Online Appendix to
Identification And Inference With Ranking Restrictions

Pooyan Amir-Ahmadi and Thorsten Drautzburg*

October 12, 2020

*Amir-Ahmadi: University of Illinois at Urbana-Champaign, pooyan[at]illinois.edu. Drautzburg: Federal Reserve Bank of Philadelphia, tdrautzburg[at]gmail.com. A previous draft circulated as "Identification Through Heterogeneity". We would like to thank Jonás Arias, Mark Bognanni, Jesús Fernández-Villaverde, James Nason, Michele Piffer, Frank Schorfheide, Mark Watson, Jonathan Wright, and various seminar and conference audiences for comments and Catherine O'Donnell and Nick Zarra for research assistance. Any remaining errors are our own. 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.

A Nonempty identified set and direct sampling

A.1 Proof for Proposition 1: Chebychev Center criterion with a single shock

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

$\Rightarrow$: Let $\tilde{\mathbf{q}}$ be a vector on the unit sphere. Then $\mathbf{y} = \mathbf{z}_c + \sqrt{r}\tilde{\mathbf{q}}$ satisfies $||\mathbf{y} - \mathbf{z}_c|| = r$.

Thus the identified set has a measure at least as large as:

$$\pi_q \left\{ \frac{\mathbf{y}}{||\mathbf{y}||} \big| \mathbf{y} = \mathbf{z}_c + \sqrt{r}\tilde{\mathbf{q}}, \tilde{\mathbf{q}}'\tilde{\mathbf{q}} = 1 \right\}.$$

$\Leftarrow$: We first show that the identified set cannot have positive measure if it imposes an equality restriction on any element of $\mathbf{q}$. 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 $\mathbf{q}$ such that $\mathbf{W}\mathbf{q} \leq 0$ and $||\mathbf{q}|| = 1$.

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

Now we construct, by induction, a $\mathbf{q}$ such that $\mathbf{W}\mathbf{q} < 0$. Pick $\mathbf{q}$ such that $\mathbf{W}_1\mathbf{q} < 0$ and $\mathbf{W}_2\mathbf{q} \leq 0$. Let $\boldsymbol{\varepsilon} = \delta \times (-\mathbf{W}_1\mathbf{q})$ and define $\hat{\mathbf{q}} = [\mathbf{q}_\ell - \delta n^{-1}\boldsymbol{\varepsilon}\,\mathrm{sgn}(\mathbf{W}_{2,\ell})\mathbf{W}_{2,\ell}]_\ell$ and $\mathbf{q}' = \frac{\hat{\mathbf{q}}}{||\hat{\mathbf{q}}||}$. Note that $\mathbf{W}_2\mathbf{q}' \propto \mathbf{W}_2\hat{\mathbf{q}} = \mathbf{W}_2\mathbf{q} - \delta n^{-1}\epsilon \sum_\ell |\mathbf{W}_{2,\ell}| < 0$. Also, for δ small enough, $\mathbf{W}_1\mathbf{q}' < 0$. Now assume that $\mathbf{W}_i\mathbf{q} < 0$ for $i = 1, \ldots, n$ and $\mathbf{W}_{n+1}\mathbf{q} = 0$. Proceed as before but with $\boldsymbol{\varepsilon} = \max_i \delta|\mathbf{W}_i\mathbf{q}|$. Going through the same argument shows that we can then also generate a $\mathbf{q}'$ such that $\mathbf{W}_i\mathbf{q}' < 0$ for all $i = 1, \ldots, n+1$.

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

□

A.2 Relationship of Chebychev center criterion to Granziera, Moon, and Schorfheide (2018)

Relationship between sufficient conditions. We now state results from the literature in our notation, whenever we have introduced the analogous notation before. In particular, we use $\boldsymbol{\theta} \equiv (\boldsymbol{\beta}, \boldsymbol{\Sigma}^{\mathrm{tr}})$ to collect the reduced form parameters, rather than ρ. We use $\mathbf{W}$ to denote the restrictions on the IRF, rather than Φ_p', and we use $n_J = \sum_{h=\underline{h}}^{H} J_h$ to refer to the number of restrictions, rather than r. We also use $\mathcal{Q}(\mathbb{H}; \boldsymbol{\theta})$ to refer to the identified set over $\mathbf{q}$, rather than F^q.

Corollary A.1 (Gordan's Alternative, Corollary 14 in Border (2013)). *Let* $\mathbf{W}$ *be an* $m \times n$ *matrix. Exactly one of the following alternatives holds. Either there exists a* $\mathbf{x} \in \mathbb{R}^n$ *satisfying* $\mathbf{W}\mathbf{x} > \boldsymbol{0}$ *(with strict inequality elementwise), or there exists a* $\mathbf{v} \in \mathbf{R}^m$ *satisfying* $\mathbf{W}'\mathbf{v} = \boldsymbol{0}$ *with* $\mathbf{v} \geq \boldsymbol{0}$ *and* $\mathbf{v} \neq \boldsymbol{0}$*.*

Assumption 1 (Assumption 1, Granziera, Moon, and Schorfheide (2018)). *There exists a compact reduced-form parameter set* $\mathcal{R}$ *and an* δ*-inflated superset* $\mathcal{R}^\delta$ *such that* $\mathcal{R} \subset \mathcal{R}^\delta \subset \bar{\mathcal{R}}$ *and:*

1. *For every* $\boldsymbol{\theta} \in \mathcal{R}^\delta$*, there does not exist an* $n_J \times 1$ *vector* $\mathbf{v} \geq \boldsymbol{0}, \mathbf{v} \neq \boldsymbol{0}$ *such that* $\mathbf{W}'\mathbf{v} = \boldsymbol{0}$*.*

2. *. . .*

Theorem A.1 (Theorem 1, Granziera, Moon, and Schorfheide (2018)). *Suppose Assumption 1(i) is satisfied. Then the admissible set* $AS(\mathbb{H}; \boldsymbol{\theta})$ *is non-empty and not a singleton for all* $\boldsymbol{\theta} \in \mathcal{R}^\delta$*.*

Granziera, Moon, and Schorfheide (2018) have a condition for a nonempty set that, in our notation and conditional on a draw of reduced form parameters requires that there is no vector $\mathbf{v} \in \mathbf{R}^{n_r}$ such that $\mathbf{W}'\mathbf{v} = 0$, with $\mathbf{v} \geq \mathbf{0}, \mathbf{v} \neq \mathbf{0}$. Given Gordan's Alternative (Border, 2013, Corollary 14), this condition implies the existence of a vector $\mathbf{z} \in \mathbf{R}^n$ such that $\mathbf{W}\mathbf{z} > \mathbf{0}$, with strict equality elementwise. Take $\mathbf{x}_c = \frac{\mathbf{z}}{||\mathbf{z}||}$ as the Chebychev center, scaled to lie within the unit cube. Note that since $\mathbf{W}\mathbf{x}_c > \mathbf{0}(= \mathbf{0}/||\mathbf{z}||)$ holds strictly, there is a non-degenerate open ball around $\mathbf{x}_c$ such that the inequality restrictions also hold for any $\mathbf{x}$ such that $||\mathbf{x} - \mathbf{x}_c|| < r, r > 0$. Thus, the assumption in their Theorem 1 implies the existence of a non-degenerate Chebychev center.

Vice-versa, if a non-degenerate Chebychev center $\mathbf{x}_c$ exists with $r > 0$, then all $\frac{\mathbf{x}}{||\mathbf{x}||} : ||\mathbf{x} - \mathbf{x}_c|| \leq r$ are in the admissible set. Thus, the admissible set is non-empty and not a singleton.

Thus, our Proposition is equivalent to Theorem 1 in Granziera, Moon, and Schorfheide (2018), except that we condition on a single parameter vector $\boldsymbol{\theta}$. However, a slightly stronger statement is possible. Indeed, when the admissible set is empty, our proposition also implies that their Assumption 1 is violated for a given $\boldsymbol{\theta}$.

If only a degenerate Chebychev center exists, then there exists an i such that $\mathbf{e}_i'\mathbf{W}\mathbf{x}_c = 0$. Assume, by contradiction, that there also exists a vector $\mathbf{x}$ such that $\mathbf{e}_j'\mathbf{W}\mathbf{x} > 0 \forall j = 1, \ldots, n_J$. Then this (rescaled) vector $\tilde{\mathbf{x}} = \frac{\mathbf{x}}{||\mathbf{x}||}$ would also be a Chebychev center. However, we could then construct a ball with strictly positive radius around it where the inequalities still hold, a contradiction. Thus, there does not exist a $\mathbf{x}$ such that $\mathbf{W}\mathbf{x} > \mathbf{0}$ with strict inequality. Gordan's Alternative then implies that there exists a $\mathbf{v}$ such that $\mathbf{W}'\mathbf{v} = \mathbf{0}$ with $\mathbf{v} \geq \mathbf{0}, \mathbf{z} \neq \mathbf{0}$.

A.3 Proposition A.1: Direct sampling from identified set $\mathcal{Q}$ with a single shock

Proposition A.1 (Direct draws from the admissible set.). *If* $\mathbf{z} \sim \mathcal{N}(0, \mathbf{I}_n)$ *and* $\mathbf{W}\mathbf{z} \leq 0$*, then* $\mathbf{q} = \frac{\mathbf{z}}{||\mathbf{z}||}$ *is a uniform draw from the unit n-sphere that satisfies* $\mathbf{W}\mathbf{q} \leq 0$*.*

Proof. Let $\tilde{\mathbf{z}} \sim \mathcal{N}(0, \mathbf{I}_n)$ such that $\mathbf{W}\tilde{\mathbf{z}} \leq 0$, where the inequality is elementwise. Let $\mathbf{H}$ be a given orthonormal matrix. Thus, $\tilde{\boldsymbol{\zeta}} = \mathbf{Q}\tilde{\mathbf{z}} \sim \mathcal{N}(0, \mathbf{Q}\mathbf{Q}') = \mathcal{N}(0, \mathbf{I}_n)$. Define g on $\mathbb{R}^n$ by $g(\mathbf{z}) = \frac{\mathbf{z}}{||\mathbf{z}||}$. Since

$||\mathbf{z}|| = \sqrt(\mathbf{z}'\mathbf{z})$, and $\mathbf{H}'\mathbf{H} = \mathbf{I}$, it follows that $g(\mathbf{Qz}) = \mathbf{Q}g(\mathbf{z})$. Let $\mathcal{A}$ be a Borel set on the σ-algebra on the unit n-sphere. Then:

$$\begin{aligned}
\frac{\Pr_{\tilde{\mathbf{z}}}\{g(\mathbf{z}) \in \mathcal{A} \cap \{\mathbf{z}|\mathbf{Wz} \le 0\}\}}{\Pr_{\tilde{\mathbf{z}}}\{\mathbf{Wz} \le 0\}} &= \frac{\Pr_{\tilde{\mathbf{z}}}\{g(\mathbf{z}) \in \mathcal{A} \cap \{\mathbf{z}|\mathbf{W}g(\mathbf{z}) \le 0\}\}}{\Pr_{\tilde{\mathbf{z}}}\{\mathbf{W}g(\mathbf{z}) \le 0\}} \\
&= \frac{\Pr_{\tilde{\mathbf{z}}}\{g(\boldsymbol{\zeta}) \in \mathcal{A} \cap \{\boldsymbol{\zeta}|Wg(\boldsymbol{\zeta}) \le 0\}\}}{\Pr_{\tilde{\mathbf{z}}}\{Wg(\boldsymbol{\zeta}) \le 0\}} \\
&= \frac{\Pr_{\mathbf{H}\tilde{\mathbf{z}}}\{\mathbf{H}g(\mathbf{z}) \in \mathcal{A} \cap \{\mathbf{Hz}|\mathbf{WH}g(\mathbf{z}) \le 0\}\}}{\Pr_{\mathbf{H}\tilde{\mathbf{z}}}\{\mathbf{WH}g(\mathbf{z}) \le 0\}}
\end{aligned}$$

The first equality follows because $||\mathbf{z}|| > 0$ with probability one, the second equality follows because $\tilde{\mathbf{z}}eta \stackrel{D}{=} \tilde{\mathbf{z}}$, and the third equality follows from substituting $\tilde{\mathbf{z}} = \mathbf{Q}\tilde{\mathbf{z}}$. From above and using the definition that $\mathbf{q} = g(\mathbf{z})$ and $\mathbf{Hq} = g(\mathbf{z}) = \mathbf{H}g(\mathbf{z})$, we also have that:

$$\begin{aligned}
&\frac{\Pr_{\tilde{\mathbf{z}}}\{g(\mathbf{z}) \in \mathcal{A} \cap \{\mathbf{z}|\mathbf{W}g(\mathbf{z}) \le 0\}\}}{\Pr_{\tilde{\mathbf{z}}}\{\mathbf{W}g(\mathbf{z}) \le 0\}} = \frac{\Pr_{\tilde{\zeta}}\{g(\boldsymbol{\zeta}) \in \mathcal{A} \cap \{\boldsymbol{\zeta}|\mathbf{W}g(\boldsymbol{\zeta}) \le 0\}\}}{\Pr_{\tilde{\zeta}}\{\mathbf{W}g(\boldsymbol{\zeta}) \le 0\}} \\
\Leftrightarrow \quad &\frac{\Pr_{\tilde{\mathbf{q}}}\{\mathbf{q} \in \mathcal{A} \cap \{\mathbf{q}|\mathbf{Wq} \le 0\}\}}{\Pr_{\tilde{\mathbf{q}}}\{\mathbf{Wq} \le 0\}} = \frac{\Pr_{\mathbf{H}\tilde{\mathbf{q}}}\{\mathbf{Hq} \in \mathcal{A} \cap \{\mathbf{Hq}|\mathbf{WHq} \le 0\}\}}{\Pr_{\mathbf{H}\tilde{\mathbf{q}}}\{\mathbf{WHq} \le 0\}}
\end{aligned}$$

Thus, the induced distribution is uniform on the truncated unit circle. □

B Characterization of identified sets in small scale VARs

B.1 Proposition B.1 (Bivariate VAR(0) with heterogeneity restrictions)

We impose two restrictions to identify the first shock. In a bivariate VAR, we can use 2.3 to express these restrictions as:

Standard sign restrictions	Heterogeneity restrictions	
$(\mathbf{r}^0_{\mathbf{a}})_1 \ge 0 \;\Leftrightarrow\; q_1\Sigma^{\text{tr}}_{1,1} \ge 0$	$(\mathbf{r}^0_{\mathbf{a}})_1 \ge 0 \;\Leftrightarrow\; q_1\Sigma^{\text{tr}}_{1,1} \ge 0$	(B.1a)
$(\mathbf{r}^0_{\mathbf{a}})_2 \ge 0 \;\Leftrightarrow\; q_1\Sigma^{\text{tr}}_{2,1} + q_2\Sigma^{\text{tr}}_{2,2} \ge 0$	$(\mathbf{r}^0_{\mathbf{a}})_2 \ge \lambda(\mathbf{r}^0_{\mathbf{a}})_1 \;\Leftrightarrow\; (q_1\Sigma^{\text{tr}}_{2,1} + q_2\Sigma^{\text{tr}}_{2,2}) - \lambda q_1\Sigma^{\text{tr}}_{1,1} \ge 0$	(B.1b)

Since the heterogeneity restriction nests the standard sign restriction for $\lambda = 0$, we now focus on this more general case.[1]

To analyze the implied restrictions, express the Cholesky factor $\boldsymbol{\Sigma}^{\text{tr}}$ in terms of the correlation and variances of the reduced-form errors: The elements of the Cholesky decomposition $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^{\text{tr}}(\boldsymbol{\Sigma}^{\text{tr}})'$ are: $\Sigma^{\text{tr}}_{1,1} = \sqrt{\Sigma_{11}}$, $\Sigma^{\text{tr}}_{2,1} = \frac{\Sigma_{21}}{\Sigma^{\text{tr}}_{1,1}} = \Sigma^{\text{tr}}_{2,2}\frac{\rho}{\sqrt{1-\rho^2}}$, $\Sigma^{\text{tr}}_{2,2} = \sqrt{\Sigma_{22} - (\Sigma^{\text{tr}}_{2,1})^2} = |\Sigma^{\text{tr}}_{2,1}|\sqrt{1/\rho^2 - 1}$. $\boldsymbol{\Sigma}$ is the covariance matrix of the forecast errors, and ρ is the reduced-form correlation between the forecast errors. We can then rewrite (B.1) as:

[1] An example of a VAR with one sign and one heterogeneity restriction 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 translates to the heterogeneity restriction that minus the quantities fall more than the prices within that industry.

$$q_1 \geq 0 \tag{B.2a}$$

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

In (q_1, q_2) space, q_2 lies in the plane above the ray through the origin with slope $-\frac{\rho}{\sqrt{1-\rho^2}}$ with $\lambda = 0$, i.e., with pure sign restrictions. The slope depends on correlation between the reduced-form forecast errors. Heterogeneity restrictions can ensure a positive slope for λ large enough. Intersecting the set described by (B.2) with the unit circle yields Figure 2.1, similar to Moon, Schorfheide, and Granziera (2013)

Proposition B.1 summarizes the results. Since $\Sigma^{\mathrm{tr}}_{11} = \sqrt{\Sigma_{11}}$ and $\Sigma^{\mathrm{tr}}_{21} = \frac{\Sigma_{21}}{\sqrt{\Sigma_{11}}}$, these restrictions depend only on the reduced-form variances and covariances.

Proposition B.1 (Set reduction of $IS(\cdot)$ under heterogeneity restrictions in bivariate VAR)**.** *The identified set for the structural impulse a_1 from (B.1) is strictly smaller under heterogeneity restrictions than under sign restrictions iff $\lambda\Sigma^{\mathrm{tr}}_{11} - \Sigma^{\mathrm{tr}}_{21} > 0$. The identified set for a_2 is strictly smaller unless $\lambda\Sigma^{\mathrm{tr}}_{11} = \Sigma^{\mathrm{tr}}_{21}$.*

Proof. This proposition follows directly from comparing the sets listed below for $\lambda = 0$ and $\lambda > 0$. Recall the restrictions:

$$\mathbf{a}_1 \equiv \mathbf{q}_1\Sigma^{\mathrm{tr}}_{11} \geq 0$$
$$\mathbf{a}_2 \equiv \mathbf{q}_1\Sigma^{\mathrm{tr}}_{21} + \mathbf{q}_2\Sigma^{\mathrm{tr}}_{22} \geq \lambda\mathbf{q}_1\Sigma^{\mathrm{tr}}_{11}$$

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

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

$$\mathbf{q}_1 = \frac{\Sigma^{\mathrm{tr}}_{22}}{\sqrt{(\Sigma^{\mathrm{tr}}_{22})^2 + ((\Sigma^{\mathrm{tr}}_{21})^2 - \lambda\Sigma^{\mathrm{tr}}_{11})^2}} \qquad \mathbf{q}_2 = \pm\frac{|\Sigma^{\mathrm{tr}}_{21} - \lambda\Sigma^{\mathrm{tr}}_{11}|}{\sqrt{(\Sigma^{\mathrm{tr}}_{22})^2 + ((\Sigma^{\mathrm{tr}}_{21})^2 - \lambda\Sigma^{\mathrm{tr}}_{11})^2}}$$

Case (a) $\Sigma^{\mathrm{tr}}_{21} \leq 0$.

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

Case (b) $\lambda\Sigma^{\mathrm{tr}}_{11} - \Sigma^{\mathrm{tr}}_{21} \leq 0, \Sigma^{\mathrm{tr}}_{21} \geq 0$.

- Upper bound for $\mathbf{a}_2$: $\mathbf{a}_2$ 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, $\mathbf{q}_1 = \frac{\Sigma^{\mathrm{tr}}_{21}}{\sqrt{(\Sigma^{\mathrm{tr}}_{22})^2 + (\Sigma^{\mathrm{tr}}_{21})^2}}$ and $\mathbf{q}_2 = \frac{\Sigma^{\mathrm{tr}}_{22}}{\Sigma^{\mathrm{tr}}_{21}}\mathbf{q}_1$. Thus: $\bar{\mathbf{a}}_2 = \sqrt{(\Sigma^{\mathrm{tr}}_{22})^2 + (\Sigma^{\mathrm{tr}}_{21})^2}$.
- Lower bound for $\mathbf{a}_2$: A negative $\mathbf{q}_2$ is now possible, but constrained by the heterogeneity constraint, as its RHS is increasing faster in $\mathbf{q}_1$ than its LHS. Thus, the lower bound is associated with a binding heterogeneity constraint and $\underline{a}_2 = \frac{\lambda\Sigma^{\mathrm{tr}}_{11}\Sigma^{\mathrm{tr}}_{22}}{\sqrt{(\Sigma^{\mathrm{tr}}_{22})^2 + ((\Sigma^{\mathrm{tr}}_{21})^2 - \lambda\Sigma^{\mathrm{tr}}_{11})^2}}$.

- Upper bound for $\mathbf{a}_1$: Since $\mathbf{q}_2 = 0, \mathbf{q}_1 = 1$ is possible, the upper bound is simply $\bar{\mathbf{a}}_1 = \Sigma^{\text{tr}}_{11}$.

Case (c) $\lambda\Sigma^{\text{tr}}_{11} - \Sigma^{\text{tr}}_{21} \geq 0, \Sigma^{\text{tr}}_{21} \geq 0$ or $0 \leq \rho \leq \lambda\frac{\sqrt{\mathbf{\Sigma}_{11}}}{\sqrt{\mathbf{\Sigma}_{22}}}$.

- Upper bound for $\mathbf{a}_2$: We proceed by brute force, checking whether the heterogeneity constrained is binding at the unconstrained maximum. We find that if $\lambda \leq \frac{(\Sigma^{\text{tr}}_{22})^2+(\Sigma^{\text{tr}}_{21})^2}{\Sigma^{\text{tr}}_{11}\Sigma^{\text{tr}}_{21}} = \frac{\mathbf{\Sigma}_{22}}{\mathbf{\Sigma}_{21}}$, the heterogeneity constraint is slack. Thus:

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

- Lower bound for $\mathbf{a}_2$: Since the interior extremum is always a maximum, we check the corners. Comparing the two corners, we find:

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

- Upper bound for $\mathbf{a}_1$: $\bar{\mathbf{a}}_1$ is also associated with the binding heterogeneity restriction:

$$\bar{\mathbf{a}}_1 = \frac{\Sigma^{\text{tr}}_{22}}{\sqrt{(\Sigma^{\text{tr}}_{22})^2 + ((\Sigma^{\text{tr}}_{21})^2 - \lambda\Sigma^{\text{tr}}_{11})^2}}\Sigma^{\text{tr}}_{11}.$$

$\square$

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 $[1, 0]\mathbf{y}_t$ and $[-\lambda, 1]\mathbf{y}_t$ replacing the correlation between variables 1 and 2.

B.2 Proposition B.2 (Trivariate VAR(0) with heterogeneity restrictions)

We now analyze when heterogeneity restrictions sharpen inference on variables that we do not directly restrict in the trivariate case. We show that 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.

$$(\mathbf{r}^0_{\mathbf{a}})_1 \geq 0 \;\Leftrightarrow\; q_1\Sigma^{\text{tr}}_{11} \geq 0 \tag{B.3a}$$

$$(\mathbf{r}^0_{\mathbf{a}})_2 \geq 0 \;\Leftrightarrow\; q_1\Sigma^{\text{tr}}_{21} + q_2\Sigma^{\text{tr}}_{22} \geq 0 \tag{B.3b}$$

$$(\mathbf{r}^0_{\mathbf{a}})_3 \geq \lambda(\mathbf{r}^0_{\mathbf{a}})_2 \;\Leftrightarrow\; q_1\Sigma^{\text{tr}}_{31} + q_2\Sigma^{\text{tr}}_{32} + q_3\Sigma^{\text{tr}}_{33} \geq \lambda(q_1\Sigma^{\text{tr}}_{21} + q_2\Sigma^{\text{tr}}_{22}) \tag{B.3c}$$

Proposition B.2 (Set reduction of $IS(\cdot)$ under heterogeneity restrictions in trivariate VAR)**.** *The identified set for the structural impulse* $\mathbf{a}_1$ *from* (B.3) *is strictly smaller under heterogeneity restrictions than under sign restrictions if* $\lambda\Sigma_{21}^{\text{tr}} > \Sigma_{31}^{\text{tr}}$ *and* $\Sigma_{31}^{\text{tr}} > 0$*. The identified set for* $\mathbf{a}_1$ *is equal under heterogeneity and sign restrictions if* $\lambda\Sigma_{21}^{\text{tr}} \leq \Sigma_{31}^{\text{tr}}$ *and* $\Sigma_{21}^{\text{tr}} > 0$*.*

Proof. The proposition follows from comparing the different cases below. **Identified set** Here we only consider bounds for $\mathbf{a}_1$. We seek a solution to the following problem:

$$\min_{\mathbf{q}} \text{ or } \max_{\mathbf{q}} \Sigma_{11}^{\text{tr}}\mathbf{q}_1 \tag{B.4a}$$

$$\text{s.t. } ||\mathbf{q}|| = 1 \tag{B.4b}$$

$$\Sigma_{11}^{\text{tr}}\mathbf{q}_1 \geq 0 \tag{B.4c}$$

$$\Sigma_{21}^{\text{tr}}\mathbf{q}_1 + \Sigma_{22}^{\text{tr}}\mathbf{q}_2 \geq 0$$

$$\underbrace{\left(\Sigma_{31}^{\text{tr}} - \lambda\Sigma_{21}^{\text{tr}}\right)}_{\equiv(\Sigma_{31}^{\text{tr}})^{\lambda}}\mathbf{q}_1 + \underbrace{\left(\Sigma_{32}^{\text{tr}} - \lambda\Sigma_{22}^{\text{tr}}\right)}_{\equiv(\Sigma_{32}^{\text{tr}})^{\lambda}}\mathbf{q}_2 + \Sigma_{33}^{\text{tr}}\mathbf{q}_3 \geq 0 \tag{B.4d}$$

Since $\Sigma_{ii}^{\text{tr}} > 0\forall i$, we can write equivalently:

$$\begin{aligned}
&\min_{\mathbf{q}} \text{ or } \max_{\mathbf{q}} \sqrt{1-(\mathbf{q}_2)^2-(\mathbf{q}_3)^3} \\
&\quad\text{s.t. } \Sigma_{21}^{\text{tr}}\sqrt{1-(\mathbf{q}_2)^2-(\mathbf{q}_3)^3} + \Sigma_{22}^{\text{tr}}\mathbf{q}_2 \geq 0 \\
&\quad\underbrace{\left(\Sigma_{31}^{\text{tr}} - \lambda\Sigma_{21}^{\text{tr}}\right)}_{\equiv(\Sigma_{31}^{\text{tr}})^{\lambda}}\sqrt{1-(\mathbf{q}_2)^2-(\mathbf{q}_3)^3} + \underbrace{\left(\Sigma_{32}^{\text{tr}} - \lambda\Sigma_{22}^{\text{tr}}\right)}_{\equiv(\Sigma_{32}^{\text{tr}})^{\lambda}}\mathbf{q}_2 + \Sigma_{33}^{\text{tr}}\mathbf{q}_3 \geq 0
\end{aligned}$$

Note that $\underline{\mathbf{a}}_1 = 0$ is always feasible by setting $\mathbf{q}_3 = 1$. 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

$$\begin{aligned}
\min_{\nu_{SR},\nu_{HR}} \max_{\mathbf{q}_2,\mathbf{q}_3} \mathcal{L} = \sqrt{1-(\mathbf{q}_2)^2-(\mathbf{q}_3)^3} &- \nu_{SR}(\Sigma_{21}^{\text{tr}}\sqrt{1-(\mathbf{q}_2)^2-(\mathbf{q}_3)^3} + \Sigma_{22}^{\text{tr}}\mathbf{q}_2) \\
&- \nu_{HR}((\Sigma_{31}^{\text{tr}})^{\lambda}\sqrt{1-(\mathbf{q}_2)^2-(\mathbf{q}_3)^3} + (\Sigma_{32}^{\text{tr}})^{\lambda}\mathbf{q}_2 + \Sigma_{33}^{\text{tr}}\mathbf{q}_3)
\end{aligned}$$

with the associated Kuhn-Tucker conditions as:

$$\begin{aligned}
[\mathbf{q}_2] &- \frac{\mathbf{q}_2}{\sqrt{1-(\mathbf{q}_2)^2-(\mathbf{q}_3)^3}}(1 - \nu_{SR}\Sigma_{21}^{\text{tr}} - \nu_{HR}(\boldsymbol{\Sigma}_{31}^{\text{tr},\lambda})) = \nu_{SR}\Sigma_{22}^{\text{tr}} + \nu_{HR}\boldsymbol{\Sigma}_{32}^{\text{tr},\lambda} \\
&\nu_{SR}(\Sigma_{21}^{\text{tr}}\sqrt{1-(\mathbf{q}_2)^2-(\mathbf{q}_3)^3} + \Sigma_{22}^{\text{tr}}\mathbf{q}_2) = 0 \\
&\nu_{SR} \geq 0 \\
[\nu_{SR}]&\Sigma_{21}^{\text{tr}}\sqrt{1-(\mathbf{q}_2)^2-(\mathbf{q}_3)^3} + \Sigma_{22}^{\text{tr}}\mathbf{q}_2 \geq 0. \\
[\mathbf{q}_3] &- \frac{\mathbf{q}_3}{\sqrt{1-(\mathbf{q}_2)^2-(\mathbf{q}_3)^3}}(1 - \nu_{SR}\Sigma_{21}^{\text{tr}} - \nu_{HR}(\Sigma_{31}^{\text{tr}})^{\lambda}) = \nu_{HR}\Sigma_{33}^{\text{tr}} \\
&\nu_{HR}(\boldsymbol{\Sigma}_{21}^{\text{tr},\lambda}\sqrt{1-(\mathbf{q}_2)^2-(\mathbf{q}_3)^3} + (\Sigma_{22}^{\text{tr}})^{\lambda}\mathbf{q}_2 + \Sigma_{33}^{\text{tr}}\mathbf{q}_3) = 0 \\
&\nu_{HR} \geq 0 \\
[\nu_{HR}]&(\Sigma_{31}^{\text{tr}})^{\lambda}\sqrt{1-(\mathbf{q}_2)^2-(\mathbf{q}_3)^3} + (\Sigma_{32}^{\text{tr}})^{\lambda}\mathbf{q}_2 + \Sigma_{33}^{\text{tr}}\mathbf{q}_3 \geq 0.
\end{aligned}$$

Clearly, the Kuhn-Tucker conditions show that the unconstrained optimum, when the multipliers ν_{SR}, ν_{HR} are zero, involves setting $\mathbf{q}_2 = \mathbf{q}_3 = 0$.

We assume throughout that $\lambda \geq 0$. For $\lambda = 0$, the heterogeneity restrictions become standard sign restrictions. We focus on the case of $\Sigma_{21}^{\text{tr}}, \Sigma_{31}^{\text{tr}} > 0$.

1. **All (conditional) covariances positive, heterogeneity restriction weak:**
   Note that when $0 \leq \Sigma_{21}^{\text{tr}}, \Sigma_{31}^{\text{tr}}$ and $\lambda\Sigma_{21}^{\text{tr}} \leq \Sigma_{31}^{\text{tr}}$, then $\mathbf{\Sigma}_{31}^{\text{tr},\lambda} \geq 0$ and $\mathbf{q}_2 = \mathbf{q}_3 = \nu_{SR} = \nu_{HR} = 0$ 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 $\lambda\Sigma_{21}^{\text{tr}} > \Sigma_{31}^{\text{tr}} > 0$, then $\mathbf{\Sigma}_{31}^{\text{tr},\lambda} < 0$. $\mathbf{q}_2 = \mathbf{q}_3 = \nu_{SR} = \nu_{HR} = 0$ no longer satisfies the optimality conditions with $\lambda > 0$, since the HR constraint is violated at this candidate point. With $\lambda = 0$, however, $\mathbf{q}_2 = \mathbf{q}_3 = 0$ is feasible, and the unconstrained maximum attains. The bound on $\mathbf{q}_1$ is thus strictly tighter with heterogeneity restrictions.

3. **Small negative conditional covariance, weak heterogeneity restriction:**
   When $\Sigma_{31}^{\text{tr}} < 0$, it follows that $\mathbf{\Sigma}_{31}^{\text{tr},\lambda} < 0$ 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:

$$\mathbf{q}_1^{HR} = \frac{(\mathbf{\Sigma}_{32}^{\text{tr},\lambda})^2 + (\Sigma_{33}^{\text{tr}})^2}{\sqrt{((\mathbf{\Sigma}_{32}^{\text{tr},\lambda})^2 + (\Sigma_{33}^{\text{tr}})^2)^2 + \Sigma_{33}^{\text{tr}}(\mathbf{\Sigma}_{31}^{\text{tr},\lambda})^2(1 + (\mathbf{\Sigma}_{32}^{\text{tr},\lambda}/\Sigma_{33}^{\text{tr}})^2)}}$$

   It can be shown that $\left.\frac{d\mathbf{q}_1^{HR}}{d\lambda}\right|_{\lambda=0} < 0$: Introducing heterogeneity restrictions tightens the upper bound. More generally, both restrictions or only the second restriction can bind if the optimum involves $\mathbf{q}_2 < 0$. The solution for $\mathbf{q}_1$ is $\mathbf{q}_1 = \max\{\mathbf{q}_1^{HR}, \mathbf{q}_1^{HR,SR}, \mathbf{q}_1^{SR}\}$ where:

$$\mathbf{q}_1^{HR,SR} = \frac{\Sigma_{33}^{\text{tr}}}{\sqrt{((\Sigma_{31}^{\text{tr}} - (\Sigma_{21}^{\text{tr}}/\Sigma_{22}^{\text{tr}})\Sigma_{32}^{\text{tr}})^2 + (\Sigma_{33}^{\text{tr}})^2)^2 + (\Sigma_{33}^{\text{tr}})^2(1 + (\Sigma_{21}^{\text{tr}}/\Sigma_{22}^{\text{tr}})^2)}}$$
$$\mathbf{q}_1^{SR} = \frac{1}{\sqrt{1 + (\Sigma_{21}^{\text{tr}}/\Sigma_{22}^{\text{tr}})^2}}.$$

□

The intuition from Proposition B.1 also explains Proposition B.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 of the reduced-form correlation.

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

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

B.3 Proposition B.3 (Bivariate VAR(1) with slope restrictions)

We now discuss slope restrictions as another type of ranking restrictions. Unlike heterogeneity restrictions which may replace standard sign restrictions, we impose this restriction in addition to the standard sign restrictions.

To consider slope restrictions, we have to introduce dynamics. We do so in the simplest case possible, a bivariate VAR(1):

$$\mathbf{y}_t = \mathbf{B}\mathbf{y}_{t-1} + \mathbf{A}\boldsymbol{\varepsilon}_t, \qquad \mathbf{A} = \boldsymbol{\Sigma}^{\mathrm{tr}}[\mathbf{q}, \mathbf{q}_\perp].$$

We analyze three sign restrictions and one ranking restriction. For $\lambda = 1$, the ranking restriction requires the response of the second variable to the identified shock to increase initially.

$$\begin{array}{llcr}
[SR_1] & (\mathbf{r}^0_{\mathbf{a}})_1 > 0 & \Leftrightarrow & \mathbf{e}_1\boldsymbol{\Sigma}^{\mathrm{tr}}\mathbf{q} > 0 \\
[SR_2] & (\mathbf{r}^0_{\mathbf{a}})_2 > 0 & \Leftrightarrow & \mathbf{e}_2\boldsymbol{\Sigma}^{\mathrm{tr}}\mathbf{q} > 0 \\
[SR_3] & (\mathbf{r}^1_{\mathbf{a}})_2 > 0 & \Leftrightarrow & \mathbf{e}_2\mathbf{B}\boldsymbol{\Sigma}^{\mathrm{tr}}\mathbf{q} > 0 \\
[RR] & (\mathbf{r}^1_{\mathbf{a}})_2 > \lambda(\mathbf{r}^0_{\mathbf{a}})_2 & \Leftrightarrow & \mathbf{e}_2\mathbf{B}\boldsymbol{\Sigma}^{\mathrm{tr}}\mathbf{q} > \lambda \times \mathbf{e}_2\boldsymbol{\Sigma}^{\mathrm{tr}}\mathbf{q}
\end{array}$$

The question whether slope restrictions sharpen inference is only meaningful when the identified set $\mathcal{Q}$ is nonempty absent slope restrictions. The crux is SR_3. It can be satisfied if the second variable is persistent ($B_{22} > 0$) or if tends to oscillate ($B_{22} < 0$), but there is positive feedback ($B_{21} > 0$). Below, we assume either condition holds and $\lambda > 0$.

Proposition B.3 (Set reduction with slope restrictions)**.** *If $B_{22} > 0$, the slope restriction binds if $B_{21} < 0$. If $B_{22} < \lambda$, this implies that the identified set for a_1 is strictly smaller with the slope restriction. If $B_{22} < 0$ and $B_{21} > 0$, the identified set for a_1 is also strictly smaller with slope restrictions.*

Proof. To arrive at these different cases, first multiply out the matrices and vectors to rewrite the restrictions as:

$$\begin{array}{lrcl}
[SR_1] & \Sigma^{\mathrm{tr}}_{11}q_1 \geq 0 & \Rightarrow & q_1 \geq 0. \\
[SR_2] & \Sigma^{\mathrm{tr}}_{21}q_1 + \Sigma^{\mathrm{tr}}_{22}q_2 \geq 0 & \Rightarrow & q_2 \geq -\dfrac{\Sigma^{\mathrm{tr}}_{21}}{\Sigma^{\mathrm{tr}}_{22}}q_1. \\
[SR_3] & (b_{22}\Sigma^{\mathrm{tr}}_{21} + b_{21}\Sigma^{\mathrm{tr}}_{11})q_1 + b_{22}\Sigma^{\mathrm{tr}}_{22}q_2 \geq 0. & \Rightarrow & q_2 \begin{cases} \geq -\frac{\Sigma^{\mathrm{tr}}_{21}}{\Sigma^{\mathrm{tr}}_{22}}q_1 - \frac{b_{21}}{b_{22}}\frac{\Sigma^{\mathrm{tr}}_{11}}{\Sigma^{\mathrm{tr}}_{22}}q_1 & b_{22} > 0 \\ \leq -\frac{\Sigma^{\mathrm{tr}}_{21}}{\Sigma^{\mathrm{tr}}_{22}}q_1 - \frac{b_{21}}{b_{22}}\frac{\Sigma^{\mathrm{tr}}_{11}}{\Sigma^{\mathrm{tr}}_{22}}q_1 & b_{22} < 0. \end{cases} \\
[RR] & ((b_{22}-\lambda)\Sigma^{\mathrm{tr}}_{21} + b_{21}\Sigma^{\mathrm{tr}}_{11})q_1 + (b_{22}-\lambda)\Sigma^{\mathrm{tr}}_{22}q_2 \geq 0. & \Rightarrow & q_2 \begin{cases} \geq -\frac{\Sigma^{\mathrm{tr}}_{21}}{\Sigma^{\mathrm{tr}}_{22}}q_1 - \frac{b_{21}}{b_{22}-\lambda}\frac{\Sigma^{\mathrm{tr}}_{11}}{\Sigma^{\mathrm{tr}}_{22}}q_1 & b_{22} > \lambda \\ \leq -\frac{\Sigma^{\mathrm{tr}}_{21}}{\Sigma^{\mathrm{tr}}_{22}}q_1 - \frac{b_{21}}{b_{22}-\lambda}\frac{\Sigma^{\mathrm{tr}}_{11}}{\Sigma^{\mathrm{tr}}_{22}}q_1 & b_{22} < \lambda. \end{cases}
\end{array}$$

As Figure B.1 shows, comparing the different cases yields:

- If $B_{22} > \lambda > 0$ and $B_{21} < 0$, then there is a strict set reduction for q_2 because the lower bound on q_2 increases. This leads to a strict set reduction in the identified set for the impact response a_1 if $\Sigma^{\mathrm{tr}}_{2,1} < 0$ or $0 \leq \frac{\Sigma^{\mathrm{tr}}_{21}}{\Sigma^{\mathrm{tr}}_{22}}q_1 < \frac{-B_{21}}{B_{22}-\lambda}\frac{\Sigma^{\mathrm{tr}}_{11}}{\Sigma^{\mathrm{tr}}_{22}}$.

- If $B_{22} > \lambda > 0$ and $B_{21} > 0$, there is no set reduction because $[RR]$ is less restrictive than $[SR_2]$.
- If $0 \leq B_{22} < \lambda$ and $B_{21} < 0$, the identified set is empty because the $[RR]$ upper bound lies below the $[SR_2]$ lower bound.
- If $0 \leq B_{22} < \lambda$ and $B_{21} > 0$ then there is a strict set reduction for q_2 because there is an upper bound for q_2 – and the identified set is non-empty because the upper bound lies above the lower bound implied by $[SR_2]$. By bounding q_2 from above, this case always bounds q_1 from below and thus leads to a strict set reduction for q_1 and thus for the identified set for the impact response a_1.
- If $B_{22}, B_{21} < 0$, then the ranking restriction is always slack. If $\Sigma_{21}^{\mathrm{tr}} > 0$, the identified set is empty already because $[SR_2]$ and $[SR_3]$ are incompatible.

$\square$

The ranking restriction binds if there is negative feedback between the two variables but persistence (or a tendency to oscillate and positive feedback). Since we require both impulse-responses to increase on impact, the negative feedback limits at least one of the two impulse-responses. The negative feedback from the first variable by itself depresses the response of the second variable over time. When $B_{22} < \lambda$, the persistence is weak enough relative to the feedback effect that the restriction shrinks the identified set for q_1 and thus the identified set for the impulse-response of the first variable. Figure B.1 illustrates the identified set $\mathcal{Q}$.

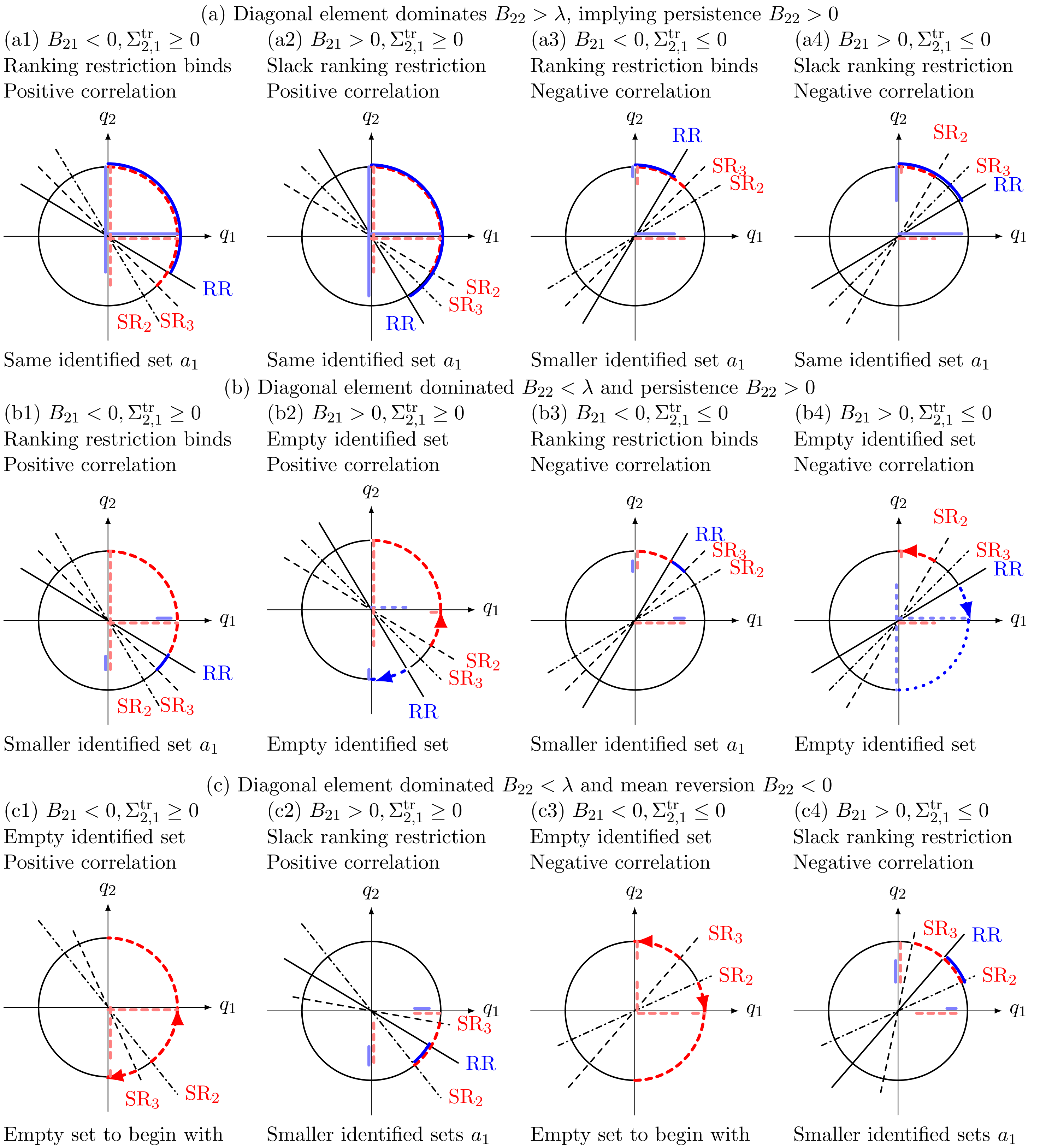


Figure B.1: Identified sets for q_1, q_2 with slope and sign restrictions.

C Forecast error variance decomposition

The total forecast error variance (FEV) for $\mathbf{y}_{t+H}$ given information up to time t is given by:

$$FEV_H = \sum_{h=0}^{H}((\mathbf{B}_X^h \boldsymbol{\Sigma}^{\text{tr}})(\mathbf{B}_X^h \boldsymbol{\Sigma}^{\text{tr}})').$$

We can decompose the FEV into the contribution due to an identified shock with impulse-vector $\boldsymbol{\Sigma}^{\text{tr}} q$. We call this the conditional forecast error variance (CFEV):

$$CFEV_H(\mathbf{q}) = \sum_{h=0}^{H}((\mathbf{B}_X^h \boldsymbol{\Sigma}^{\text{tr}} \mathbf{q})(\mathbf{B}_X^h \boldsymbol{\Sigma}^{\text{tr}} \mathbf{q})').$$

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:

$$CFEV_{i,\underline{H},\bar{H}}(\mathbf{q}) = \sum_{h=\underline{H}}^{\bar{H}} \sum_{k=0}^{h}((\mathbf{B}_X^k \boldsymbol{\Sigma}^{\text{tr}} \mathbf{q})(\mathbf{B}_X^k \boldsymbol{\Sigma}^{\text{tr}} \mathbf{q})')_{(ii)} = \mathbf{q}' \mathbf{S}_{i,\underline{H},\bar{H}} \mathbf{q}, \tag{C.1}$$

$$\mathbf{S}_{i,\underline{H},\bar{H}} \equiv \sum_{h=0}^{\bar{H}} (\bar{H} + 1 - \max\{\underline{H}, h\})(\mathbf{e}_i \mathbf{B}_X^h \boldsymbol{\Sigma}^{\text{tr}})'(\mathbf{e}_i \mathbf{B}_X^h \boldsymbol{\Sigma}^{\text{tr}}). \tag{C.2}$$

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

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

D Data and empirical results

D.1 VAR with heterogeneity restrictions and micro 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): See `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`. 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.

Application 1: Industry data and heterogeneity restrictions

Variable	Level	Impact	2 years	6 years	Min	Median	Max
Output	98	13.5	30.7	29.5	10.4	29.0	35.2
	90	13.6	35.4	25.8	13.6	29.9	36.5
	68	16.8	39.2	31.9	16.3	35.7	42.2
TFP	98	12.4	20.8	24.0	11.2	24.3	27.1
	90	12.1	23.5	31.1	12.1	27.7	31.4
	68	13.0	27.0	37.3	13.0	30.6	38.4
Confidence	98	14.0	33.6	27.4	14.0	23.2	37.2
	90	14.0	36.3	28.7	14.0	28.7	39.1
	68	15.4	39.1	29.9	15.4	29.9	42.1
Employment	98	12.3	32.8	26.7	12.2	28.5	38.7
	90	13.0	35.4	24.6	13.0	29.1	37.9
	68	16.0	38.4	27.6	15.0	30.7	42.1

Table D.1: Reduction of prior-robust credible set for macro variables in the nine variable VAR with ranking restrictions

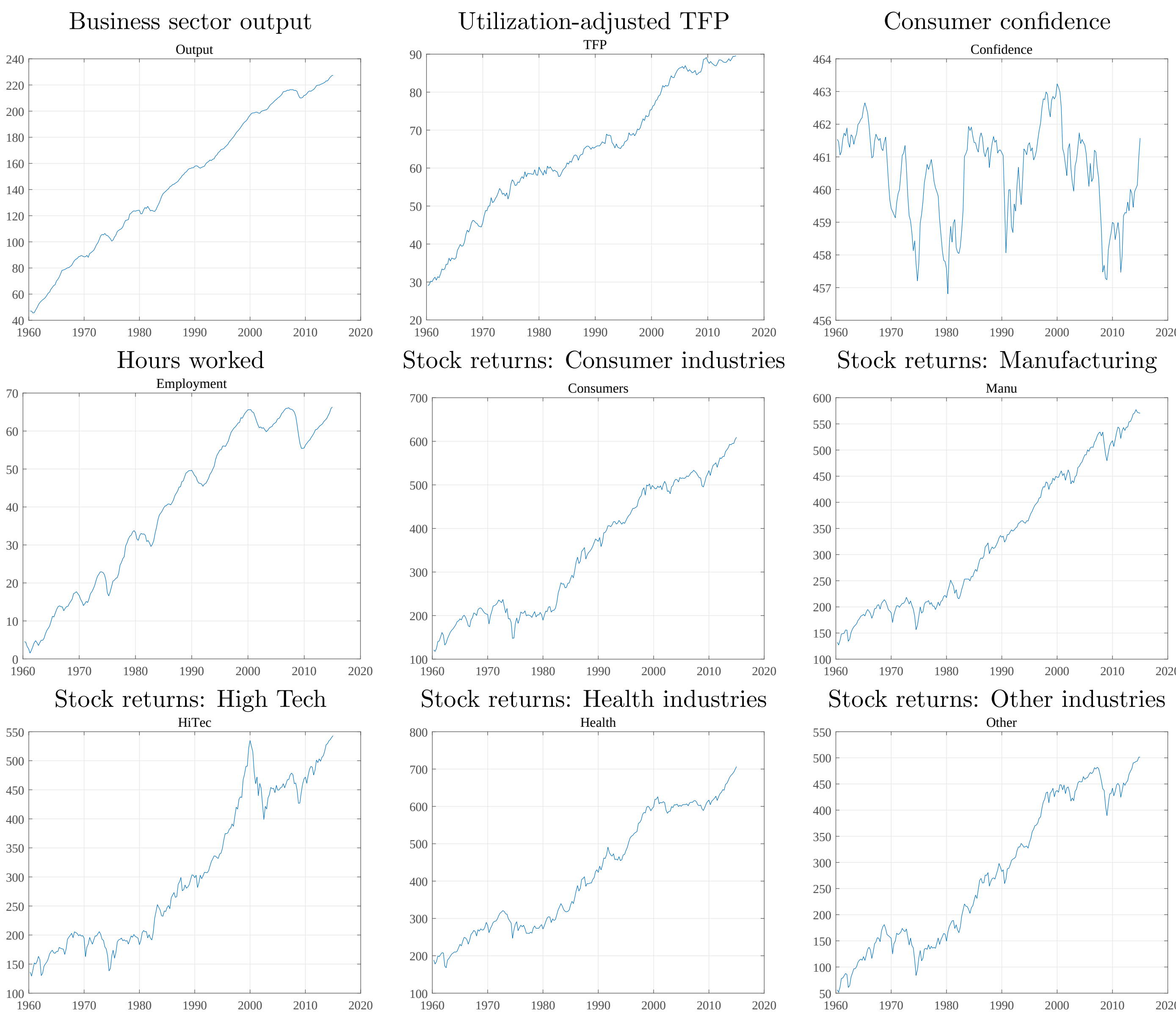


Figure D.1: Raw data: News application with five Fama-French industries

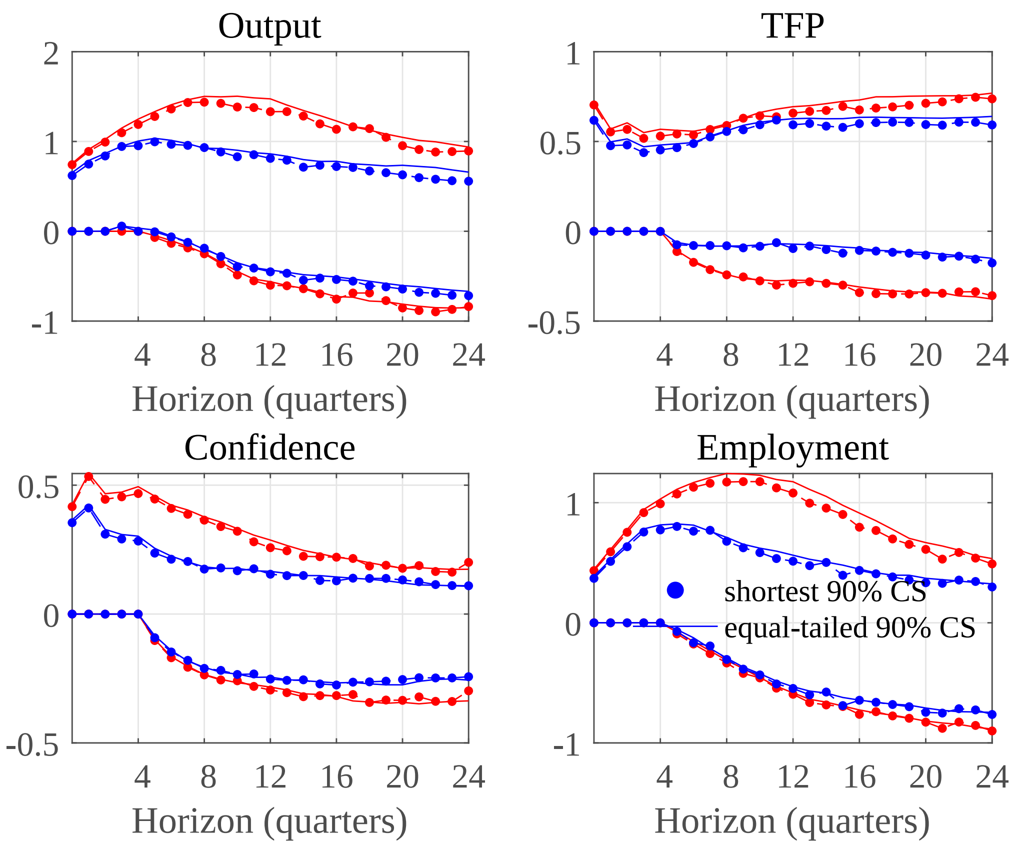
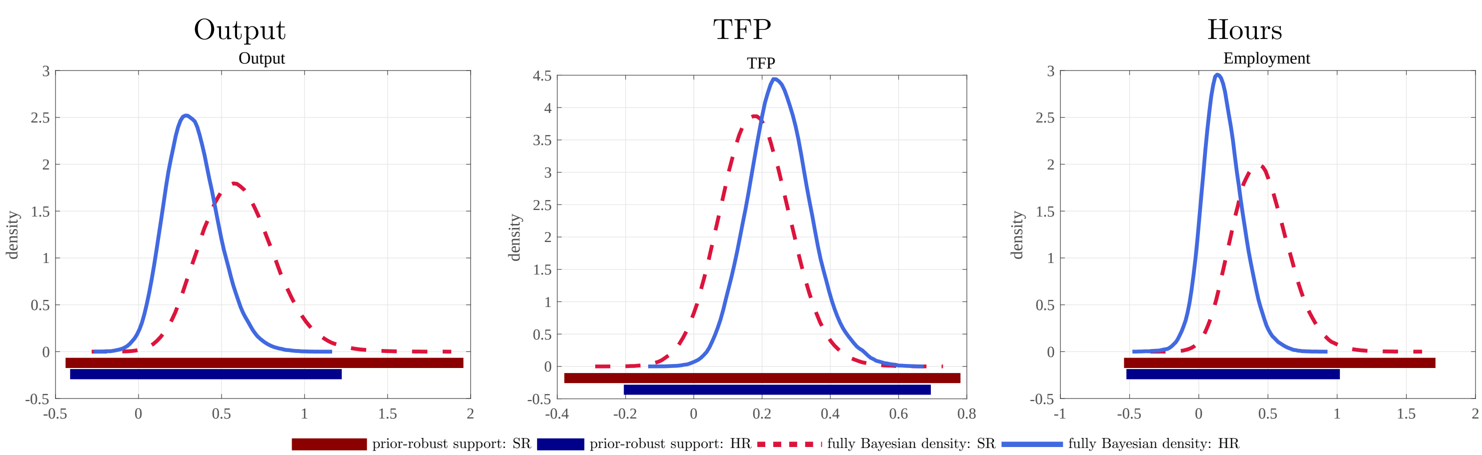


Figure D.2: Comparing shortest and equal-tailed prior-robust credible sets



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 lead to less dispersed distributions of responses and for TFP shift mass away from zero.

Figure D.3: Distribution of the two-year-responses of macro variables to a one standard deviation productivity news shock.

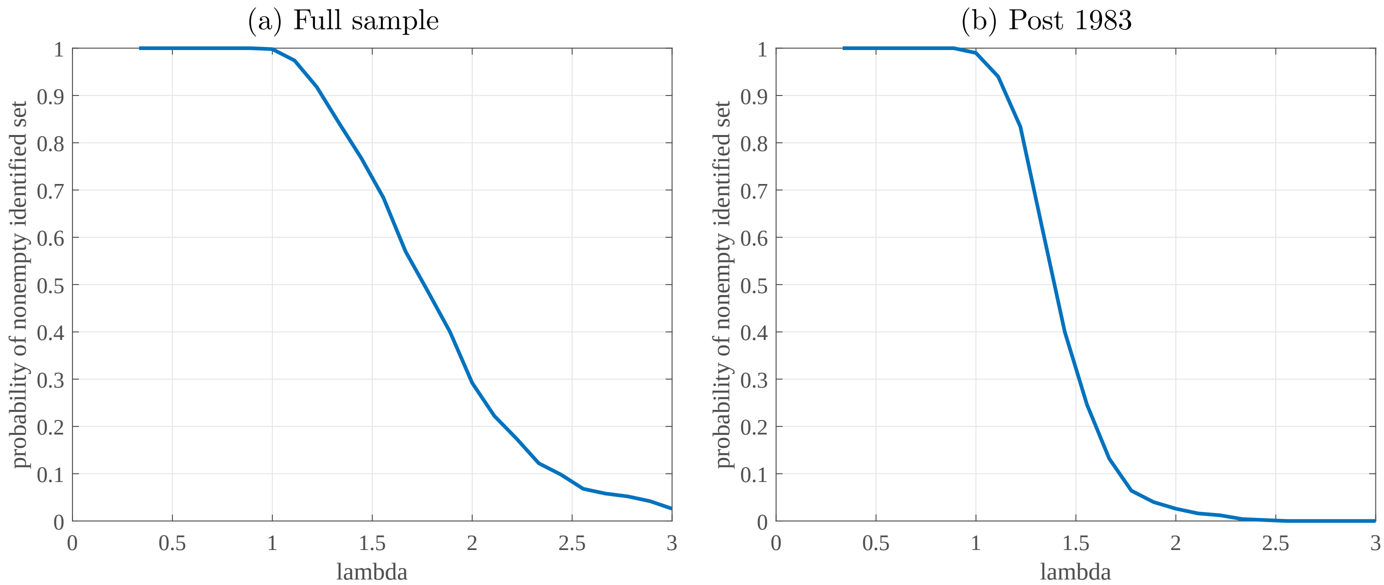


Figure D.4: Posterior plausibility of heterogeneity restrictions as a function of the intensity parameter λ

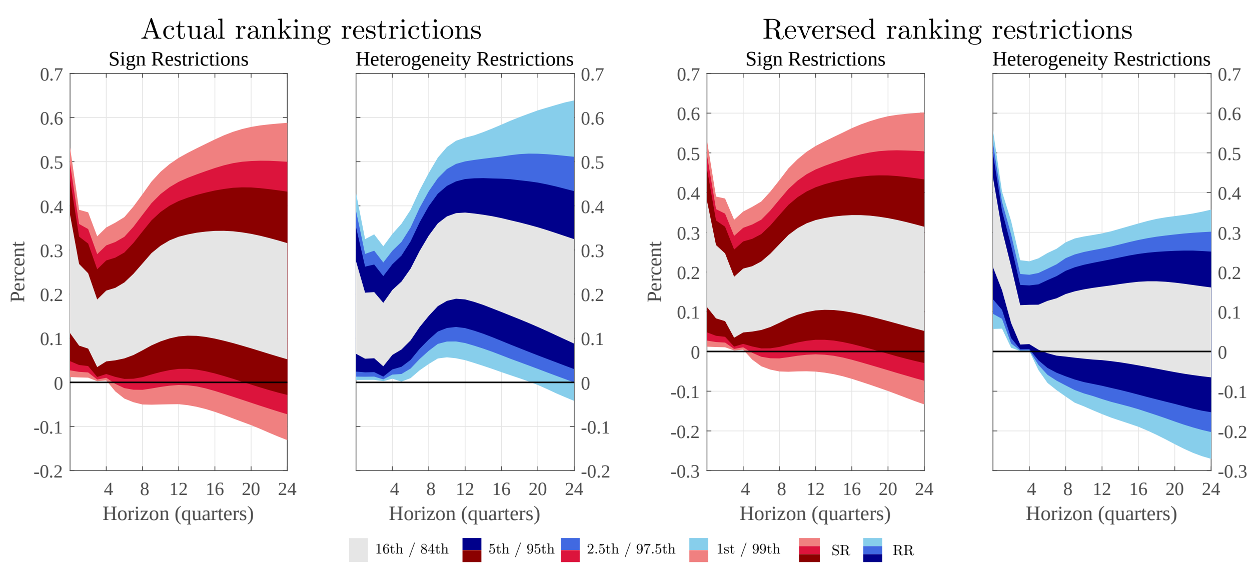


Figure D.5: Fully Bayesian responses of TFP with the actual ranking restrictions and reversed ranking restrictions.

D.2 VAR with slope restrictions and only macro data

We use the following macro variables:

- The average of business sector GDP and GDI: BEA via Fernald (2014) (accumulated growth rates).
- Utilization adjusted TFP: Fernald (2014) (accumulated growth rates).
- The real SP500 index from Robert Shiller's website `http://www.econ.yale.edu/~shiller/data.htm`.
- CPI price index from Robert Shiller's website `http://www.econ.yale.edu/~shiller/data.htm`.
- The 2-year Treasury constant maturity rate `GS2` from the St. Louis Fed FRED website.

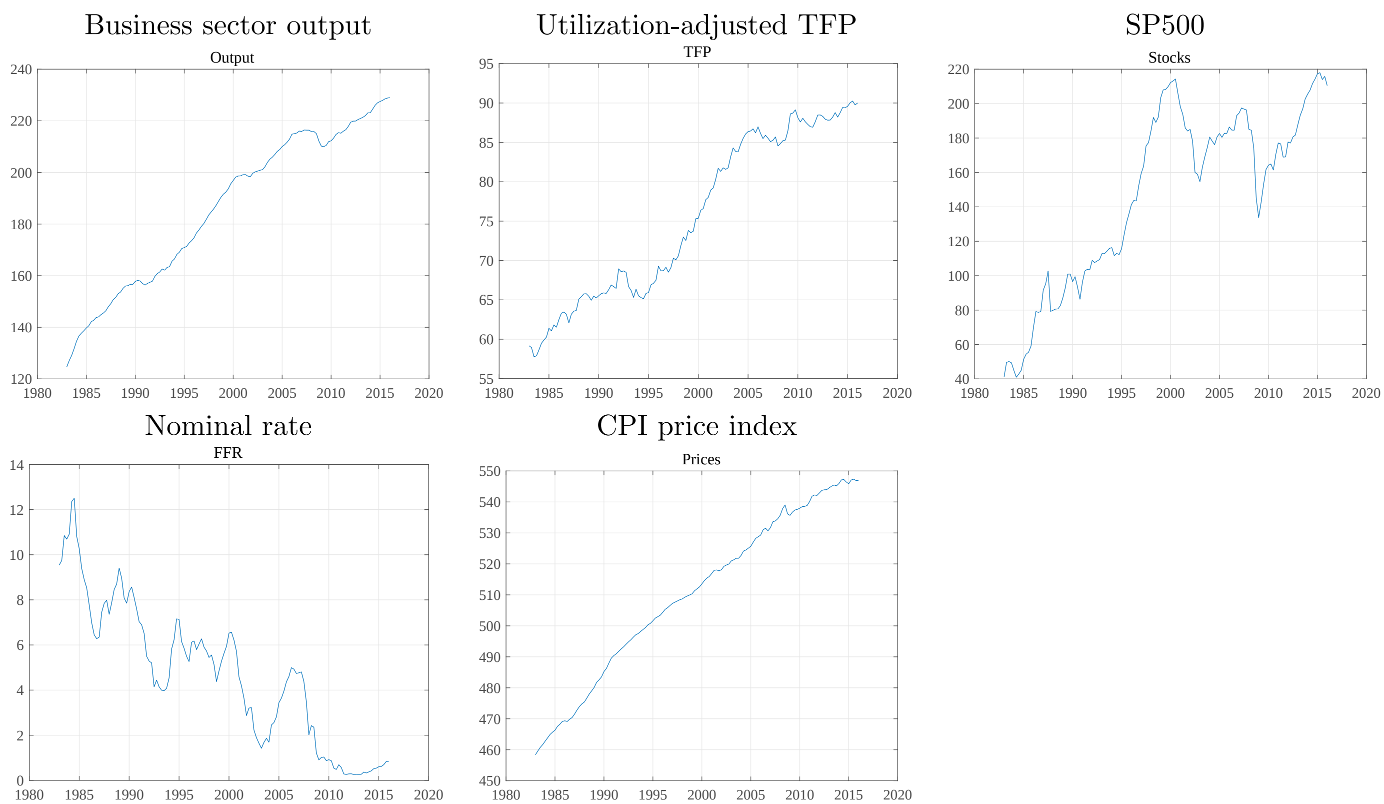


Figure D.6: Raw data: 5-variable VAR news shock application

Application 2: Macro data and slope restrictions							
Variable	Level	Impact	2 years	6 years	Min	Median	Max
Output	98	9.5	19.9	14.4	4.2	18.5	22.7
	90	14.6	22.9	27.3	7.9	23.0	28.6
	68	18.2	33.6	28.7	12.8	32.4	34.2
TFP	98	10.6	11.8	25.1	0.4	15.2	25.1
	90	11.3	18.2	27.4	1.4	23.9	28.5
	68	20.3	21.8	34.2	5.7	32.9	36.7
Stocks	98	24.3	17.5	-0.2	-0.2	11.4	24.3
	90	29.8	17.1	17.9	9.9	16.6	29.8
	68	38.7	26.0	24.4	15.9	24.6	38.7
FFR	98	7.3	12.8	11.7	4.8	12.8	18.3
	90	10.4	20.1	18.2	10.4	16.9	26.5
	68	16.6	23.8	25.1	16.6	22.7	30.8
Prices	98	14.0	12.5	21.4	9.6	19.5	28.8
	90	25.4	22.9	20.7	19.5	24.4	30.9
	68	36.4	29.4	25.4	25.4	30.1	38.9

Table D.2: Reduction of prior-robust credible set for macro variables in the five-variable VAR with slope restrictions

References

Border, K. C. (2013): "Alternative Linear Inequalities," Discussion paper, Caltech Division of the Social Sciences.

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

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

Granziera, E., H. Moon, and F. Schorfheide (2018): "Inference for VARs Identified with Sign Restrictions," Quantitative Economics.

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

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