WORKING PAPERS

RESEARCH DEPARTMENT

WORKING PAPER NO. 17-11
IDENTIFICATION THROUGH HETEROGENEITY

Pooyan Amir-Ahmadi
University of Illinois at Urbana–Champaign

Thorsten Drautzburg
Research Department
Federal Reserve Bank of Philadelphia

April 26, 2017

RESEARCH DEPARTMENT, FEDERAL RESERVE BANK OF PHILADELPHIA

Ten Independence Mall, Philadelphia, PA 19106-1574 • www.philadelphiafed.org/research-and-data/

Identification through Heterogeneity

Pooyan Amir-Ahmadi and Thorsten Drautzburg*

April 26, 2017

Abstract

We analyze set identification in Bayesian vector autoregressions (VARs). Because set identification can be challenging, we propose to include micro data on heterogeneous entities to sharpen inference. First, we provide conditions when imposing a simple ranking of impulse-responses sharpens inference in bivariate and trivariate VARs. Importantly, we show that this set reduction also applies to variables not subject to ranking restrictions. Second, we develop two types of inference to address recent criticism: (1) an efficient fully Bayesian algorithm based on an agnostic prior that directly samples from the admissible set and (2) a prior-robust Bayesian algorithm to sample the posterior bounds of the identified set. Third, we apply our methodology to U.S. 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.

Keywords: Structural VAR; set-identification; heterogeneity and sign restrictions; posterior bounds; Bayesian inference; sampling methods; productivity news; government spending.

1 Introduction

Since the seminal paper by Sims (1980), the structural vector autoregressive (SVAR) model has been the workhorse for analyzing the dynamics caused by macroeconomic shocks. The first generation of the literature focused on zero short-run, medium-run, or long-run restrictions on impulse responses for identification (e.g., Sims, 1980; Uhlig, 2004; Blanchard and Quah, 1989). More recent contributions have attempted to work with weaker assumptions. One popular strand of the literature, going back to Uhlig (2005), Faust (1998), and Canova and De Nicolo (2002), abandons point identification. Most prominently, Uhlig (2005) introduces agnostic beliefs with qualitative sign restrictions on impulse response functions to identify a set of impulse response functions. Here, we build on this agnostic approach and

*Amir-Ahmadi: University of Illinois at Urbana-Champaign, pooyan[at]illinois.edu. Drautzburg: Federal Reserve Bank of Philadelphia, tdrautzburg[at]gmail.com. We would like to thank Jonás Arias, Mark Bognanni, Jesús Fernández-Villaverde, Michele Piffer, Frank Schorfheide, Mark Watson, Jonathan Wright, and audiences at the Boston College, 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. This paper is available free of charge at `www.philadelphiafed.org/research-and-data/publications/working-papers/`.

show how to use easily available micro data in the analysis to sharpen inference for macro variables.

We make three contributions to the class of set-identified dynamic time series models: First, we make a conceptual contribution by showing how to formally use micro data for the sharper identification of macroeconomic shocks by ranking impulse responses. Second, we derive conditions under which restrictions on the relative ranking of responses sharpen inference relative to standard sign restrictions. Importantly, this includes inference about responses of variables not subject to extra restrictions. Third, we contribute theoretically well founded algorithms for inference. Specifically, we propose two inference methods: (1) We design an algorithm to directly sample from the posterior over the admissible set, satisfying all identifying restrictions under agnostic beliefs. The algorithm uses the insight that we can normalize draws from the appropriately truncated multivariate distribution to generate draws from the truncated Haar measure that satisfy all restrictions. (2) We also propose an algorithm for prior robust inference under ambiguous beliefs that generates a distribution over the bounds of the identified set. One byproduct of this algorithm is a novel way to characterize the importance of individual constraints. Additionally, both algorithms draw on the insight that we can use a simple linear program to determine whether the identified set has positive measure.

How do we rank responses? Consider heterogeneity in an a priori shock elasticities of different micro aggregates such as industries, such as in the context of a defense spending shock. Manufacturing industry A might, a priori, be more exposed to these shocks relative to sector B if the military is a key client of the former industry but not of the latter industry. We would then restrict industry A to respond more than industry B to a defense spending shock. We label such restrictions *heterogeneity restrictions*. While earlier papers such as De Graeve and Karas (2014) have used heterogeneity restrictions, we show systematically how such restrictions affect inference. First, we derive analytically conditions under which our approach sharpens inference. Second, we provide two quantitative applications, to productivity news shocks and to defense spending shocks, to show that such restrictions using micro data also lead to set reduction in practice. In both applications, we also show how to impose additional heterogeneity restrictions on macro data that reflect beliefs on macro elasticities.

Our analysis builds on the recent literature on set-identified SVARs. Moon and Schorfheide (2012) 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. As a consequence, Moon, Schorfheide, and Granziera (2013) and others (Giacomini and Kitagawa, 2014; Gafarov, Meier, and Montiel Olea, 2016a,b)[1] have developed methods for inference directly over the identified set. Our prior robust algorithm samples from the distribution over the bounds of the identified set. To understand the determinants of the identified set and building on Moon, Schorfheide, and Granziera (2013), we derive conditions when heterogeneity restrictions lead to a strict set reduction relative to standard sign restrictions in bivariate and trivariate cases. While the logic of the bivariate case carries over to the trivariate case, the latter matters because we need three dimensions to analyze variables not directly involved in a ranking of restrictions. Intuitively, sign restrictions shrink the identified set if the conditional covariance required by sign restrictions on impulse responses has the opposite sign of the unconditional covariance: The identified shock then cannot account for all of the variation in the data. Heterogeneity restrictions impose sign restrictions on linear combinations of variables. This leads to a set reduction relative to sign restrictions if the linear combination of covariances following a shock is more at odds with the unconditional covariances.

Our algorithm for agnostic Bayesian inference is designed to handle even tight identified sets of strictly positive measure. Its novelty is that it samples directly from the set of admissible rotation vectors. These rotation vectors map reduced-form parameters into structural parameters. We show that sampling from the truncated multivariate standard normal distribution and normalizing these draws provides draws from the Haar measure over rotation vectors, truncated to the admissible set. To implement this insight, we build on the Gibbs sampler for the truncated multivariate normal distribution from Li and Ghosh (2015). While the rejection samplers in Uhlig (2005) and Rubio-Ramírez, Waggoner, and Zha (2010) can work well in applications with relatively few restrictions, we show that in our medium-sized VAR applications, the corresponding acceptance rates are impractically low.

Instead of working with an agnostic prior, some researchers prefer a procedure that is robust to the choice of prior over rotation vectors. We also consider this case. Our algorithm for prior robust Bayesian inference is related to the ones in Giacomini and Kitagawa (2014) when applied to impulse response functions and to Faust (1998) when applied to variance decompositions. Like the algorithms in Faust (1998) and Giacomini and Kitagawa (2014), our algorithm provides a distribution of the bounds over the identified set. Giacomini and Kitagawa (2014) provide a detailed numerical recipe and elegantly derive conditions when a prior robust Bayesian posterior is asymptotically equivalent to frequentist inference. The key differences of our algorithm compared with Giacomini and Kitagawa (2014) are: (1) We avoid rejection sampling to determine whether the identified is empty. Instead, we show that

[1] See also Faust (1998) and Faust, Swanson, and Wright (2004) for two early different and early applications to variance decompositions and high-frequency identification, respectively.

solving a simple linear programming problem allows us to identify empty sets. (2) We always use optimization to find the boundaries of the identified set rather than rejection sampling. Given our finding that rejection samplers perform poorly in our quantitative applications, these differences matter. Instead, for each set of reduced-form parameters, we optimize the bounds on variance decompositions or impulse responses also in our quantitative applications. (3) As a byproduct of our algorithm, we compute Lagrange multipliers for each restriction. These multipliers are a natural measure of the importance of restrictions for set reduction.

Heterogeneity restrictions can also address recent criticism that seemingly uninformative priors used in applications can have implications that researchers do not desire. Baumeister and Hamilton (2015) show that the uniform prior over rotation matrices and dominates the posterior over the structural parameters. They propose to elicit priors directly on the structural representation of the SVAR. Instead, we show how, besides micro heterogeneity restrictions, we can impose beliefs on macro elasticities while working with an agnostic prior or being prior robust. Technically, this is another heterogeneity restriction. In one application, we use this device to discipline the posterior distribution over fiscal multipliers. Conceptually, this is similar to the elasticity bounds Kilian and Murphy (2012) introduce in their analysis of oil markets. Arias, Rubio-Ramirez, and Waggoner (2014) analyze agnostic Bayesian inference with sign and zero restrictions. They provide a sophisticated framework for such inference and provide examples of how existing papers imposed beliefs that were only implicit in their analysis. Our algorithm for agnostic inference is efficient also in large VARs and does not require special treatment of zero restrictions. Instead, we approximate zero restrictions as "soft" zero restrictions that are just another heterogeneity restriction. This approach is less dogmatic than precise zero restrictions and works well in our news shock application.

We show in two quantitative applications that heterogeneity restrictions on micro variables sharpen inference about macro variables substantially in practice. In our first application we identify the dynamic effects of productivity news shocks. These shocks raise the value of stocks, are expansionary, and lead to higher TFP in the future. Beaudry and Portier (2006) and most recently Kurmann and Sims (2017) have argued that productivity news shocks are an important determinant of output, accounting for 60-80% of the output forecast error — despite disagreement about the relevant horizon. We revisit this topic under the heterogeneity restriction that productivity news move the stock returns of R&D-intensive industries more. We measure the R&D intensity using Compustat data for five Fama and French (1997) industries. The restriction that more innovative sectors respond more sharpens our inference substantially: Heterogeneity restrictions lower the maximum forecast error variance (FEV) in output from 70-90% of the total variance to about 30-60%

at the posterior median, depending on the horizon. The peak prior robust impulse response of output to productivity news also falls, by about 30% or 0.6pp at the 99th percentile. Heterogeneity restrictions are more important than standard sign restrictions for sharpening the upper bound on impulse responses. We also use a macro heterogeneity restriction to show that our results are robust to a (soft) zero restriction on initial TFP, similar to Barsky and Sims (2011).

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 under the fully Bayesian posterior. We also find that heterogeneity restrictions halve the width of the 90% fully Bayesian credible set, so that the 10-year present value multiplier (Mountford and Uhlig, 2009) lies between 0.5 and 3 with 68% posterior probability. This compares to a range of -1 to 4 with standard sign restrictions.

We see at least two avenues for future research based on our work. First, our contributions naturally carry over to factor-augmented VARs and dynamic factor models in general as in Amir-Ahmadi and Uhlig (2015), to panel VARs as in De Graeve and Karas (2014), or to time-varying parameter VARs with stochastic volatility 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 from the admissible set becomes even more important than in our VAR application. Second, we could potentially combine heterogeneity restrictions with sign restrictions on shocks, developed by Antolin-Diaz and Rubio-Ramírez (2016) and Ludvigson, Ma, and Ng (2017), to combine our approach with approaches that use proxy variables for inference.

This paper is structured as follows. First, we set up the general statistical model and identification problem with sign and heterogeneity restrictions. In this general framework, we characterize the identified sets in bivariate and trivariate models. Second, we show how to detect whether identified sets are non-empty for given restrictions, derive our direct sampler from the admissible set of rotation vectors, and develop two algorithms for inference. Third, we provide two quantitative applications to productivity news shocks and defense spending shocks in the US. An extensive appendix contains all proofs and additional empirical results.

2 Model

Here we set up the standard Bayesian VAR framework and define the sign and heterogeneity restrictions that we analyze. We discuss that heterogeneity restrictions can narrow the identified set and provide sufficient conditions for identified sets to have positive measure. We illustrate the concept of heterogeneity restrictions through examples. Last, we provide conditions when heterogeneity restrictions, compared with pure sign restrictions, lead to strict set 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-to-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 as follows:

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

In addition to this generic VAR restriction, we 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 (Impulse vector)**.** *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.*

Formally, let $\tilde{A}$ be the lower Cholesky matrix and take any rotation matrix $Q = [q_1, \dots, q_p]$. Then the columns of $\tilde{A}Q$ are impulse vectors. We can thus express impulse vectors without loss of generality as:

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

We use $||\cdot||$ to denote the Euclidean norm throughout. Generally, these impulse vectors do not have an economic interpretation, unless they satisfy economic restrictions that we introduce below. In general, we can then write our full model as:

$$p(Y^T, B, \Sigma, Q) = \ell(B, \Sigma|Y^T)\pi_0(B, \Sigma)\pi_Q(Q|B, \Sigma), \tag{2.4}$$

where Y^T collects the history of observables, ℓ is the likelihood function, π_0 denotes the prior over the identifiable reduced-form parameters, and π_Q denotes the prior over Q that incorporates restrictions on impulse responses. We later assume a standard conjugate prior over (B, Σ) and take these parameters for now as given. We discuss estimation in section 3. Now we focus on what we can learn about Q from the reduced-form parameters and beliefs about impulse responses. We assume that π_Q has full support over the identified set.

2.2 Sign and heterogeneity restrictions

To identify structural impulse vectors, we impose qualitative restrictions on the impulse-responses Q induces. To define these restrictions, we need extra notation. Given the companion form $Y_t = B_Y Y_{t-1} + A\epsilon_t$ of the VAR (2.1) the impulse-response at horizon h is:

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

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

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

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

Sign or heterogeneity restrictions shape the admissible set of q and the identified set of structural parameters that we now define:

[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 use a common H here.

Definition 4 (Admissible set)**.** *The admissible set* $AS(\mathbb{L}, H)$ *is the collection of all* q*, with* L_2-*norm* 1 *for which* $a = \tilde{A}q$ *satisfies the restrictions in* $(\mathbb{L}, H)$*.*

Definition 5 (Identified set)**.** *The identified set* $IS(f \mid \mathbb{L}, H)$ *is the set of all* $f(q)$ *with* q *in the admissible set, where* $f(q)$ *is some objective function.*

$f(q)$ can be any nonlinear function of parameters such as impulse response functions or FEV decompositions.

With a specific prior over Q and for given percentiles of the posterior distribution, the heterogeneity restrictions can produce more dispersed posterior percentiles: The tighter restrictions can shift mass away from the center of the prior toward the tails of the distribution, as we show 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 (Weakly smaller identified set with heterogeneity restrictions)**.** *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* $j = 1, \ldots, J$ $n^{(j)}_{SR} = n^{(j)}_{HR}$ *and* $\lambda^{(j)} \geq 0$*, then the identified 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 $\mathbb{L}^{(j)}_{HR}$ imply the restrictions in $\mathbb{L}^{(j)}_{SR}$ given that $n^{(j)}_{SR} = n^{(j)}_{HR}$. □

Below we provide conditions under which the identified sets are also strictly smaller than with pure sign restrictions in two- and three-dimensional VARs.

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.

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

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 (Non-empty identified set under rank condition)**.** *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, $j = 1, \dots, J$. Let E be a $J \times p$ matrix of zeros, except for ones in the $(j, n^{(j)})$ positions, $j = 1, \dots, J$. If $M \equiv E - \Lambda$ is of rank J, then the identified set for a induced by $\mathbb{L}_{HR}$ has positive measure given that π_q has full support.*

Proof. See (A.1) □

This Lemma is also useful to guide the design of heterogeneity restrictions: 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 this 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 nondegenerate 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 (Non-empty identified set if heterogeneity restrictions do not overlap)**.** *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.:*

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

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

Proof. See (A.2) □

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. We discuss in Section 3 how to detect empty sets.

2.3 Equivalence to change of variables

There is an equivalence between heterogeneity restrictions and sign restrictions with an appropriate change of variables in simple settings. 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 Different variations of heterogeneity restrictions

While we focus on heterogeneity restrictions in the form of qualitative short-run restrictions on responses in most of our applications, our approach also applies to several variations:

Strength of the restrictions. How do we choose the strength of the heterogeneity restrictions λ? If we have a notion that we want to rank the responses of different sectors qualitatively, the case of $\lambda = 1$ is the most natural. However, λ can also express confidence in the measured heterogeneity. For example, setting $\lambda \in (0, 1)$ expresses a weaker ranking.

Qualitative beliefs about macroeconomic relationships can generate bounds that translate to heterogeneity restrictions of varying strengths. For example, to restrict impact multipliers for government spending to be smaller than two, write:

$$r^0_{a,\text{output}} \leq \lambda \times r^0_{a,\text{government spending}} \quad \text{with} \quad \lambda = 2 \times \frac{\bar{G}}{\bar{Y}}.$$

Thus, our framework allows us to use prior information on some elasticities, as advocated by Baumeister and Hamilton (2015), without specifying full priors for all elasticities. In prior work, Kilian and Murphy (2012) introduce elasticity bounds in the context of oil markets.

Soft zero restrictions. We can also use varying ranking intensities λ to impose approximate zero restrictions (i.e., soft zero restrictions). For example, Christiano, Eichenbaum, and Evans (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:

$$-\lambda \times r^0_{a,\text{FFR}} \leq r^0_{a,\text{GDP}} \leq \lambda \times r^0_{a,\text{FFR}} \quad \text{with} \quad \lambda = 0.01.$$

Long run heterogeneity restrictions. Long run identification 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:

$$r^\infty_{a,\text{Productivity in A}} \geq \lambda r^\infty_{a,\text{Productivity in B}} \quad \text{with} \quad \lambda = 1.$$

Long run zero restrictions. Zero restrictions can also be useful in the long run. For example, consider long run monetary neutrality. To enforce approximate monetary neutrality after a monetary policy shock, impose:

$$-\lambda \times r^0_{a,\text{FFR}} \leq r^\infty_{a,\text{GDP}} \leq \lambda \times r^0_{a,\text{FFR}} \quad \text{with} \quad \lambda = 0.01.$$

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

2.5 Characterizing the identified set analytically

When do heterogeneity restrictions sharpen inference relative to sign restrictions? Here we first follow Moon, Schorfheide, and Granziera (2013) to characterize the identified set analytically in a bivariate VAR with impact restrictions only. We show that for the common restrictions associated with $\lambda^{(j)}_{HR} = 0$, the identified set for $a_{n^{(j)}}$ can be strictly or weakly

[4]Existing Bayesian analyses of sign and zero restrictions often inadvertently impose non-stated beliefs in the identification. See Arias, Rubio-Ramirez, and Waggoner (2014), who analyze sign and hard zero restrictions in a fully Bayesian fashion.

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 shrinks the set more if the reduced-form correlation of forecast errors is more negative. Trivially, heterogeneity restrictions have the more bite, the stronger the known degree of 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 variables subject only to sign restrictions.

2.5.1 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]

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

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, express 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. The 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, an extension

[5]As written, we impose one sign and one heterogeneity restriction. An example is identifying a cost shock in a competitive industry with decreasing returns for which we observe prices and quantities. The restriction that demand is elastic translate to the heterogeneity restriction that minus the quantities fall more strongly than the prices within that industry.

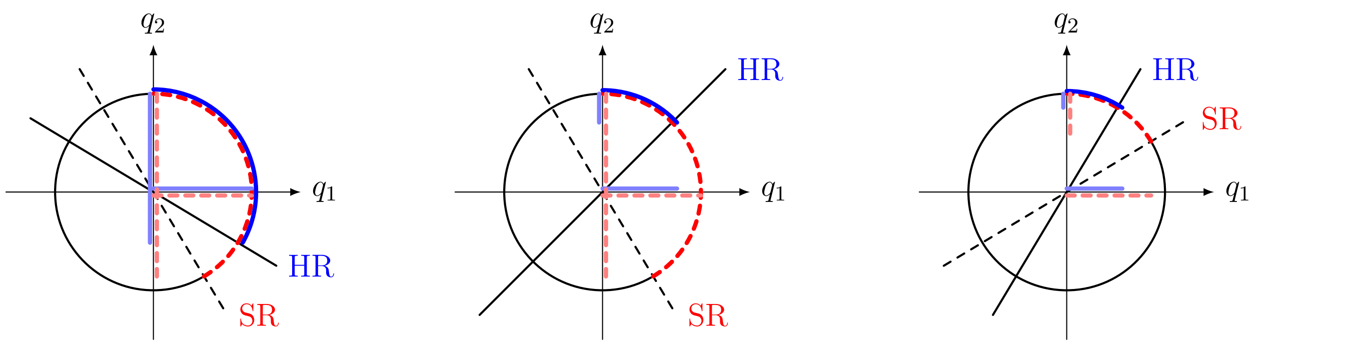
[6]The elements of the Cholesky decomposition $\Sigma = \tilde{A}\tilde{A}'$ are: $\tilde{A}_{1,1} = \sqrt{\Sigma_{11}}$, $\tilde{A}_{2,1} = \frac{\Sigma_{21}}{\tilde{A}_{1,1}} = \tilde{A}_{2,2}\frac{\rho}{\sqrt{1-\rho^2}}$, $\tilde{A}_{2,2} = \sqrt{\Sigma_{22} - (\tilde{A}_{2,1})^2} = |\tilde{A}_{2,1}|\sqrt{1/\rho^2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

of Moon, Schorfheide, and Granziera (2013): First, q_1 is positive. Second, q_2 lies above the straight line through the origin, which 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, or (c) the correlation is negative so that both contributions are positive. In cases (b) and (c), the admissible set for q_1 is strictly smaller with heterogeneity restrictions. In case (a), the admissible set for q_1 is $[0, 1]$ in both cases, but the admissible 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



Note: The admissible set is given by the intersection of the unit circle with the $q_1 \geq 0$ plane and the plane above the heterogeneity restrictions (HR) and sign restrictions (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 HR and in dashed and blue lines for the traditional 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 admissible 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.3 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 (Set reduction of $IS(\cdot)$ under heterogeneity restrictions in bivariate VAR)**.** *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. See (A.3) $\square$

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 with 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 or else the VAR could not generate the observed reduced-form correlation. This intuition also applies to the case of heterogeneity restrictions, with the reduced-form correlation between the linear combinations $[1,0]Y_t$ and $[-\lambda,1]Y_t$ replacing the correlation between variables 1 and 2.

When are the two sets equal? When there is only one common shock to both variables and an idiosyncratic shock to the second variable, the identified sets in response to the common shock will be 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.

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 a maximal 180° angle. Standard sign restrictions as defined above imply an angle of 90°, but heterogeneity restrictions imply an angle of more than 90°.[7] 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 $\tilde{A}_{22}$. This case arises when we impose a soft zero restriction: For large λ, we are constraining the response of variable one, that is, $r_a(0)_1$, to lie in $[0, \lambda^{-1} r_a(0)_2]$. Given that $r_a(0)_2 \geq 0$, the limit of $\lambda \nearrow \infty$ is point identification.

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 $a_{1,1}(Q) = \tilde{A}_{11}[q_{1,1}, q_{1,2}]$ subject to $||[q_{1,1}, q_{1,2}]|| = 1$. 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.

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

2.5.2 Trivariate VAR with impact restrictions

Proposition 1 shows that the 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 with 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 restrictions, set $\lambda = 0$. In the Appendix, we allow for restrictions of different signs.

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

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

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

Proposition 2 (Set reduction of $IS(\cdot)$ under heterogeneity restrictions in trivariate VAR)**.** *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 (A.4) □

The intuition from Proposition 1 also explains Proposition 2: Consider a case in which shock identification calls for positive comovements between the variables. The sufficient condition applies to the case in which 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.[8]

In Appendix B, 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

[8]The same logic generalizes to the case of a p dimensional VAR in which $\Sigma_{1i} \geq 0$ for $i = 1, \ldots, p$ with up to $p-3$ positivity restrictions on the extra variables appended to (2.8).

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 3 to both shocks satisfy the heterogeneity restriction in population? We show in Appendix B.2 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 responses to all aggregate shocks satisfy the heterogeneity restrictions, heterogeneity restrictions have no bite.

3 Estimation

The uncertainty about the identified impulse response functions stems from two sources – the size of the identified set given reduced form parameters and the estimation 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.

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

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

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

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. The resultant draws are independent realizations of the reduced form posterior; see Uhlig (1994). We drop reduced form draws with admissible sets of measure zero.

To find out which reduced form parameters to drop, if any, we first re-write the restrictions in a convenient matrix form. We use this matrix form below to check whether the admissible has positive measure. Specifically, given draws $B^{(d)}$ and $\tilde{A}^{(d)} = \mathrm{chol}(\Sigma^{(d)})$, compute the following matrix:

$$W^{(d)} \equiv \begin{bmatrix} S(E-\Lambda)\mathcal{B}_0^{(d)}\tilde{A}^{(d)} \\ S(E-\Lambda)\mathcal{B}_1^{(d)}\tilde{A}^{(d)} \\ \cdots \\ S(E-\Lambda)\mathcal{B}_H^{(d)}\tilde{A}^{(d)} \end{bmatrix}, \tag{3.3}$$

where $\mathcal{B}_h^{(d)} = \sum_{s=0}^{h}(B^{(d)})^s$ if estimated in growth rates and $\mathcal{B}_h^{(d)} = (B^{(d)})^h$ if estimated in levels. We then write the restrictions on q as $W^{(d)}q \leq 0$. In the previous section, we focused on impact restrictions only with $h = 0$ so that $W^{(d)}$ simplified to $W^{(d)} = S(E-\Lambda)\tilde{A}^{(d)}$.

We provide a novel condition for assessing whether identified sets have positive measure: We check that the Chebychev center of the constrained set (prior to normalization) is non-degenerate. Intuitively, the Chebychev center is the center of the largest ball inscribed in the constrained set. We prove below that for continuous prior distributions, existence of a Chebychev center x_c with a ball of radius $r > 0$ ball around it is equivalent to an identified set with positive measure. To ensure the problem is well defined, we additionally restrict the solution to $[-1, 1]^n$, the unit n-cube. If $r > 0$, then we can construct a set of positive measure on the constrained unit n-sphere. And if the identified set has positive measure, it must lie strictly in the interior, so that we can construct a candidate Chebychev center that is nondegenerate.

Definition 6 (Chebychev center)**.** *The Chebychev center of the constrained set $Wx \leq 0$ is given by x_c that solves:*

$$\{x_c, r\} = \arg\max_{x_c, r\geq 0} r \quad \textit{s.t. } W(x_c+u) \leq 0, \quad x_c + u \leq 1, \quad -(x_c+u) \leq 1, \quad \forall u : ||u|| \leq r,$$

where the inequalities are element by element.

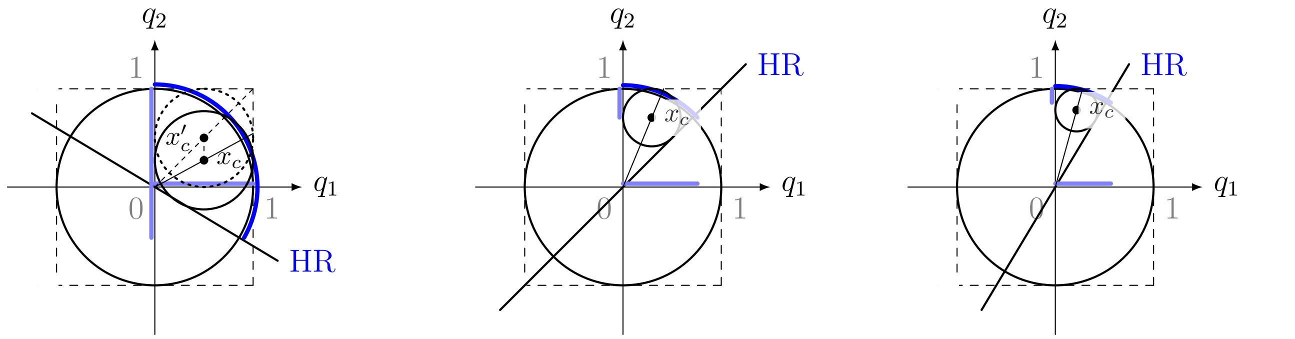
The center point x_c need not be unique, as Figure 3.1 shows. However, the constrained set is convex and the objective linear. Any local maximum r^* is therefore a global maximum.

Proposition 3 (Non-empty admissible set)**.** *Take any continuous prior π_q over q with strictly positive support on the unit sphere. The identified set for reduced form parame-*

(a) $\lambda\tilde{A}_{1,1} - \tilde{A}_{2,1} \leq 0, \tilde{A}_{2,1} \geq 0$
Chebychev center x_c not unique
Largest radius r

(b) $\lambda\tilde{A}_{1,1} - \tilde{A}_{2,1} \geq 0, \tilde{A}_{2,1} \geq 0$
Chebychev center x_c unique
Smaller radius r

(c) $\lambda\tilde{A}_{1,1} - \tilde{A}_{2,1} \geq 0, \tilde{A}_{2,1} \leq 0$
Chebychev center x_c unique
Smallest radius r



Note: The plot shows the Chebychev centers x_c for the cases considered in Figure 2.1. While the center is not unique in case (a), the radius r of the largest ball inscribed in the intersection of $Wx \leq 0$ and the unit n-cube is well defined. If r is strictly positive, the identified set has positive measure under any strictly positive prior on the unit n-sphere.

Figure 3.1: Chebychev center: Examples in $\mathbb{R}^2$.

ters (B, Σ) has positive measure under π_q iff the Chebychev center x_c of the constrained set in $[-1, 1]^n$ satisfies $r > 0$ with strict inequality.

Proof. See (A.5) □

In our algorithms below, we use a standard reformulation of the Chebychev center problem that transforms it to a simple linear programming problem whose solution is easy to compute numerically.[9] The condition that guarantees a non-empty identified set in Giacomini and Kitagawa (2014), in contrast, is stated in terms of a vector they search for by Monte Carlo integration.

3.2 Prior-robust inference

In a standard BVAR with sign restrictions such as Uhlig (2005), the posterior distribution of impulse response functions (IRF) results from integrating out both the rotation matrix Q and the reduced form parameter uncertainty. However, the many possible prior distributions over Q may imply different shapes for the resultant IRF: Baumeister and Hamilton (2015) point out that the commonly used prior that Q is uniformly distributed in the space of orthonormal matrices does not translate to a uniform distribution over the identified set of the structural parameters. We also find this in our applications below. Additionally, Arias,

[9] Specifically, Definition 6 is equivalent to solving $\max_{x_c, r} r$ s.t. $W_{i,\circ} x_c + r||W_{i,\circ}|| \leq 0 \forall i$ and $-x_n + r \leq 1, +x_n + r \leq 1 \forall n$ given that $W_{i,\circ} u \leq ||W_{i,\circ}|| ||u|| \leq ||W_{i,\circ}|| r$.

Rubio-Ramirez, and Waggoner (2014) argue that practitioners have combined sign and zero restrictions in ways that introduced unnoticed prior information.

One can address the criticism by Baumeister and Hamilton (2015) and Arias, Rubio-Ramirez, and Waggoner (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 distribution over the reduced-form parameters (β, Σ), and we should use this information that transparently reflects the data.

Thus, we follow Moon, Schorfheide, and Granziera (2013) to compute the infimum and supremum over all admissible 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), similar to Faust (1998) and Giacomini and Kitagawa (2014). 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 tell 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). 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.

As a byproduct of the optimization problem in the algorithm, we obtain Lagrange multipliers on the constraints. These multipliers serve as a measure of the importance of restrictions, as we illustrate in our empirical applications.

Algorithm 1 Prior robust inference

For $d = 1, \dots, D$ do:

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 $W^{(d)}$ according to (3.3).
3. Calculate the bounds if the identified set it is non-empty:
   (a) Solve for the Chebychev center x^c of the set $W^{(d)}x \leq 0$:

$$\{x^c, r\} = \arg\max_{x,r} r \quad \text{s.t. } W^{(d)}_{i,\circ}x + r||W^{(d)}_{i,\circ}|| \leq 0 \forall i,$$
$$e_i'x + r \leq 1, -e_i'x + r \leq 1 \forall i.$$

   (b) Verify that the identified set is non-empty, i.e., proceed if $r > 0$. Otherwise, go back to Step 1.
   (c) For each variable $i = 1, \dots, p$ and for each horizon $s = 0, \dots, S$ solve:

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

Save the resulting values as upper and lower bounds as well as the Lagrange multipliers on the constraints $Wq \leq 0$.

Our Algorithm 1 is close to those of Faust (1998) and Giacomini and Kitagawa (2014). Faust (1998) focuses just on upper bounds for the variance decomposition. Giacomini and Kitagawa (2014) focus on impulse response functions. They differences to Giacomini and Kitagawa (2014) are three: (1) We avoid simulation to determine whether the identified set is non-empty and use the Chebychev center instead. (2) We always use optimization to compute bounds rather than approximating the bounds using stochastic integration in complex cases: We have found our algorithm to work well in high-dimensional problems when the dimension of q was 19 and the number of restrictions above 200. Our applications below operate in 10 to 14 dimensions. 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`[10] solves the problem efficiently. For high-dimensional problems, we can run the algorithm in parallel, given independent posterior

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

draws for $B^{(d)}$ and $\tilde{A}^{(d)} = \text{chol}(\Sigma^{(d)})$.[11] (3) We use the optimization to compute Lagrange multipliers as a natural measure of the importance of restrictions.

Moon, Schorfheide, and Granziera (2013) show that, under some conditions, the identified set for IRFs is convex and bounded. Our setup maps into theirs[12] so that under their Assumption 1, convexity follows. Giacomini and Kitagawa (2014) show that under convexity, the prior-robust Bayesian algorithm restores asymptotic equivalence between frequentist and Bayesian inference, which does not hold for fully Bayesian inference on set-identified models (Moon and Schorfheide, 2012).

We can adapt the prior-robust algorithm to compute bounds on any well-defined moment of the reduced form parameters and rotation vectors. For, example, the admissible set over q also has implications for policy rules (see Arias, Caldara, and Rubio-Ramirez, 2015), multipliers, or the FEV decomposition. Besides impulse responses, we focus here on the FEV decomposition. In Appendix C we follow Uhlig (2003) to show that the FEV for variable i at horizon H associated with the orthonormal vector q can be expressed as $q'S_{i,H}q$ where $S_{i,H} \equiv \sum_{h=0}^{H}(e_i\mathcal{B}_h^{(d)}\tilde{A})'(e_i\mathcal{B}_h^{(d)}\tilde{A})$. We can now compute bounds on the FEV 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 $q'S_{i,H}q$. 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 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 admissible set. Because the admissible 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 iden-

[11] In the language of Giacomini and Kitagawa (2014) and Kline and Tamer (forthcoming), we find that for the applications reported here, the posterior plausibility of our restrictions is 100% unless we introduce soft zero restrictions. With soft zero restrictions, the posterior plausibility was 97% or higher.

[12] Our notation maps to Moon, Schorfheide, and Granziera (2013) as follows: $R^v = [\mathcal{B}_0\tilde{A}; \ldots; \mathcal{B}_{\max_{\tilde{h},H}}\tilde{A}]$, where $\mathcal{B}_h$ is defined below (3.3), H is the restriction horizon, and $\tilde{h}$ is the impulse response horizon of interest. $M^{S,1} = [e_{\tilde{h}}' \otimes e_j'; I_H \otimes (S(E - \Lambda))]$, $M^{S,2} = I$ and $j \in \{1, \ldots, n\}$ is the impulse response variable of interest. Their equation (24) then becomes $\tilde{S}_{(q)}\phi = ((M^{S,1})' \otimes q')S_\phi\phi$ where $\phi = \text{vec}(R^v)$. Note that $S_\phi, \tilde{S}(\cdot)$ are Moon, Schorfheide, and Granziera (2013) notation, unrelated to our matrix S. Furthermore, this setup already imposes $\tilde{k} = 1$ (one structural parameter of interest) and that $\tilde{S}_R(q)$ includes the response of interest if it is sign-restricted, as assumed in their Lemma B1.

tified set is non-empty, the rotation vector q is distributed uniformly on the unit n-sphere corresponds to the following complete Bayesian model:[13]

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

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

In practice, we found that it can be extremely difficult to sample from $\pi_q(q|B, \Sigma; R(\cdot))$ when R has many restrictions. We therefore devise 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 unit norm is uniformly distributed on the unit n-sphere.[14] We formally state these facts in Proposition 4.

Proposition 4 (Condition for admissible set to represent valid draw from Haar measure.)**.** *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. See (A.6) □

Proposition 4 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]

[13]Our formulation of the agnostic prior is the same as the one in Del Negro and Schorfheide (2011). While our prior without restrictions is also agnostic in the sense of Arias, Rubio-Ramirez, and Waggoner (2014) it may not be conditionally agnostic in their language, because the size of the identified set enters the probability of q via $\int_{\mathbb{Q}\cap\{\tilde{q}|R(B,\Sigma)\tilde{q}\leq 0\}} d\tilde{q}$ in our setup. Our prior implies that the marginal data density is unaffected by the prior over q when the identified set is never empty: $p(Y) = \int\int\int p(Y|B, \Sigma)p(B, \Sigma)p(q|B, \Sigma; R(\cdot))dqdBd\Sigma = \int\int p(Y|B, \Sigma)p(B, \Sigma)\int p(q|B, \Sigma; R(\cdot))dqdBd\Sigma = \int\int p(Y|B, \Sigma)p(B, \Sigma)dBd\Sigma$.

[14]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 on 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

For $d = 1, \ldots, D$: do:

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 $W^{(d)}$ according to (3.3).
3. Draw from $p(q|B^{(d)}, \Sigma^{(d)}; R)$
   (a) Verify that the identified set is non-empty, i.e., proceed if $r > 0$. Otherwise, go back to Step 1.
   (b) Initialize $x^{(d,0)} = \frac{x^c}{||x^c||}$ where x^c is the following Chebychev center:
   $$\{x^c, r\} = \arg\max_{x,r} r \quad \text{s.t. } W^{(d)}_{i,\circ} x + r||W^{(d)}_{i,\circ}|| \leq 0 \forall i,$$
   $$e_i' x + r \leq 1, -e_i' x + r \leq 1 \forall i.$$
   (c) Draw $\bar{\ell}$ realizations of $q^{(d,\ell)}$ using the following Gibbs sampler:
      i. For $\ell = 1, \ldots, \hat{\ell} + f \times \bar{\ell}$: For $m = 1, \ldots n$, draw $x_m^{(d,\ell)}$ from the univariate truncated normal distribution truncated to $[l_m^{(d,\ell)}, u_m^{(d,\ell)}]$, where $u_m^{(d,\ell)} = \min\left\{\infty, \min_{\{j:W_{jn}^{(d)}>0\}} -\frac{W_{jn}^{(d)} x_{-m}^{(d,\ell-\mathbf{1}\{n>m\})}}{W_{jn}^{(d)}}\right\}$ and $l_n^{(d,\ell)} = \max\left\{-\infty, \max_{\{j:W_{jn}^{(d)}<0\}} -\frac{W_{jn}^{(d)} x_{-m}^{(d,\ell-\mathbf{1}\{n>m\})}}{W_{jn}^{(d)}}\right\}$.
      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

We are now equipped to analyze whether heterogeneity restrictions also sharpen inference in practice in two applications. The first application analyzes productivity news shocks; the second application analyzes defense spending shocks.

4.1 Productivity news shocks

In our first application, we ask: Is news about future fundamentals an important source of economic fluctuations? Beaudry and Portier (2006) prominently argue that news shocks about future productivity constitute one of the main drivers of business cycles. Beaudry and Portier (2014) report that news shocks account for 50%-80% of the variance in consumption, investment, and GDP at the two-year horizon. Kurmann and Sims (2017), in contrast, use a different identification scheme in the spirit of Barsky and Sims (2011) to argue that news shocks account for 60% of the variation in these variable but not at horizons of less than five years. These approaches all use methods to point identify news shocks. Alternatively, Beaudry, Nam, and Wang (2011) use sign restrictions as an identification strategy. However, Arias, Rubio-Ramirez, and Waggoner (2014) show that their approach uses prior information that is not acknowledged and, when implemented only with the stated prior, inference becomes imprecise.

We show here how adding and exploiting heterogeneity combined with our suggested inference method sharpens inference substantially but is impractical with standard inference methods because of the larger dimension of the model and the larger cardinality of the set of restrictions. To incorporate heterogeneity restrictions, we include and rank the exposure of different industry returns to news shocks. The novel identifying assumption is that we restrict productivity news shocks to move the stock returns of the most innovative sectors the most.

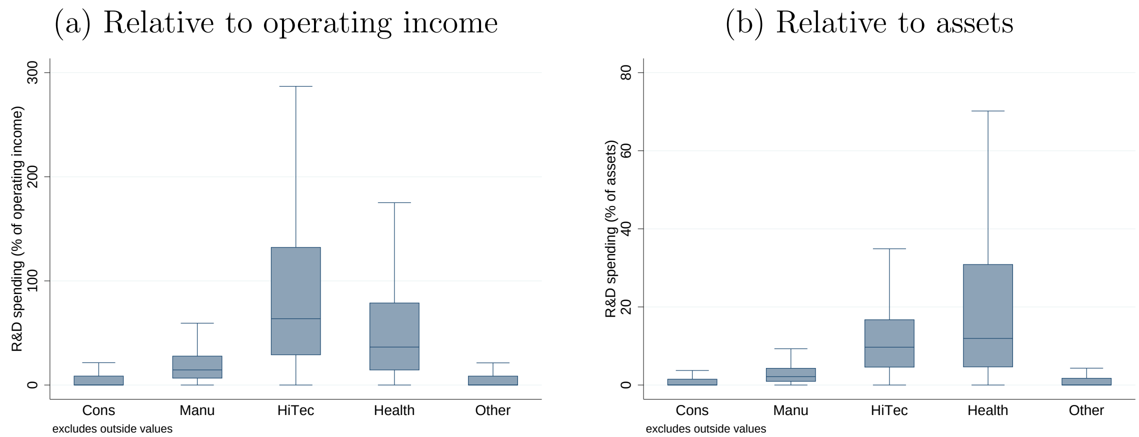
4.1.1 Data, specification, and identification

Data. Our benchmark data set consists of nine variables in total in quarterly frequency. We include four macro variables, namely real gross value added in the business sector (output), utilization-adjusted TFP, hours worked in the business sector (all taken from Fernald, 2014) and consumer confidence. For the micro series, we use readily available real industry stock returns. To keep the estimation simple, we focus on the five-industry classification by Fama and French (1997), namely consumers, manufacturing, high tech, health and other[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 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 five industries to keep the model parsimonious, we show below that our results hold using the finer ten industry classification. Table 4.2 summarizes the benchmark data. The sample period is from 1960:Q1 to 2015:Q4.

Specification. We use quarterly data in log-levels and allow for four lags. Our benchmark specification uses a flat prior. When we use a flat prior, we also include a deterministic

[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 U.S. and whose records are denominated in U.S. dollars. For our analysis, we winsorize the data at the 1st and 99th percentile year by year.



Note: 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 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. Source: Compustat. U.S. firms, 1960–2015.

Figure 4.1: R&D intensity by industry in the five industry classification

quadratic trend, as recommended by Ramey (2016). Below we also report results with a Minnesota prior. Throughout, we take 500 reduced-form draws from the posterior. For each reduced-form draw, we generate $10,000$ draws of the rotation vector q from the Gibbs sampler, with a thinning parameter of 10 that leaves us $1,000$ draws of q.

Identification. We require a news shock to raise real GDP, hours, 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) and Barsky and Sims (2011).

Turning to the results, we discuss the impulse-responses first because this is where we impose the restrictions. We then turn to the FEV decomposition. Next, we analyze which restrictions are the most important and then conclude with various robustness checks. In what follows, we focus on a select number of results, but we provide a full set of results for each subsection in Appendix E.1. Throughout, we show results for heterogeneity and sign restrictions for the same reduced-form draws.

	Benchmark data	Sign restrictions	Heterogeneity restrictions
Macro	Real output	Real output ≥ 0	Real output ≥ 0
	TFP	TFP ≥ 0	TFP ≥ 0
	Confidence	Confidence ≥ 0	Confidence ≥ 0
	Hours worked	Hours worked ≥ 0	Hours worked ≥ 0
Micro	FF-5 Consumers	FF-5 Consumers ≥ 0	FF-5 Manu $\geq$ FF-5 Consumers ≥ 0
	FF-5 Manufacturing	FF-5 Manufacturing ≥ 0	FF-5 Manufacturing $\geq$ FF-5 Other ≥ 0
	FF-5 High Tech	FF-5 High Tech ≥ 0	FF-5 Health $\geq$ FF-5 Manu
	FF-5 Health	FF-5 Health ≥ 0	FF-5 High Tech $\geq$ FF-5 Manu
	FF-5 Other	FF-5 Other ≥ 0	FF-5 Other ≥ 0

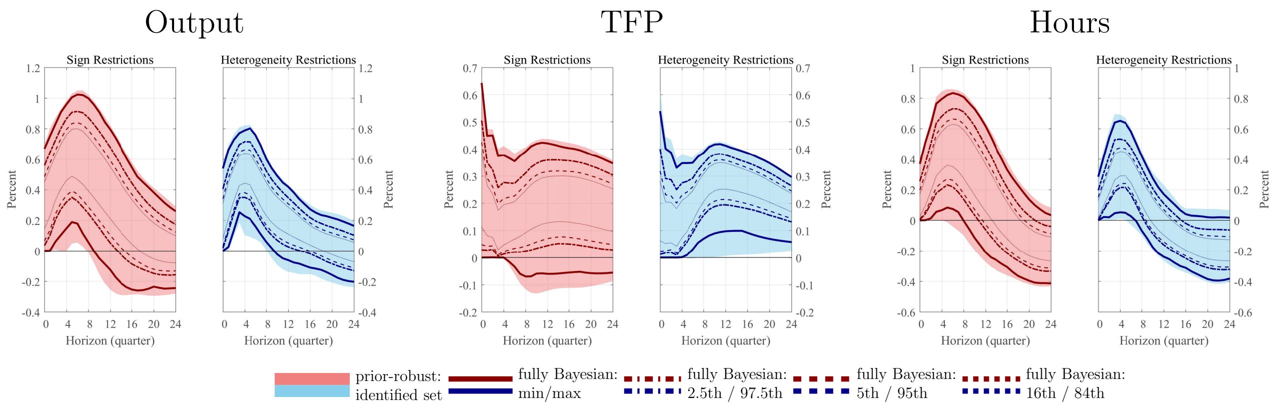
Table 4.1: Benchmark data and identifying restrictions

4.1.2 Impulse response functions

We begin our discussion of impulse-response functions by abstracting from reduced-form parameter uncertainty. Figure 4.2 shows the corresponding impulse-response function at the posterior mean of the reduced-form parameters. Specifically, the shaded areas are the identified set, computed using the prior-robust Algorithm 1. The identified set with sign restrictions is on the left in shades of red, while the set with heterogeneity restrictions is on the right in shades of blue. The lines on top of the identified set show the percentiles of the fully Bayesian posterior, conditional on the reduced-form parameters. Two conclusions emerge: First, the fully Bayesian algorithm has good coverage over the identified set: The numerical min and max according to the fully Bayesian algorithm come close to the bounds of the identified set. Second, heterogeneity restrictions sharpen the inference substantially: The identified set for TFP after six years is 30% smaller than with sign restrictions and bounded away from zero. Heterogeneity restrictions reveal a more pronounced hump-shaped increase in output. Peak responses shrink by 0.1-0.2pp relative to sign restrictions.[19]

The set reduction due to heterogeneity restrictions increases when we consider parameter uncertainty. Figure 4.3 shows the impulse-response functions for three macro variables (for micro responses and confidence, see the Appendix) computed according to Algorithm 1. The figure shows the posterior distribution over the bounds of the identified set (i.e., the prior-robust posterior). The shaded areas represent the inner 98%, 95%, and 68% of the posterior. After the shock impact, parameter uncertainty muddies the conclusions about the shape of most responses based on the identified set, indicated with blacked dashed lines for comparison. The exception is output, whose hump-shaped response is significant across parameters with heterogeneity restrictions. Overall, parameter uncertainty amplifies the set reduction due to heterogeneity restrictions: The peak output and hours responses shrink

[19] The Appendix shows the response of consumer confidence, which is unremarkable: Confidence rises initially and then reverts back to zero. Heterogeneity restrictions sharpen inference.



Note: Heterogeneity restrictions sharpen the inference with sign restrictions in economically meaningful ways: TFP is found to increase persistently, consistent with a one standard deviation news shock leading to a strictly positive increase within $(0, 0.3\%]$ after six years. With sign restrictions, in contrast, the range is $[-0.1\%, 0.35\%]$, 50% wider. Heterogeneity restrictions also reveal a well-defined hump-shaped rise in output. On the technical side, the close match of the sampled outer bounds of the fully Bayesian algorithm with the shaded identified set confirms that our sampling algorithm samples virtually from the entire support.

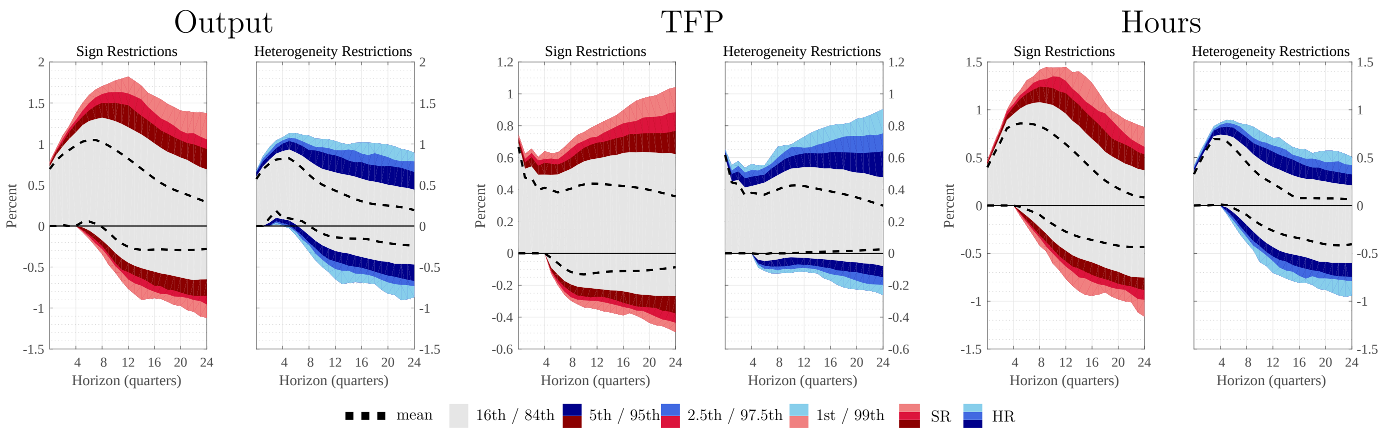
Figure 4.2: Plug-In: Responses of macro variables to a productivity news shock.

by 0.6pp, or 30-40%, with heterogeneity restrictions. Also the conceivable drop in TFP six years after an initially positive news shock is cut by 0.25pp or 50%. Table E.1(a) in the Appendix shows the set reduction.

What can researchers with conditionally agnostic priors learn from sign or heterogeneity restrictions? Figure 4.4 shows that agnostic beliefs about q sharpen inference substantially: The pointwise 98% credible sets 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 that place the peak increase in TFP between 0.1% and 0.5% about three years after the initial shock with 95% probability. This causes a hump-shaped expansion in output, peaking one year out between 0.3-0.8% with 95% confidence, according to the model with heterogeneity restrictions or 0.3-1.1% with sign restrictions only. Hours worked peak at 0.15% to 0.6% (0.15% to 0.9% with sign restrictions) with 95% probability and then may turn negative: In the long term, the wealth effect on labor supply seems to offset productivity growth, returning output to trend. Overall, we see economically sensitive responses that are much sharper with heterogeneity restrictions.

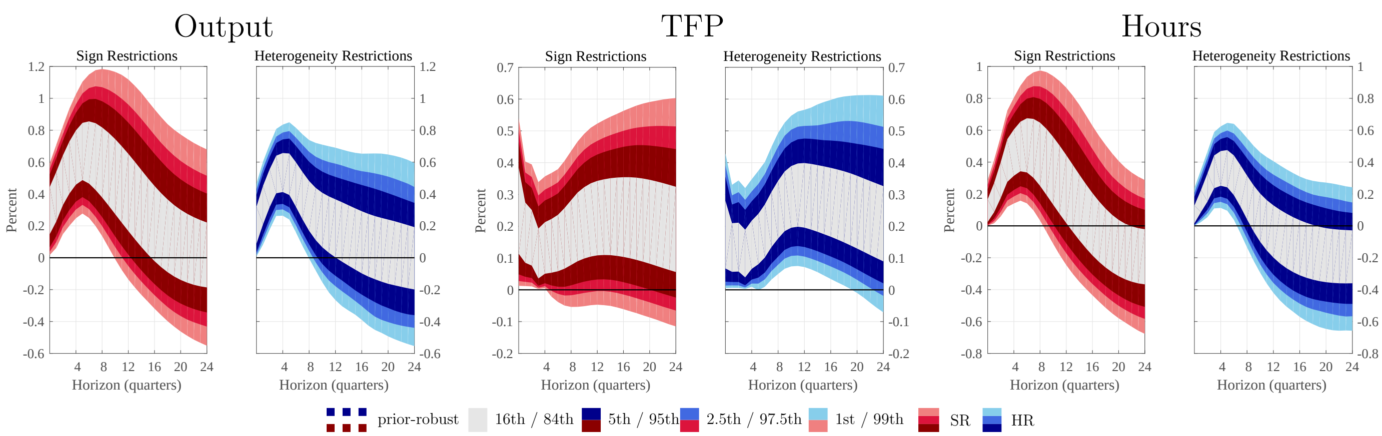
The credible sets alone could obscure irregular posterior distributions, but we show that they do not in Figure 4.5.[20] The posterior densities are unimodal and largely symmetric.

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



Parameter uncertainty is pervasive after the impact of the shock. Taking it into account we can only bound the responses with the restrictions. Heterogeneity restrictions do, however, sharpen the bounds: The peak upper bounds with heterogeneity restrictions are a third smaller for output and hours at the 99th percentile. Heterogeneity restrictions cut the lowest lower bound for TFP also by half.

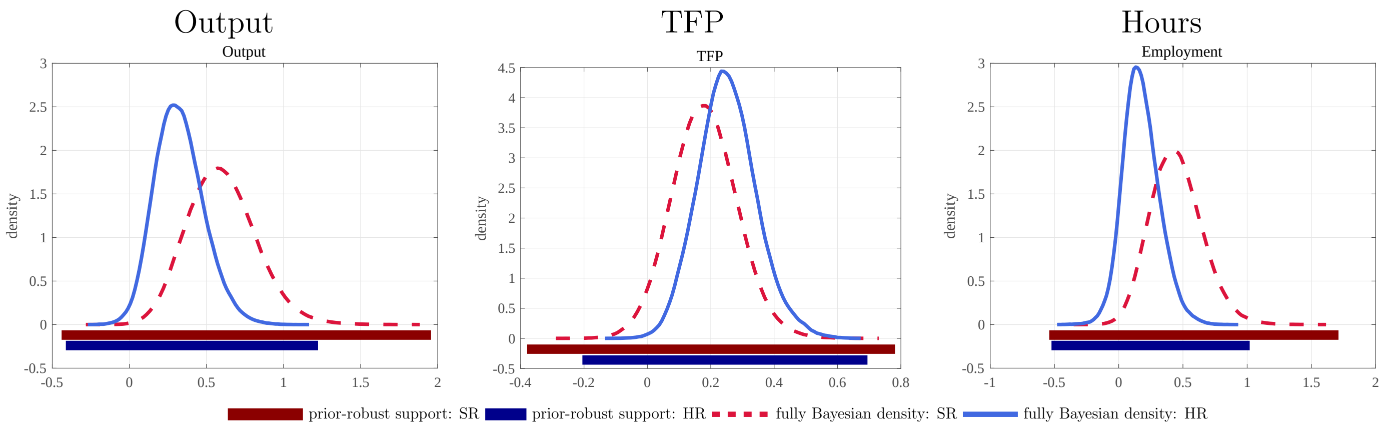
Figure 4.3: Prior-robust responses of output, TFP, and hours worked to a productivity news shock



Note: Our fully Bayesian posterior implies that TFP increases in a hump-shaped fashion in response to a productivity news shock. The hump shape in TFP is well defined only with heterogeneity restrictions. Heterogeneity restrictions also visibly reduce the credible sets for hours and TFP. In economic terms, negative wealth effects due to the permanent TFP increase on labor supply may explain the high probability put on a decline in hours worked and the zero response of output after six years.

Figure 4.4: Prior-robust responses of output, TFP, and hours worked to a productivity news shock

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 shifts toward zero using heterogeneity restrictions. In contrast, the mass shifts toward positive values for TFP. For all variables, the densities are more concentrated with heterogeneity restrictions.



Note: 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, shown 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: The TFP response puts 2.5-5% probability on zero with sign restrictions but less than 1% with heterogeneity restrictions.

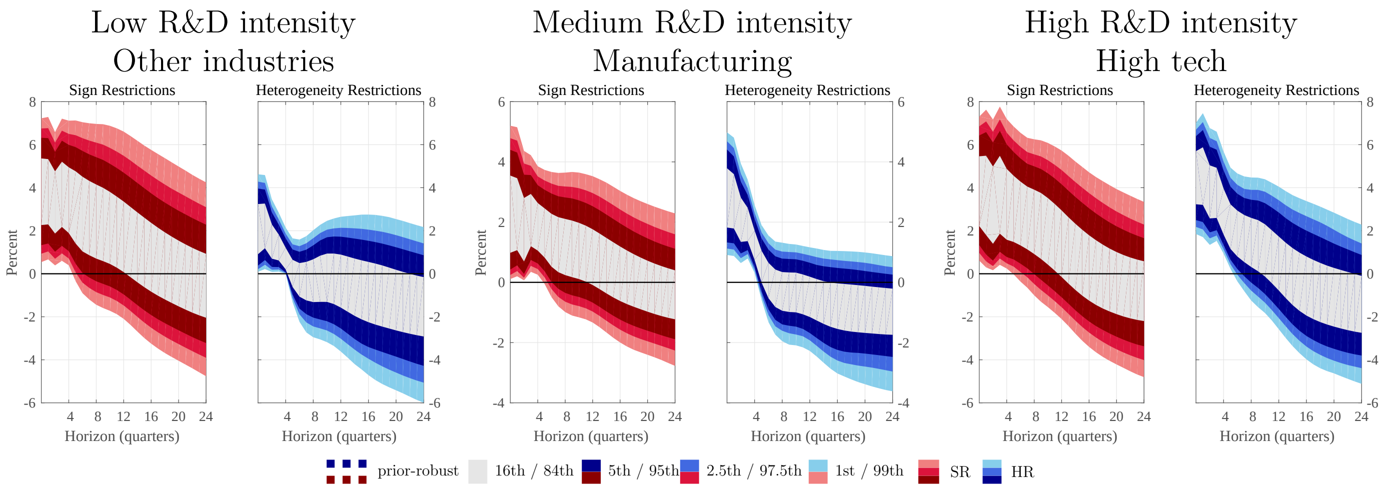
Figure 4.5: Distribution of responses to productivity news: Macro variables two years out.

We now turn to the micro-responses of cumulative industry stock returns in Figure 4.6. Heterogeneity restrictions on these responses yield tighter bounds on the macro variables. Mechanically, they also sharpen inference about the micro responses themselves. The initial gains erode more quickly under heterogeneity restrictions in the other industries and the manufacturing industry than sign restrictions would indicate. We find similar patterns for all five industries and show here one of each category: the low R&D intensive other industries, manufacturing, and high tech. Heterogeneity restrictions reduce the magnitudes of responses up to 43% on impact and 67% one year out (Table E.1(b)).

4.1.3 Forecast error variance decomposition

The literature on news shocks often highlights that news shocks explain a significant part of the FEV in macro variables (e.g., Beaudry and Portier, 2014). Barsky and Sims (2011), and Kurmann and Sims (2017) advocate identifying productivity news shocks by maximizing the FEV. Here we decompose the FEV using both prior-robust and agnostic beliefs. The substantive conclusion that emerges is that under some beliefs news shocks could indeed drive most of the FEV in macro variables over some horizons, but an agnostic belief points to an important but more modest role.

Figure 4.7 summarizes the variance reduction relative to pure sign restrictions for each variable at horizons of up to six years in its upper panel. The red line, the posterior median bound with sign restrictions, shows that sign restrictions alone are uninformative because the FEV contribution ranges from 60% to 100% of the total FEV. Heterogeneity restrictions,



Note: We rank the responses of stock returns of industries from zero to two quarters according to their R&D intensity. Our goal was to sharpen inference about macro variables, but the restrictions also helps to sharpen inference about the size of returns differences. With sign restrictions, there is little difference between the response of other industries and the high tech industry but heterogeneity restrictions mechanically deliver this difference and narrow the range of possible returns.

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

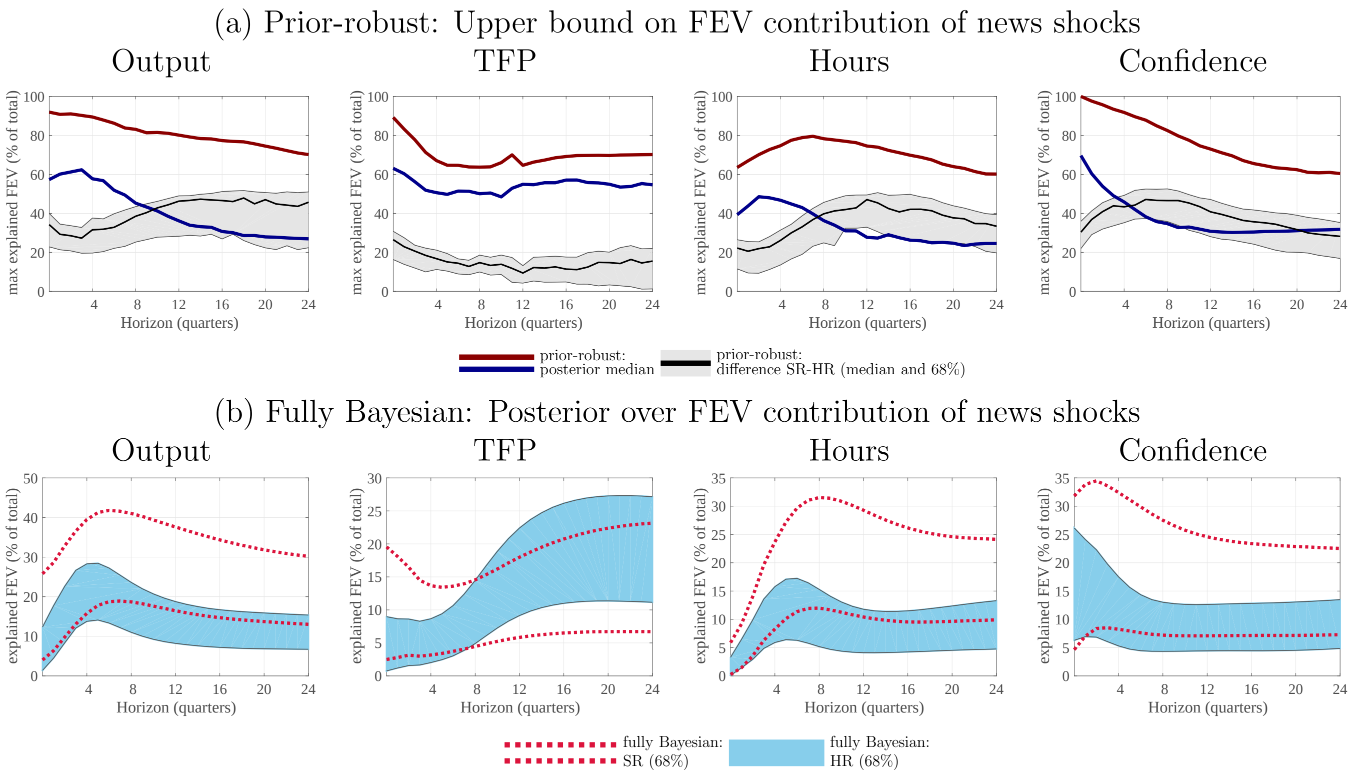
in contrast, bring the median bound for the FEV contribution to below 70% of the total throughout. At the one-year horizon, the median reduction in the FEV contribution ranges from is 32pp for output. The 68% credible set, the shaded area, ranges from 20pp to 38pp. The set reductions are persistent.

While our prior-robust results show that news shocks can play a very large role even under heterogeneity restrictions, they play a more modest role under agnostic beliefs. The bottom panel of Figure 4.7 shows the variance decomposition in this case. News shocks are still important but far less so than the upper bounds indicate: For output, news explains 14-28% of the total FEV at the one-year peak, and 11-27% of TFP after six years with heterogeneity restrictions. Compared with sign restrictions, heterogeneity restrictions suggest that news shocks are less important for output, hours, and confidence but more important for TFP, suggesting that heterogeneity restrictions identify a productivity news shock more sharply.

4.1.4 Important restrictions

We impose tighter restrictions to achieve sharper identification. Which of these restrictions matter? We show that as a byproduct of the prior-robust Algorithm 1, we can easily quantify the importance of individual restrictions. This makes our results even more transparent.

We use Lagrange multipliers to quantify the role of individual restrictions in sharpening our inference about IRFs. Lagrange multipliers answer the question: How would responses change if we tightened a given restriction by a small amount? For example, how would the



Note: We normalize FEV contributions relative to the total FEV. Heterogeneity restrictions significantly reduce the maximum role of news shocks (upper panel): With sign restrictions alone, news could explain all of the initial output and confidence FEV. Heterogeneity restrictions shrink the maximum FEV by about 20pp to 50pp for output. The reduction in the bound for TFP is lower but positive and ranges up to 30pp. The bounds remain wide in the short-run. With agnostic beliefs (lower panel), the importance of news peaks at below 30% for output after one year with 68% probability, compared to the prior-robust bound of 60%: News shocks can be a key driver of business cycles under some beliefs, but an agnostic belief suggests a modest role.

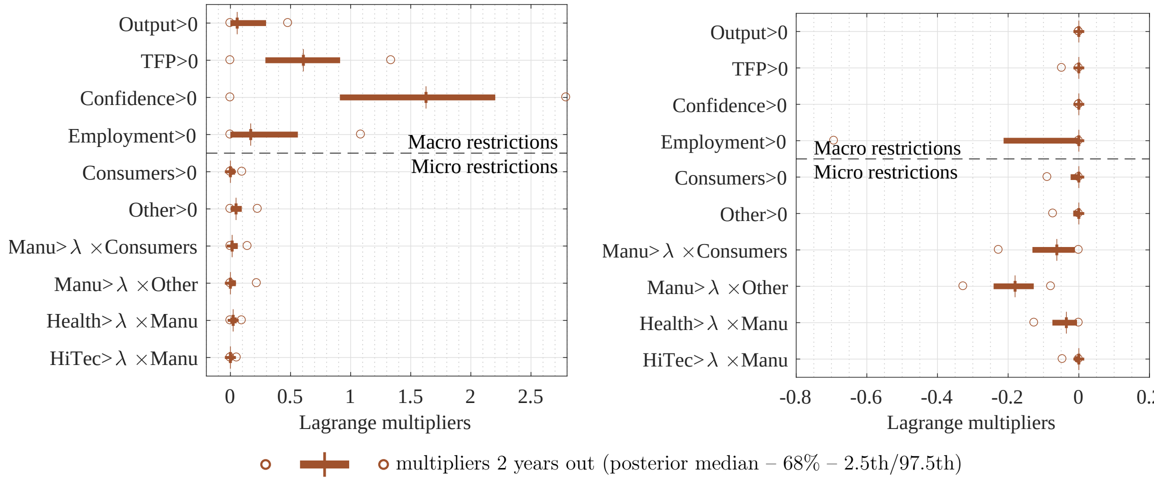
Figure 4.7: Forecast error variance contribution of productivity news shock: Macro variables.

output response two years after a news shock change if we required consumer confidence not only to be positive initially but also bigger than ϵ? Or what if manufacturing stock returns had to increase more than consumer goods stock returns plus ϵ? Figure 4.8 answers these questions for all our restrictions. The left panel shows the multipliers characterizing the lower bound, and the right panel shows those for the upper bound. Because multipliers depend on uncertain parameters, we show their distribution across reduced-form parameters. To simplify, we sum multipliers on restrictions across the entire horizon for which we impose them.[21]

Heterogeneity restrictions are the most important for tightening the upper bound on the output response but matter little for the lower bound, according to Figure 4.8. For example,

[21] Figures E.6 and Figures E.5 in the Appendix show different horizons and more macro variables.

(a) Lower bound multipliers　　(b) Upper bound multiplier



This figure quantifies the importance of all sign and heterogeneity restrictions that we use to identify the news shock for narrowing the identified set of the output response at the two-year horizon. It shows the distribution of Lagrange multipliers on all restrictions over all reduced-form draws, separately for multipliers on the lower and upper bound. We sum the multipliers across the restrictions horizons $0, \ldots, \bar{H}$. Multipliers on upper bounds are negative because tighter restrictions reduce the upper bound. We find that for pinning down the upper bound on output, the heterogeneity restrictions on stock return micro data matter more than the macro sign restrictions. Restrictions on manufacturing relative to other industries are particularly important, reducing the upper bound by 0.11pp to 0.23pp. For the lower bound on responses, the opposite pattern emerges: Sign restrictions on macro variables dominate.

Figure 4.8: Importance of restrictions for output responses to news shock two years after impact: Lagrange multipliers on restrictions.

the upper bound would not decrease if we imposed that consumer confidence had to rise more initially. In contrast, tightening the heterogeneity restriction that stock returns in the manufacturing industry rise relative to returns in the other industry would lower the upper bound of output by 0.11 to 0.23 pp. The restriction that manufacturing stocks rise more than consumer industry stocks still would lower the upper bound by 0.02 to 0.14 pp. When we look at the determinants of the lower bound, the opposite picture emerges: The initial restrictions on macro variables, and consumer confidence in particular, matter the most while the micro restrictions hardly matter.

Our analysis of multipliers in the Appendix also shows that sign restrictions on stock returns per se has little effect on the bounds of the identified set: Looking at the distribution of differences in multipliers with $\lambda = 0$ and $\lambda = 1$, we found that multipliers on micro restriction would be close to zero when we set $\lambda = 0$ and work with pure sign restrictions. In contrast, multipliers on macro restrictions do not change systematically. This highlights the importance of our heterogeneity restrictions (Figures E.6 and E.5).[22]

[22] We also show in the appendix, in Figure E.7, that the first and the last restriction horizon matter the

4.1.5 Robustness

More industries. Our results are robust to the particular industry classification we use. We double the number of industries in our VAR and again use R&D expenses to order the industry-level responses. In Figures E.8 and E.9, we contrast the responses of the macro variables across the datasets. We find significant set reductions with heterogeneity restrictions in both cases: For example, we see a reduction of the 99th percentile of the prior-robust GDP response from 1.2% to about 0.7% at the three-year horizon. One minor difference emerges when using ten industries: Instead of a hump shape, we find a constant increase in the TFP level.

Variations in the restriction horizon. Varying the horizon for which restrictions bind from five quarters ($H = 0, \ldots, 4$) to three quarters, we find that all qualitative results hold (Figures E.10 and E.11). As restrictions bind longer, impulse-responses tend to be sharper and stronger. This is consistent with the findings in Uhlig (2005, Figure 7).

Soft zero restrictions. Beaudry and Portier (2006) and Barsky and Sims (2011) impose the restriction that news cannot raise TFP immediately to identify news shocks. Here we incorporate this assumption as a "soft" zero restriction on the initial TFP response.[23] Table E.4 shows that this extra restriction yields an additional set reduction: For output, this reduces the maximal FEV by an additional 10pp to 15pp compared with heterogeneity restrictions alone. The reduction for employment is 5pp to 10pp, while consumer confidence is hardly affected. By construction, the FEV for TFP that can be explained drops dramatically at short horizons but rises with the forecast horizon. The impulse-responses change little, except for TFP (Figures E.12 and E.13).

Minnesota prior. Our benchmark is implemented with a flat prior on the identifiable reduced-form parameters as in Uhlig (2005) and Arias, Rubio-Ramirez, and Waggoner (2014). However, it is common to implement BVARs with shrinkage in the form of a Minnesota prior. As a robustness check, we implement our baseline model specifications with a Minnesota prior (Figures E.14 and E.15). As expected, impulse-responses are now smoother and more persistent. Otherwise, our qualitative results hold.

most. In absolute terms, the multipliers on the restrictions at $h = 0$ and at $h = \bar{H}$ are the largest, both for $\bar{H} = 2$ and $\bar{H} = 4$. This is intuitive when responses are monotone over the restriction horizon. Then the boundary conditions at $h = 0, \bar{H}$ largely determine the shape in between.

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

Information set. To implement heterogeneity restrictions, we need more observables. These increase both the information set and the number of unknown parameters in the VAR. Figures E.16 and E.17 show that the larger VAR with sectoral stock returns and heterogeneity restrictions sharpens inference about the macro variables also relative to a smaller VAR despite the extra parameter uncertainty. The smaller VAR has the same four macro variables, but only an aggregate stock index.[24] For output, hours, and confidence, the estimated IRFs are sharper with heterogeneity restrictions and micro data. For TFP, the credible sets are comparable at the six-year horizon but again sharper at short horizons.

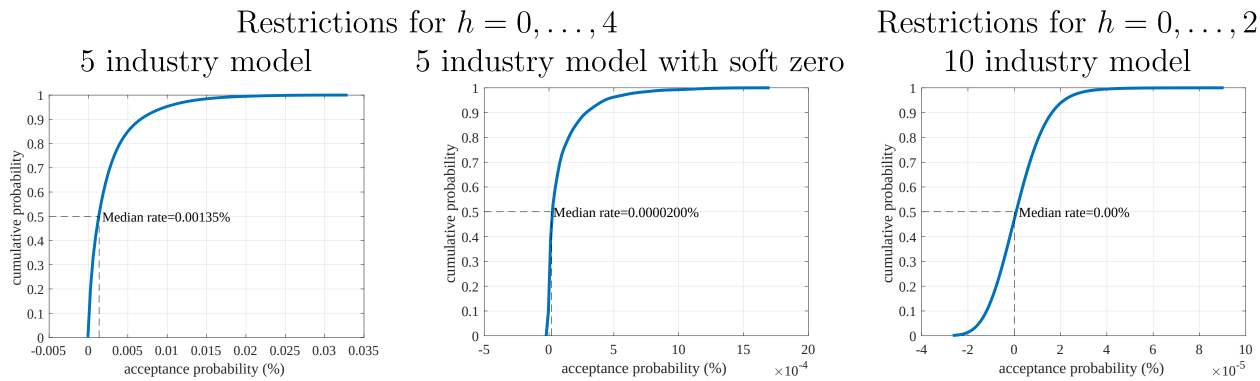
4.1.6 Comparison with the traditional rejection sampler

Sampling admissible rotations for a given set of reduced-form draws can infeasible using rejection sampling even if we now that the admissible set has positive measure: Rejection sampling becomes hard when the admissible set is relatively small. Examples include high-dimensional VARs or many restrictions for long horizons. While these case may interest practitioners, traditional samplers may fail to deliver admissible draws and underrepresent the identified set. 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 4.9 illustrates based on 5 million draws for q. To put this in perspective, Inoue and Kilian (2013) report that they used 20,000 draws for numerical stability. In our baseline application, the median acceptance rate across reduced-form draws is just 0.0014%. With the harder problem of including a soft zero restriction, the median acceptance rate drops by two orders of magnitude, to 0.00002%. Without the soft zero restriction but in the higher dimensional version with ten industry returns, the acceptance rate was zero for all reduced-form draws. In the relaxed problem that imposes the restrictions for a total of three instead of five quarters, the median acceptance rate is still 0 for the rejection sampler. We conclude that Algorithm 2 is important in practice.

4.2 Fiscal shocks

What are the effects of increased discretionary spending on the economy? A large literature that includes, among others, Blanchard and Perotti (2002), Mountford and Uhlig (2009), Barro and Redlick (2011), and Ramey (2011) debates this question. Here, we focus on defense spending as the largest component of federal government consumption and investment. We use sign restrictions for identification as in Mountford and Uhlig (2009) but incorporate

[24] In the Appendix, we show results for an equally weighted average of the five sectoral real returns. Results based on the real Wilshire 5000 index are almost identical.



Note: 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. With ten industries and restrictions for $h = 0, \ldots, 4$ (not shown), we find a zero acceptance rate for all reduced-form draws.

Figure 4.9: Distribution of acceptance probabilities for uniform proposal density over reduced-form draws

our belief that defense spending shocks benefit the defense industry more than other industries. Fisher and Peters (2010) exploit this insight to construct a proxy for defense spending shocks based on the excess returns of military contractors. Perotti (2008) and Nekarda and Ramey (2011) use this insight to analyze defense spending shocks based on shipments of manufacturing industries to the government. Here, we follow Nekarda and Ramey (2011) and use differences in defense shipments to better identify defense spending shocks.

4.2.1 Data, specification and identification

Data. Building on Nekarda and Ramey (2011), we use the NBER manufacturing database and IO-table information as the sources of our micro data. Specifically, we focus on the 20 SIC2 manufacturing industries. Our heterogeneity restriction is that shipments of all industries rise after a defense spending shock, but more so when the government is an important client of an industry.

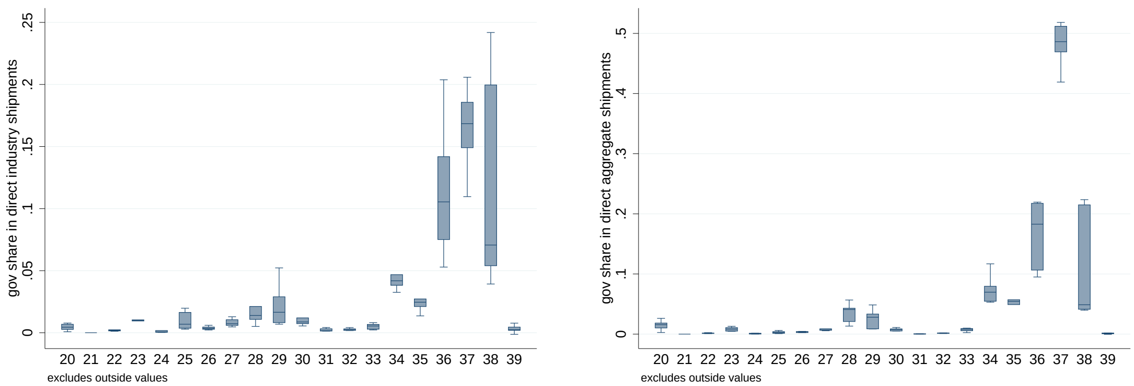
Figure 4.10 measures the importance of the government for each SIC2 industry by showing the median and distribution of the government share for the industry over time. The left panel uses direct shipments and the right panel overall shipments to the government to normalize.[25] Both measures clearly show that the electronics (SIC 36), transportation (SIC 37), and sensors (SIC 38) are the most exposed to the government. Our strategy is to pick two industries at the top of the distribution, two in the middle, and two in the bottom of the distribution. We consider different variants but choose the following six industries for our

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

baseline model. Transportation and electronics are the most exposed industries, petrol and refineries (SIC 29) and equipment (SIC 35) are the industries with an intermediate exposure, and tobacco (SIC 21) and lumber (SIC 24) are the least exposed to defense spending.

Our baseline 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), the average marginal tax rate and shipments from the six industries.[26] The sample period is 1959 to 2008.

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



Note: The boxes show the median along with the interquartile range of the importance of government shipments within 2-digit SIC manufacturing industries for the seven years between 1963 and 1992 with matching IO tables. The whiskers cover the values just outside the interquartile range ±1.5 times the interquartile range. We obtain the data from Nekarda and Ramey (2011). For our baseline, 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 4.10: Importance of government shipments by industry

Specification. We estimate an annual VAR with one lag in levels and, following Ramey (2011), include a linear-quadratic trend to remove low-frequency movements.[27] In alternative specifications we use both different sets of industries and control for expectations to rule out that our results are driven by fiscal foresight.

Identification. 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

[26] 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 (`https://www.dallasfed.org/research/econdata/govdebt`) and deflate it by the CPI.

[27] 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 make sign restrictions more competitive by reducing the incidence of explosive roots.

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.

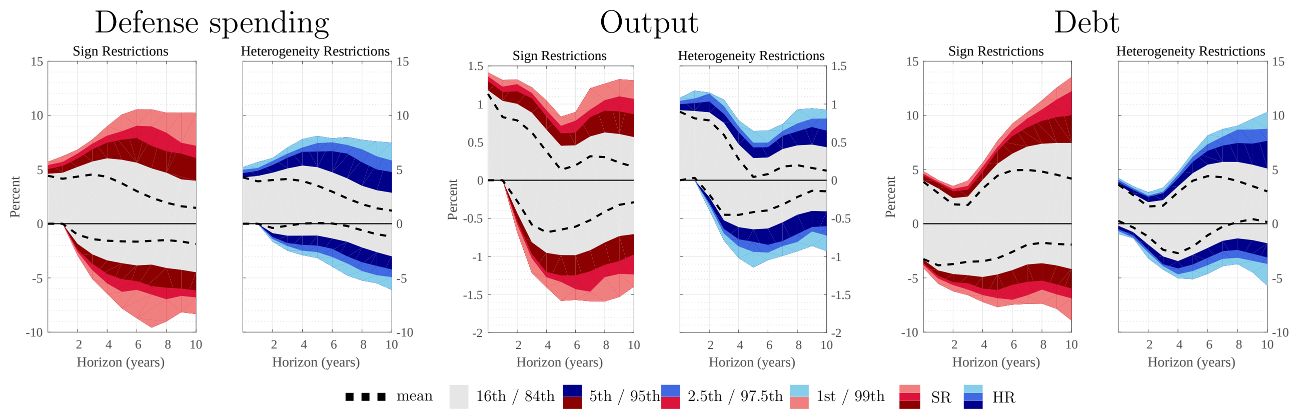
For brevity, we focus on a few results only. We provide the full results, ordered by subsection, in Appendix E.2. Let us now turn to the implied responses.

	Benchmark data	**Sign restrictions**	**Heterogeneity restrictions**
Macro	Real defense spending	Real defense spending ≥ 0	Real defense spending ≥ 0
	Real GDP	Real GDP ≥ 0	Real GDP ≥ 0
	Real federal debt	none	none
	Total hours worked	Total hours worked ≥ 0	Total hours worked ≥ 0
	Marginal tax rate	Marginal tax rate≥ 0	Marginal tax rate≥ 0
Micro	SIC21 Tobacco	Tobacco ≥ 0	Tobacco ≥ 0
	SIC24 Lumber	Lumber ≥ 0	Lumber ≥ 0
	SIC29 Petrol	Petrol ≥ 0	Petrol $\geq$ Tobacco
			Petrol $\geq$ Lumber
	SIC35 Equipment	Equipment ≥ 0	Equipment $\geq$ Tobacco
			Equipment $\geq$ Lumber
	SIC36 Electronics	Electronics ≥ 0	Electronics $\geq$ Petrol
			Electronics $\geq$ Equipment
	SIC37 Transportation	Transportation ≥ 0	Transportation $\geq$ Petrol
			Transportation $\geq$ Equipment

Table 4.2: Benchmark data and identifying restrictions

4.2.2 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 4.11 shows the results for three macro variables: defense spending, output, and federal debt. As in the analysis of the news shocks, at short horizons, the uncertainty is modest and mostly due to the width of the identified set 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 eight-year horizon); see Table E.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 (not shown) remain persistently high, while output quickly reverts back toward zero. Figure E.18 in the Appendix shows all responses.



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

A Bayesian with an agnostic prior would find that her prior sharpens inference significantly because there is little mass in the extremes of the identified set.[28] Already with sign restrictions, a Bayesian can infer 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 95% credible set for a one standard deviation defense spending increase is wide, around 0.25-3.5%, with initial increases in GDP of 0.1-0.6%. See Figure 4.12.

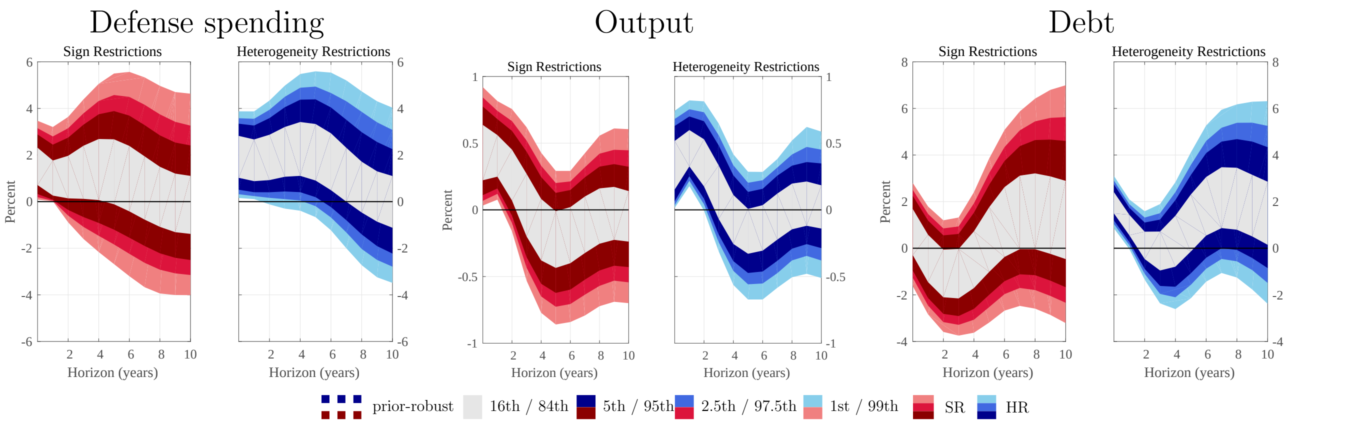
Shipments in the industries used for heterogeneity restrictions exhibit a sensible pattern; see Figure E.19 in the Appendix. As for aggregate output, industry shipments revert back to zero after three years. Heterogeneity restrictions cut the prior-robust upper bounds by 17% to 63% and almost shrink the width of the inner 95% credible sets; see Table E.5.

4.2.3 Defense spending multipliers

What does our model imply about how effectively defense spending stimulates the economy? Following Mountford and Uhlig (2009), we compute cumulative present discounted value multipliers to answer this question.

Without additional beliefs, the multiplier can be implausibly high – and virtually unbounded on impact: Our identifying assumptions are consistent with large GDP increases that coincide with tiny defense spending increases. This scenario may be caused by shocks

[28] The tails of the numerically computed posterior confidence sets cover the identified set well. See Figure E.21 in the Appendix for a comparison at the posterior mean of the reduced for parameters.



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

Figure 4.12: Prior-robust responses to defense spending shock: Macro variables

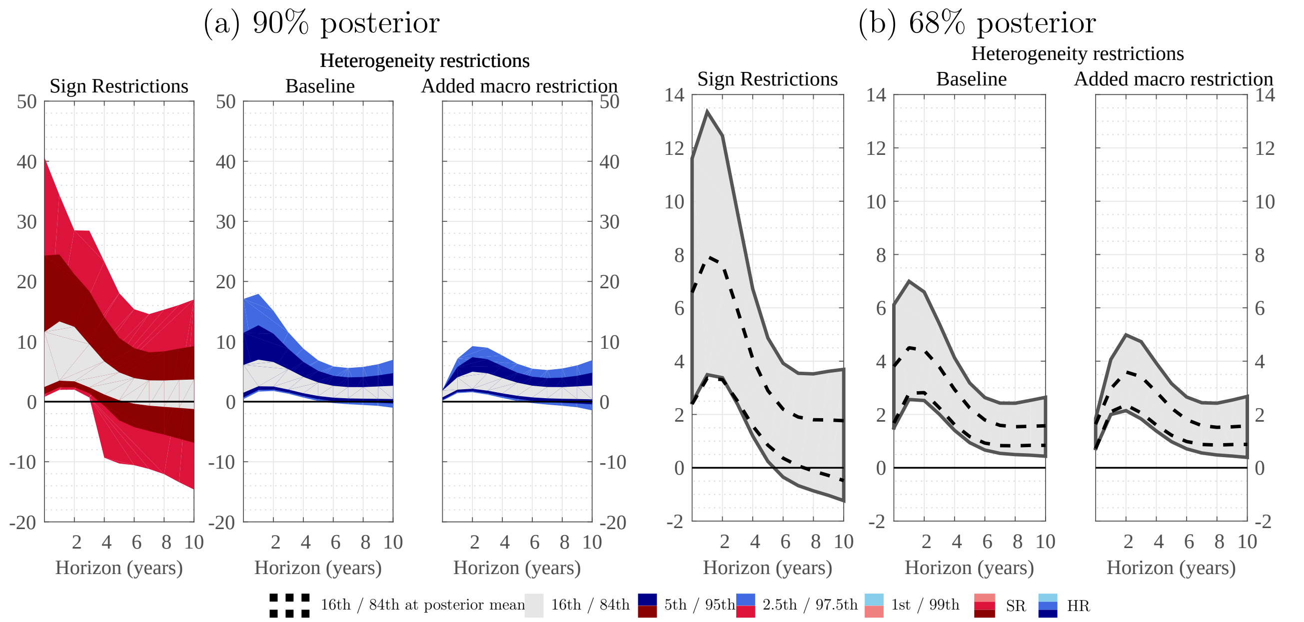
that are not spending shocks. To rule this out, we add our belief that the impact multiplier is no larger than two: a heterogeneity restriction on defense spending and output. Unlike the micro heterogeneity restrictions, this macro heterogeneity restriction has little effects on impulse-responses, so we show the corresponding impulse-responses only in the Appendix.[29]

Figure 4.13 shows the distribution of the cumulative multipliers over time: With sign restrictions, the impact multiplier ranges up to 40 with 90% posterior probability. With 68% it still ranges from 3 to 11. With our baseline heterogeneity restrictions the 68% credible set covers 1.5 to 6 – an improvement but still implausible. Restricting the impact multiplier to lie below two mechanically does just that on impact, yielding a 68% range of 0.75 to just below 2. The tighter restriction on impact multipliers gives way to increasing multipliers at the two-year horizon that converge to a range of 0.5 to 2.5 after ten years with 68% posterior probability compared with a range of -1 to 4 with sign restrictions.

4.2.4 FEV decomposition

Pure sign restrictions would allow defense spending shocks to be a main driver of business cycles: According to Table E.6 in the Appendix, the posterior mean of the prior-robust bounds on the forecast error explained by the defense shock range from 64% to 85% of the variance on impact for macro variables. After 10 years, the shock could still account for up to 33 to 43%. As Table 4.3 shows, heterogeneity restrictions tighten these bounds by 13%

[29]The added assumption shrinks the 95th percentile of the impact response from 0.70% to 0.25%. Subsequent responses of output and the other macro variables are, however, virtually unchanged.



Note: We compute PDV multipliers using a 5% discount rate and a defense share in GDP of 5.5%. With sign restrictions, credible sets are very wide. Heterogeneity restrictions on industry shipments cut the range for impact multipliers in half, but still puts significant probability on unreasonable multipliers. With the added belief that impact multipliers do not exceed two as a macro heterogeneity restriction, we obtain a reasonable distribution: With 68% probability, multipliers lie between 0.5 and 2 on impact, rise to 2 to 5 after two years, and then converge toward a range of 0.5 and 2.5. After impact, about half of the uncertainty stems from uncertain parameters.

Figure 4.13: Distribution of present discounted value (PDV) output multipliers.

to 35% on impact and 11% to 22% after 10 years. The reduction is particularly pronounced for taxes in which case the posterior median of the upper bound is essentially cut in half.

4.2.5 Important restrictions

To assess which restrictions are important, we focus on the output response. We conclude from the Lagrange multipliers that the heterogeneity restrictions on micro data are more important for bounding the output response from below after three years than the macro sign restrictions. For the upper bound on output, the opposite is true.

Specifically, the posterior median of the Lagrange multipliers on the macro sign restrictions is zero at the three-year horizon, according to Figure E.25 in the Appendix. In contrast, the heterogeneity restrictions on transportation shipments have median Lagrange multipliers around 0.1, implying that marginally tightening these restrictions would raise the upper bound on output by 0.1pp. Also, the sign restrictions on tobacco and lumber are important, with Lagrange multipliers around 0.15. In contrast, the median upper-bound multipliers on the micro data are close to zero but -0.4 for the sign restriction on tax rates. This implies

Difference between HR and SR (% of total FEV)

Variable	Horizon H (year) 0	1	2	3	10
Defense spending	14 (6,19)	10 (2,15)	8 (2,13)	8 (1,12)	11 (2,14)
Output	35 (21,42)	19 (9,25)	11 (6,16)	9 (4,14)	13 (7,15)
Debt	13 (5,18)	15 (7,23)	26 (10,34)	26 (11,32)	16 (5,22)
Hours	17 (8,22)	10 (3,15)	10 (3,14)	9 (2,13)	12 (6,14)
Tax rate	27 (12,32)	25 (12,31)	24 (12,32)	26 (11,33)	22 (7,36)

Note: We find a moderate to large reduction in the maximum FEV of macro variables when comparing heterogeneity restrictions with sign restrictions, even though the direct restrictions on macro variables are the same. Set reductions at the posterior median peak at more than 25% of the total FEV for output, debt, and the tax rate at some horizons. The reduction for defense spending and hours worked is 8% to 17% of the total FEV. The table shows the posterior median along with 68% credible sets across reduced-form draws.

Table 4.3: FEV of macro variables due to defense spending shocks: Prior-robust bounds

that marginally strengthening restriction that tax rates rise would lower the upper bound on the output response after three years by 0.4pp.

We also find that moving from pure sign restrictions on all variables to sign and heterogeneity restrictions does not diminish the importance of macro sign restrictions: The 68% posterior for the change in multipliers always includes zero. Indeed, imposing heterogeneity restrictions on some industries makes the sign restrictions on tobacco and lumber more important.

4.2.6 Robustness

Other industries. To address specification uncertainty, we consider different sets of SIC2 industries. We still find set reductions but sometimes for different macro responses than in our baseline. In one variant, we keep electronics and transportation (SIC 36 and 37) as the industries most exposed to defense but use chemicals and metal (SIC 28 and 34) as industries with an intermediate exposure and furniture and leather (SIC 25 and 31) as the industries with the lowest exposure. In another variant, we use just four industries: We impose that transportation reacts more than instruments (SIC 38), instruments react more than metals, and metals react more than apparel. The first variant leads to sharper inference on output than our baseline but is ambiguous about the initial debt response. The second variant is similar to the baseline in terms of allowing us to sign the initial debt response, but for output, the difference to sign restrictions is less pronounced. See Figures E.29 and E.30.

Fiscal foresight. To address fiscal foresight, we control for expectations by including forecast of industrial production (IP) from the Livingston Survey in the VAR and the proxy for defense spending news from Ramey (2011) as an exogenous regressor, following Anderson, Inoue, and Rossi (forthcoming). Controlling for general expectations and defense news hardly changes the results: Hours worked increase slightly less both with sign and heterogeneity restrictions. Defense spending may become slightly more persistent. See Figures E.29 and E.30. Interestingly, the forecast for IP increases one year after the shock, although we do not restrict the IP response. This increase is sharper with heterogeneity restrictions.[30] If IP forecasts and the news proxy in Ramey (2011) capture fiscal foresight well, our results are robust to fiscal foresight.[31]

5 Conclusion

Our paper shows how to design and implement heterogeneity restrictions to sharpen inference in VARs. Heterogeneity restriction impose a ranking on the relative magnitude of impulse-responses. For most of our paper, we impose them based on easily available micro data, but we also provide two applications with heterogeneity restrictions on macro variables. We derive conditions under which these restrictions sharpen our inference about variables not subject to extra restrictions. These conditions say that heterogeneity restrictions sharpen inference if they require the conditional covariance to be more at odds with the unconditional forecast error covariance than the standard sign restrictions.

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 use a simple linear program to determine whether the identified set has positive measure.

Using these tools, we demonstrate how useful heterogeneity restrictions are in 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 het-

[30]Figure E.31 shows that heterogeneity restrictions shrink the identified set at the identified mean sufficiently to exclude negative responses and to exclude negative responses with 97.5% posterior probability when being fully Bayesian.

[31]We have also explored an alternative quarterly dataset that includes less industry detail but includes the Korean war (results available on request). For this dataset, we found that responses differ in the early and late sample period with significant effects of heterogeneity restrictions in the second half of the sample.

erogeneity 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 relative to sign restrictions, with peak reductions of bounds on variance decompositions and impulse-responses for output in the order of 30% in both applications. We verify the importance of heterogeneity restrictions for set reduction using Lagrange multipliers. These show, for example, that heterogeneity restrictions on micro data can be more important for the bounds on output responses than sign restrictions on macro variables.

Heterogeneity restrictions also help to sharpen fully Bayesian inferences. Interestingly, the extra restrictions do not simply shrink the response toward zero. For example, in the fiscal application, we find that we cannot sign the debt response with sign restrictions but find that debt increases significantly after a spending shock with heterogeneity restrictions.

References

Amir-Ahmadi, P., and H. Uhlig (2015): "Sign Restrictions in Bayesian FaVARs with an Application to Monetary Policy Shocks," Working Paper 21738, National Bureau of Economic Research.

Anderson, E., A. Inoue, and B. Rossi (forthcoming): "Heterogeneous consumers and fiscal policy shocks," *Journal of Money Credit and Banking*.

Antolin-Diaz, J., and J. F. Rubio-Ramírez (2016): "Narrative Sign Restrictions for SVARs," CEPR Discussion Papers 11517, C.E.P.R. Discussion Papers.

Arias, J. E., D. Caldara, and J. Rubio-Ramirez (2015): "The Systematic Component of Monetary Policy in SVARs: An Agnostic Identification Procedure," unpublished, Duke University.

Arias, J. E., J. F. Rubio-Ramirez, and D. F. Waggoner (2014): "Inference Based on SVARs Identified with Sign and Zero Restrictions: Theory and Applications," Dynare Working Papers 30, CEPREMAP.

Barro, R. J., and C. J. Redlick (2011): "Macroeconomic Effects From Government Purchases and Taxes," *The Quarterly Journal of Economics*, 126(1), 51–102.

Barsky, R. B., and E. R. Sims (2011): "News shocks and business cycles," *Journal of Monetary Economics*, 58(3), 273 – 289.

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

Beaudry, P., D. Nam, and J. Wang (2011): "Do Mood Swings Drive Business Cycles and is it Rational?," NBER Working Papers 17651, National Bureau of Economic Research.

Beaudry, P., and F. Portier (2006): "Stock Prices, News, and Economic Fluctuations," *American Economic Review*, 96(4), 1293–1307.

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

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

Blanchard, O. J., and D. Quah (1989): "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, 79(4), 655–73.

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

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

Christiano, L. J., M. Eichenbaum, and C. L. Evans (1999): "Monetary policy shocks: What have we learned and to what end?," in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and M. Woodford, vol. 1 of *Handbook of Macroeconomics*, chap. 2, pp. 65–148. Elsevier.

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

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

Del Negro, M., and F. Schorfheide (2011): "Bayesian Macroeconometrics," in *The Oxford Handbook of Bayesian Econometrics*, ed. by J. Geweke, G. Koop, and H. van Dijk, pp. 293–289. Oxford University Press.

DiTraglia, F. J., and C. García-Jimeno (2016): "A Framework for Eliciting, Incorporating, and Disciplining Identification Beliefs in Linear Model," unpublished, University of Pennsylvania.

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

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

Faust, J., E. T. Swanson, and J. H. Wright (2004): "Identifying VARs based on high frequency futures data," *Journal of Monetary Economics*, 51(6), 1107 – 1131.

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

Fisher, J. D., and R. Peters (2010): "Using Stock Returns to Identify Government Spending Shocks," *Economic Journal*, 120(544), 414–436.

Gafarov, B., M. Meier, and J. Montiel Olea (2016a): "Delta-Method Inference For A Class of Set-Identified SVARs," Discussion paper, New York University.

——— (2016b): "Projection Inference For Set-Identified SVARs," Discussion paper, New York University.

Giacomini, R., and T. Kitagawa (2014): "Inference about Non-Identified SVARs," CeMMAP working papers CWP45/14, Centre for Microdata Methods and Practice, Institute for Fiscal Studies.

Inoue, A., and L. Kilian (2013): "Inference on impulse response functions in structural VAR models," *Journal of Econometrics*, 177(1), 1 – 13.

Kilian, L., and D. P. Murphy (2012): "Why Agnostic Sign Restrictions Are Not Enough: Understanding The Dynamics of Oil Market VAR Models," *Journal of the European Eco-*

nomic Association, 10(5), 1166–1188.

Kline, B., and E. Tamer (forthcoming): "Bayesian Inference In A Class Of Partially Identified Models," *Quantitative Economics*.

Kurmann, A., and E. Sims (2017): "Revisions in Utilization-Adjusted TFP and Robust Identification of News Shocks," NBER Working Papers 23142, National Bureau of Economic Research.

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

Ludvigson, S. C., S. Ma, and S. Ng (2017): "Shock Restricted Structural Vector-Autoregressions," Working Paper 23225, National Bureau of Economic Research.

Moon, H. R., and F. Schorfheide (2012): "Bayesian and Frequentist Inference in Partially Identified Models," *Econometrica*, 80(2), 755–782.

Moon, H. R., F. Schorfheide, and E. Granziera (2013): "Inference for VARs Identified with Sign Restrictions," Discussion paper, University of Southern California.

Mountford, A., and H. Uhlig (2009): "What are the effects of fiscal policy shocks?," *Journal of Applied Econometrics*, 24(6), 960–992.

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

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

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

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

——— (2016): "Macroeconomic Shocks and Their Propagation," NBER Working Papers 21978, National Bureau of Economic Research.

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

Schorfheide, F. (2016): "Macroeconometrics - A Discussion," Discussion.

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

Uhlig, H. (1994): "What Macroeconomists Should Know about Unit Roots: A Bayesian Perspective," *Econometric Theory*, 10(3/4), pp. 645–671.

——— (2003): "What moves GNP?," *draft, Humboldt Universität zu Berlin*.

——— (2004): "Do Technology Shocks Lead To a Fall in Total Hours Worked?," *Journal of the European Economic Association*, 2(2/3), 361–371.

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

A Proofs

A.1 Proof of Lemma 2

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 rewrite $M = U[D, 0_{J \times (p-J)}]V$, where U, V are orthogonal matrices of dimension J and p, respectively. D is a $J \times J$ diagonal matrix with nonzero 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 nonzero. 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 q and the identified set for a have positive measure given that π_q has full support. ☐

A.2 Proof of Corollary 1

Proof. To keep the notation simpler, assume that the variables with restrictions are ordered first in the VAR, such that $n^{(j)} = j$ for $j = 1, \ldots, J$. 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 $\{(j1, j1), \ldots, (\bar{j}, \bar{j})\}$ in E, which are unity, and (2) positions $\{(j1, m^{(j1)}), \ldots, (\bar{j}, m^{(\bar{j})})\}$ in Λ, which equal $\lambda^{(j1)}, \ldots, \lambda^{(\bar{j})}$, 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 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 $M_{j1,\circ}, \ldots, M_{\bar{j},\circ}$ 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. ☐

A.3 Proof of Proposition 1 (Bivariate VAR(0))

Proof. This follows directly from comparing the sets listed below for $\lambda = 0$ and $\lambda > 0$. 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.4 Proof of Proposition 2 (Trivariate VAR(0))

Proof. **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:

$$\begin{aligned} &\min_q \text{ or } \max_q \sqrt{1-(q_2)^2-(q_3)^3} \\ &\text{s.t. } \tilde{A}_{21}\sqrt{1-(q_2)^2-(q_3)^3} + \tilde{A}_{22}q_2 \geq 0 \\ &\underbrace{(\tilde{A}_{31} - \lambda\tilde{A}_{21})}_{\equiv\tilde{A}_{31}^\lambda}\sqrt{1-(q_2)^2-(q_3)^3} + \underbrace{(\tilde{A}_{32} - \lambda\tilde{A}_{22})}_{\equiv\tilde{A}_{32}^\lambda} q_2 + \tilde{A}_{33}q_3 \geq 0 \end{aligned}$$

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

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

with the associated Kuhn-Tucker conditions as:

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

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

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 $\lambda \geq 0$. For $\lambda = 0$, the heterogeneity restrictions become standard sign restrictions. We focus on the case of $\tilde{A}_{21}, \tilde{A}_{31} > 0$.

1. **All (conditional) covariances positive, heterogeneity restriction weak:**
   Note that when $0 \leq \tilde{A}_{21}, \tilde{A}_{31}$ and $\lambda\tilde{A}_{21} \leq \tilde{A}_{31}$, 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 $\lambda\tilde{A}_{21} > \tilde{A}_{31} > 0$, 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. The bound on q_1 is thus strictly tighter with heterogeneity restrictions.

3. **Small negative conditional covariance, weak heterogeneity restriction:**
   When $\tilde{A}_{31} < 0$, it follows that $\tilde{A}^{\lambda}_{31} < 0$ for all λ. For λ, A_{31} close enough to zero, the optimum involves $\nu_{HR} > 0 = \nu_{SR}$. In this case, the solution is given by:

   $$q_1^{HR} = \frac{(\tilde{A}^{\lambda}_{32})^2 + (\tilde{A}_{33})^2}{\sqrt{((\tilde{A}^{\lambda}_{32})^2 + (\tilde{A}_{33})^2)^2 + \tilde{A}_{33}(\tilde{A}^{\lambda}_{31})^2(1 + (\tilde{A}^{\lambda}_{32}/\tilde{A}_{33})^2)}}$$

   It can be shown that $\left.\frac{dq_1^{HR}}{d\lambda}\right|_{\lambda=0} < 0$: Introducing heterogeneity restrictions tightens the upper bound.

   More generally, both restrictions or only the second restriction can bind if the optimum involves $q_2 < 0$. The solution for q_1 is $q_1 = \max\{q_1^{HR}, q_1^{HR,SR}, q_1^{SR}\}$ where:

   $$
   \begin{aligned}
   q_1^{HR,SR} &= \frac{\tilde{A}_{33}}{\sqrt{((\tilde{A}_{31} - (\tilde{A}_{21}/\tilde{A}_{22})\tilde{A}_{32})^2 + (\tilde{A}_{33})^2)^2 + (\tilde{A}_{33})^2(1 + (\tilde{A}_{21}/\tilde{A}_{22})^2)}} \\
   q_1^{SR} &= \frac{1}{\sqrt{1 + (\tilde{A}_{21}/\tilde{A}_{22})^2}}.
   \end{aligned}
   $$

□

A.5 Proof of Proposition 3

Proof. Suppose without loss of generality that there are no zero rows of W: $\forall i = 1, \ldots, n_r$: $||W_{i,\circ}|| > 0$.

$\Rightarrow$: Let $y_{1,2} = x_c \pm \alpha\sqrt{\frac{r}{n}}\mathbf{1}$, where $\alpha \in [0,1]$. Let $q_i = \frac{y_i}{||y_i||}$. Then

$$q_1 - q_2 = ||y_1||^{-1}\left[(x_c^{(i)} - \alpha\sqrt{r/n})(1 - ||y_1||/||y_2||)\right]_i \neq 0,$$

for (almost) all α. Thus the identified set has a measure at least as large as:

$$\pi\left\{\frac{y}{||y||}|y = x_c + \alpha\sqrt{\frac{r}{n}}\mathbf{1}, -1 \leq \alpha \leq 1\right\}.$$

$\Leftarrow$: We first show that the identified set cannot have positive measure if it imposes an equality restriction on any element of q. Then we we use this to construct a candidate solution to the problem of finding the Chebychev center.

Note that if the identified set has positive measure, then there exists a q such that $Wq \leq 0$ and $||q|| = 1$.

By means of contradiction, assume $\exists i \in \{1, \dots, n_r\}$ such that $\forall q$ satisfying $||q|| = 1$, $Wq \leq 0$, and $W_{i,\circ}q = 0$. Let j denote a non-zero entry W_{ij}. Then $q_j = -\frac{1}{W_{ij}}\sum_{\ell\neq j} W_{i\ell}q_l \forall q$. However, such a q has zero π measure for continuous π, contradicting the assumption that the identified set has positive measure. Thus, for each i there exists a $\tilde{q}_i$ such that $W_{i,\circ}\tilde{q}_i < 0$ and $Wq_i \leq 0$. By continuity, we can also find a nearby q_i such that $Wq_i < 0$.

Now we construct, by induction, a q such that $Wq < 0$. Pick q such that $W_1q < 0$ and $W_2q \leq 0$. Let $\epsilon = \delta \times (-W_1q)$ and define $\hat{q} = [q_\ell - \delta n^{-1}\epsilon \,\mathrm{sgn}(W_{2,\ell})W_{2,\ell}]_\ell$ and $q' = \frac{\hat{q}}{||\hat{q}||}$. Note that $W_2q' \propto W_2\hat{q} = W_2q - \delta n^{-1}\epsilon\sum_\ell |W_{2,\ell}| < 0$. Also, for δ small enough, $W_1q' < 0$. Now assume that $W_iq < 0$ for $i = 1, \dots, n$ and $W_{n+1}q = 0$. Proceed as before but with $\epsilon = \max_i \delta|W_iq|$. Going through the same argument shows that we can then also generate a q' such that $W_iq' < 0$ for all $i = 1, \dots, n+1$.

Thus, $Wq < 0$. First, this implies that $q \neq 0$. Second, by continuity, there exists an $\tilde{r} > 0$ small enough such that $\forall u$ satisfying $||u|| < \tilde{r}$ also $W(q+u) \leq 0$. Because $||q|| = 1, q \in [-1,1]^n$. Thus, there exists a feasible solution to the Chebychev problem with $r > 0$.

□

A.6 Proof of Proposition 4

Proof. Let $\tilde{x} \sim \mathcal{N}(0, I_n)$ such that $W\tilde{x} \leq 0$, where the inequality is elementwise. Let Q be a given orthonormal matrix. Thus, $\tilde{y} = Q\tilde{x} \sim \mathcal{N}(0, QQ') = \mathcal{N}(0, I_n)$. Let $\mathcal{A}$ be a Borel set on the σ-algebra on the unit n-sphere. Then:

$$\begin{aligned}
\frac{\Pr_{\tilde{x}}\{x \in \mathcal{A} \cap \{x'|Wx' \leq 0\}\}}{\Pr_{\tilde{x}}\{Wx \leq 0\}} &= \frac{\Pr_{\tilde{x}}\{x \in \mathcal{A} \cap \{x'|Wx'/||x'|| \leq 0\}\}}{\Pr_{\tilde{x}}\{Wx/||x|| \leq 0\}} \\
&= \frac{\Pr_{\tilde{y}}\{y \in \mathcal{A} \cap \{y'|Wy'/||y'|| \leq 0\}\}}{\Pr_{\tilde{y}}\{Wy/||y|| \leq 0\}} \\
&= \frac{\Pr_{Q\tilde{x}}\{Qx \in \mathcal{A} \cap \{Qx'|WQx'/||Qx'|| \leq 0\}\}}{\Pr_{Q\tilde{x}}\{WQx/||Qx|| \leq 0\}} \\
&= \frac{\Pr_{\tilde{q}}\{q \in \mathcal{A} \cap \{q'|Wq' \leq 0\}\}}{\Pr_{\tilde{q}}\{Wq \leq 0\}}.
\end{aligned}$$

The first equality follows because $||x|| > 0$ with probability one, the second equality follows because $\tilde{y} \stackrel{D}{=} \tilde{x}$, the third equality follows from substituting $\tilde{y} = Q\tilde{x}$, and the last equality follows from the definition of $\tilde{q}$. Thus, the resultant distribution has the desired rotation-invariant measure on the truncated unit n-sphere.

$\square$

B Trivariate VAR: Further analysis and empirical examples

B.1 General signs

To take Proposition 2 to the data, we need to allow for general sign restrictions. We thus generalize the setup from Appendix A.4 by introducing signs for each restriction $s_i \in \{-1, 1\}$.

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

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

$$s_1 \tilde{A}_{11} q_1 \geq 0 \tag{B.1c}$$

$$s_2(\tilde{A}_{21} q_1 + \tilde{A}_{22} q_2) \geq 0 \tag{B.1d}$$

$$s_3 \left(\underbrace{(\tilde{A}_{31} - \lambda \tilde{A}_{21})}_{\equiv \tilde{A}_{31}^{\lambda}} q_1 + \underbrace{(\tilde{A}_{32} - \lambda \tilde{A}_{22})}_{\equiv \tilde{A}_{32}^{\lambda}} q_2 + \tilde{A}_{33} q_3 \right) \geq 0 \tag{B.1e}$$

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 (B.1) as

$$\min_q \text{ or } \max_q s_1 \tilde{A}_{11} \hat{q}_1 \tag{B.2a}$$

$$\text{s.t. } ||q|| = 1 \tag{B.2b}$$

$$\tilde{A}_{11} q_1 \geq 0 \tag{B.2c}$$

$$\hat{A}_{21} \hat{q}_1 + \tilde{A}_{22} \hat{q}_2 \geq 0 \tag{B.2d}$$

$$\underbrace{(\hat{A}_{31} - \hat{\lambda} \hat{A}_{21})}_{\equiv \hat{A}_{31}^{\lambda}} \hat{q}_1 + \underbrace{(\hat{A}_{32} - \hat{\lambda} \hat{A}_{22})}_{\equiv \hat{A}_{32}^{\lambda}} \hat{q}_2 + \tilde{A}_{33} \hat{q}_3 \geq 0, \tag{B.2e}$$

whose constraints are of the same form as (B.1). 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 from Appendix A.4 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. $s_1 = s_3 = -1$. $s_2 = 1$: As the FFR rises, inflation and real activity fall.

(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.

Cholesky 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.

Cholesky 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 B.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. $s_1 = s_3 = -1$. $s_2 = 1$: As the FFR rises, inflation and real activity fall.

   (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.

   Cholesky 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 B.2.

If we replace the median house price with the Case-Shiller index, we find that the following Cholesky 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 B.2 displays the results and shows that the set reduction happens but is negligible. In both cases, 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: $s_2 = +1$.

   (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 $\lambda\tilde{A}_{21} > \tilde{A}_{31} > 0$.

   Cholesky 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 B.3 shows the corresponding set reduction. Note the nonlinear scale of λ that shows that for small λ, there is no set reduction, confirming our theoretical analysis.

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

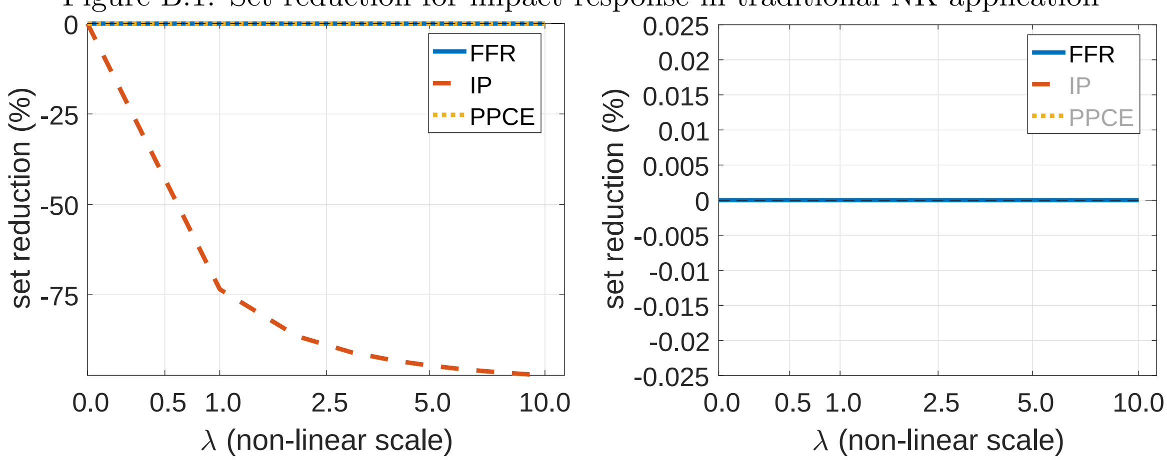


Figure B.2: Set reduction for impact response in NK housing application

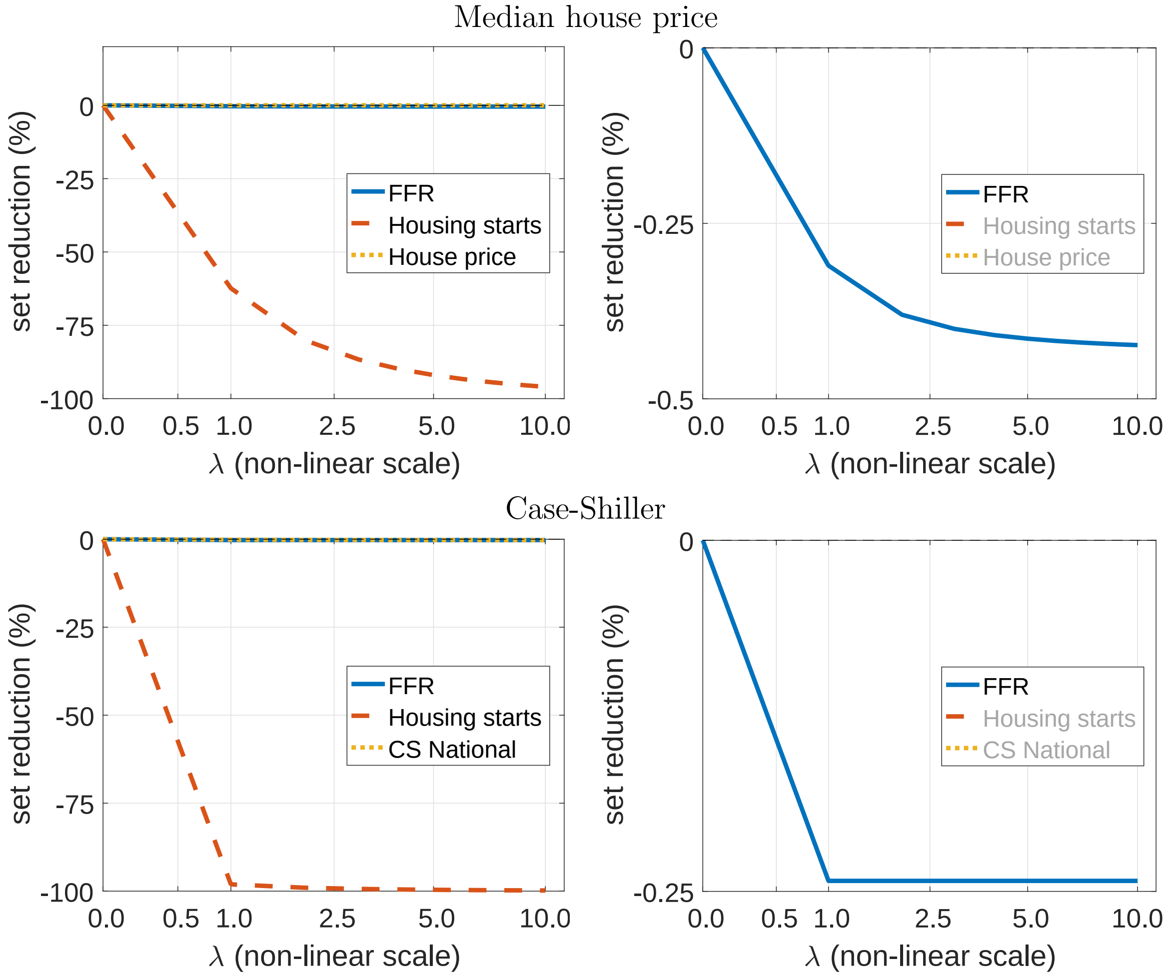
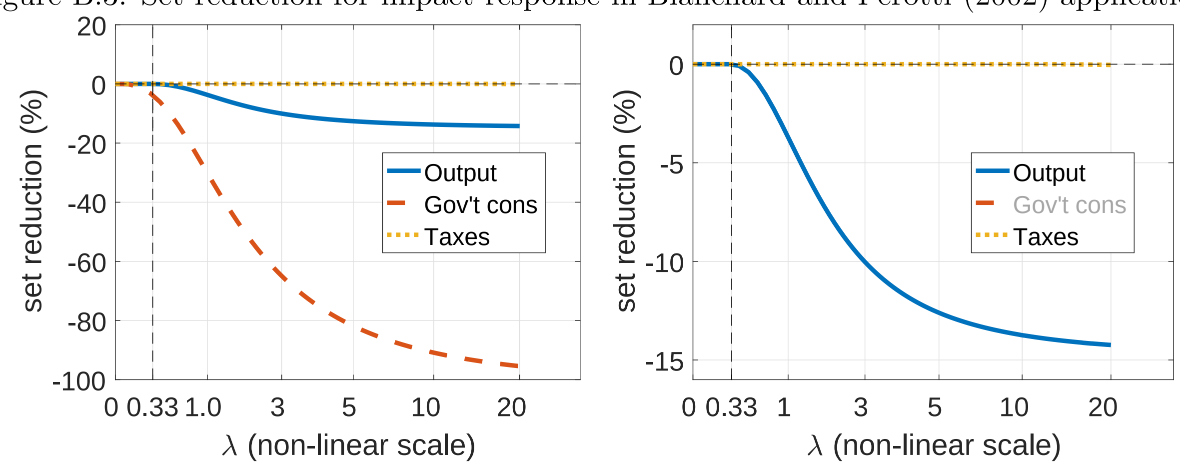


Figure B.3: Set reduction for impact response in Blanchard and Perotti (2002) application



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

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 B.4 shows the corresponding set reduction. Note the nonlinear scale of λ that shows again that for small λ, there is no set reduction.

B.2 Redundant restrictions

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

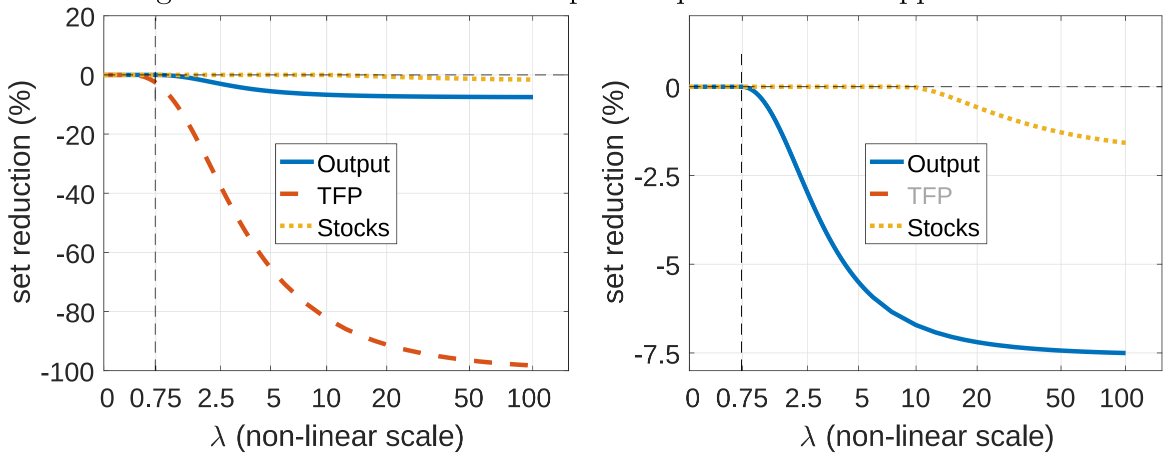
$$A = \begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \Rightarrow AA' = \begin{bmatrix} a_{11}^2 + a_{12}^2 & a_{11}a_{21} + a_{22}a_{12} & a_{11}a_{31} + a_{12}a_{32} \\ a_{21}a_{11} + a_{22}a_{12} & a_{21}^2 + a_{22}^2 & a_{21}a_{31} + a_{22}a_{32} \\ a_{31}a_{11} + a_{12}a_{32} & a_{31}a_{21} + a_{32}a_{22} & a_{31}^2 + a_{33}^2 + a_{32}^2 \end{bmatrix} \tag{B.3}$$

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.

The lower-triangular Cholesky decomposition is given by:

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

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



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

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

$$\tilde{A}_{31} = \kappa \tilde{A}_{21}, \tag{B.4a}$$

$$\tilde{A}_{32} = \kappa \tilde{A}_{22}, \tag{B.4b}$$

$$\tilde{A}_{33} = |a_{33}|. \tag{B.4c}$$

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:[32]

$$\max_q e_i' \tilde{A} q, \qquad i \in \{1, 2\}, \tag{B.5a}$$

$$\text{s.t. } ||q|| = 1 \tag{B.5b}$$

$$e_1' \tilde{A} q \geq 0 \tag{B.5c}$$

$$e_2' \tilde{A} q \geq 0 \tag{B.5d}$$

$$(e_3 - \lambda e_2)' \tilde{A} q \geq 0, \tag{B.5e}$$

where $\tilde{A}$ is the Cholesky factor of AA' in (B.3) that satisfies (B.4). 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:

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

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

The necessary conditions are:

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

$$\nu_j e_j' \tilde{A} q = 0 \qquad\qquad \nu_j \geq 0 \qquad , j = 1, 2$$

$$\nu_3 (e_3 - \lambda e_2)' \tilde{A} q = 0 \qquad\qquad \nu_3 \geq 0.$$

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. This is, in turn, because responses to both shocks satisfy the heterogeneity restrictions in the data-generating process. They are, thus, redundant.

C Forecast error variance decomposition

The total forecast error variance (FEV) 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 FEV 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 the cumulative conditional forecast error variance from horizon $\underline{H}$ to $\bar{H}$, $CFEV_{i,\underline{H},\bar{H}}(q)$, as:

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

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

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 interpret the FEV explained by the identified shock, we normalize $CFEV_{i,H}(q)$ by the total FEV for variable i up to horizon H.

D Data

D.1 News data

We use the following macro variables:

- The average of business sector GDP and GDI: BEA via Fernald (2014) (accumulated growth rates)
- Consumer confidence `CSCICP03USM665S` from the St. Louis Fed FRED website
- PCE price index `PCEPI` from the St. Louis Fed FRED website
- Utilization adjusted TFP: Fernald (2014) (accumulated growth rates)
- Business sector hours worked: BLS via Fernald (2014) (accumulated growth rates)

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 percentiles, 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.

D.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[33] that we then deflate by the CPI from Ramey (2011). All variables enter the VAR in log-levels relative to population.

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

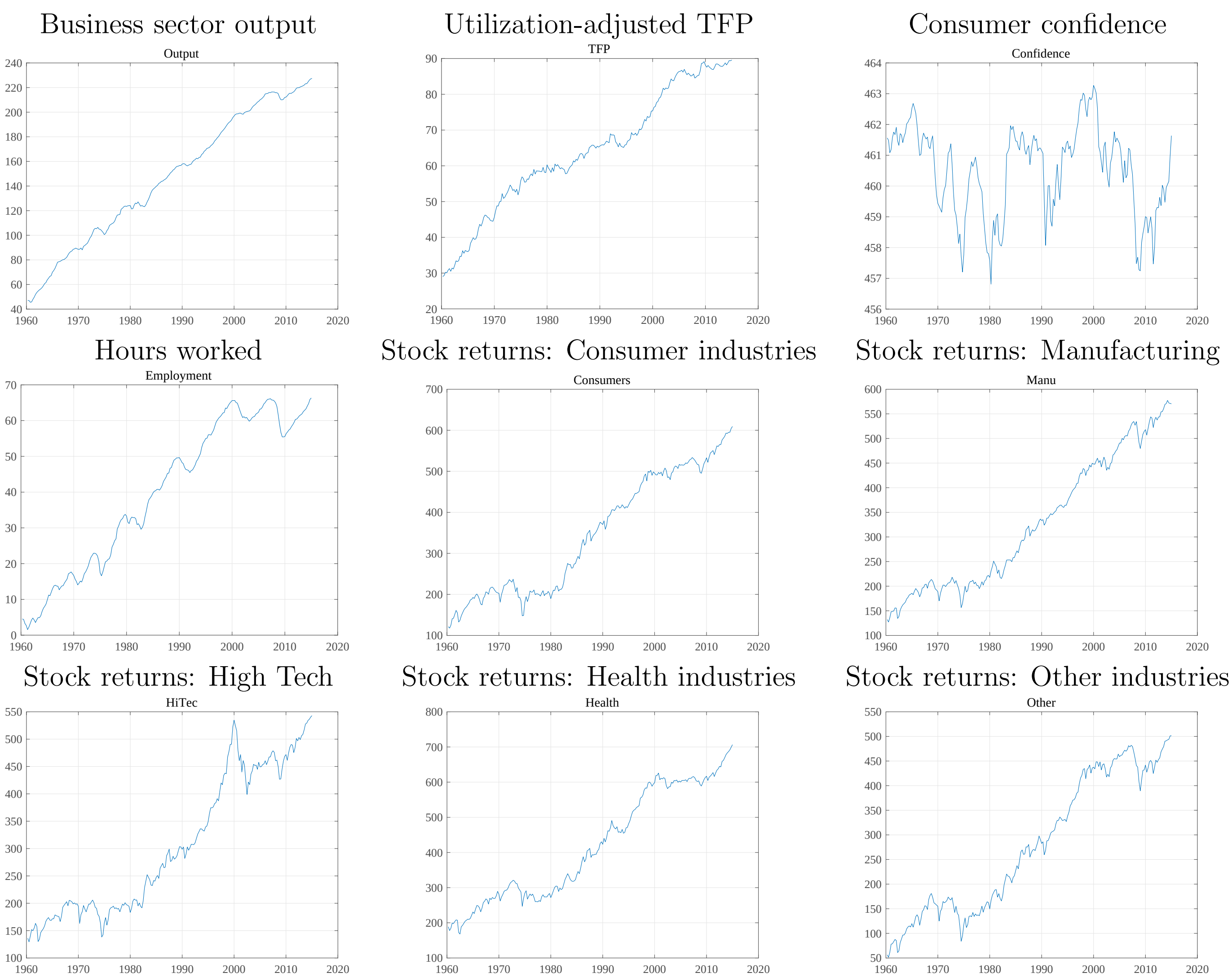


Figure D.1: Raw data: News application with five Fama-French industries

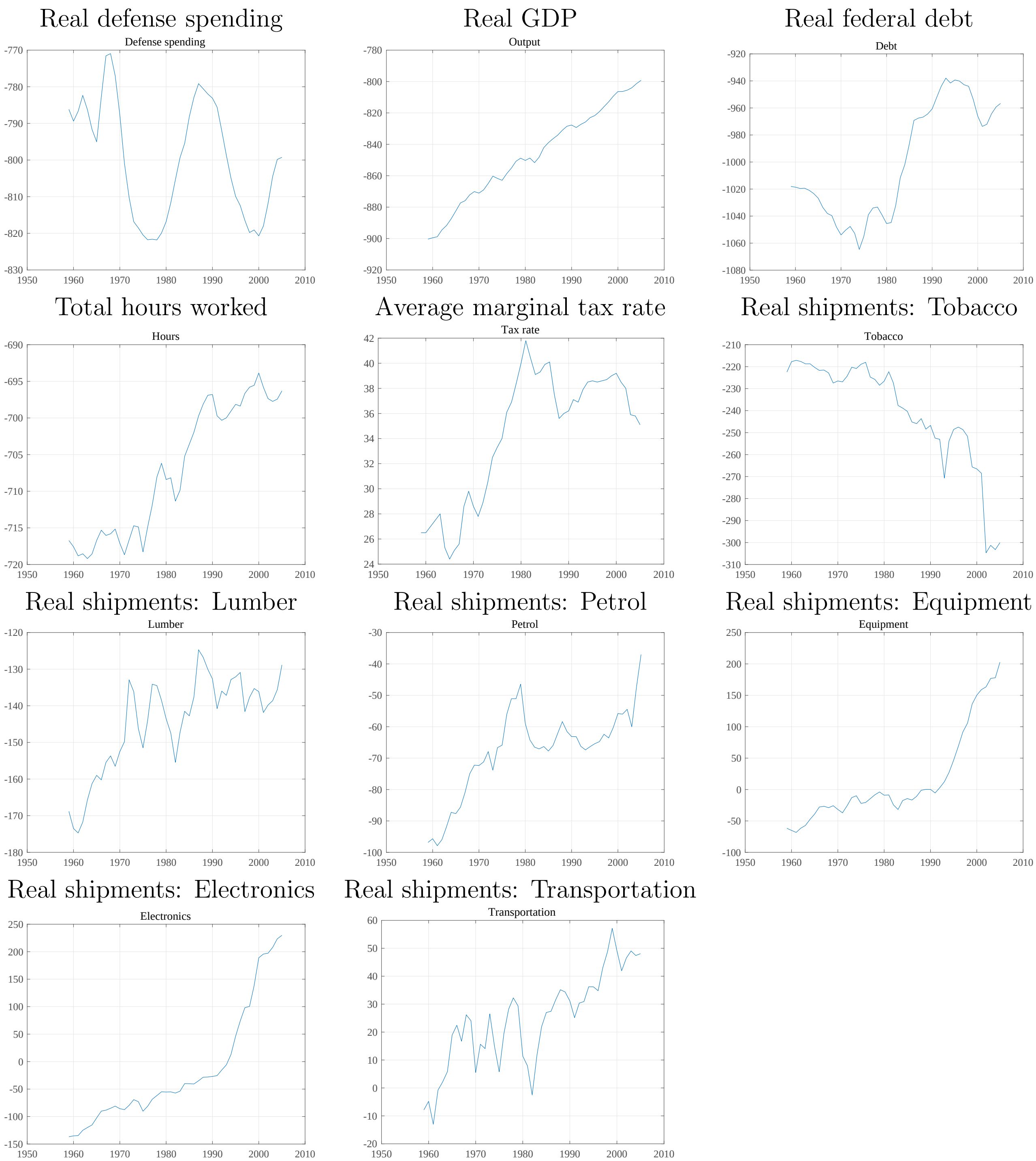


Figure D.2: Raw data: Defense spending application with six industries

E Additional results

E.1 News shocks

E.1.1 IRFs

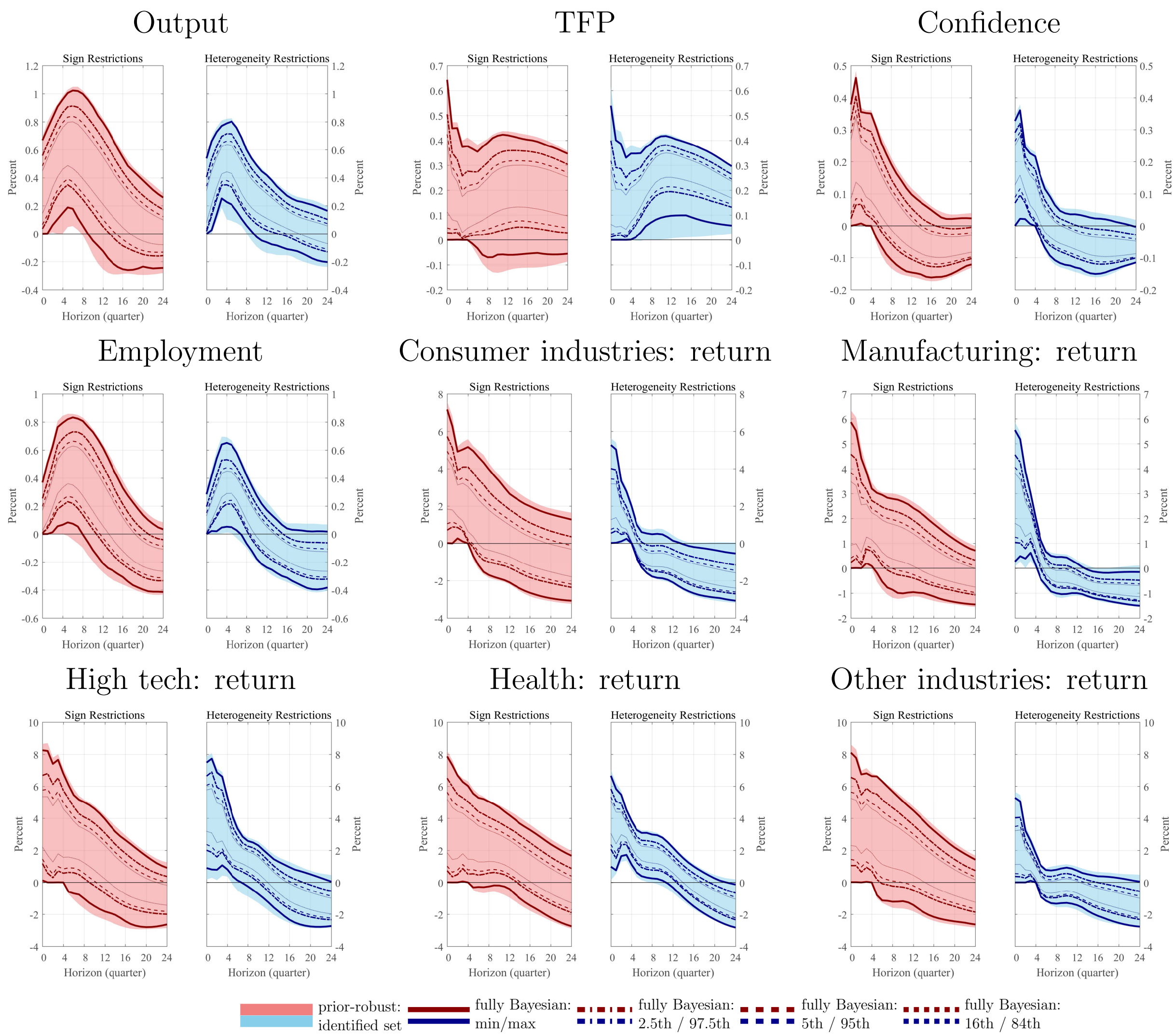


Figure E.1: Identified set of responses of all variables to productivity news shock: Baseline

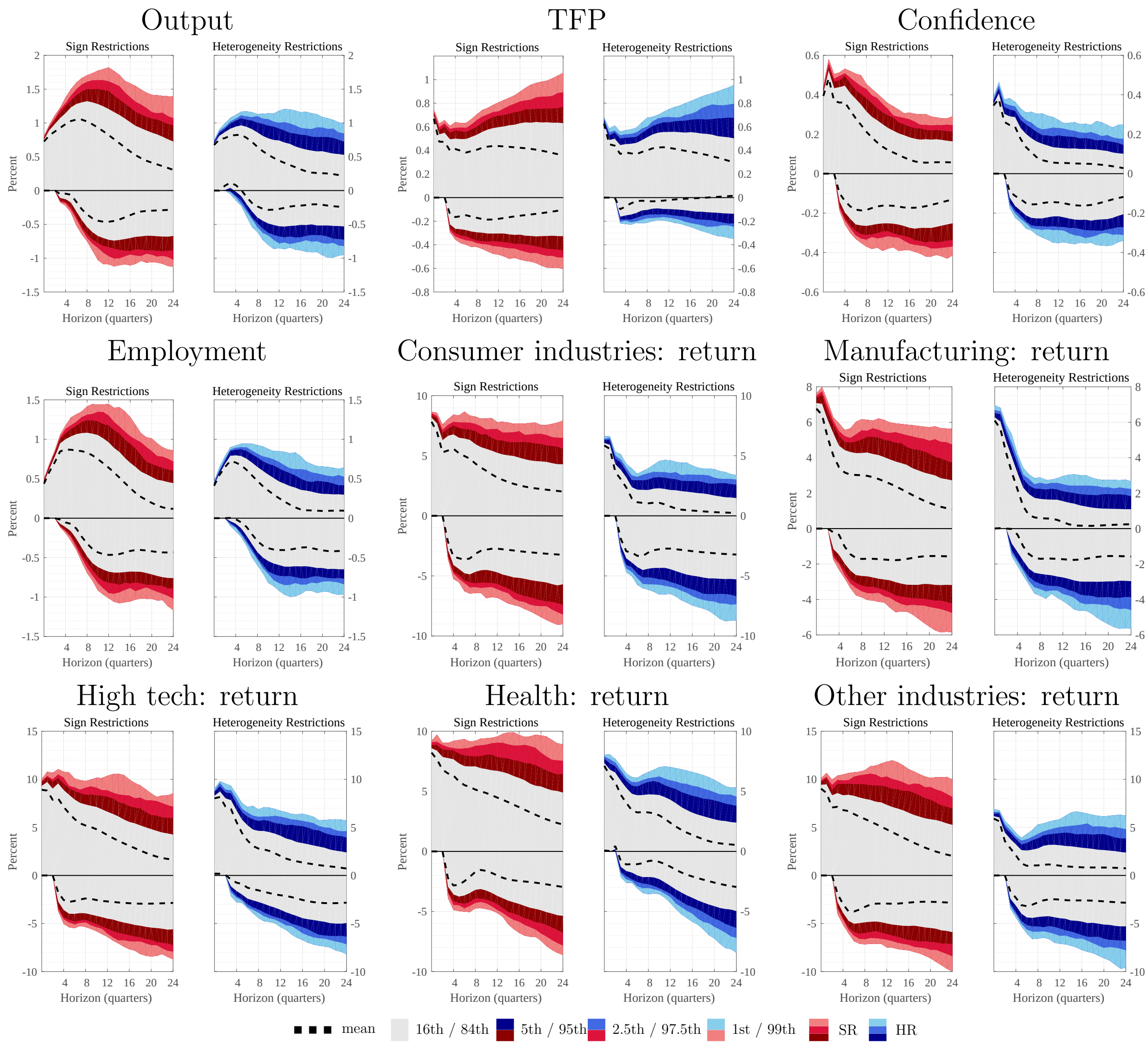


Figure E.2: Prior-robust responses of all variables to productivity news shock: Baseline

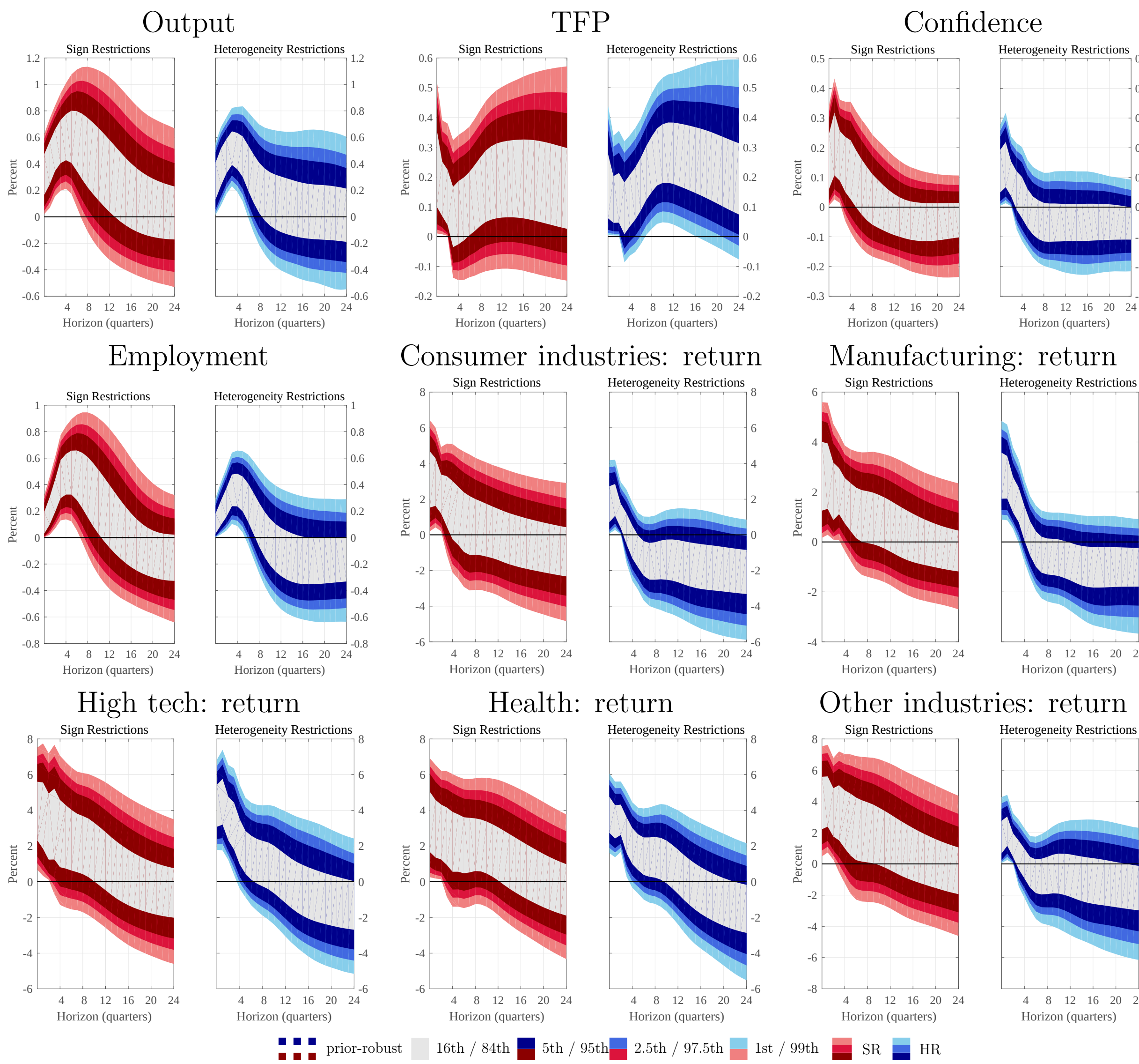


Figure E.3: Fully Bayesian responses of all variables to productivity news shock

(a) Prior-robust

Variable	Horizon H (quarters) 0	1	2	3	4	8	12	16	20	24
Output	8.3	7.1	10.5	12.3	18.7	33.6	35.2	26.9	23.2	28.6
TFP	9.0	6.7	7.0	8.0	10.8	4.1	7.1	9.0	9.9	9.6
Confidence	12.1	20.1	25.5	25.4	31.5	33.2	23.7	19.1	10.2	14.5
Employment	6.3	5.5	8.4	15.2	21.7	34.6	41.4	38.8	35.2	25.2
Consumers	23.6	22.8	29.9	36.4	48.8	52.1	40.8	45.5	48.7	57.3
Manu	8.8	14.6	17.9	15.3	29.1	53.2	51.7	48.5	52.0	52.6
HiTec	9.9	9.2	10.6	15.8	21.7	30.1	35.2	35.7	32.5	32.9
Health	12.6	17.0	15.0	21.1	29.5	28.2	26.6	31.9	38.3	40.3
Other	31.7	36.4	43.0	50.7	57.8	51.6	50.6	40.8	39.4	37.5

(b) Fully Bayesian

Variable	Horizon H (quarters) 0	1	2	3	4	8	12	16	20	24
Output	10.5	8.5	10.6	11.8	17.8	38.4	38.8	26.2	12.5	9.6
TFP	16.2	13.1	6.4	1.4	0.2	-13.0	-9.1	-6.6	-5.4	-4.1
Confidence	19.1	26.7	34.4	39.3	43.0	52.7	29.2	6.4	6.2	13.0
Employment	6.9	8.8	13.2	17.1	21.9	45.1	55.2	48.0	28.5	9.7
Consumers	35.0	30.0	39.9	50.5	64.3	75.0	61.1	58.9	64.8	70.7
Manu	13.4	15.7	19.8	23.6	35.4	66.8	67.3	65.2	60.8	61.1
HiTec	9.0	4.7	9.0	17.3	23.2	29.8	29.2	27.1	27.9	30.7
Health	12.5	14.4	9.8	17.7	26.1	26.5	29.9	36.5	39.8	42.8
Other	43.1	42.0	52.1	60.6	67.4	68.7	56.5	50.1	47.8	47.1

Note: 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 E.1: Reduction of 99th percentile of IRF relative to sign restrictions

E.1.2 FEVD

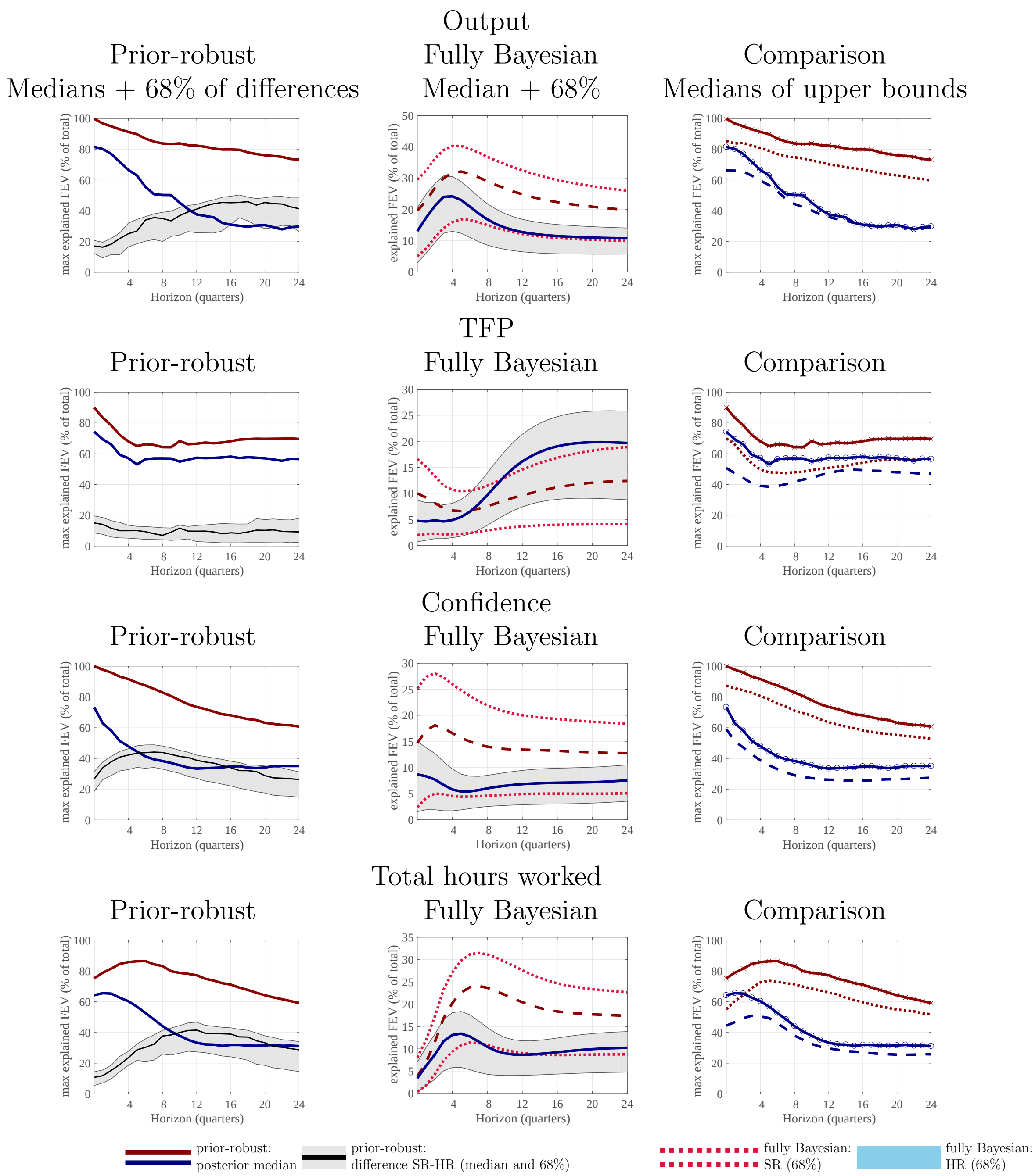


Figure E.4: FEVD of news shock: Macro variables

E.1.3 Importance of restrictions

Sign restrictions

Variable	Horizon H (year) 0	4	8	12	16	24
Output	100 (97,100)	91 (89,92)	84 (79,87)	82 (72,83)	80 (69,82)	73 (57,78)
TFP	90 (85,92)	68 (63,71)	64 (57,70)	67 (57,73)	68 (59,72)	70 (60,71)
Confidence	100 (100,100)	92 (88,93)	83 (77,86)	74 (64,78)	68 (57,72)	61 (54,66)
Employment	75 (65,78)	86 (81,89)	83 (77,85)	77 (69,80)	71 (62,75)	59 (51,65)

Heterogeneity restrictions

Variable	Horizon H (year) 0	4	8	12	16	24
Output	82 (70,84)	67 (55,70)	50 (39,54)	38 (31,46)	31 (24,37)	30 (21,33)
TFP	74 (63,77)	57 (48,62)	57 (46,59)	57 (45,60)	58 (47,60)	57 (45,62)
Confidence	73 (64,77)	48 (40,55)	38 (29,42)	33 (26,37)	35 (26,36)	35 (26,39)
Employment	64 (50,67)	60 (52,65)	44 (36,48)	33 (27,40)	32 (25,36)	31 (24,35)

Difference between SR and HR

Variable	Horizon H (year) 0	4	8	12	16	24
Output	17 (12,21)	25 (16,32)	35 (20,39)	41 (26,44)	45 (31,49)	41 (26,48)
TFP	15 (8,20)	10 (5,13)	7 (4,12)	10 (3,13)	9 (2,14)	9 (2,18)
Confidence	27 (19,31)	42 (33,46)	44 (33,48)	39 (27,42)	34 (22,38)	26 (15,31)
Employment	11 (6,14)	24 (19,28)	38 (26,41)	42 (27,47)	39 (24,43)	29 (15,34)

Note: Each cell of the table for a given variable and horizon shows the following three percentiles across reduced-form draws respectively: Median and 68% credible set ($16^{th}, 84^{th}$).

Table E.2: FEVD of news shocks: Prior-robust bounds

(a) Sign restrictions

Variable	Horizon H (year) 0	4	8	12	16	24
Output	(4,26)	(16,39)	(19,41)	(16,38)	(15,34)	(13,30)
TFP	(2,20)	(3,14)	(4,15)	(6,18)	(7,21)	(7,23)
Confidence	(5,32)	(8,32)	(7,28)	(7,25)	(7,23)	(7,23)
Employment	(0,6)	(8,24)	(12,32)	(10,29)	(10,26)	(10,24)

(b) Heterogeneity restrictions

Variable	Horizon H (year) 0	4	8	12	16	24
Output	(1,12)	(14,28)	(11,24)	(8,19)	(7,17)	(7,15)
TFP	(1,9)	(2,9)	(5,14)	(9,22)	(11,26)	(11,27)
Confidence	(6,26)	(5,18)	(4,13)	(4,13)	(4,13)	(5,13)
Employment	(0,3)	(6,16)	(5,15)	(4,12)	(4,11)	(5,13)

(c) Difference between SR and HR 16th percentile

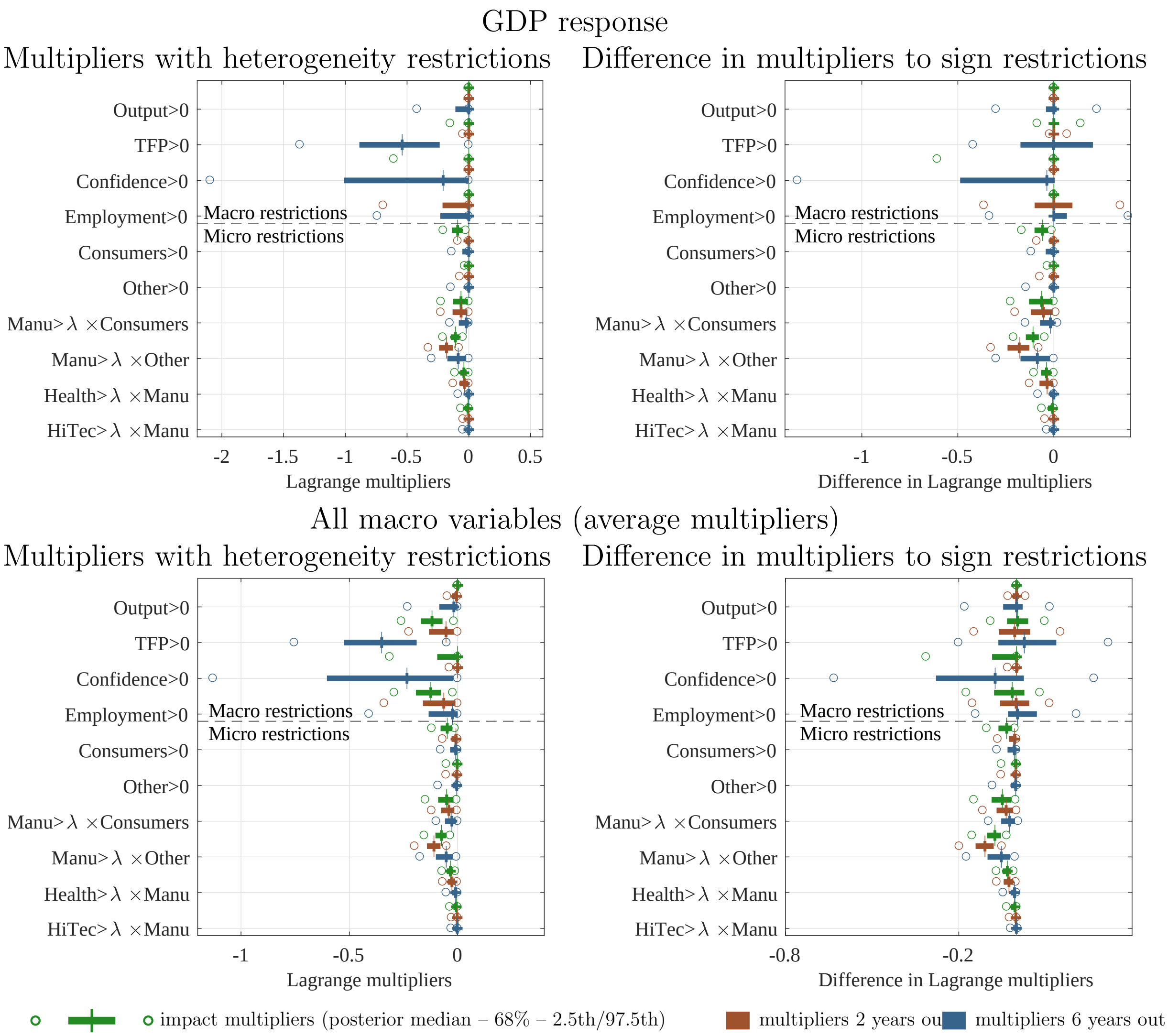
Variable	Horizon H (year) 0	4	8	12	16	24
Output	2 (1,3)	1 (-2,3)	6 (2,8)	7 (3,8)	6 (3,7)	5 (2,6)
TFP	1 (0,2)	1 (-0,1)	-1 (-5,-0)	-5 (-10,-3)	-6 (-13,-4)	-6 (-16,-3)
Confidence	-3 (-9,-1)	2 (-1,4)	2 (-1,3)	2 (-2,3)	2 (-2,3)	1 (-2,3)
Employment	0 (-0,0)	1 (-1,3)	5 (2,7)	5 (2,7)	4 (1,6)	3 (-0,5)

(d) Difference between SR and HR 68th percentile

Variable	Horizon H (year) 0	4	8	12	16	24
Output	13 (8,16)	12 (6,14)	18 (12,21)	19 (13,21)	18 (12,20)	15 (9,17)
TFP	11 (6,12)	5 (2,6)	0 (-3,3)	-4 (-10,-1)	-5 (-12,-1)	-3 (-14,2)
Confidence	7 (2,11)	15 (10,18)	14 (9,17)	12 (7,14)	11 (6,13)	10 (4,12)
Employment	2 (0,3)	9 (3,11)	17 (9,19)	17 (11,21)	14 (8,18)	12 (5,15)

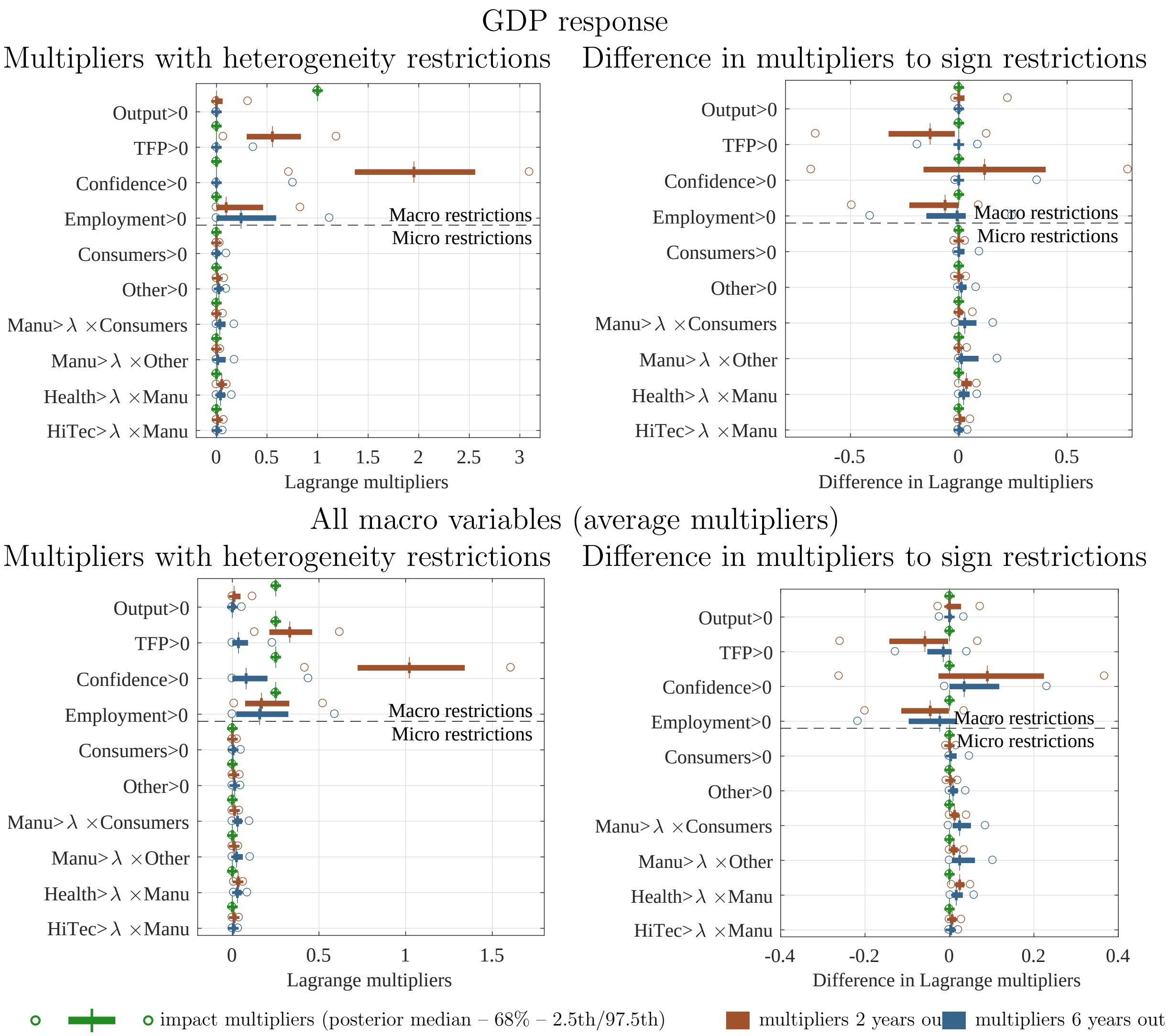
Note: Each cell of the table for a given variable and horizon shows the full posterior 68% credible set ($16^{th}, 84^{th}$) for panels (a) and (b) and the posterior median and 68% credible set across reduced-form draws for panels (c) and (d).

Table E.3: FEVD of news shocks: Fully Bayesian



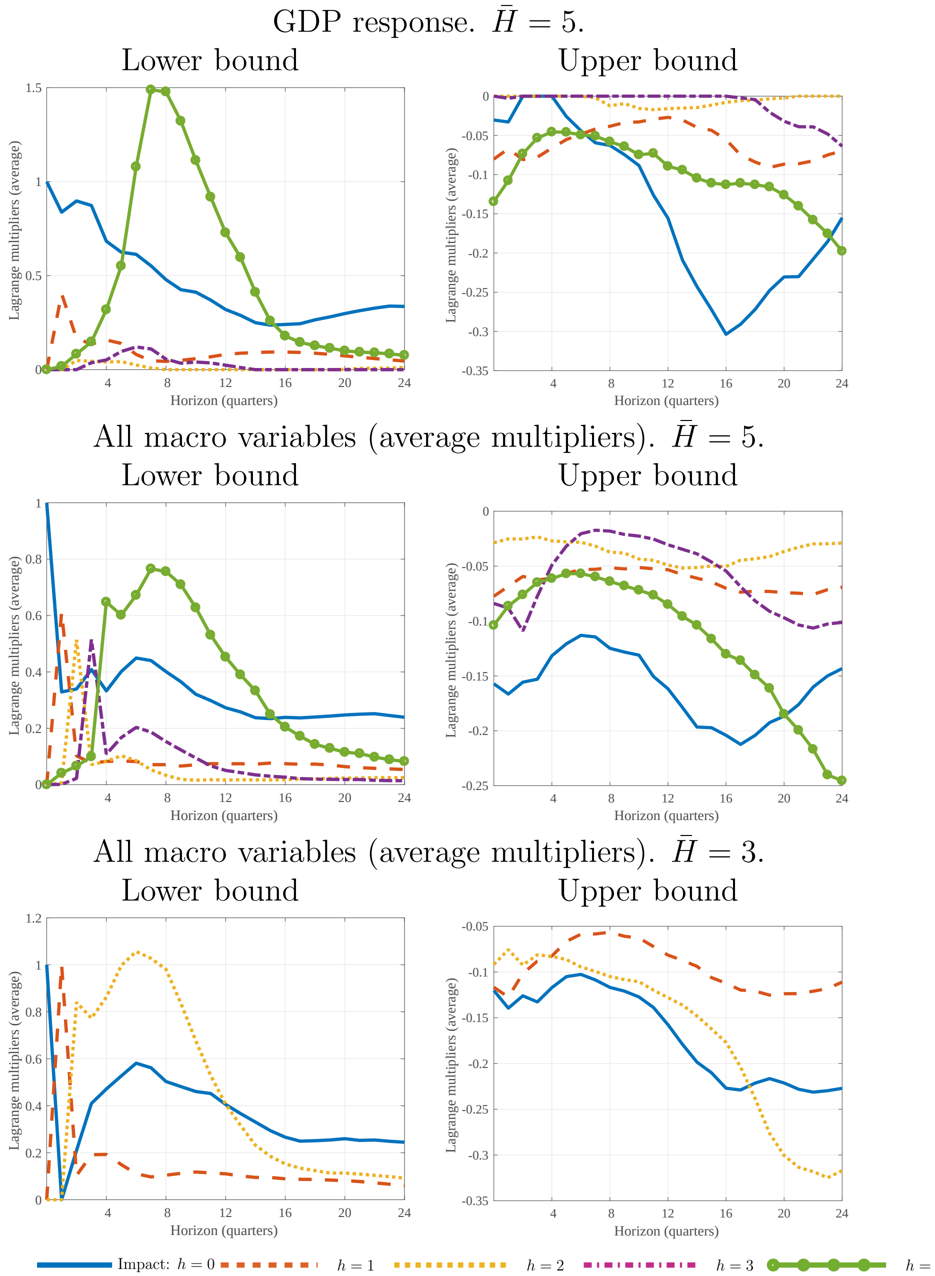
This figure quantifies the importance of all sign and heterogeneity restrictions that we use to identify the news shock by showing the distribution of Lagrange multipliers on these restrictions over all reduced-form draws. We sum the multipliers across the restrictions horizons $0, \ldots, \bar{H}$. Multipliers are negative because tighter restrictions reduce the upper bound. We find that for pinning down the upper bound on both the GDP response and for the macro variables as a whole, the heterogeneity restrictions on stock return micro data matter little as much or more as the macro restrictions. Restrictions on manufacturing are particularly important. Comparing multipliers under heterogeneity restrictions ($\lambda = 1$) with those with sign restrictions ($\lambda = 0$) reveals that the large multipliers on the micro heterogeneity restrictions rarely reduce multipliers on macro restrictions. The distribution of multipliers is asymmetric: For a few reduced-form draws, the restriction that TFP rise is very important on impact, but it is small for the median reduced-form draw.

Figure E.5: Importance of restrictions for upper bounds of responses: Lagrange multipliers on restrictions. Upper bound.



This figure quantifies the importance of all sign and heterogeneity restrictions that we use to identify the news shock by showing the distribution of Lagrange multipliers on these restrictions over all reduced-form draws. We sum the multipliers across the restrictions horizons $0, \ldots, \bar{H}$. We find that for pinning down the lower bound on both the GDP response and for the macro variables as a whole, the restrictions on stock return micro data matter little – the corresponding multipliers are close to zero. Intuitively, the Lagrange multiplier on the non-negativity restriction on GDP on impact is unity: Tightening the non-negativity constraint to a positive value would raise the impact-response one to one. On impact, this restriction on GDP is the only binding restriction. At longer horizons, the restrictions on initial confidence are more important. The importance of these restrictions changes little with and without heterogeneity restrictions.

Figure E.6: Importance of restrictions for upper bounds of responses: Box plot of Lagrange multipliers on restrictions. Lower bound.



The Lagrange multipliers show the importance of all the restrictions at different restriction horizons. Because all Lagrange multipliers have the same sign or are zero, we sum them across all restrictions. Because all restrictions force responses to be positive, the impact multiplier for lower bounds for each of the own restrictions on macro variables is unity. When restrictions are imposed for three or five quarters, i.e., at quarters $h = 0, \ldots, \bar{H} - 1$, we find that the impact restriction and the restriction at $h = \bar{H} - 1$ are the most important, measured by the absolute size of the multipliers. This is natural with largely smooth responses – the restrictions at after impact and before $\bar{H} - 1$ mostly lie between those at the extremes and make the restrictions at the extremes the most important.

Figure E.7: Importance of restrictions for upper bounds of responses: Lagrange multipliers on restrictions by restriction horizon plotted over IRF horizon.

E.1.4 Robustness

(a) Difference between sign and baseline heterogeneity restrictions

Variable	Horizon H (quarters) 0	4	8	12	16	24
Output	34 (23,40)	32 (20,38)	39 (25,43)	46 (28,49)	47 (27,51)	46 (22,51)
TFP	26 (16,31)	17 (11,21)	15 (10,19)	9 (4,13)	11 (5,18)	15 (1,22)
Confidence	30 (22,36)	43 (35,49)	47 (36,53)	41 (28,47)	36 (24,42)	28 (17,35)
Employment	22 (12,26)	26 (14,31)	40 (25,46)	47 (33,49)	42 (26,50)	33 (20,39)

(b) Difference between sign and heterogeneity restrictions with soft zero restriction

Variable	Horizon H (quarters) 0	4	8	12	16	24
Output	46 (34,55)	41 (28,45)	46 (34,52)	51 (35,55)	52 (32,55)	47 (28,55)
TFP	89 (84,91)	59 (52,60)	46 (37,48)	35 (22,42)	28 (19,37)	28 (13,37)
Confidence	32 (23,38)	46 (37,51)	47 (36,54)	44 (30,49)	39 (26,45)	33 (19,40)
Employment	29 (17,35)	37 (23,41)	46 (32,52)	50 (36,55)	47 (34,53)	39 (26,45)

(c) Level with baseline heterogeneity restrictions

Variable	Horizon H (quarters) 0	4	8	12	16	24
Output	57 (42,63)	58 (47,64)	45 (35,50)	36 (23,40)	31 (19,36)	27 (18,32)
TFP	63 (53,69)	51 (38,55)	50 (33,53)	55 (41,59)	57 (43,60)	55 (39,61)
Confidence	70 (59,73)	46 (37,51)	35 (27,39)	31 (24,35)	30 (22,34)	32 (22,37)
Employment	39 (22,45)	47 (36,53)	36 (22,43)	28 (16,34)	26 (18,30)	25 (17,29)

(d) Level with sign and heterogeneity restrictions with soft zero restriction

Variable	Horizon H (quarters) 0	4	8	12	16	24
Output	44 (27,51)	49 (38,54)	36 (27,42)	30 (19,35)	24 (16,31)	20 (14,24)
TFP	0 (0,1)	8 (6,10)	17 (8,20)	30 (22,35)	37 (28,42)	41 (28,49)
Confidence	67 (56,72)	44 (32,50)	33 (23,40)	28 (21,34)	27 (20,31)	29 (19,32)
Employment	31 (14,40)	39 (23,45)	29 (14,39)	25 (10,31)	20 (12,26)	20 (11,26)

Note: Compared with the results with heterogeneity restrictions but without the soft zero restriction (upper panel), the soft zero restriction causes a significant further reduction in the maximal FEV attributable to the news shock (lower panel). 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%. We normalize variance contributions by the total FEV. The table shows the posterior median and 68%.

Table E.4: Maximum forecast error variance explained by productivity news: Heterogeneity restrictions without and with soft zero restriction

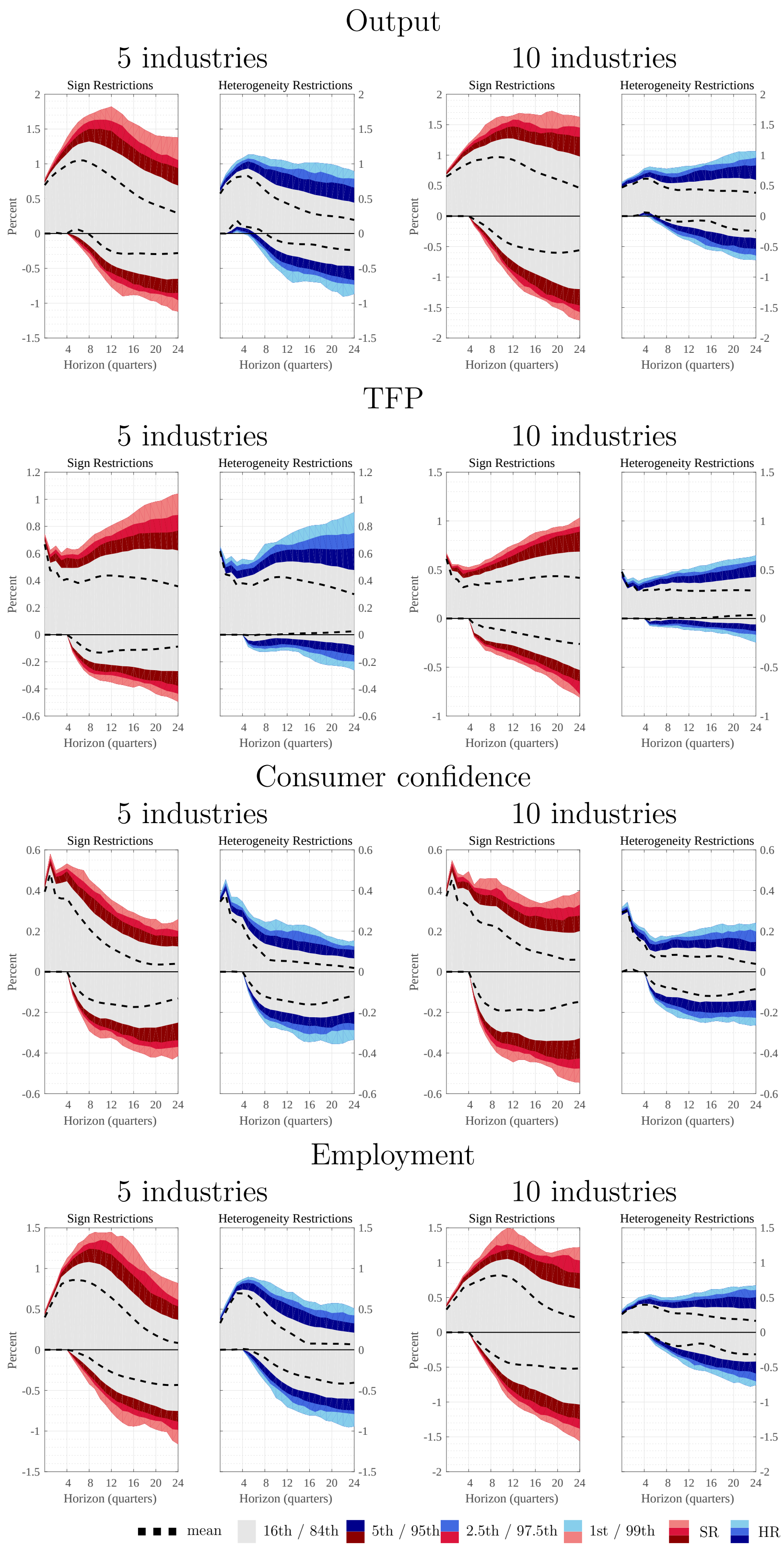


Figure E.8: Prior-robust responses to productivity news shock for macro variables: Five vs. ten FF industries

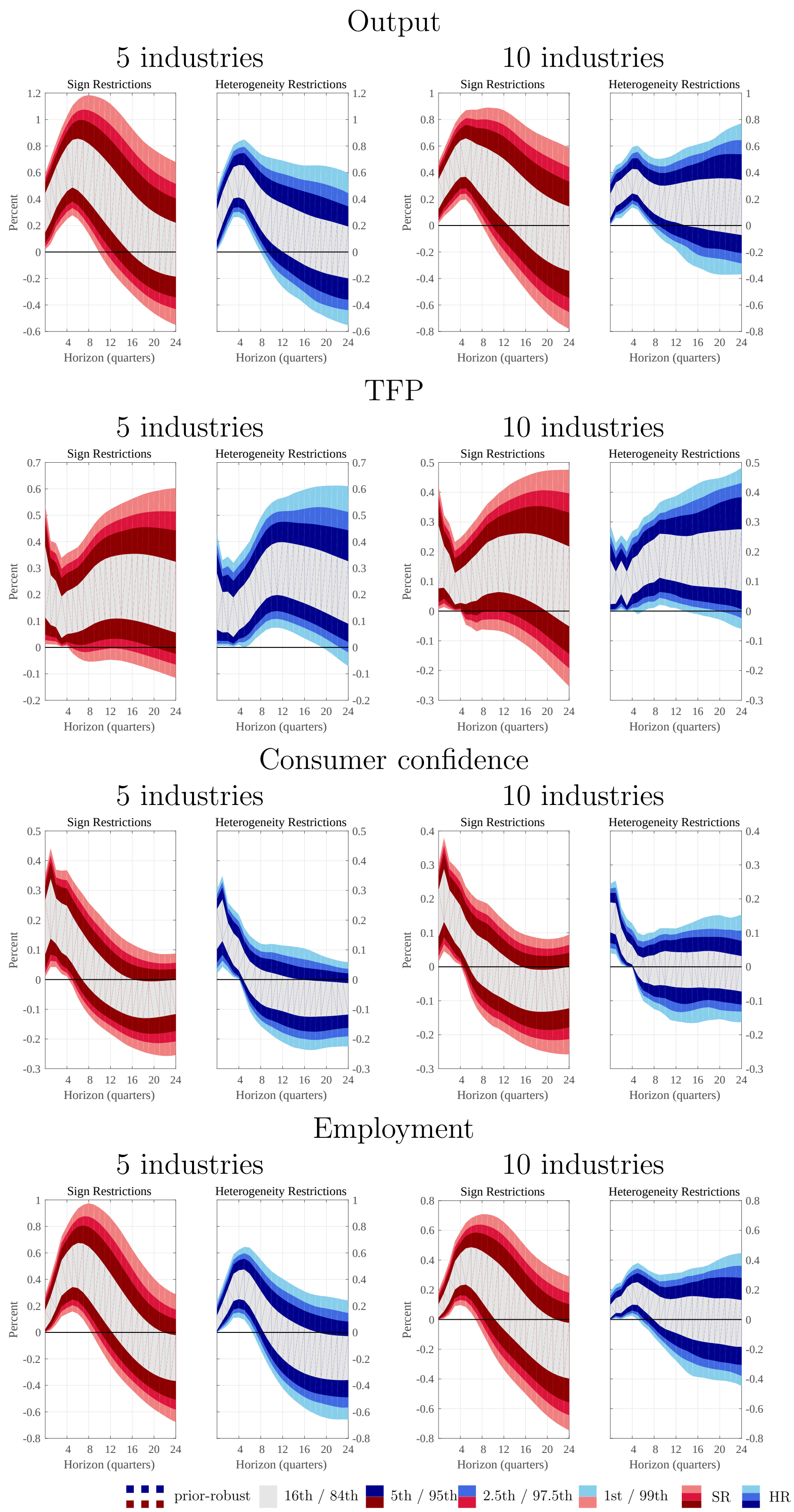


Figure E.9: Responses to productivity news shock for macro variables: Five vs. ten FF industries

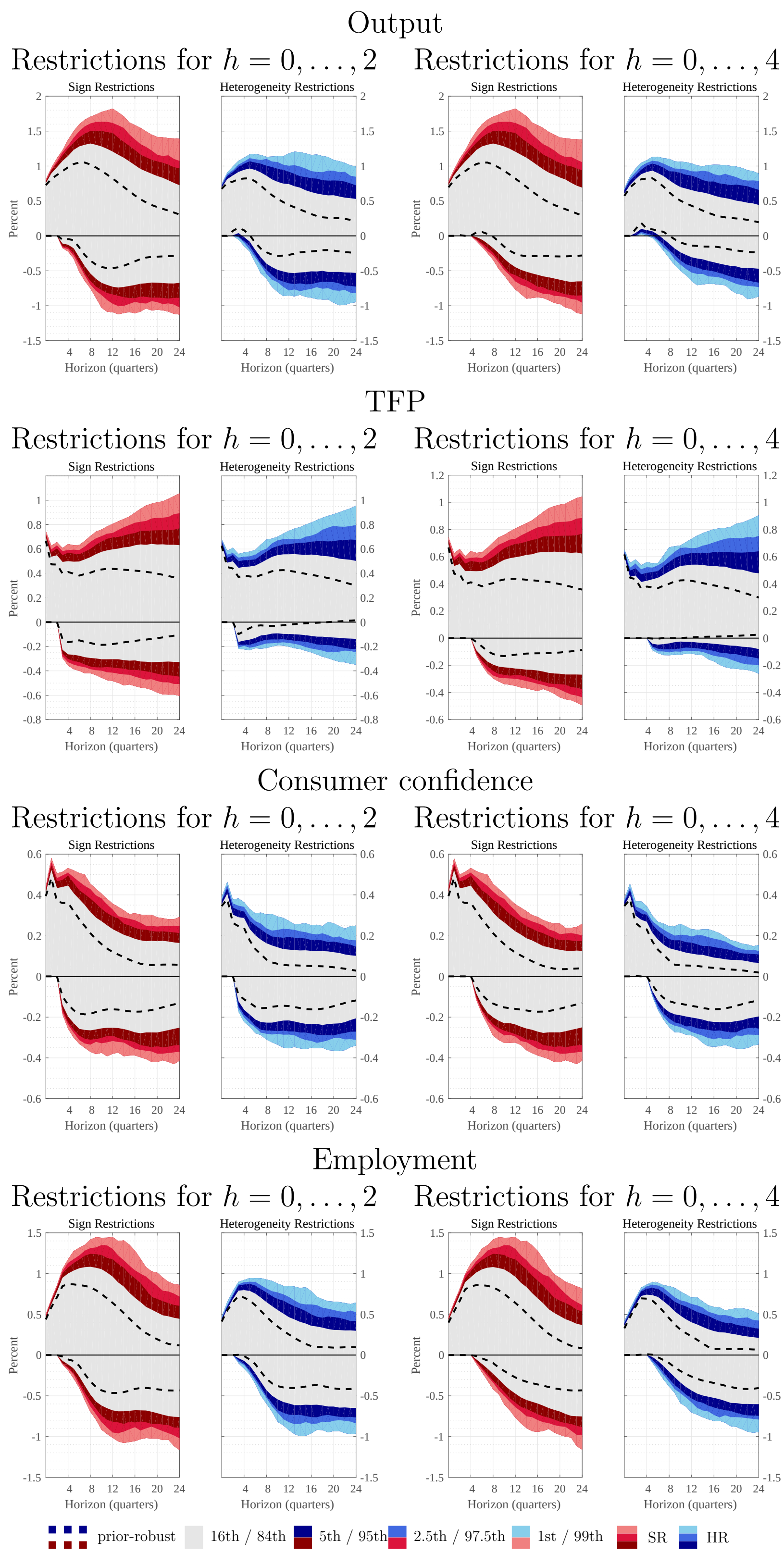


Figure E.10: Prior-robust responses to productivity news shock for macro variables: Restrictions for three vs. five quarters

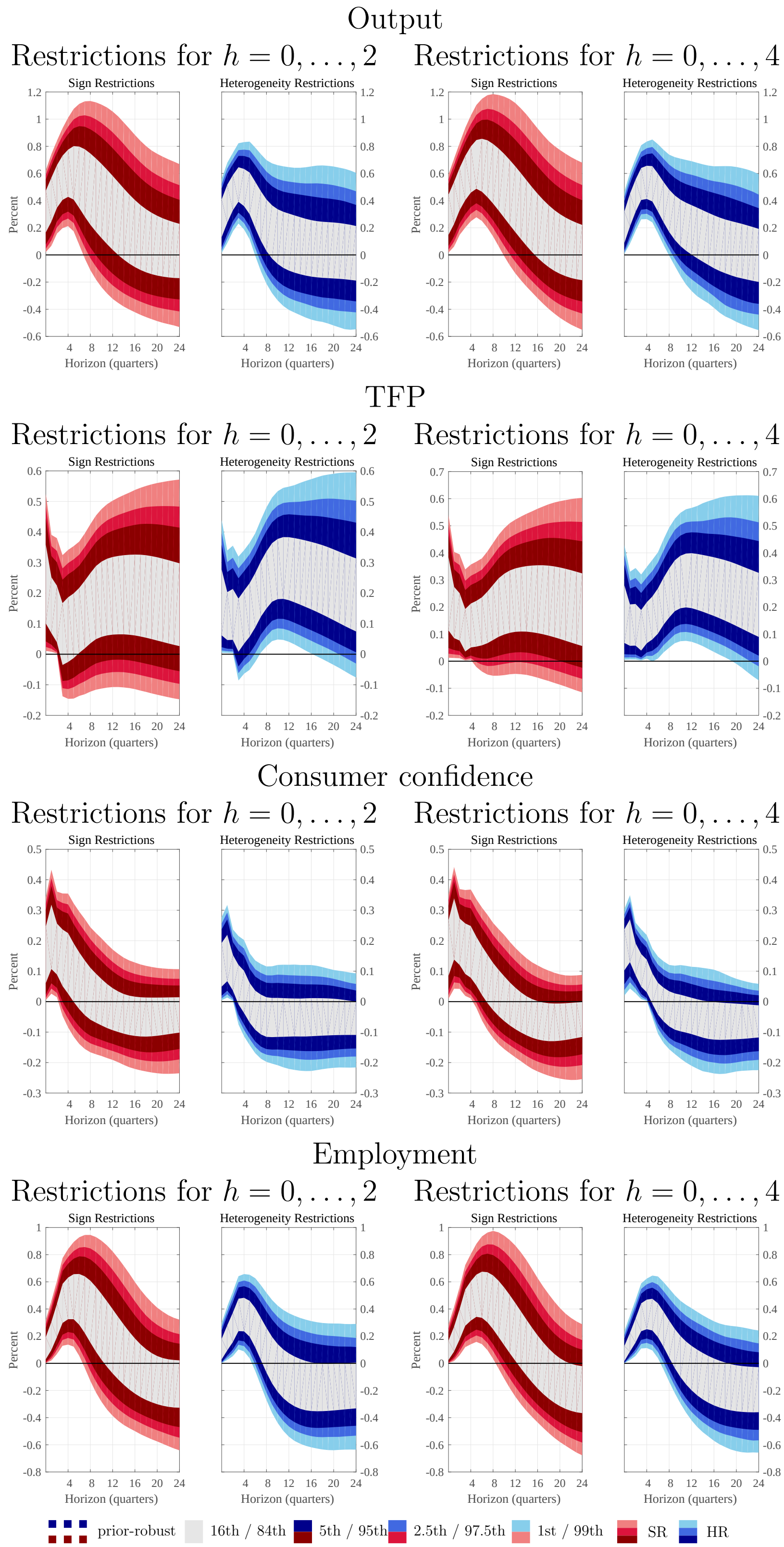


Figure E.11: Fully Bayesian responses to productivity news shock for macro variables: Restrictions for three vs. five quarters

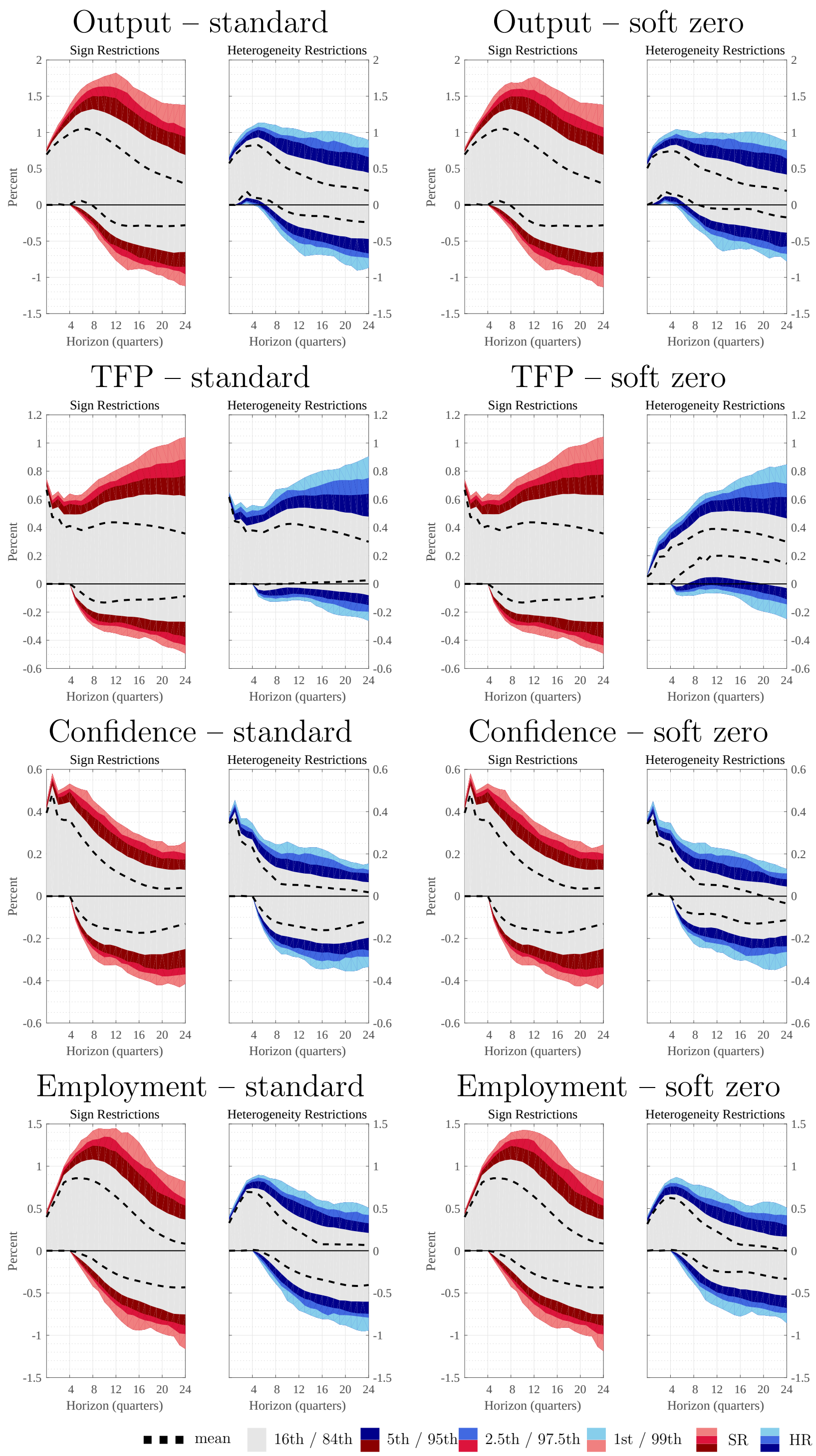


Figure E.12: Prior-robust responses of macro variables and industry returns to productivity news shock with soft zero restriction.

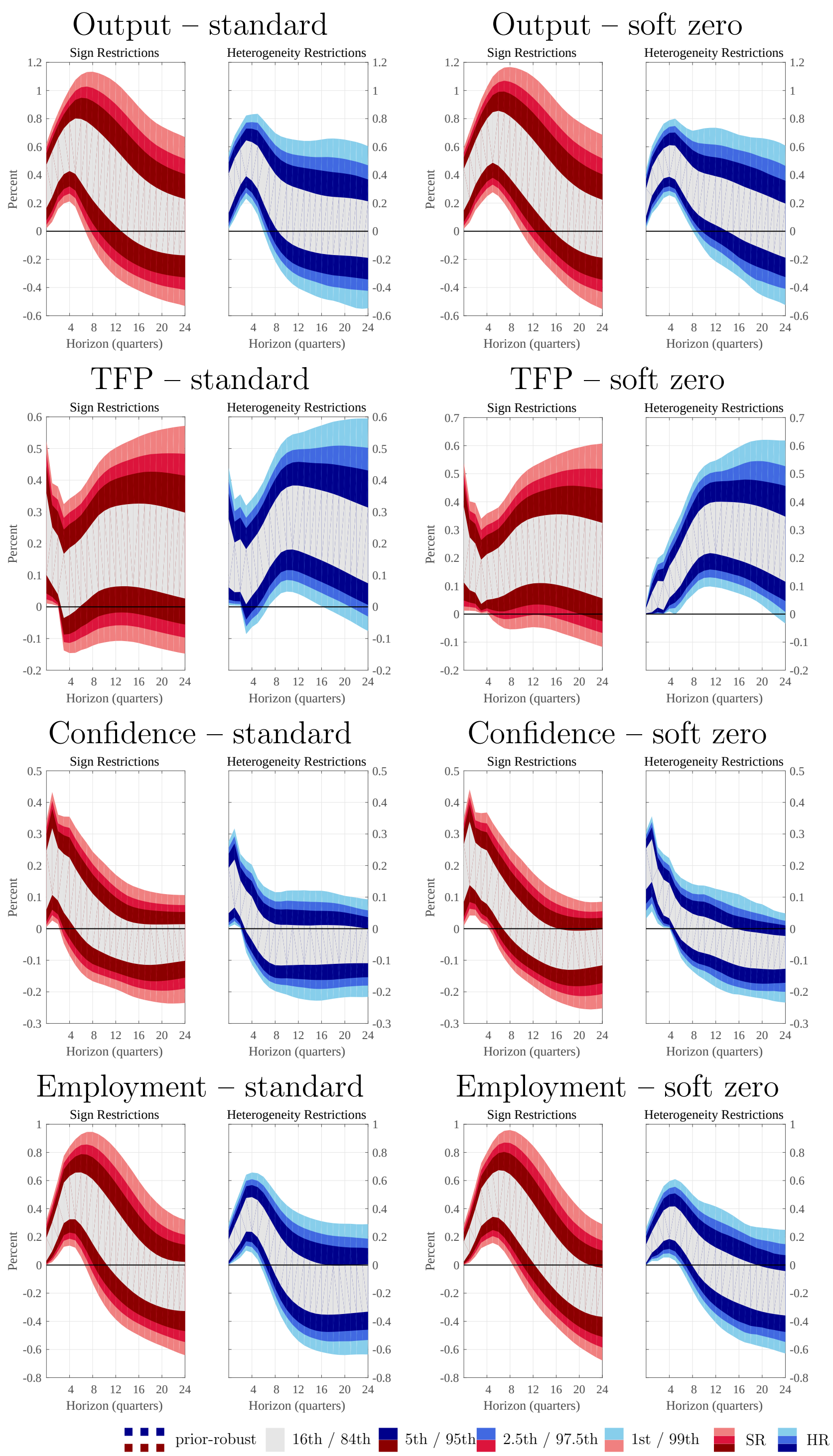


Figure E.13: Fully Bayesian responses of macro variables and industry returns to productivity news shock with soft zero restriction.

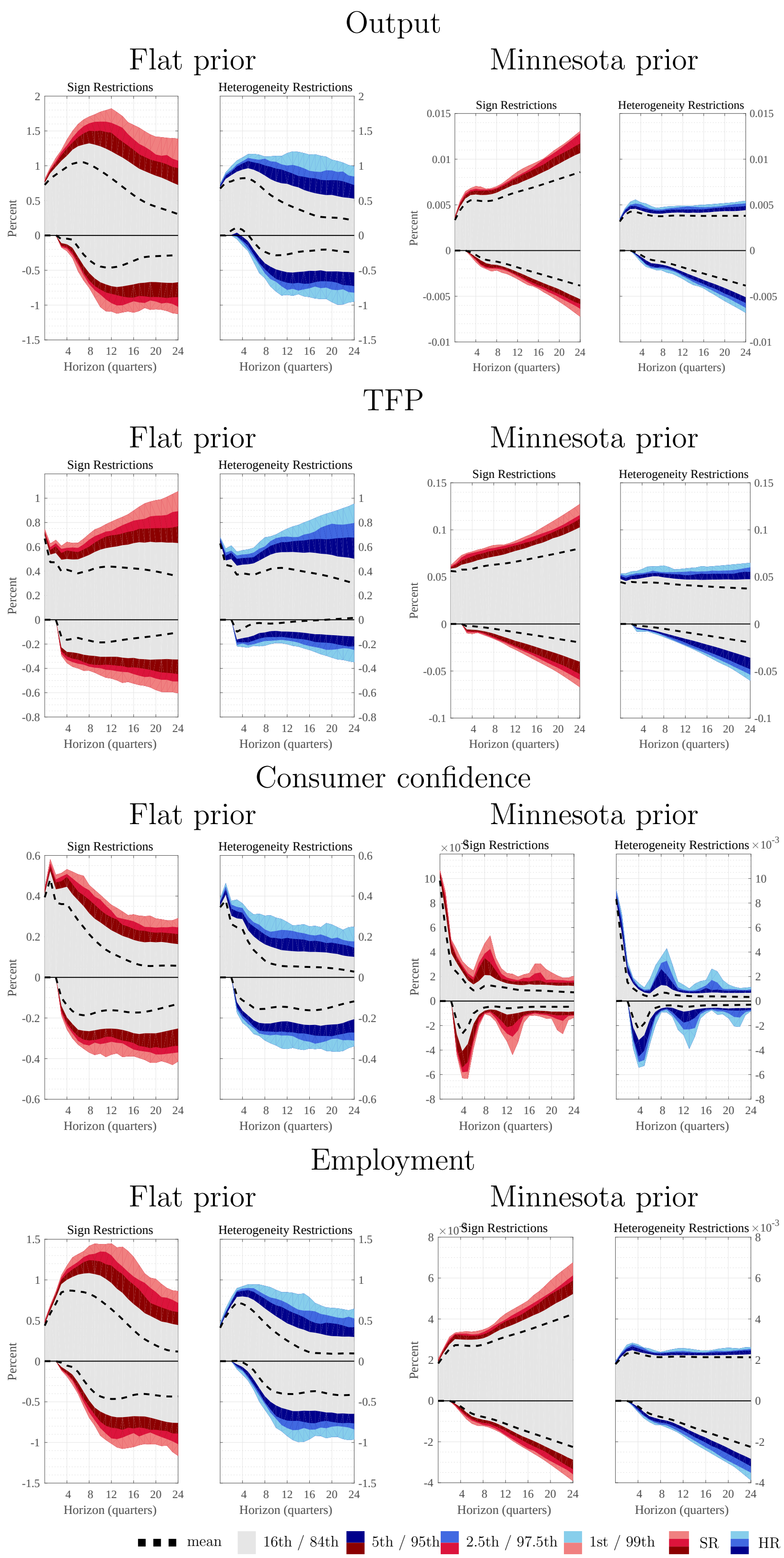


Figure E.14: Prior-robust responses to productivity news shock for macro variables: Flat vs. Minnesota reduced-form prior

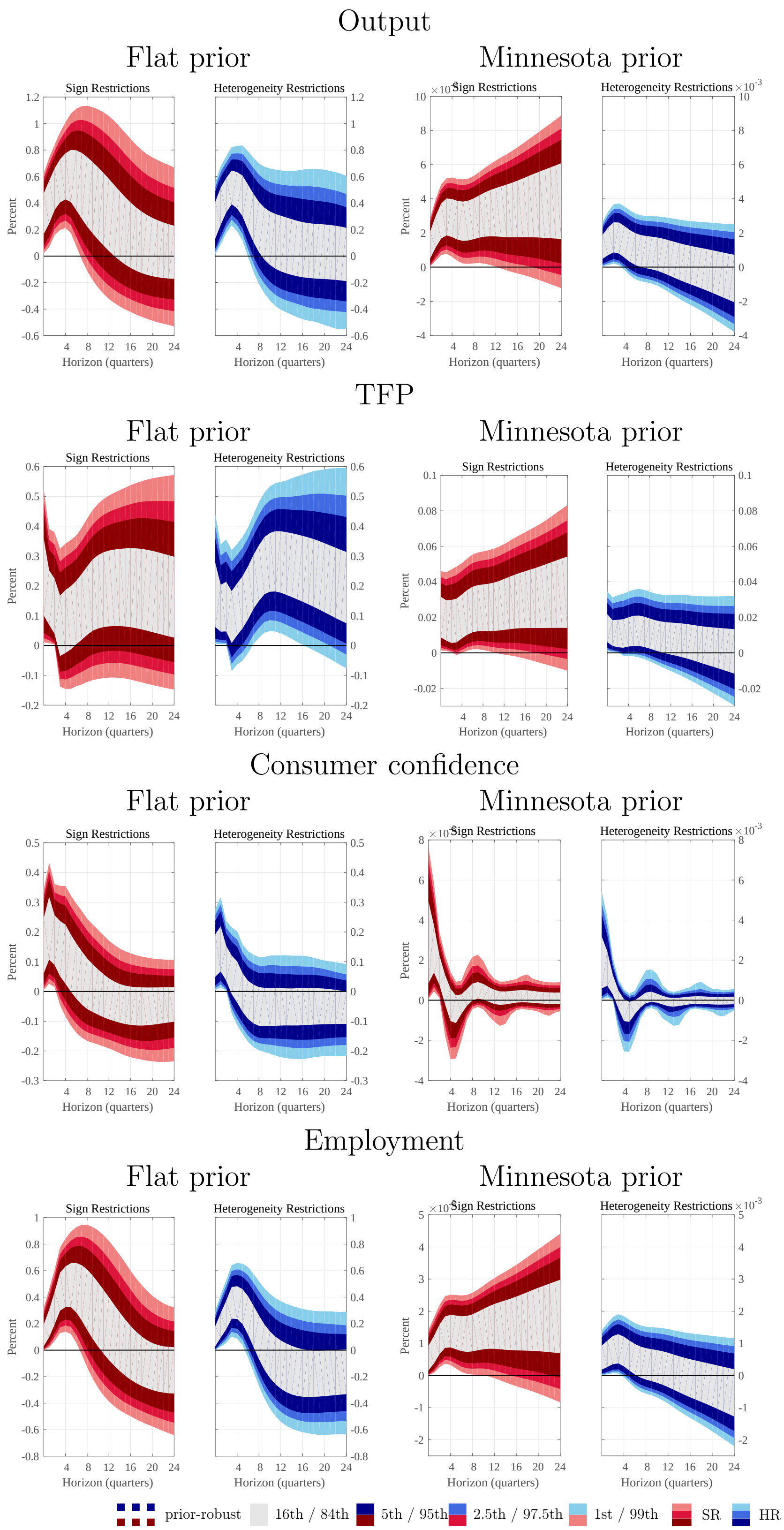


Figure E.15: Fully Bayesian responses to productivity news shock for macro variables: Flat vs. Minnesota reduced-form prior

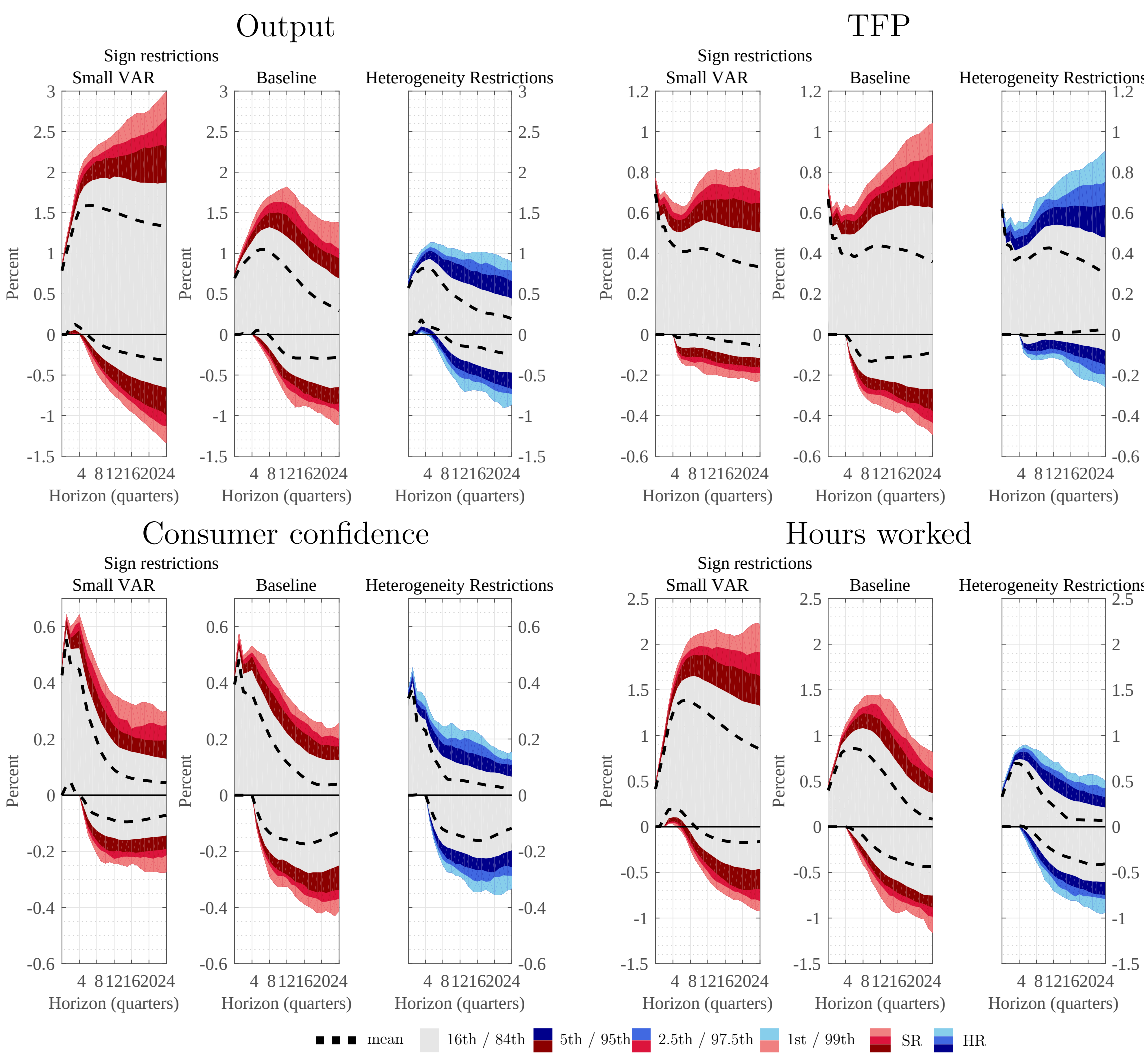


Figure E.16: Prior-robust Bayesian responses of macro variables and industry returns to productivity news shock: Effect of information set and parameter uncertainty.

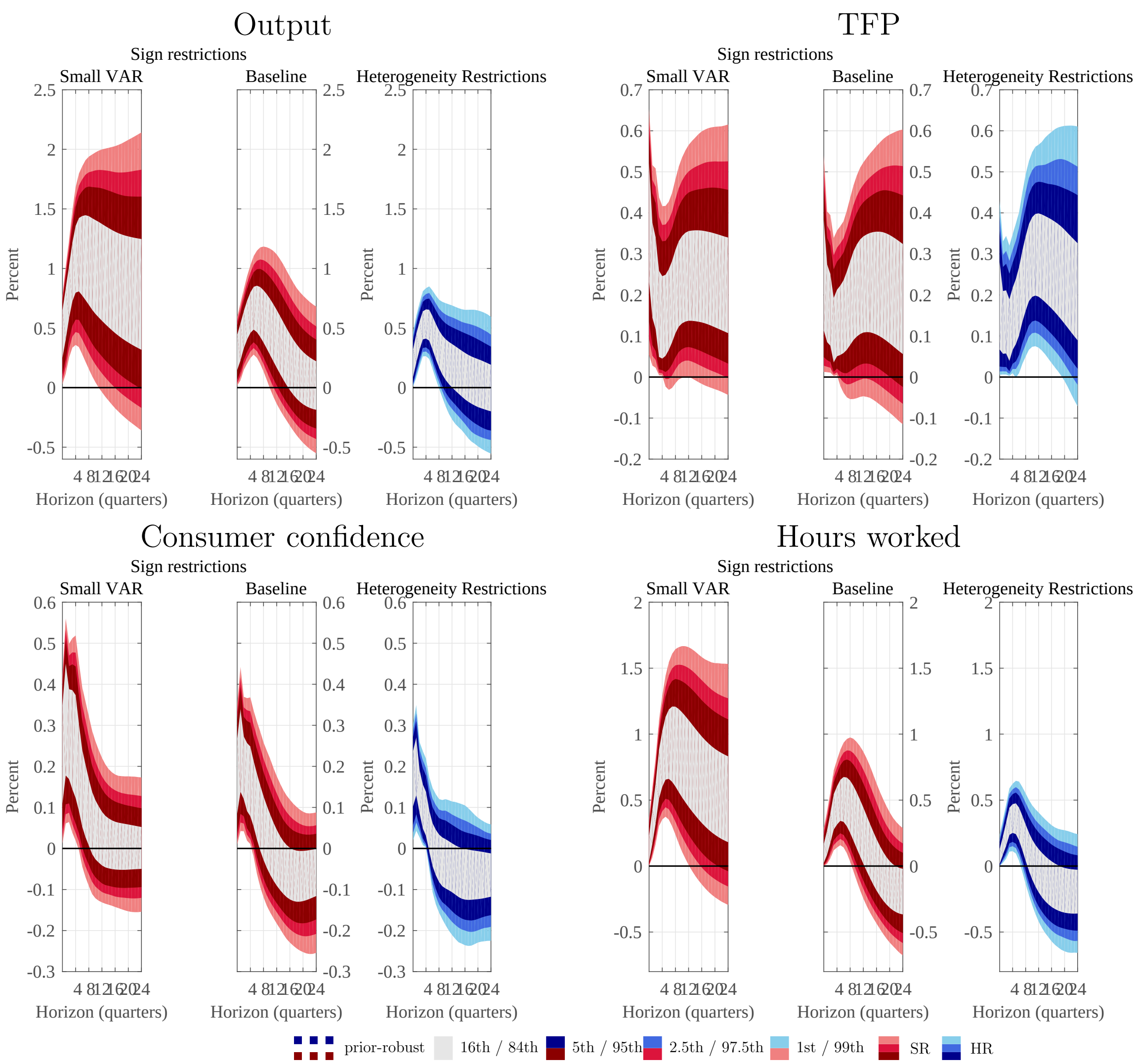


Figure E.17: Fully Bayesian responses of macro variables and industry returns to productivity news shock: Effect of information set and parameter uncertainty.

E.2 Fiscal shocks

E.2.1 IRFs

Prior-robust

Variable	Horizon H (year) 0	1	2	3	4	8	10
Defense spending	8.5	8.7	11.3	9.4	13.2	24.4	26.8
Output	23.6	10.7	12.9	13.3	24.9	26.8	29.4
Debt	12.9	14.3	16.7	16.3	15.1	20.6	23.9
Hours	11.1	12.6	11.8	21.0	25.2	22.9	24.0
Tax rate	18.7	27.5	32.7	36.8	43.3	44.4	37.9
Tobacco	62.8	56.5	43.8	37.4	43.5	40.2	26.9
Lumber	48.2	34.2	32.5	38.4	39.5	28.5	30.9
Petrol	35.0	33.3	34.6	34.3	34.2	28.6	30.9
Equipment	46.7	40.1	28.7	30.6	33.3	44.7	35.3
Electronics	33.0	33.5	31.1	34.0	42.0	49.8	32.3
Transportation	22.5	13.5	15.9	24.9	32.6	29.0	20.0

Fully Bayesian

Variable	Horizon H (year) 0	1	2	3	4	8	10
Defense spending	-11.7	-20.9	-19.3	-13.6	-8.3	4.8	13.0
Output	19.3	-0.7	-7.7	-4.5	2.7	6.5	3.0
Debt	-10.8	-17.0	-32.5	-43.7	-17.9	3.8	9.8
Hours	0.7	-0.8	-5.6	0.4	10.2	7.4	11.4
Tax rate	15.5	30.6	34.4	38.3	41.9	43.6	37.7
Tobacco	65.4	57.5	42.0	36.6	45.8	21.2	7.6
Lumber	58.3	19.5	10.6	24.0	31.7	2.4	4.4
Petrol	34.8	34.0	33.7	33.4	32.3	10.6	19.2
Equipment	51.3	38.4	22.1	22.8	33.8	30.1	14.2
Electronics	35.6	31.4	24.0	24.2	31.5	49.5	15.8
Transportation	16.6	-4.3	-1.6	3.4	3.7	-4.0	0.5

Note: 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 E.5: Reduction of 99th percentile of IRF to defense spending shocks relative to sign restrictions

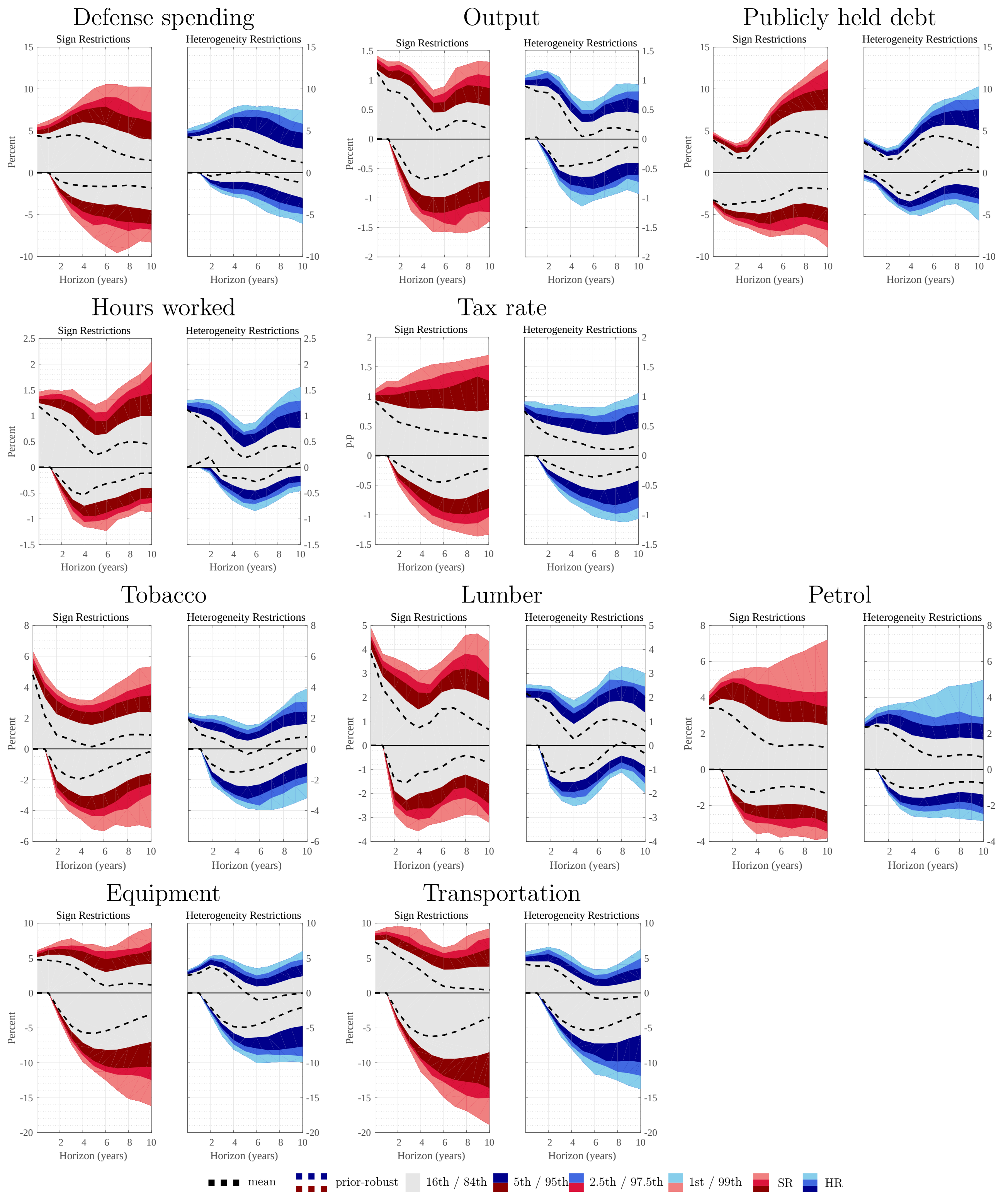


Figure E.18: Prior-robust responses to defense spending shock: All variables, baseline model

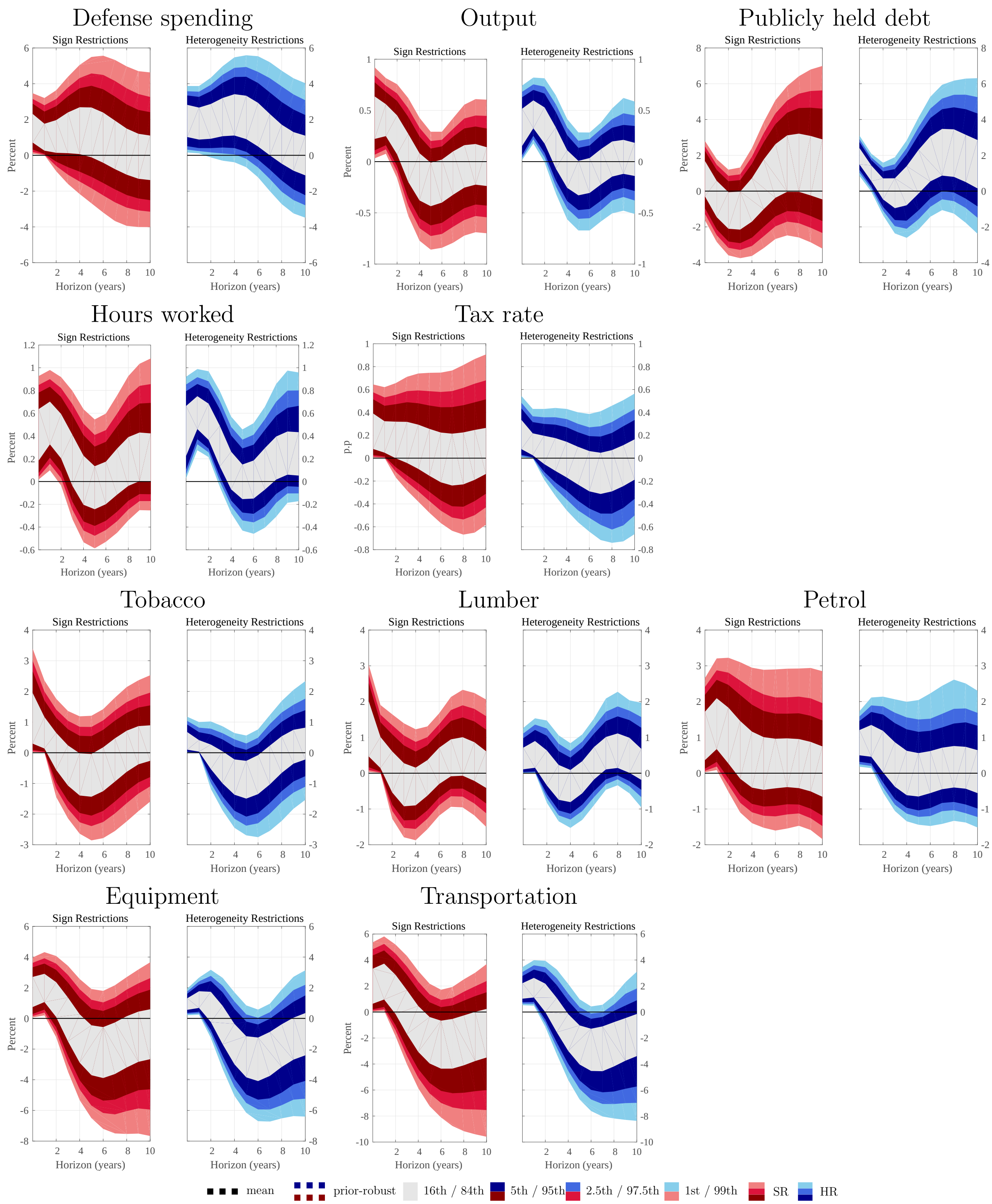


Figure E.19: Fully Bayesian responses to defense spending shock: All variables, baseline model

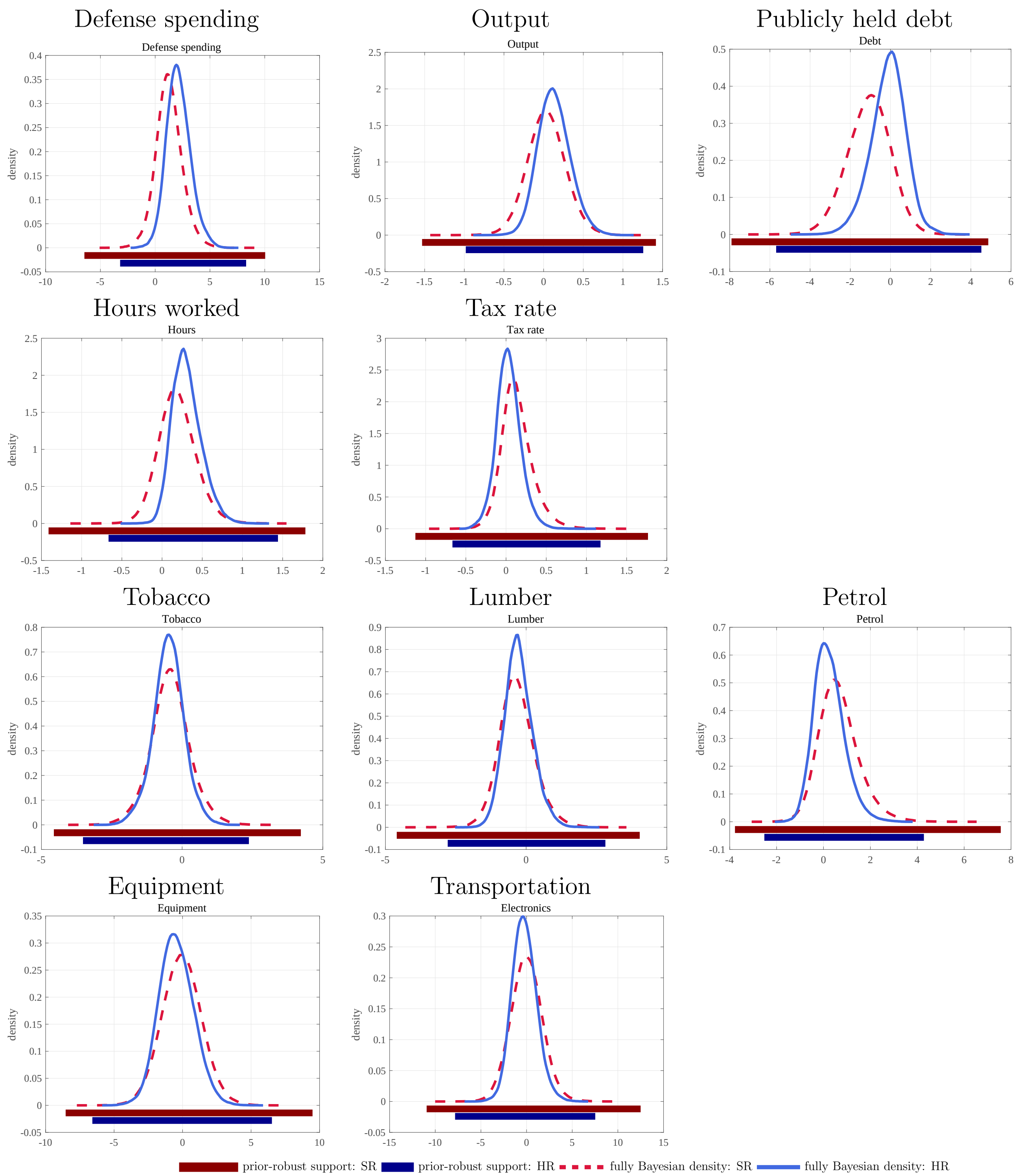


Figure E.20: Fully Bayesian posterior density of responses to defense spending shock three years out: All variables, baseline model

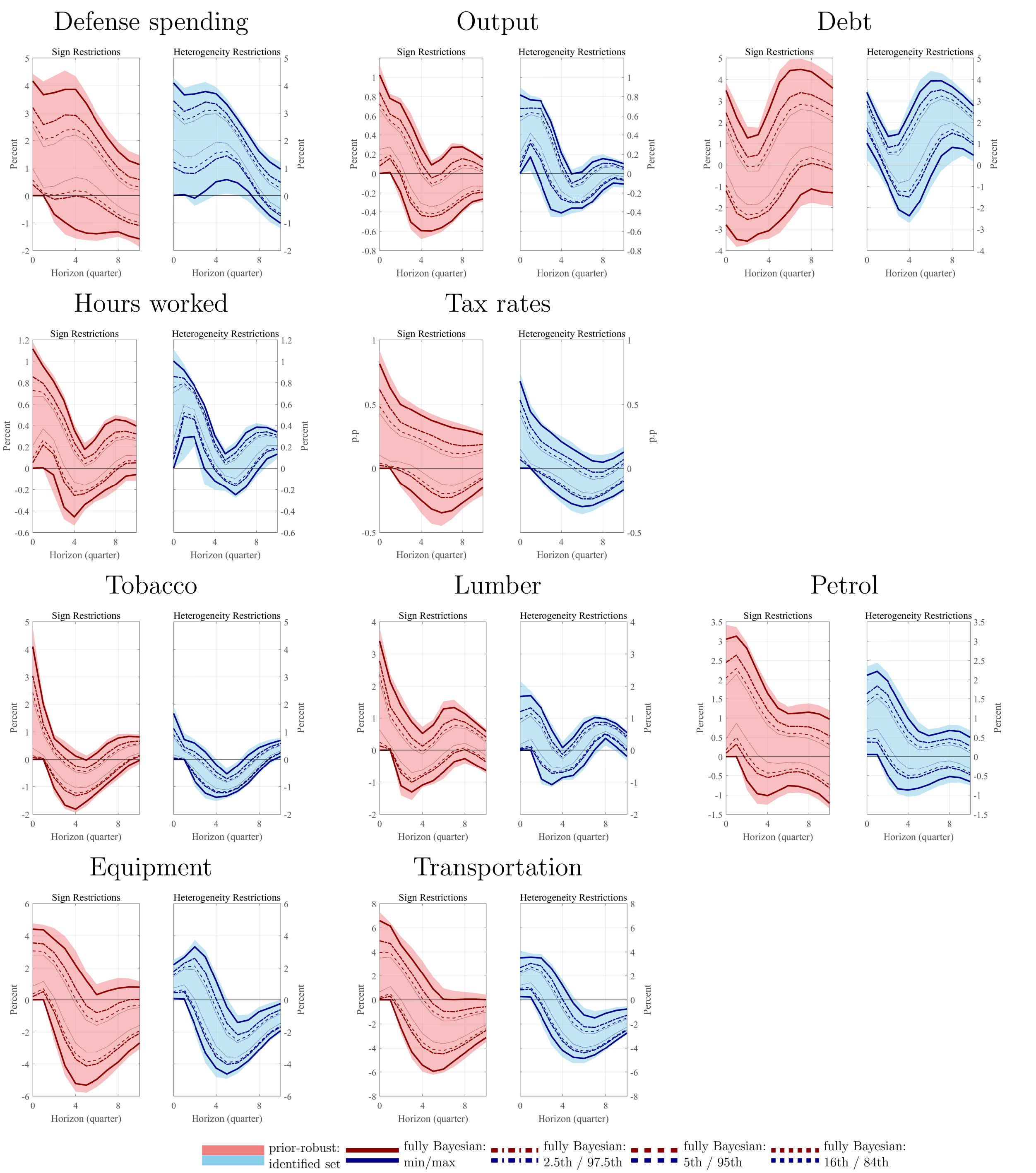


Figure E.21: Plug-In: Identified set and conditional posterior of responses to defense spending shock.

E.2.2 IRFs with restricted multipliers

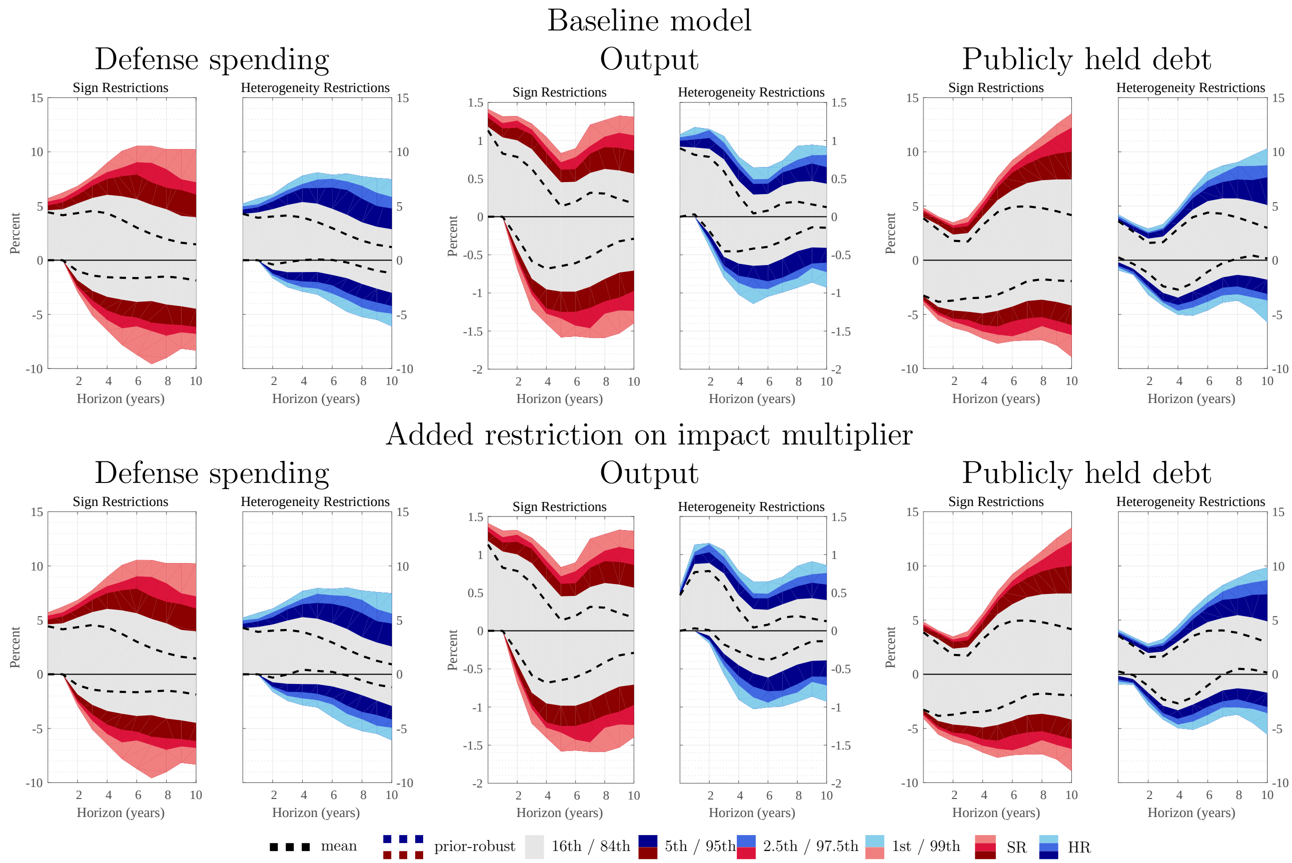


Figure E.22: Prior-robust responses to defense spending shock: Select macro variables, with and without restrictions on multipliers

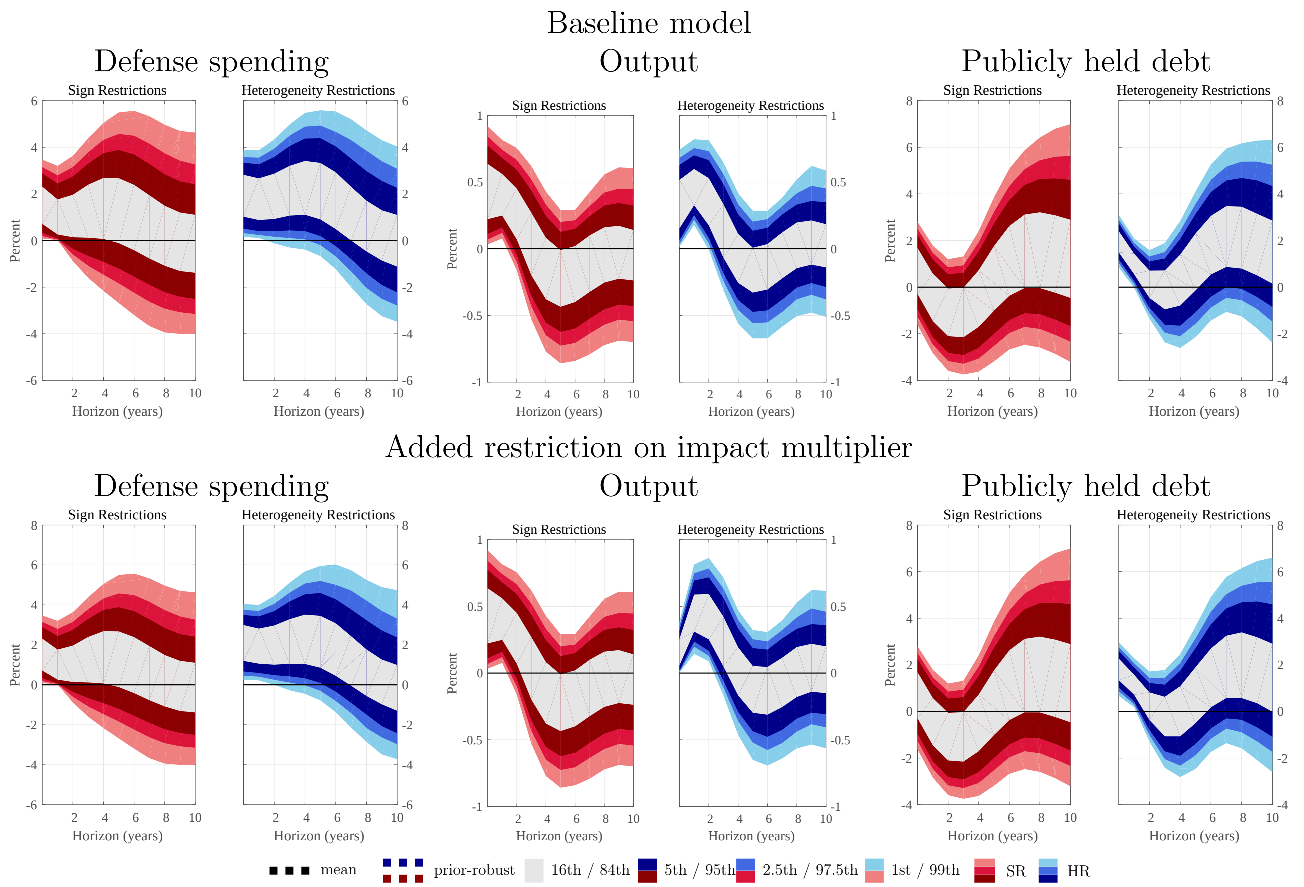


Figure E.23: Fully Bayesian responses to defense spending shock: Select macro variables, with and without restrictions on multipliers

E.2.3 FEVD

Sign restrictions

Variable	0	1	2	3	4	8	10
	Horizon H (year)						
Defense spending	79 (59,84)	58 (44,67)	51 (36,60)	48 (31,55)	46 (28,54)	42 (27,50)	40 (27,46)
Output	85 (71,90)	57 (45,62)	44 (33,48)	38 (28,42)	36 (28,40)	32 (26,36)	33 (27,36)
Debt	67 (52,73)	50 (39,56)	51 (40,56)	45 (36,50)	37 (29,43)	35 (23,39)	36 (23,44)
Hours	84 (70,89)	66 (55,74)	60 (45,65)	52 (37,55)	45 (32,50)	40 (27,44)	43 (29,46)
Tax rate	64 (46,73)	56 (42,62)	54 (37,62)	50 (33,58)	48 (28,56)	41 (24,51)	38 (21,49)

Heterogeneity restrictions

Variable	0	1	2	3	4	8	10
	Horizon H (year)						
Defense spending	61 (36,71)	47 (30,56)	40 (26,49)	38 (24,44)	35 (23,44)	30 (18,38)	27 (17,34)
Output	47 (33,53)	33 (24,40)	30 (23,35)	25 (20,32)	25 (18,29)	19 (9,24)	20 (12,23)
Debt	53 (37,59)	34 (24,38)	25 (17,28)	20 (14,23)	17 (14,21)	19 (12,24)	19 (11,25)
Hours	65 (48,72)	54 (38,62)	47 (35,51)	37 (30,42)	32 (27,36)	29 (18,34)	31 (18,35)
Tax rate	35 (25,45)	28 (18,33)	24 (15,31)	20 (14,25)	18 (11,21)	13 (9,17)	14 (9,17)

Difference between SR and HR

Variable	0	1	2	3	4	8	10
	Horizon H (year)						
Defense spending	14 (6,19)	10 (2,15)	8 (2,13)	8 (1,12)	8 (2,12)	11 (2,14)	11 (2,14)
Output	35 (21,42)	19 (9,25)	11 (6,16)	9 (4,14)	10 (4,15)	13 (7,15)	13 (7,15)
Debt	13 (5,18)	15 (7,23)	26 (10,34)	26 (11,32)	19 (7,25)	11 (4,17)	16 (5,22)
Hours	17 (8,22)	10 (3,15)	10 (3,14)	9 (2,13)	10 (2,13)	10 (5,13)	12 (6,14)
Tax rate	27 (12,32)	25 (12,31)	24 (12,32)	26 (11,33)	27 (11,33)	27 (6,37)	22 (7,36)

Note: Each cell of the table for a given variable and horizon shows the following three percentiles across reduced-form draws respectively: Median and 68% credible set $(16^{th}, 84^{th})$.

Table E.6: FEVD of defense spending shocks: Prior-robust bounds

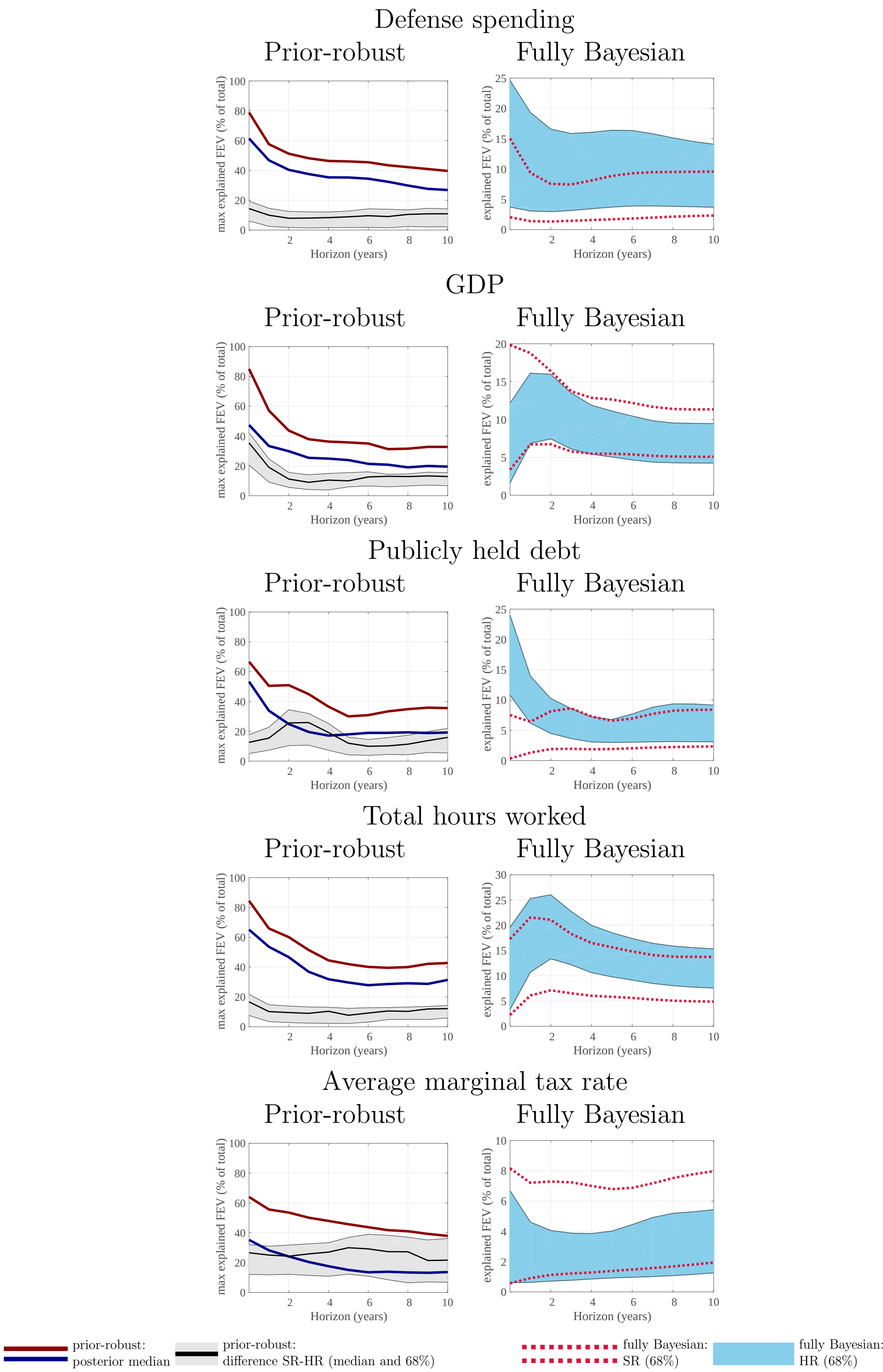


Figure E.24: FEVD of defense spending shock: Macro variables

(a) Sign restrictions

Variable	Horizon H (year) 0	1	2	3	4	8	10
Defense spending	(2,15)	(1,9)	(1,8)	(1,7)	(2,8)	(2,10)	(2,10)
Output	(3,20)	(7,19)	(7,16)	(6,14)	(5,13)	(5,11)	(5,11)
Debt	(0,8)	(1,6)	(2,8)	(2,9)	(2,7)	(2,8)	(2,8)
Hours	(2,17)	(6,22)	(7,21)	(7,18)	(6,16)	(5,14)	(5,14)
Tax rate	(1,8)	(1,7)	(1,7)	(1,7)	(1,7)	(2,8)	(2,8)

(b) Heterogeneity restrictions

Variable	Horizon H (year) 0	1	2	3	4	8	10
Defense spending	(4,25)	(3,19)	(3,17)	(3,16)	(3,16)	(4,15)	(4,14)
Output	(2,12)	(7,16)	(7,16)	(6,14)	(5,12)	(4,10)	(4,9)
Debt	(11,24)	(6,14)	(5,10)	(4,9)	(3,7)	(3,9)	(3,9)
Hours	(3,20)	(11,25)	(13,26)	(12,23)	(11,20)	(8,16)	(8,15)
Tax rate	(1,7)	(1,5)	(1,4)	(1,4)	(1,4)	(1,5)	(1,5)

(c) Difference between SR and HR 16th percentile

Variable	Horizon H (year) 0	1	2	3	4	8	10
Defense spending	-5 (-15,-1)	-4 (-13,-1)	-3 (-11,-2)	-3 (-10,-2)	-4 (-10,-2)	-3 (-10,-1)	-3 (-9,-1)
Output	1 (-2,2)	-2 (-4,0)	-1 (-5,-0)	-1 (-3,-0)	-1 (-3,0)	-0 (-2,1)	0 (-2,1)
Debt	-13 (-19,-11)	-6 (-10,-5)	-3 (-6,-2)	-2 (-5,-1)	-2 (-4,-1)	-2 (-5,-1)	-1 (-4,-0)
Hours	-2 (-7,-0)	-6 (-10,-4)	-8 (-13,-6)	-7 (-13,-5)	-6 (-11,-4)	-4 (-8,-3)	-4 (-7,-2)
Tax rate	-0 (-2,0)	0 (-1,0)	0 (-1,1)	0 (-1,1)	0 (-1,1)	0 (-3,1)	0 (-3,1)

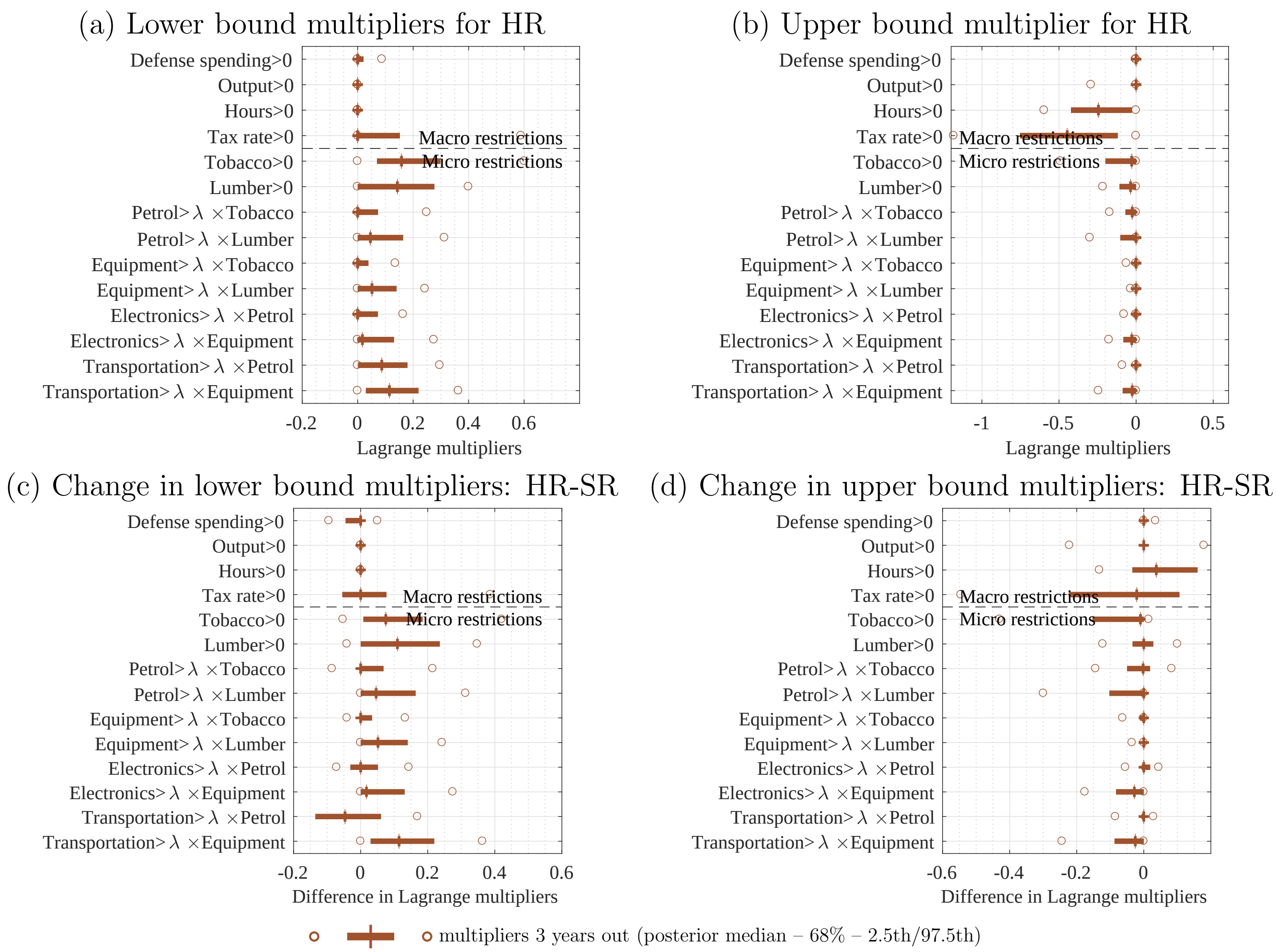
(d) Difference between SR and HR 68th percentile

Variable	Horizon H (year) 0	1	2	3	4	8	10
Defense spending	-6 (-20,-1)	-8 (-17,-3)	-7 (-14,-4)	-7 (-14,-4)	-6 (-14,-3)	-4 (-11,-1)	-3 (-9,-1)
Output	8 (1,11)	3 (-2,5)	1 (-4,3)	0 (-3,2)	1 (-2,2)	2 (-1,4)	2 (-1,3)
Debt	-14 (-21,-12)	-7 (-11,-5)	-3 (-6,-0)	-0 (-4,2)	-0 (-4,2)	-1 (-5,1)	-0 (-5,2)
Hours	-2 (-10,2)	-4 (-10,1)	-4 (-11,-1)	-3 (-10,-2)	-3 (-9,-1)	-1 (-7,-0)	-1 (-7,0)
Tax rate	2 (-2,4)	2 (-1,4)	2 (-0,5)	3 (-0,5)	2 (-1,5)	2 (-2,4)	3 (-2,5)

Note: Each cell of the table for a given variable and horizon shows the full posterior 68% credible set ($16^{th}, 84^{th}$) for panels (a) and (b) and the posterior median and 68% credible set across reduced-form draws for panels (c) and (d).

Table E.7: FEVD of defense spending shocks: Fully Bayesian

E.2.4 Importance of shocks



This figure quantifies the importance of all sign and heterogeneity restrictions that we use to identify the news shock for narrowing the identified set of the responses of output at the two year horizon. It shows the distribution of Lagrange multipliers on all restrictions over all reduced-form draws, separately for multipliers on the lower and upper bound. We sum the multipliers across the restrictions horizons $0, \ldots, \bar{H}$. Multipliers on upper bounds are negative because tighter restrictions reduce the upper bound. We find that for pinning down the upper bound on the average macro variable, the heterogeneity restrictions on stock return micro data matter more than the macro sign restrictions. Restrictions on manufacturing are particularly important, reducing the upper bound by 0.10 to 0.17 percentage points. For the lower bound on responses, the opposite pattern emerges: Sign restrictions on macro variables dominate, even though multipliers on some heterogeneity restrictions are significantly positive.

Figure E.25: Importance of restrictions for output responses to defense spending shocks two years after impact: Lagrange multipliers on restrictions.

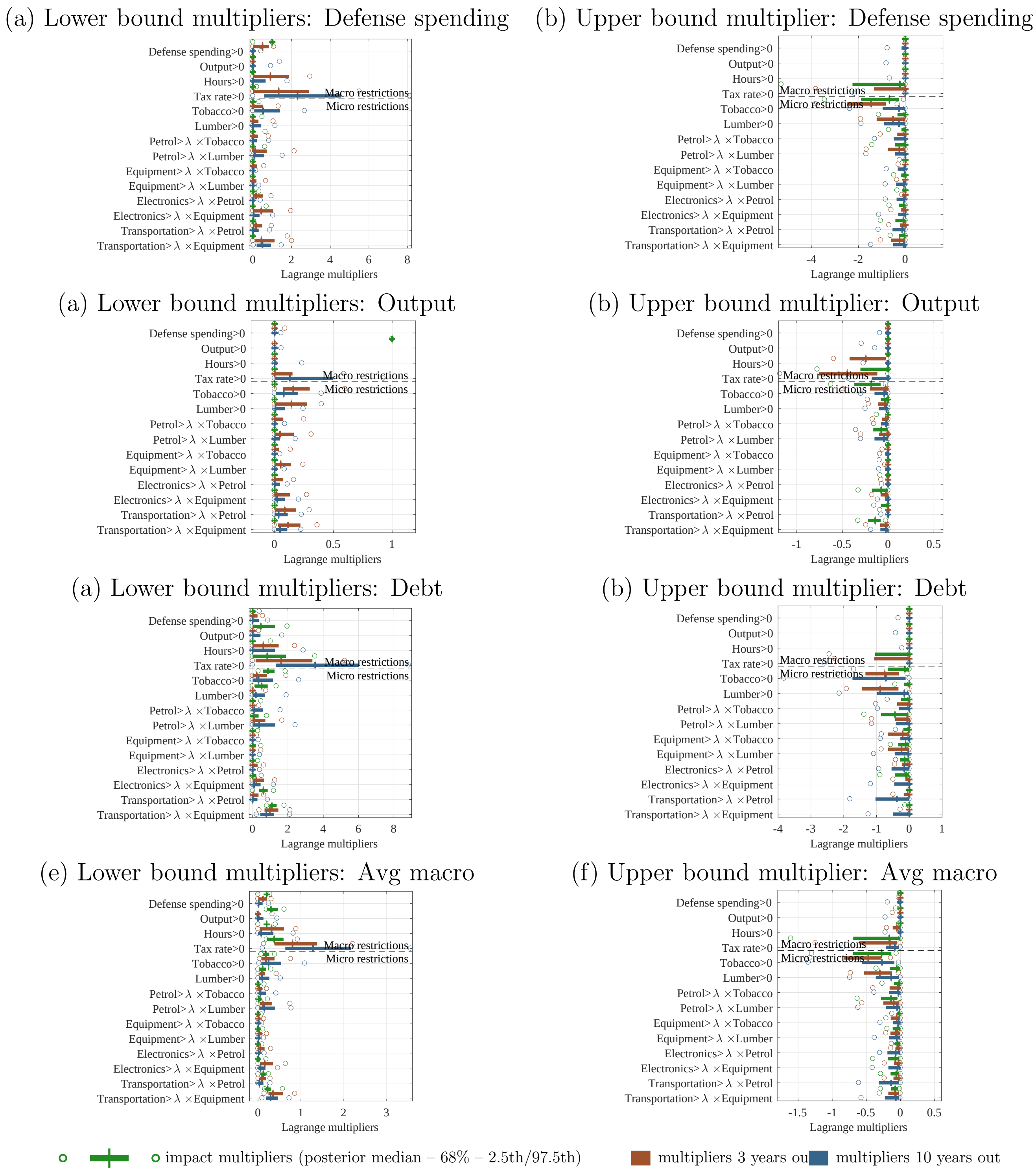


Figure E.26: Importance of restrictions for macro responses to defense spending shocks two years after impact: Lagrange multipliers on restrictions.

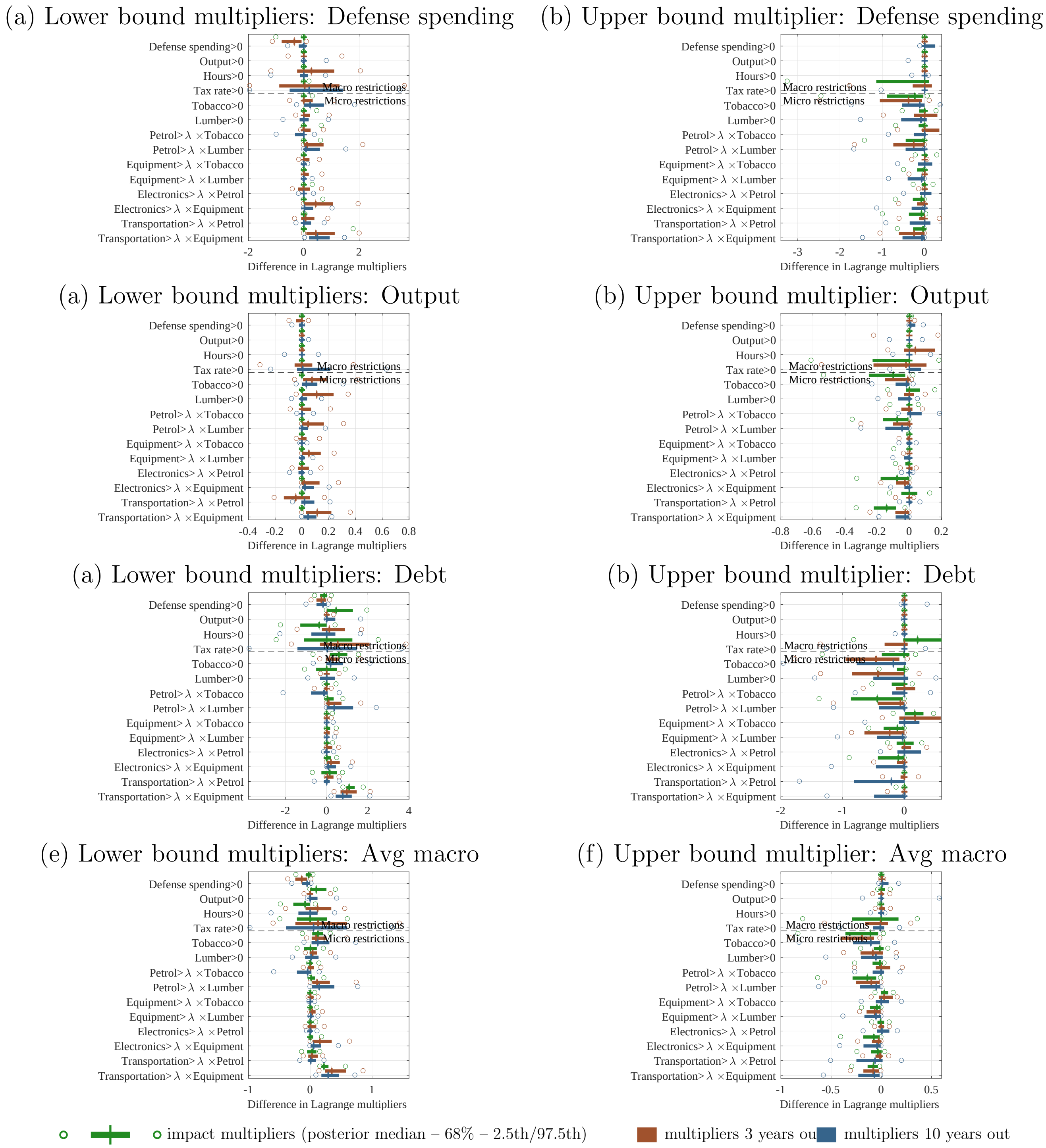
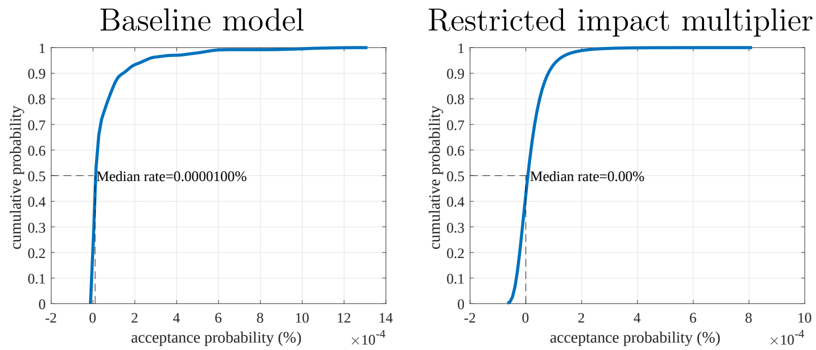


Figure E.27: Change in the importance of restrictions for macrp responses to defense spending shocks two years after impact: Change in lagrange multipliers on restrictions with heterogeneity restrictions relative to sign restrictions

E.2.5 Acceptance rates with standard algorithm



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 E.28: Distribution of acceptance probabilities for uniform proposal density over reduced-form draws: Fiscal application

E.2.6 Robustness

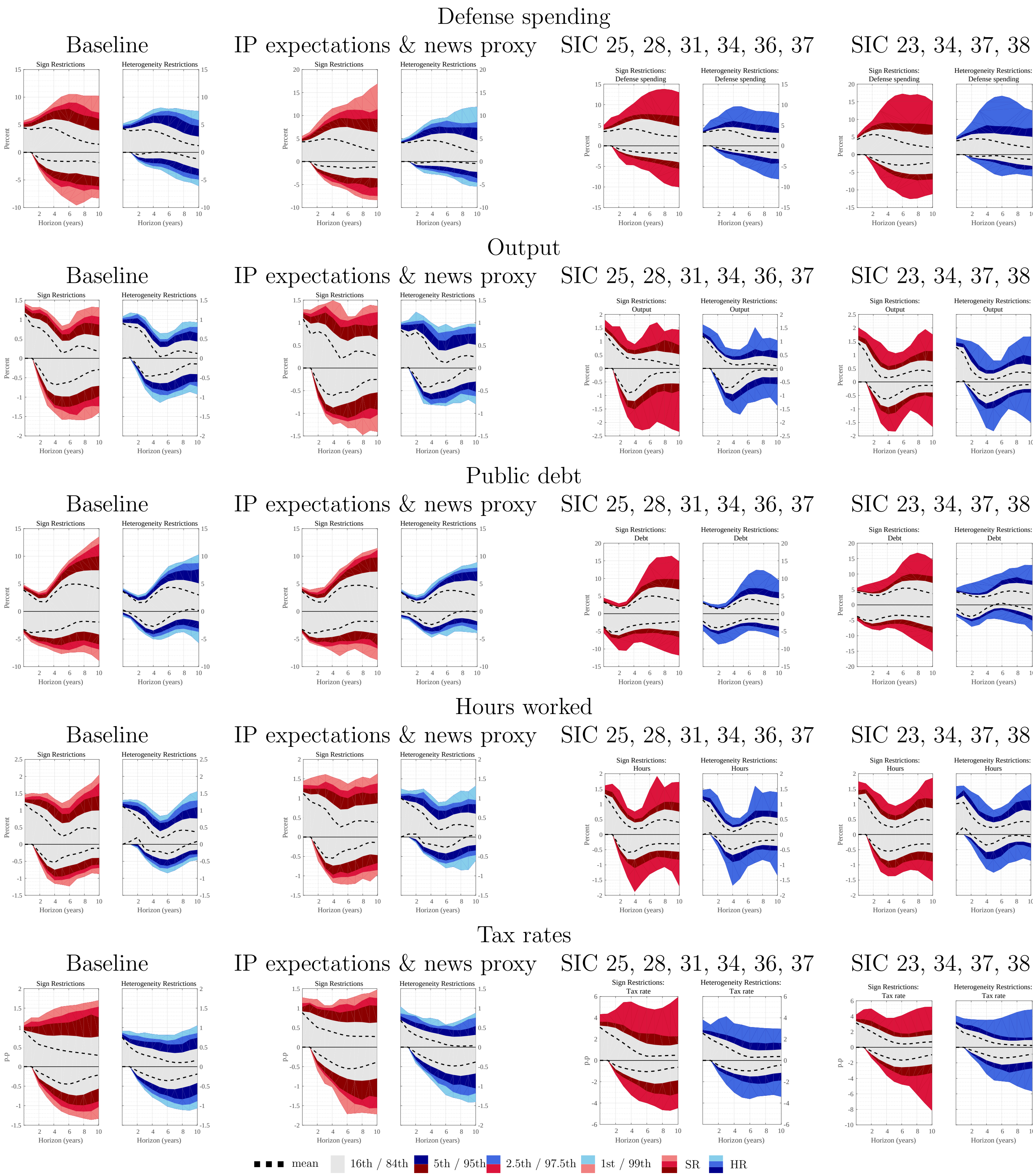


Figure E.29: Prior-robust responses to defense spending shock: All variables, baseline model

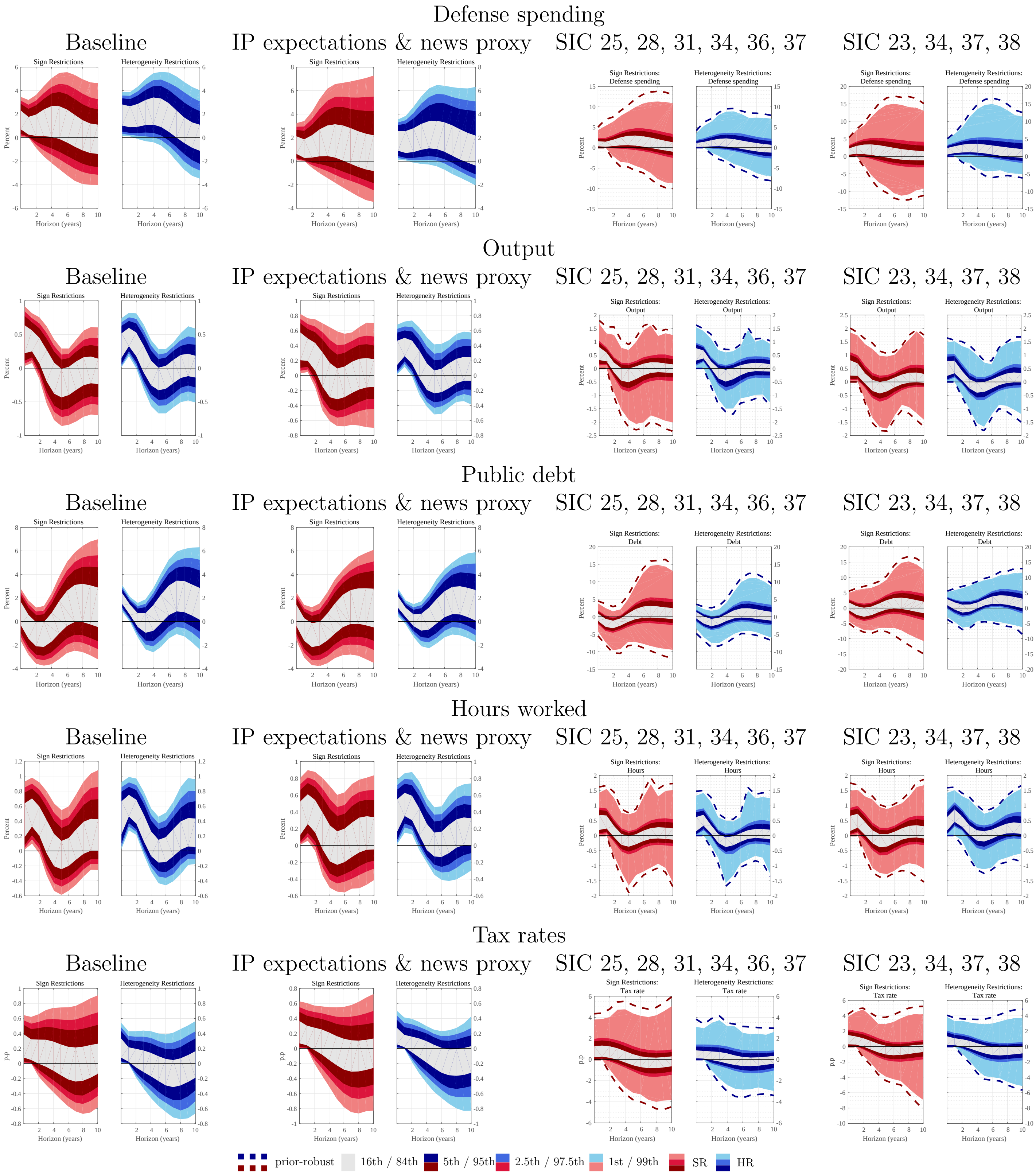


Figure E.30: Fully Bayesian responses to defense spending shock: All variables, robustness

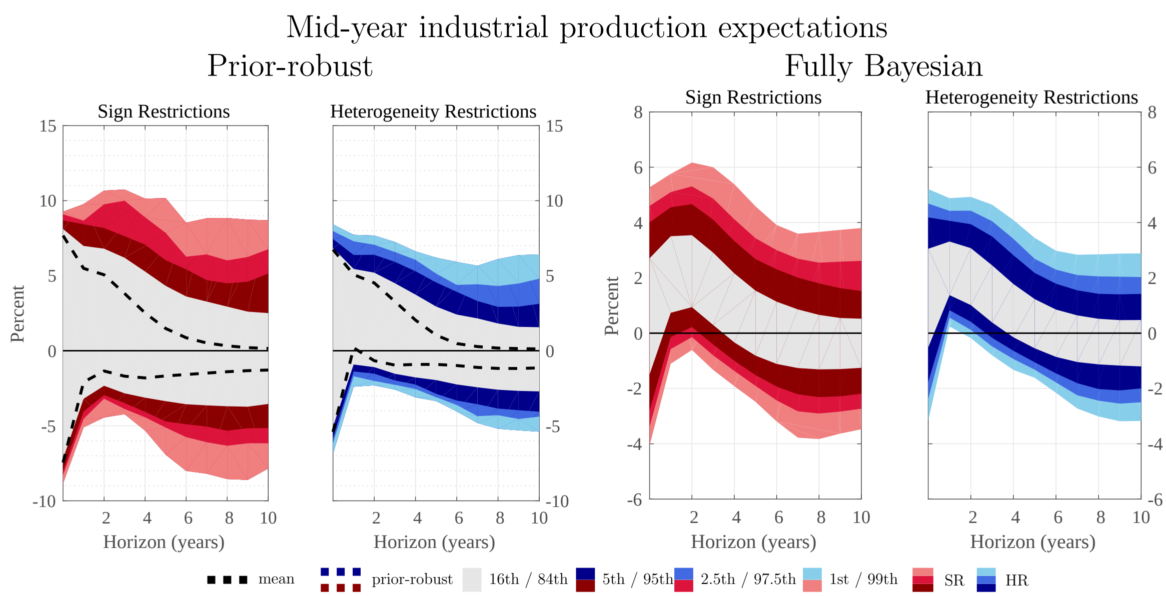


Figure E.31: Prior-robust and fully Bayesian responses of IP expectations to defense spending shock: Controlling for news proxy