{32} - \lambda\tilde{A}_{22})}_{\equiv\tilde{A}_{32}^\lambda} q_2 + \tilde{A}_{33}q_3 \geq 0$$

Note that $\underline{a}_1 = 0$ is always feasible by setting $q_3 = 1$. We therefore focus on the maximization problem

Using Lagrange multipliers ν_{SR} and ν_{HR} to denote the inequality constraints we can equivalently write the Lagrangian as

$$\min_{\nu_{SR},\nu_{HR}} \max_{q_2,q_3} \mathcal{L} = \sqrt{1-(q_2)^2-(q_3)^3} - \nu_{SR}(\tilde{A}_{21}\sqrt{1-(q_2)^2-(q_3)^3} + \tilde{A}_{22}q_2)$$
$$- \nu_{HR}(\tilde{A}_{31}^\lambda\sqrt{1-(q_2)^2-(q_3)^3} + \tilde{A}_{32}^\lambda q_2 + \tilde{A}_{33}q_3)$$

with the associated Kuhn-Tucker conditions as:

$$[q_2] - \frac{q_2}{\sqrt{1-(q_2)^2-(q_3)^3}}(1 - \nu_{SR}\tilde{A}_{21} - \nu_{HR}\tilde{A}_{31}^\lambda) = \nu_{SR}\tilde{A}_{22} + \nu_{HR}\tilde{A}_{32}^\lambda$$
$$\nu_{SR}(\tilde{A}_{21}\sqrt{1-(q_2)^2-(q_3)^3} + \tilde{A}_{22}q_2) = 0$$
$$\nu_{SR} \geq 0$$
$$[\nu_{SR}]\tilde{A}_{21}\sqrt{1-(q_2)^2-(q_3)^3} + \tilde{A}_{22}q_2 \geq 0.$$

$$[q_3] - \frac{q_3}{\sqrt{1-(q_2)^2-(q_3)^3}}(1 - \nu_{SR}\tilde{A}_{21} - \nu_{HR}\tilde{A}^{\lambda}_{31}) = \nu_{HR}\tilde{A}_{33}$$
$$\nu_{HR}(\tilde{A}^{\lambda}_{21}\sqrt{1-(q_2)^2-(q_3)^3} + \tilde{A}^{\lambda}_{22}q_2 + \tilde{A}_{33}q_3) = 0$$
$$\nu_{HR} \geq 0$$
$$[\nu_{HR}]\tilde{A}^{\lambda}_{31}\sqrt{1-(q_2)^2-(q_3)^3} + \tilde{A}^{\lambda}_{32}q_2 + \tilde{A}_{33}q_3 \geq 0.$$

Clearly, the Kuhn-Tucker conditions show that the unconstrained optimum, when the multipliers ν_{SR}, ν_{HR} are zero, involves setting $q_2 = q_3 = 0$.

We assume throughout that $\lambda > 0$.

1. **All (conditional) covariances positive, heterogeneity restriction weak:**
   Note that when $0 \leq \tilde{A}_{21} \leq \frac{1}{\lambda}\tilde{A}_{31}$, then $\tilde{A}^{\lambda}_{31} \geq 0$ and $q_2 = q_3 = \nu_{SR} = \nu_{HR} = 0$ is a local extremum – specifically, an optimum. All conditions are trivially satisfied at zero. This equals the unconstrained optimum.

2. **All (conditional) covariances positive, heterogeneity restriction strong:**
   Note that when $\tilde{A}_{21} > \frac{1}{\lambda}\tilde{A}_{31} > 0$, then $\tilde{A}^{\lambda}_{31} < 0$. $q_2 = q_3 = \nu_{SR} = \nu_{HR} = 0$ no longer satisfies the optimality conditions with $\lambda > 0$, since the HR constraint is violated at this candidate point. With $\lambda = 0$, however, $q_2 = q_3 = 0$ is feasible and the unconstrained maximum attains.

3. **$\tilde{A}_{21} < 0 < \tilde{A}_{31}$:**
   In this case, the first sign restriction is binding, and the second restriction is slack for any $\lambda \geq 0$ that still satisfies $\tilde{A}^{\lambda}_{32} \geq 0$. The constrained maximum with sign restrictions is attained at:

   $$q_1^* = \sqrt{1 + \left(\frac{\tilde{A}_{21}}{\tilde{A}_{22}}\right)^2}^{-1}.$$

   Note that $\tilde{A}^{\lambda}_{31} > 0$ by construction.

4. **$\tilde{A}_{31} < 0 < \tilde{A}_{21}$:**

   In this case, the second sign restriction is binding, and the first restriction is slack. The heterogeneity restriction leads to a tighter bound because the slope $\tilde{A}^{\lambda}_{31}$ (the penalty for larger q_1) increases, while the ability to compensate via higher q_2 decreases ($\tilde{A}^{\lambda}_{32} < \tilde{A}_{32}$).

A.2.2 General signs

$$\min_q \text{ or } \max_q \tilde{A}_{11}q_1 \tag{A.2a}$$
$$\text{s.t. } ||q|| = 1 \tag{A.2b}$$
$$s_1\tilde{A}_{11}q_1 \geq 0 \tag{A.2c}$$
$$s_2(\tilde{A}_{21}q_1 + \tilde{A}_{22}q_2) \geq 0 \tag{A.2d}$$
$$s_3\left(\underbrace{(\tilde{A}_{31} - \lambda\tilde{A}_{21})}_{\equiv\tilde{A}^{\lambda}_{31}}q_1 + \underbrace{(\tilde{A}_{32} - \lambda\tilde{A}_{22})}_{\equiv\tilde{A}^{\lambda}_{32}}q_2 + \tilde{A}_{33}q_3\right) \geq 0 \tag{A.2e}$$

Now, define $\hat{A}_{ij} = \tilde{A}_{ij}$ if $i = j$ and $\hat{A}_{ij} = \tilde{A}_{ij} \times s_j \prod_{\ell=1}^{j} s_\ell$ for $i = 1, 2, 3$ and $j \leq i$. Also define $\hat{q}_i = s_i q_i$ and $\hat{\lambda} = s_3 \lambda$. Then we can re-write problem (A.2) as

$$\min_q \text{ or } \max_q s_1 \tilde{A}_{11} \hat{q}_1 \tag{A.3a}$$

$$\text{s.t. } ||q|| = 1 \tag{A.3b}$$

$$\tilde{A}_{11} q_1 \geq 0 \tag{A.3c}$$

$$\hat{A}_{21} \hat{q}_1 + \tilde{A}_{22} \hat{q}_2 \geq 0 \tag{A.3d}$$

$$\underbrace{(\hat{A}_{31} - \hat{\lambda} \hat{A}_{21})}_{\equiv \hat{A}_{31}^{\lambda}} \hat{q}_1 + \underbrace{(\hat{A}_{32} - \hat{\lambda} \hat{A}_{22})}_{\equiv \hat{A}_{32}^{\lambda}} \hat{q}_2 + \tilde{A}_{33} \hat{q}_3 \geq 0, \tag{A.3e}$$

whose constraints are of the same form as (A.2). Thus, the previous solution applies to the transformed vector $\hat{q}$ in terms of the transformed coefficients $\hat{A}_{ij}$. However, if $s_1 = -1$, maximization and minimization are interchanged.

Thus, the sufficient condition for set reduction becomes $\hat{\lambda} \hat{A}_{21} > \hat{A}_{31} > 0$. In terms of the original components:

$$\lambda s_3 s_2 s_1 \tilde{A}_{21} > s_3 s_1 \tilde{A}_{31} > 0.$$

For these sufficient conditions to apply we need that $\hat{A}_{21} > 0$ and $\hat{A}_{31} > 0$. In terms of the original components:

$$s_2 s_1 \tilde{A}_{21} > 0, \qquad s_3 s_1 \tilde{A}_{31} > 0.$$

Examples include:

1. Traditional New Keynesian example: Variable 1 is the funds rate. Variable 2 is a real activity measure. Variable 3 is the measure of prices. $s_1 = s_3 = -1$. $s_2 = 1$: Inflation and real activity fall, the FFR rises.

   (a) Interest rates rise: $s_1 = +1$

   (b) Industrial production or PCE falls: $s_2 = -1$, and

   (c) Prices fall (and more than $\lambda \times$ output): $s_3 = -1$.

   Thus, the sufficient condition here becomes: $\lambda \tilde{A}_{21} > -\tilde{A}_{31} > 0$. Equivalently: $-\lambda \tilde{A}_{21} < \tilde{A}_{31} < 0$. For this condition to apply we also need that $\hat{A}_{21}, \hat{A}_{31} > 0$, or $-\tilde{A}_{21} > 0$ and $\tilde{A}_{31} > 0$ in this example.

   Choleski of covariance matrix MLE estimate for FFR, PCE prices, and PCE quantities:

$$\begin{bmatrix} 0.5086 & 0 & 0 \\ 0.0610 & 0.4899 & 0 \\ -0.0022 & 0.0425 & 0.1493 \end{bmatrix}$$

   Thus, $\tilde{A}_{31} < 0$ and our theorem does not apply.

Choleski of covariance matrix MLE estimate for FFR, PCE prices, and IP quantities:

$$\begin{bmatrix} 0.4990 & 0 & 0 \\ 0.1645 & 0.5877 & 0 \\ 0.0002 & 0.0052 & 0.1564 \end{bmatrix}$$

Thus, $\tilde{A}_{21} > 0$ and our theorem does not apply. Because our conditions are only sufficient, we also verify the lack of set reduction numerically. As the right panel of Figure A.1 shows, there is no set reduction in the Federal Funds Rate response, nor in the response of prices. By construction, higher λ enables us to impose soft zero restrictions.

2. New Keynesian housing example: Variable 1 is the interest rate. Variable 2 is the measure of housing starts, variable 3 of house price inflation. $s_1 = s_3 = -1$. $s_2 = 1$: Inflation and real activity fall, the FFR rises.

   (a) Interest rates rise: $s_1 = +1$
   (b) Housing starts: $s_2 = -1$, and
   (c) House prices fall (and more than $\lambda\times$ output): $s_3 = -1$.

   Thus, the sufficient condition here becomes: $\lambda\tilde{A}_{21} > -\tilde{A}_{31} > 0$. Equivalently: $-\lambda\tilde{A}_{21} < \tilde{A}_{31} < 0$. For this condition to apply we also need that $\hat{A}_{21}, \hat{A}_{31} > 0$, or $-\tilde{A}_{21} > 0$ and $\tilde{A}_{31} > 0$ in this example.

   Choleski of covariance matrix MLE estimate for FFR, housing prices, and median house prices:

$$\begin{bmatrix} 0.5086 & 0 & 0 \\ -0.6119 & 6.5659 & 0 \\ -0.0567 & 0.0538 & 2.6529 \end{bmatrix}$$

   Thus, $\tilde{A}_{31} < 0$ and our theorem does not apply. However, we still find a very modest set reduction, see the upper panel in Figure A.2.

   If we replace the median house price with the Case-Shiller index we find that the following Choleski factor of the covariance matrix MLE estimate:

$$\begin{bmatrix} 0.5035 & 0 & 0 \\ -0.4178 & 6.5342 & 0 \\ 0.0105 & 0.0090 & 2.1261 \end{bmatrix}$$

   Now our theoretical results also apply formally and we expect a set reduction. The bottom panel of Figure A.2 displays the results and shows that the set reduction is there, but negligible. In both case it is clear how the large value for λ imposes a soft zero restriction on housing starts, as intended.

3. Blanchard and Perotti (2002) example: Variable 1 becomes output. Variable 2 is government consumption. Variable 3 is the tax rate. $s_1 = 1$ (arbitrary), $s_2 = +1, s_3 = +1$.

   (a) Output rises: $s_1 = +1$,
   (b) G rise: $s_2 = +1$.

Figure A.1: Set-reduction for impact response in traditional NK application

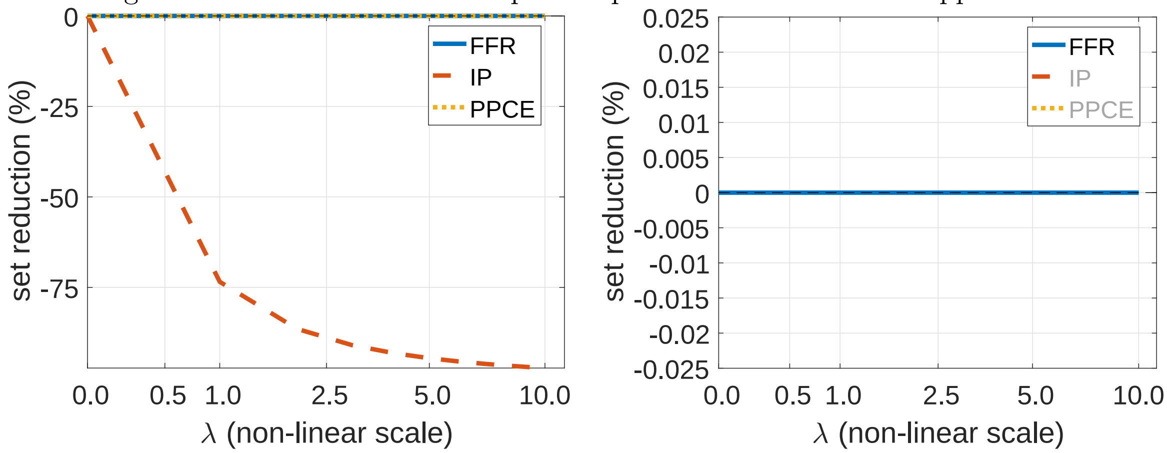
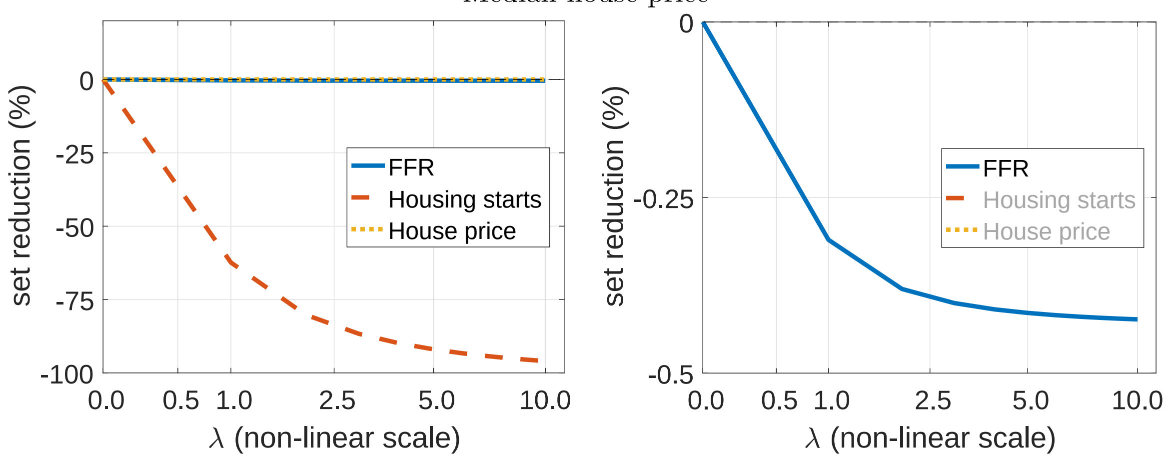
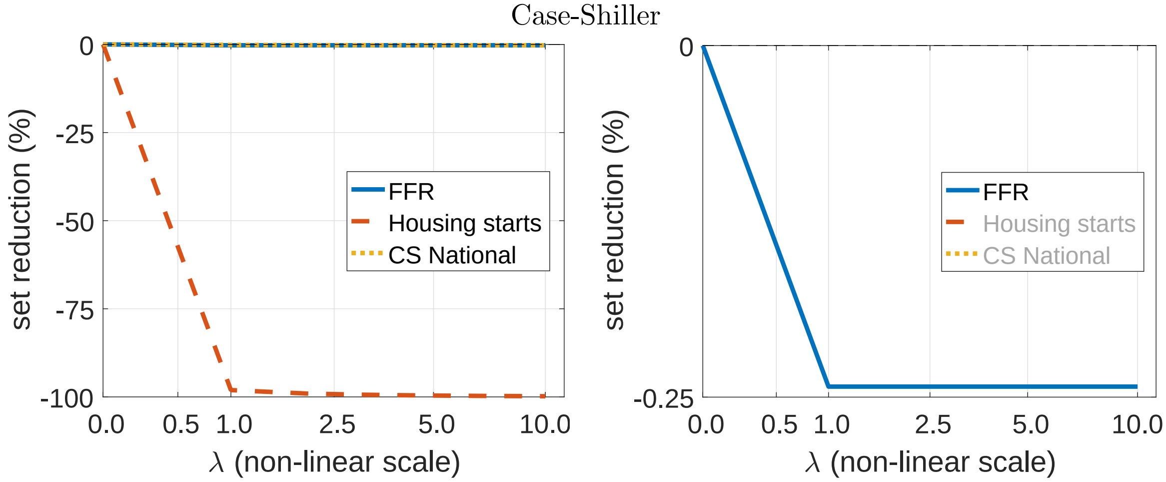


Figure A.2: Set-reduction for impact response in NK housing application

Median house price



Case-Shiller



(c) Taxes τ rise (and more than $\lambda\times$ government spending): $s_3 = +1$, and

For high values of λ, this restriction imposes a "soft" zero restriction on government spending: Spending does not rise (significantly) on impact in response to tax shocks.

The sufficient condition is thus simply: $\lambda \tilde{A}_{21} > \tilde{A}_{31} > 0$.

Choleski of Blanchard and Perotti (2002) covariance matrix MLE estimate, after ordering:

$$\begin{bmatrix} 0.0086 & 0 & 0 \\ 0.0135 & 0.0220 & 0 \\ 0.0044 & -0.0007 & 0.0232 \end{bmatrix}$$

Thus, $\tilde{A}_{31} = 0.0044 > 0$ and $\lambda \tilde{A}_{21} > \tilde{A}_{31}$ iff $\lambda > \frac{0.44}{1.35} \approx \frac{1}{3}$. Figure A.3 shows the corresponding set reduction. Note the non-linear scale of λ that shows that for small λ there is not set reduction, confirming our theoretical analysis.

4. Productivity news example (inspired by Beaudry and Portier, 2006): Variable 1 is output growth. Variable 2 is utilization-adjusted TFP growth (Fernald, 2014). Variable 3 is the real growth of the Wilshire 5000 index.

   (a) Output rises: $s_1 = +1$,

   (b) TFP τ does not fall: $s_2 = +1$, and

   (c) The stock market rises (and more than $\lambda\times$ TFP): $s_3 = +1$.

   For high values of λ, this restriction imposes a "soft" zero restriction on TFP: TFP does not rise (significantly) on impact in response to positive news.

   The Choleski of the covariance matrix MLE estimate, after ordering:

$$\begin{bmatrix} 0.44 & 0 & 0 \\ 1.02 & 2.47 & 0 \\ 0.75 & 0.68 & 4.61 \end{bmatrix}$$

   Here, $\tilde{A}_{31} = 0.75 > 0$ and $\lambda \tilde{A}_{21} > \tilde{A}_{31}$ iff $\lambda > \frac{0.75}{1.02} \approx \frac{3}{4}$. Figure A.4 shows the corresponding set reduction. Note the non-linear scale of λ that shows again that for small λ there is not set reduction.

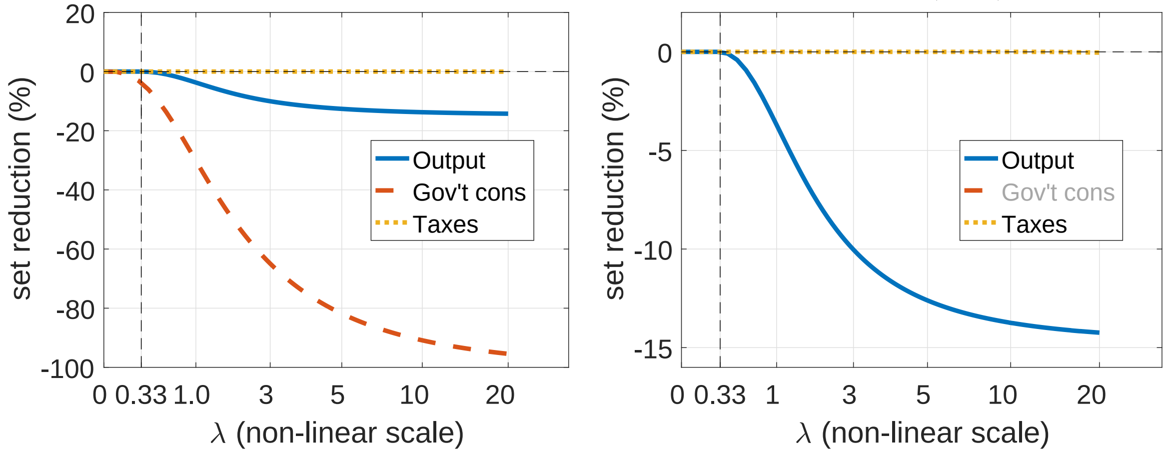
A.2.3 Redundant restrictions

Consider a three variable, three-shock case where the true impulse matrix is given by:

$$A = \begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \Rightarrow AA' = \begin{bmatrix} a_{11}^2 + a_{12}^2 & a_{11}a_{21} + a_{22}a_{12} & a_{11}a_{31} + a_{12}a_{32} \\ a_{21}a_{11} + a_{22}a_{12} & a_{21}^2 + a_{22}^2 & a_{21}a_{31} + a_{22}a_{32} \\ a_{31}a_{11} + a_{12}a_{32} & a_{31}a_{21} + a_{32}a_{22} & a_{31}^2 + a_{33}^2 + a_{32}^2 \end{bmatrix} \quad \text{(A.4)}$$

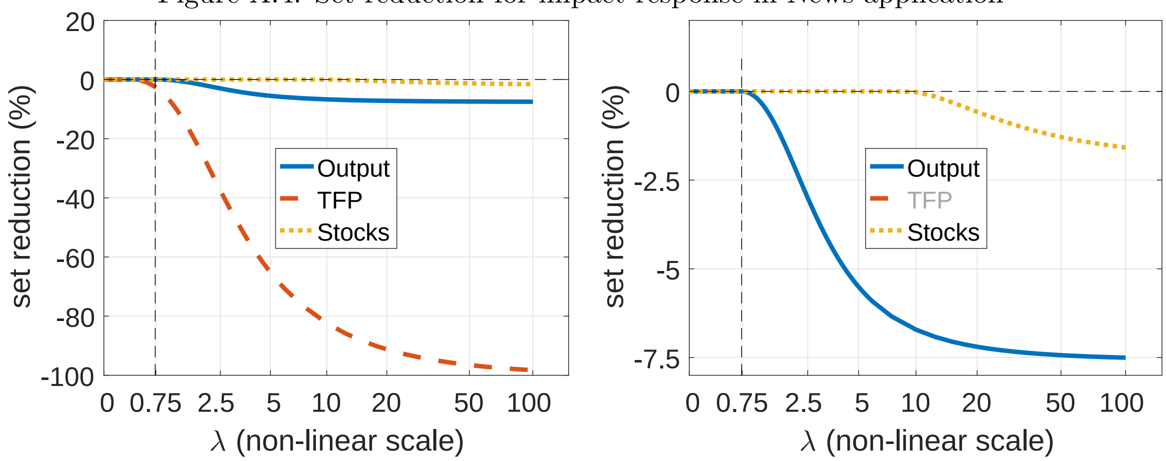
One interpretation of this structure is that there are only two aggregate shocks. These aggregate shocks affect all three variables while the third variables also contains a third idiosyncratic shock.

Figure A.3: Set-reduction for impact response in Blanchard and Perotti (2002) application



Note: The dashed vertical line marks the treshold for λ above which there is set reduction for the output response, i.e. $\bar{\lambda} = \tilde{A}_{31}/\tilde{A}_{21}$.

Figure A.4: Set-reduction for impact response in News application



Note: The dashed vertical line marks the treshold for λ above which there is set reduction for the output response, i.e. $\bar{\lambda} = \tilde{A}_{31}/\tilde{A}_{21}$.

The lower-triangular Choleski decomposition is given by:

$$\begin{aligned}
\tilde{A}_{11} &= \sqrt{a_{11}^2 + a_{12}^2} \\
\tilde{A}_{21} &= \frac{a_{21}a_{11} + a_{22}a_{12}}{\tilde{A}_{11}} \\
\tilde{A}_{22} &= \sqrt{a_{21}^2 + b_{22}^2 - \tilde{A}_{21}^2} \\
\tilde{A}_{31} &= \frac{a_{31}a_{11} + a_{12}a_{32}}{\tilde{A}_{11}} \\
\tilde{A}_{32} &= \frac{a_{31}a_{21} + a_{32}a_{22} - \tilde{A}_{31}\tilde{A}_{21}}{\tilde{A}_{22}} \\
\tilde{A}_{33} &= \sqrt{a_{31}^2 + a_{33}^2 + a_{32}^2 - \tilde{A}_{31}^2 - \tilde{A}_{32}^2}
\end{aligned}$$

Now consider the case that $a_{31} = \kappa a_{21}$ and $a_{32} = \kappa a_{22}$. In that case:

$$\tilde{A}_{31} = \kappa \tilde{A}_{21}, \tag{A.5a}$$

$$\tilde{A}_{32} = \kappa \tilde{A}_{22}, \tag{A.5b}$$

$$\tilde{A}_{33} = |a_{33}|. \tag{A.5c}$$

We now show that if the heterogeneity restrictions are weaker than those of the data generating process, i.e. $\lambda \leq \kappa$, then adding the heterogeneity restrictions does not change the identified set for variables 1 and 3. We start by stating the problem:[29]

$$\max_q e_i' \tilde{A} q, \qquad i \in \{1, 2\}, \tag{A.6a}$$

$$\text{s.t. } ||q|| = 1 \tag{A.6b}$$

$$e_1' \tilde{A} q \geq 0 \tag{A.6c}$$

$$e_2' \tilde{A} q \geq 0 \tag{A.6d}$$

$$(e_3 - \lambda e_2)' \tilde{A} q \geq 0, \tag{A.6e}$$

where $\tilde{A}$ is the Cholesky factor of AA' in (A.4) that satisfies (A.5). e_i denotes a selection vector with zeros except for a one in the ith position.

We derive the Kuhn-Tucker conditions using a Lagrangean:

$$\min_{\mu, \nu_i \geq 0} \max_q \mathcal{L} = e_i' \tilde{A} q + \mu(1 - ||q||) - \sum_{j=1}^{2} \nu_j e_j' \tilde{A} q - \nu_3 (e_3 - \lambda e_2)' \tilde{A} q$$

The necessary conditions are:

$$\left(e_i' - \sum_{j=1}^{2} \nu_j e_j' \tilde{A} - \nu_3 (e_3 - \lambda e_2)' \right) \tilde{A} - 2\mu q' = 0$$

[29] We focus on the upper bounds because we can always attain the lower bound of zero.

$$\nu_j e_j' \tilde{A} q = 0 \qquad\qquad \nu_j \geq 0 \qquad ,j = 1,2$$
$$\nu_3 (e_3 - \lambda e_2)' \tilde{A} q = 0 \qquad\qquad \nu_3 \geq 0.$$

Note that $\nu_{3-i} = 0$ for $i = 1,2$, by the complementary slackness condition. We now guess and verify that we can ignore the heterogeneity restrictions, i.e., the third set of restrictions. Simplifying:

$$q_{SR}' = \frac{1}{|| \left(e_i' - \nu_{3-i} e_{3-i}' \tilde{A} \right) ||} \left(e_i' - \nu_{3-i} e_{3-i}' \tilde{A} \right)$$
$$\nu_{3-i} e_{3-i}' \tilde{A} q = 0.$$

Note that $e_3' q_{SR} = 0$. Now, does this solution satisfy the heterogeneity restriction?

$$\begin{aligned}
(e_3 - \lambda e_2)' \tilde{A} q_{SR} &= e_3' \tilde{A} q_{SR} - \lambda e_2' \tilde{A} q_{SR} \\
&= [\kappa e_2' \tilde{A} + e_3' \tilde{A}] q_{SR} - \lambda e_2' \tilde{A} q_{SR} \\
&= \kappa e_2' \tilde{A} q_{SR} - \lambda e_2' \tilde{A} q_{SR} \\
&= (\kappa - \lambda) e_2' \tilde{A} q_{SR} \geq 0,
\end{aligned}$$

where the last inequality follows from $\kappa \geq \lambda$ and the sign restriction $e_2 \tilde{A} q \geq 0$. Thus, the solution without heterogeneity restriction is also a solution with heterogeneity restriction. Thus, the upper bound coming from the heterogeneity restriction is not binding when $\lambda \leq \kappa$, i.e. the imposed restriction is weaker than the one implied by the data generating process.

Intuitively, in this case the heterogeneity restrictions have no bite because they do not help to tell the first shock from the second shock, because in the data generating process, responses to both shocks satisfy the heterogeneity restrictions. They are, thus, redundant.

B Forecast error variance decomposition

The total forecast error variance for Y_{t+H} given information up to time t is given by:

$$FEV_H = \sum_{h=0}^{H}((B_X^h \tilde{A})(B_X^h \tilde{A})').$$

We can decompose the forecast error variance into the contribution due to an identified shock with impulse-vector $\tilde{A}q$. We call this the conditional forecast error variance (CFEV):

$$CFEV_H(q) = \sum_{h=0}^{H}((B_X^h \tilde{A}q)(B_X^h \tilde{A}q)').$$

Let $CFEV_{i,H}(q)$ by the (i,i)th element of the CFEV. As shown by Uhlig (2003), we can rewrite $CFEV_{i,H}(q)$ as:

$$CFEV_{i,H}(q) = \sum_{h=0}^{H}((B_X^h \tilde{A}q)(B_X^h \tilde{A}q)')_{(ii)} = q' S_{i,H} q, \tag{B.1}$$

$$S_{i,H} \equiv \sum_{h=0}^{H}(H+1-h)(e_i B_X^h \tilde{A})'(e_i B_X^h \tilde{A}). \tag{B.2}$$

We can compute the upper and lower bound on $CFEV_{i,H}$ simply by replacing the objective function algorithm in Section 4 by $q'S_{i,H}q$ and keeping the same set of constraints.

To convert the forecast error variance explained by the identified shock, we normalize $CFEV_{i,H}(q)$ by the total forecast error variance for variable i up to horizon H.

C Data

C.1 News data

We use the following macro variables, taken, unless otherwise stated, from the St. Louis Fed FRED website:

- Real GDP `GDPC1`
- Hours worked (nonfarm, business sector) `PRS85006032` (growth, accumulated)
- Consumer confidence `CSCICP03USM665S`
- PCE price index `PCEPI`
- Utilization adjusted TFP: Fernald (2014) (accumulated)

All variables enter the VAR in log-levels.

We use industry data from Ken French's data library, based on Fama and French (1997). Specifically, we use the FF5 industry returns, and convert them to real ex post returns using the change in the log of the PCE price index.

To compute industry R&D intensities, we use Compustat data. We drop all firms not headquartered in the U.S. and all observations with negative sales or assets. For each year, we winsorize the data at the 1st and 99th percentile, although our results do not depend on this. We then compute the R&D intensity as the ratio of the three-month moving average of R&D expenditures `xrd` relative to the three year moving average of operating income before depreciation `oibdp`, net sales `sales`, or total assets `at`. We tabulate the data pooling firm-calendar year observations and drop observations with multiple fiscal years in a given calendar year.

C.2 Fiscal data

We merge the datasets of Ramey (2011) and Nekarda and Ramey (2011). To this, we add information on the market value of publicly held federal debt from the Dallas Fed website[30] that we then deflate by the CPI from Ramey (2011). All variables enter the VAR in log-levels relative to population.

[30] See `https://www.dallasfed.org/research/econdata/govdebt`.

D Additional results

D.1 News shocks

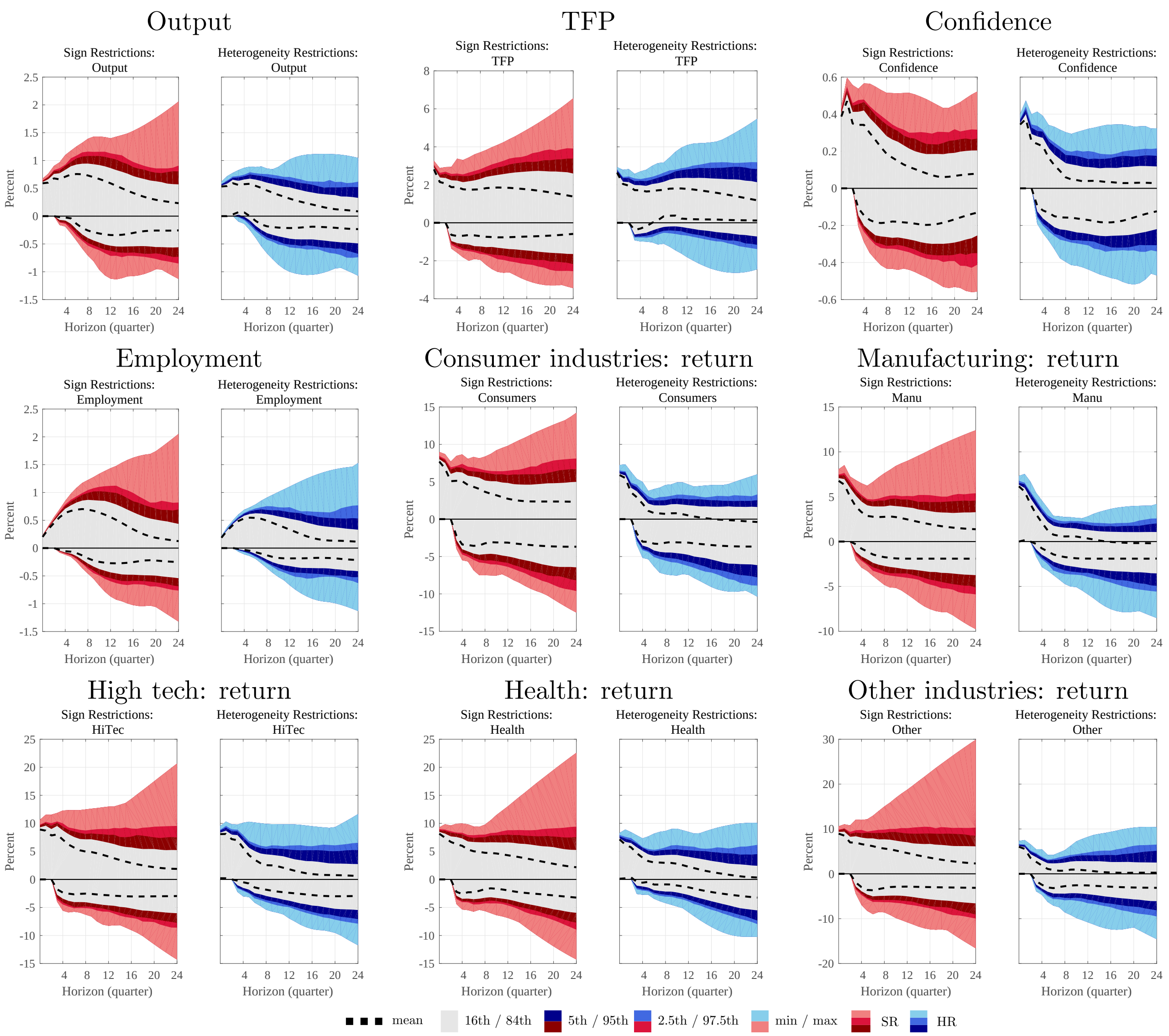


Figure D.1: Prior-robust responses of all variables to productivity news shock

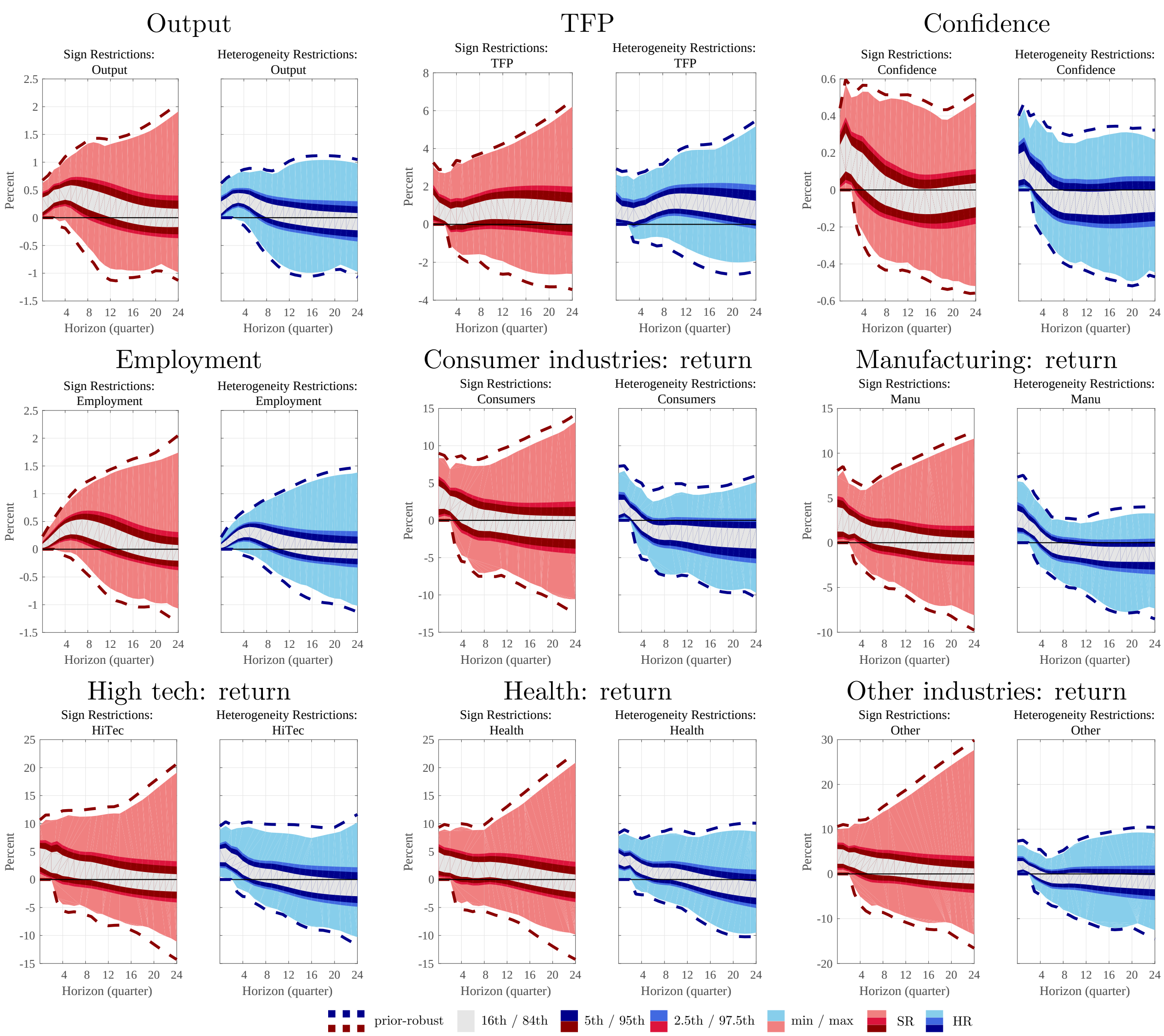


Figure D.2: Fully Bayesian responses of all variables to productivity news shock

(a) Prior-robust

Variable	Horizon H (quarters) 0	1	2	3	4	8	12	16	20	24
Output	9.7	10.7	15.1	15.4	20.4	34.9	38.9	36.0	34.0	35.2
TFP	8.4	6.7	8.8	10.6	11.2	5.9	5.3	10.4	10.5	16.1
Confidence	11.8	20.0	26.6	31.8	34.3	47.8	42.2	31.3	34.0	34.7
Employment	7.7	10.2	11.7	14.7	18.4	30.9	35.7	30.6	24.3	23.3
Consumers	23.5	20.6	29.1	38.7	50.0	59.3	55.0	58.0	62.0	61.9
Manu	11.1	13.6	16.5	21.1	31.2	56.4	60.1	60.5	61.0	55.9
HiTec	9.5	8.2	10.7	15.2	20.6	33.2	32.9	31.8	32.8	30.0
Health	12.2	16.3	16.4	22.4	29.7	30.0	34.7	37.0	40.1	40.3
Other	32.7	36.5	43.6	52.4	60.7	61.8	54.2	53.2	48.3	39.7

(b) Fully Bayesian

Variable	Horizon H (quarters) 0	1	2	3	4	8	12	16	20	24
Output	16.9	11.8	13.8	15.2	22.1	44.0	45.8	38.0	31.6	34.1
TFP	18.5	14.8	8.7	1.6	-0.1	-16.6	-11.7	-7.7	-6.3	-6.6
Confidence	19.5	28.9	36.2	42.7	48.3	71.3	55.4	30.2	33.3	45.8
Employment	7.9	8.7	8.8	12.8	17.3	33.6	38.5	28.1	7.1	-7.7
Consumers	37.1	31.3	42.5	55.3	73.7	101.7	98.8	105.0	109.8	111.2
Manu	12.3	15.5	19.7	25.4	40.9	82.9	92.3	95.3	94.7	90.6
HiTec	5.9	1.1	7.0	15.7	23.2	34.2	37.3	41.3	42.1	44.2
Health	8.5	10.8	5.0	14.7	24.5	30.6	38.4	49.3	56.2	60.1
Other	44.9	44.8	55.9	66.0	76.3	82.5	72.1	68.5	66.8	64.9

The contribution is expressed in percent of the 95th percentile of the IRF using sign restrictions only. A negative number implies a higher IRF with heterogeneity restrictions. Here, this happens in the Fully Bayesian case and indicates that the heterogeneity restriction shifts posterior mass up. By construction, this cannot happen with prior-robust bounds.

Table D.1: Reduction of 95th percentile of IRF relative to sign restrictions

Maximum FEV under pure sign restrictions (% of maximal FEV)

Variable	Horizon H (quarters) 0	1	2	3	4	8	12	16	20	24
Output	99.0	96.1	93.8	91.8	89.6	80.7	75.7	72.8	70.6	68.8
TFP	99.2	98.2	95.7	92.8	90.4	85.0	82.2	80.1	78.4	77.1
Confidence	100.0	98.1	97.2	95.8	94.3	88.6	83.0	77.8	73.7	70.5
Employment	92.3	91.6	90.2	89.7	89.9	89.1	85.5	82.4	79.8	77.6
Consumers	96.8	96.4	95.6	93.5	90.7	81.7	76.2	71.9	67.8	63.8
Manu	93.2	93.6	93.6	93.1	92.0	85.9	80.1	75.4	71.6	68.5
HiTec	96.0	95.8	94.6	92.8	91.5	87.3	83.0	79.2	75.9	73.0
Health	99.9	98.5	97.0	94.7	92.9	89.5	88.1	86.6	84.6	82.2
Other	95.8	96.0	95.8	94.2	92.4	86.6	85.0	84.3	83.3	82.0

Maximum FEV under heterogeneity restrictions (% of maximal FEV)

Variable	Horizon H (quarters) 0	1	2	3	4	8	12	16	20	24
Output	80.6	77.6	74.9	72.1	68.8	57.7	49.7	44.6	41.3	39.1
TFP	87.7	87.5	86.1	84.0	81.9	77.8	75.9	74.0	72.2	70.6
Confidence	79.0	71.2	67.3	64.2	61.4	53.1	47.9	44.6	42.9	41.9
Employment	79.6	79.0	78.6	76.2	73.4	64.3	56.5	51.4	48.2	46.0
Consumers	56.8	58.7	58.0	55.4	52.1	41.5	35.2	31.1	27.9	25.4
Manu	77.1	75.9	74.7	73.3	71.6	64.8	58.9	54.4	50.9	48.1
HiTec	79.2	80.4	79.5	77.6	75.9	68.8	62.5	57.5	53.7	50.7
Health	76.8	74.1	72.3	69.3	66.3	58.5	54.0	50.3	47.1	44.4
Other	43.6	42.6	40.3	37.4	34.5	26.4	22.0	19.3	17.5	16.2

The contribution is expressed in percent of the total forecast error variance up to horizon H. All contributions are computed at the posterior mean.

Table D.2: Maximum forecast error variance explained by productivity news: Results for sign restrictions

Binding heterogeneity restrictions (as a share of all cases) – all IRFs

Restrictions	Lower Bound SR ($\lambda = 0$)	Lower Bound HR ($\lambda = 1$)	Upper Bound SR ($\lambda = 0$)	Upper Bound HR ($\lambda = 1$)
Output> 0	0.06	0.05	0.19	0.22
TFP> 0	0.12	0.10	0.59	0.60
Confidence> 0	0.09	0.10	0.21	0.37
Employment> 0	0.23	0.19	0.32	0.25
Consumers> 0	0.20	0.30	0.02	0.12
Other> 0	0.24	0.42	0.02	0.10
Manu¿Consumers	0.17	0.36	0.06	0.46
Manu¿Other	0.17	0.34	0.06	0.60
Health¿Manu	0.17	0.29	0.19	0.52
HiTec¿Manu	0.12	0.23	0.13	0.32

For sign restrictions, we set the right-hand-side of the heterogeneity restrictions to zero, i.e., “Manu>Other” and “Manu>Consumers” both become “Manu>0”.

Table D.3: Maximum forecast error variance explained by productivity news: Results for sign restrictions and heterogeneity restrictions

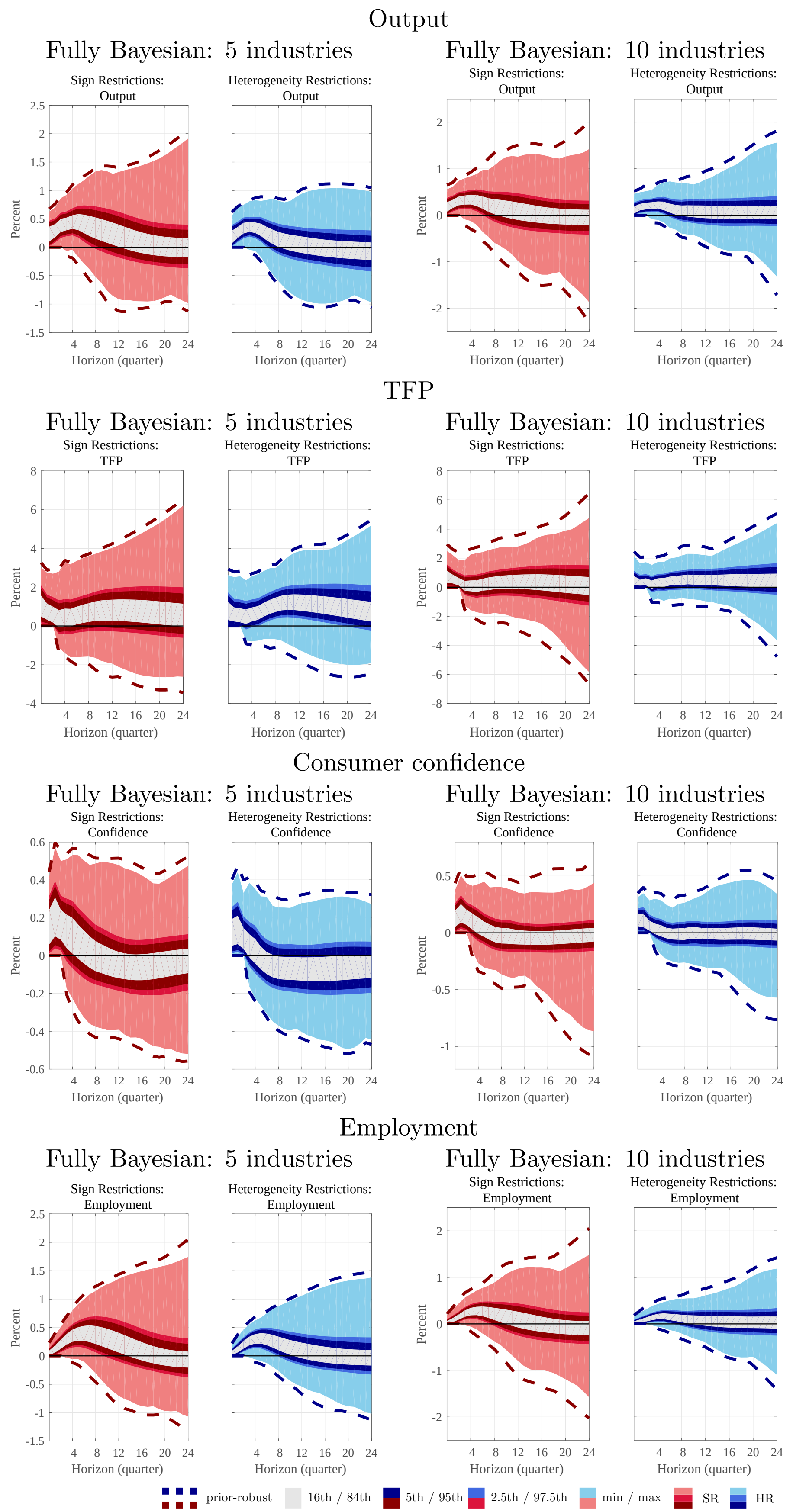


Figure D.3: Responses to productivity news shock for macro variables: Five vs. ten FF industries

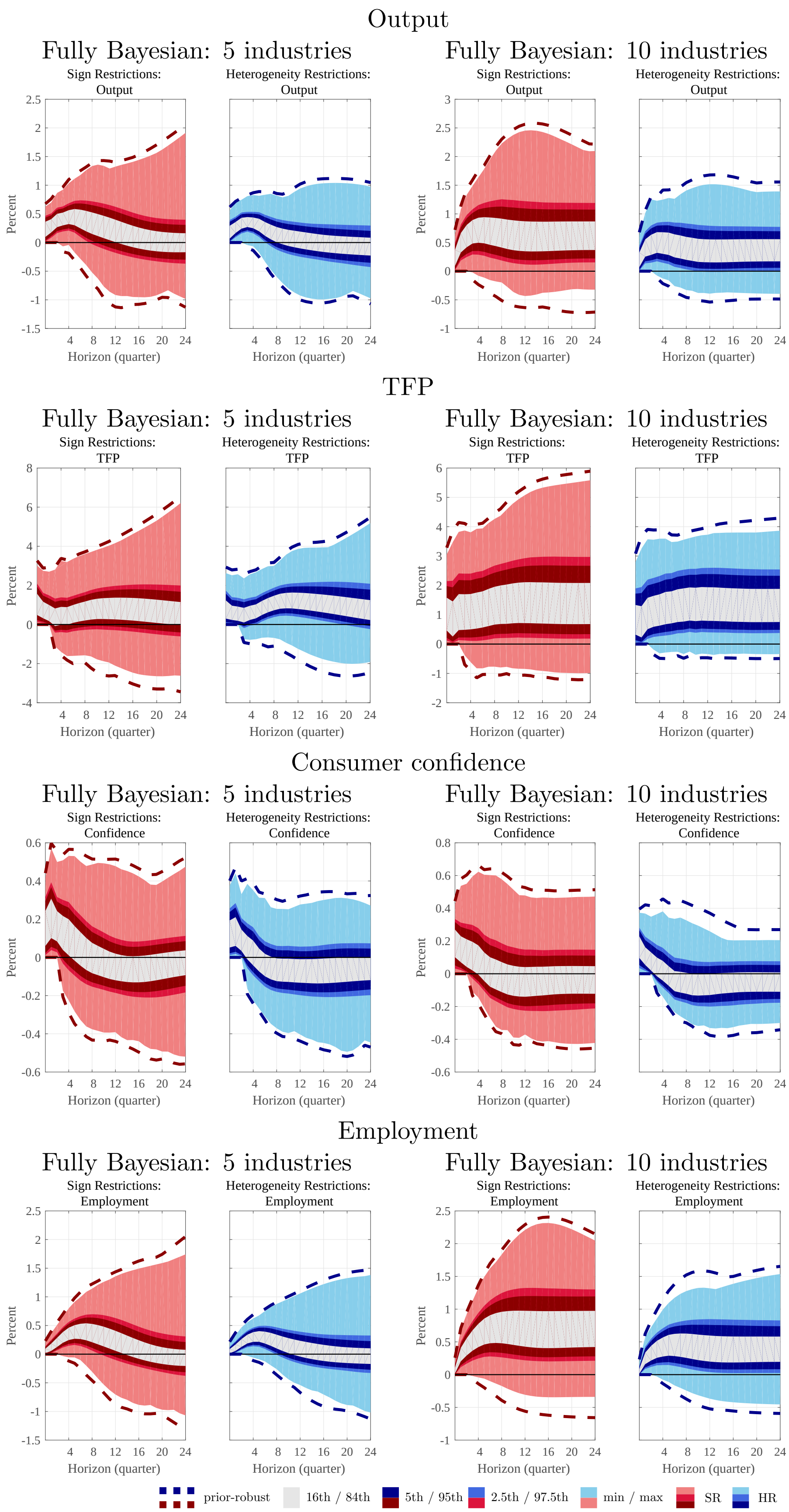


Figure D.4: Responses to productivity news shock for macro variables: Level vs. first difference specification

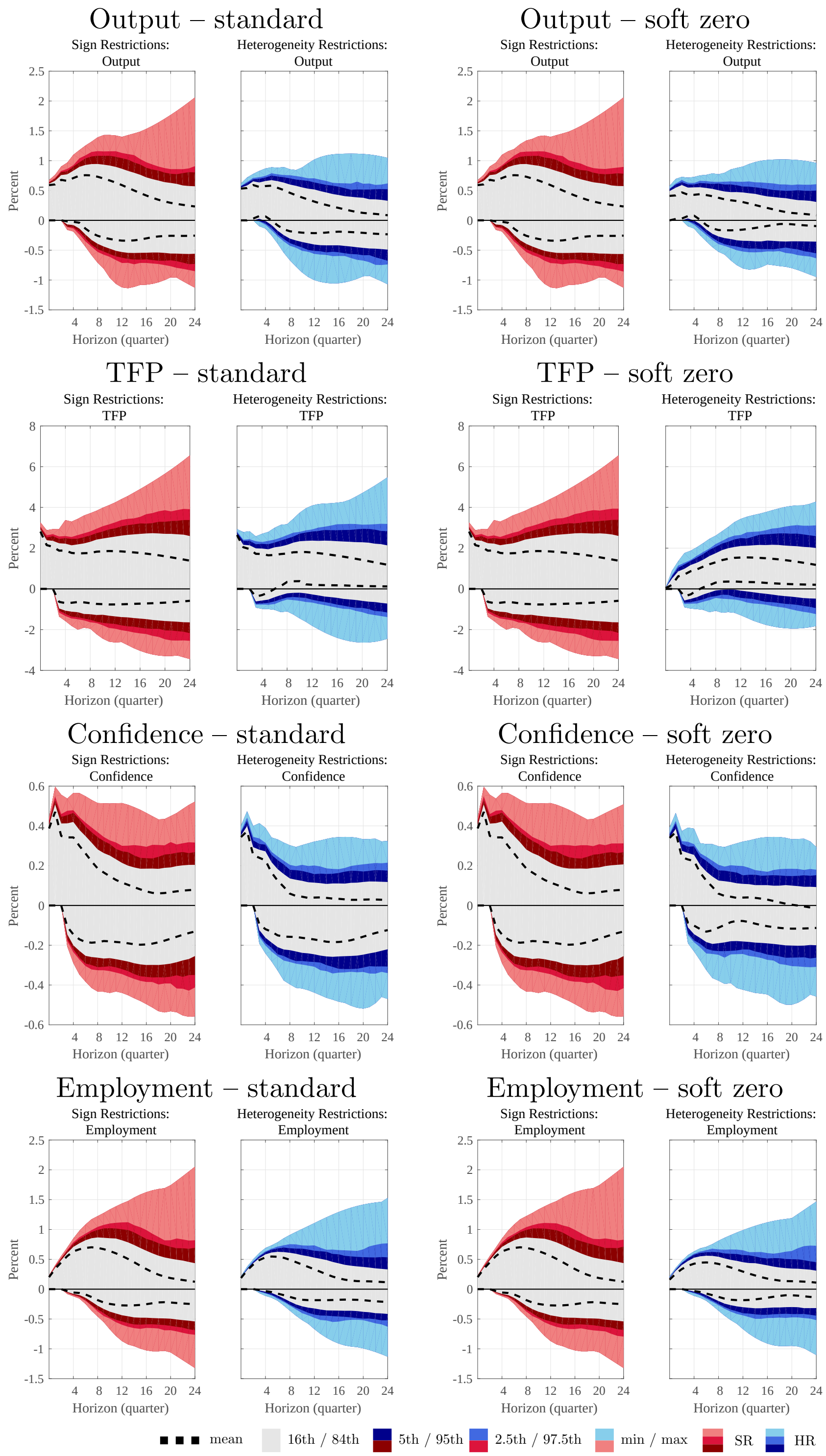


Figure D.5: Prior-robust responses of macro variables and industry returns to productivity news shock with soft zero restriction.

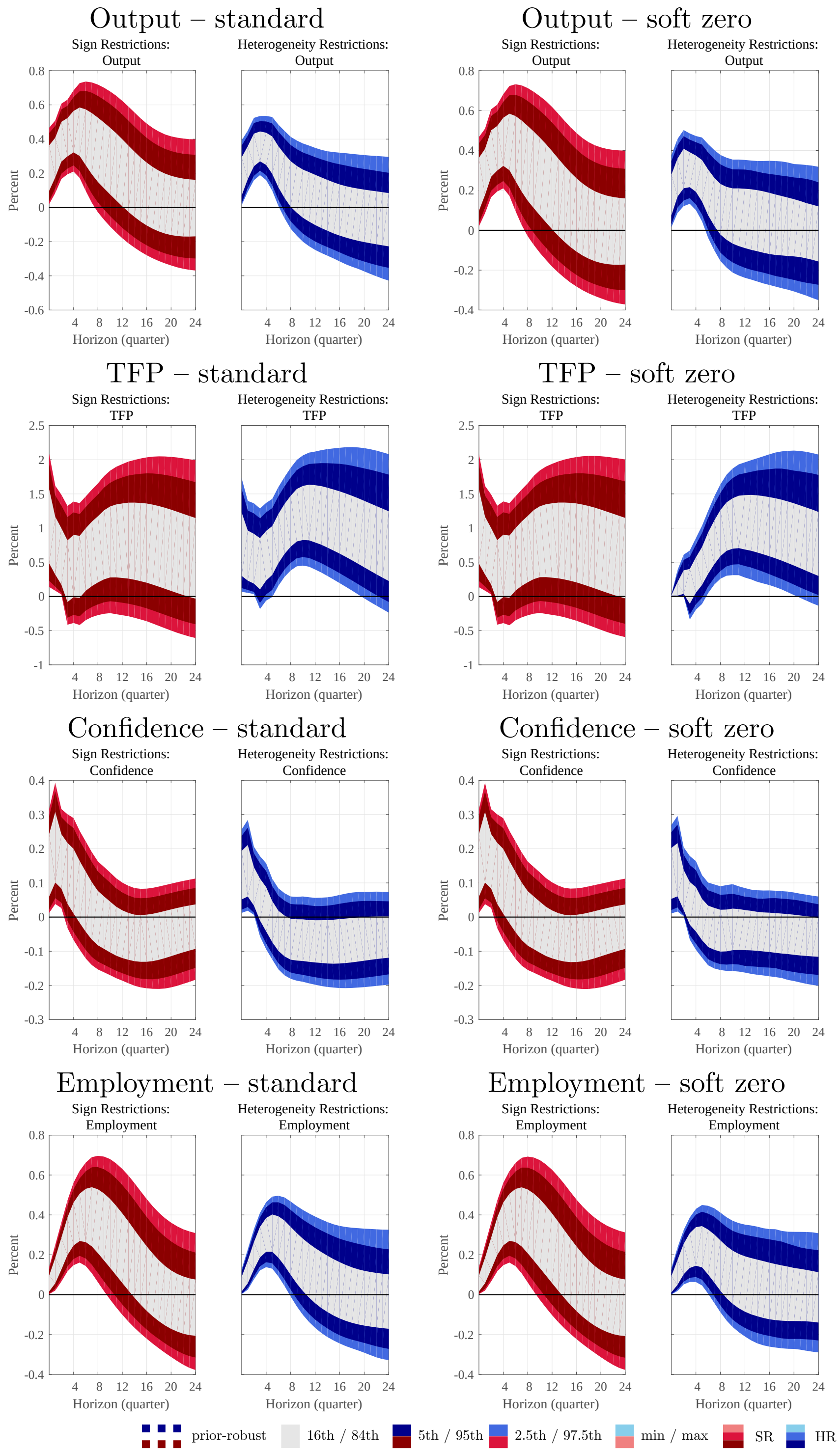


Figure D.6: Fully Bayesian responses of macro variables and industry returns to productivity news shock with soft zero restriction.

Maximum FEV under pure sign restrictions (% of maximal FEV)

Variable	Horizon H (quarters) 0	1	2	3	4	8	12	16	20	24
Output	99.0	96.1	93.8	91.8	89.6	80.7	75.7	72.8	70.6	68.8
TFP	99.2	98.2	95.7	92.8	90.4	85.0	82.2	80.1	78.4	77.1
Confidence	100.0	98.1	97.2	95.8	94.3	88.6	83.0	77.8	73.7	70.5
Employment	92.3	91.6	90.2	89.7	89.9	89.1	85.5	82.4	79.8	77.6
Consumers	96.8	96.4	95.6	93.5	90.7	81.7	76.2	71.9	67.8	63.8
Manu	93.2	93.6	93.6	93.1	92.0	85.9	80.1	75.4	71.6	68.5
HiTec	96.0	95.8	94.6	92.8	91.5	87.3	83.0	79.2	75.9	73.0
Health	99.9	98.5	97.0	94.7	92.9	89.5	88.1	86.6	84.6	82.2
Other	95.8	96.0	95.8	94.2	92.4	86.6	85.0	84.3	83.3	82.0

Reduction in Maximum FEV due to heterogeneity restrictions (% of maximal FEV) plus soft zero restriction on initial TFP

Variable	Horizon H (quarters) 0	1	2	3	4	8	12	16	20	24
Output	51.2	50.4	50.4	50.2	50.0	48.1	47.3	46.5	45.8	45.0
TFP	99.2	98.1	94.6	90.7	87.1	76.0	66.1	57.4	50.4	44.9
Confidence	22.9	28.8	31.9	33.8	35.1	37.3	36.7	35.0	33.1	31.4
Employment	34.3	36.5	37.2	39.0	41.4	46.3	46.3	45.5	44.5	43.6
Consumers	61.4	59.2	58.8	58.4	57.7	55.1	53.6	51.9	49.7	47.2
Manu	44.0	44.6	45.1	45.2	45.0	43.0	41.0	39.3	37.8	36.5
HiTec	36.0	32.1	30.1	28.8	28.2	28.0	28.3	28.4	28.2	27.7
Health	24.1	25.0	25.1	25.7	26.8	31.1	34.3	36.5	37.8	38.2
Other	68.4	68.9	70.0	70.3	70.2	69.5	70.6	71.6	71.7	71.2

The contribution is expressed in percent of the total forecast error variance up to horizon H. All contributions are computed at the posterior mean.

Table D.4: Reduction in maximum forecast error variance explained by productivity news: Heterogeneity restrictions without and with soft zero restriction

D.2 Fiscal shocks

Prior-robust

Variable	Horizon H (year) 0	1	2	3	4	8	10
Defense spending	11.1	8.8	9.7	11.8	11.2	13.3	26.1
Output	23.9	13.3	10.6	20.2	24.4	26.7	16.9
Debt	13.2	21.2	19.8	11.1	12.5	18.4	28.1
Hours	8.2	12.5	15.4	20.9	34.4	23.0	21.0
Tax rate	19.7	33.5	36.1	40.9	44.4	48.6	45.2
Tobacco	63.9	51.0	46.8	44.2	52.3	24.5	37.9
Lumber	47.0	29.3	26.4	31.0	35.3	16.7	19.7
Petrol	36.7	36.9	37.6	41.5	40.5	29.8	31.7
Equipment	48.4	40.3	23.7	30.7	40.7	45.0	37.5
Electronics	37.1	36.4	29.3	32.7	35.4	49.2	33.6
Transportation	22.6	16.0	17.3	27.6	36.3	16.7	24.6

Fully Bayesian

Variable	Horizon H (year) 0	1	2	3	4	8	10
Defense spending	-15.8	-32.6	-30.6	-21.2	-16.3	-8.8	-0.4
Output	19.1	-0.3	-11.9	-9.6	-6.8	-3.9	-8.3
Debt	-23.6	-39.7	-84.1	-93.4	-35.6	3.8	6.4
Hours	-2.1	-1.6	-9.5	-6.7	3.8	7.8	4.5
Tax rate	13.7	32.3	38.4	41.0	42.5	49.4	33.0
Tobacco	65.5	60.2	44.2	39.1	59.5	20.8	7.3
Lumber	61.6	15.6	2.1	22.2	28.4	-3.0	-2.4
Petrol	33.2	34.3	34.0	36.8	37.0	17.1	12.4
Equipment	51.9	37.9	21.7	26.2	50.0	60.0	17.7
Electronics	35.5	30.6	21.7	27.8	46.4	94.5	36.9
Transportation	13.6	-9.9	-8.4	-5.4	-13.6	-12.3	-9.0

The contribution is expressed in percent of the 95th percentile of the IRF using sign restrictions only. Negative entries imply a larger response under heterogeneity restrictions.

Table D.5: Reduction of 95th percentile of IRF to defense spending shocks relative to sign restrictions

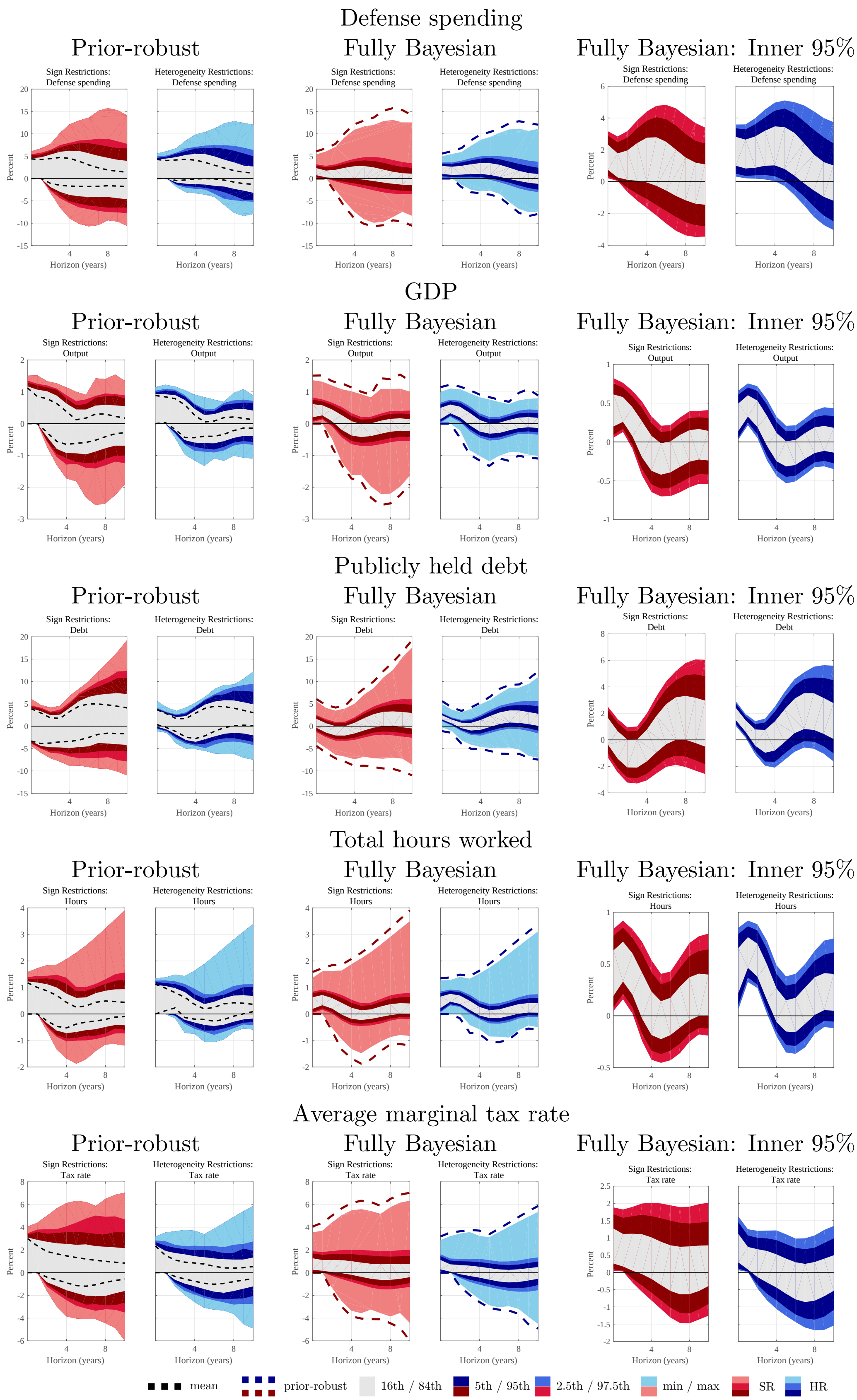


Figure D.7: Responses to defense spending shock: Macro variables

Maximum FEV under pure sign restrictions (% of maximal FEV)

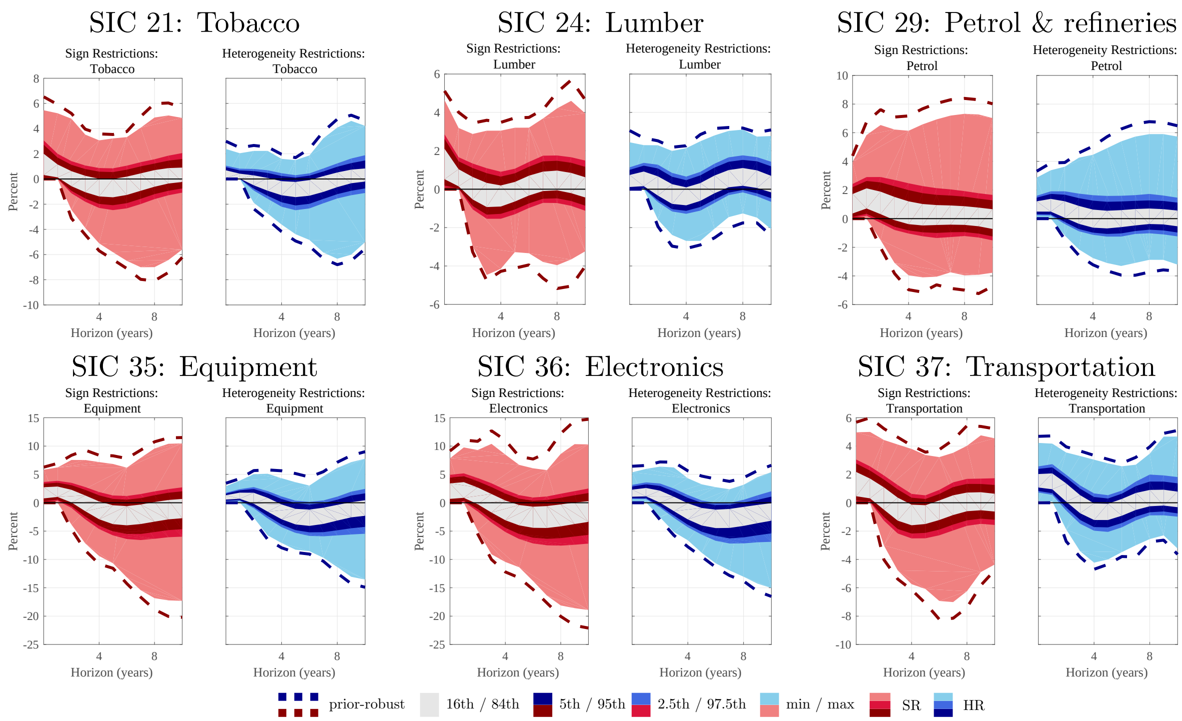
Variable	Horizon H (year) 0	1	2	3	4	8	10
Defense spending	86.9	72.5	63.9	59.6	57.4	54.5	53.3
Output	92.1	72.4	60.4	52.8	48.6	39.9	37.6
Debt	72.7	58.2	49.9	43.3	36.8	27.4	31.5
Hours	88.7	81.7	73.9	67.3	62.7	54.1	52.5
Tax rate	78.0	73.5	71.5	69.9	68.1	60.2	57.7
Tobacco	59.0	56.4	54.0	51.4	48.9	43.5	42.1
Lumber	97.1	81.8	71.7	65.7	62.2	51.3	48.1
Petrol	89.5	84.9	81.5	78.7	76.5	70.2	66.9
Equipment	86.2	83.1	73.5	62.0	52.1	34.8	37.3
Electronics	88.5	83.8	77.9	70.3	62.1	41.5	37.1
Transportation	77.5	58.5	48.7	42.8	39.7	33.9	32.5

Maximum FEV under heterogeneity restrictions (% of maximal FEV)

Variable	Horizon H (year) 0	1	2	3	4	8	10
Defense spending	81.5	68.1	59.7	55.4	53.2	50.1	49.0
Output	57.3	50.1	46.5	42.4	39.4	33.1	31.3
Debt	64.7	51.9	44.9	38.9	32.9	26.3	29.0
Hours	80.4	75.4	69.3	63.8	59.6	51.4	49.6
Tax rate	48.2	42.2	39.1	36.7	34.5	28.0	26.5
Tobacco	9.5	8.7	8.4	8.3	8.4	9.6	10.0
Lumber	32.1	29.8	27.3	25.6	24.6	21.2	20.4
Petrol	43.0	41.5	39.8	38.6	38.0	36.2	34.9
Equipment	24.7	24.1	23.1	23.9	21.9	16.1	15.1
Electronics	29.6	26.4	24.1	21.5	19.1	20.7	23.2
Transportation	48.0	41.7	38.1	34.9	32.9	29.2	28.2

The contribution is expressed in percent of the total forecast error variance up to horizon H. All contributions are computed at the posterior mean.

Table D.6: Maximum forecast error variance explained by defense spending shocks: Results for sign restrictions and heterogeneity restrictions



Algorithm 2 plus bounds from Algorithm 1.

Figure D.8: Responses of industry shipments to defense spending shock.

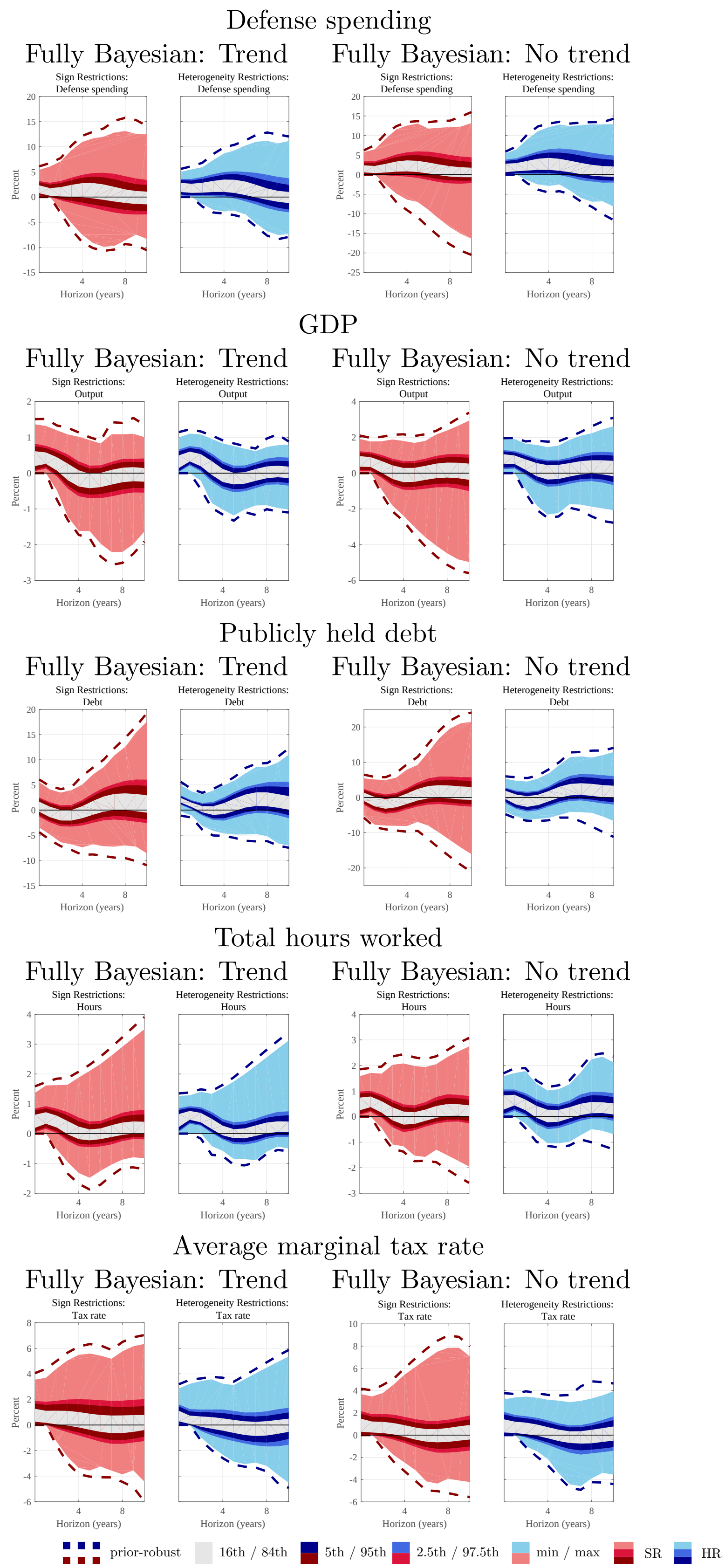


Figure D.9: Responses to defense spending shock: Macro variables robustness