Supplement to:

Innovations meet Narratives

–improving power and credibility in macro inference–

Régis Barnichon[(a)] *and* *Geert Mesters*[(b)]

[(a)] Federal Reserve Bank of San Francisco and CEPR

[(b)] Universitat Pompeu Fabra, Barcelona School of Economics and CREI

January 16, 2025

Abstract

In this web-appendix we present

S1. Supporting lemmas

S2. Deterministic selection

S3. Compatible underlying structural models.

S4. Details simulation study

S5. Additional empirical results

JEL classification: C14, C32, E32, E52, N10.

Keywords: Narrative methods, Time series, Prediction powered, Monetary policy.

Throughout this document, references to lemmas, equations etc. which start with a "S" are references to this document. Those which consist of just a number refer to the main text.

S1 Supporting Lemmas

Lemma S1. *Given Assumptions 1-2 and Assumption R we have that*

$$\widehat{\boldsymbol{\delta}} \xrightarrow{p} \boldsymbol{\delta} \ ,$$

and if Assumption I also holds we have for any $\boldsymbol{\gamma}_n \xrightarrow{p} \boldsymbol{\gamma}$ that

$$\widehat{\boldsymbol{\psi}}(\boldsymbol{\gamma}_n) \xrightarrow{p} \boldsymbol{\psi} \ .$$

Proof. For both claims we verify the conditions of Theorem 2.1 in Newey and McFadden (1994). Starting with part 1. Assumption R-1 ensures that the ergodic theorem implies that $\frac{1}{n}\sum_{t=1}^{n} m(\mathbf{q}_t; \tilde{\boldsymbol{\delta}}) \xrightarrow{p} \mathbb{E}m(\mathbf{d}_t, \tilde{\boldsymbol{\delta}})$. Since by Assumption R-4-(i) this expectation is only zero at $\boldsymbol{\delta}$, $\|\mathbb{E}m(\mathbf{d}_t, \tilde{\boldsymbol{\delta}})\|^2$ is uniquely minimized at $\boldsymbol{\delta}$. By R-2 the parameter space Δ is compact. By R-4-(ii) $m(\mathbf{d}_t, \tilde{\boldsymbol{\delta}})$ is continuous with respect to $\tilde{\boldsymbol{\delta}}$. Further,

$$\mathbb{E}\left(\sup_{\tilde{\boldsymbol{\delta}}\in\Delta} \|m(\mathbf{q}_t; \tilde{\boldsymbol{\delta}})\|\right) \leq \mathbb{E}\left(\sup_{\tilde{\boldsymbol{\delta}}\in\Delta} \|\boldsymbol{\kappa}^{(1)}(\mathbf{q}_t; \tilde{\boldsymbol{\delta}})\|\right)(1+c_2)/(c_1(1-c_2)) < \infty \ . \tag{S1}$$

which follows from R-4-(iii),(iv). Combining the previous two statements Lemma 2.4 in Newey and McFadden (1994) allows us to conclude that $\mathbb{E}m(\mathbf{d}_t, \tilde{\boldsymbol{\delta}})$ is continuous and uniform convergence holds $\sup_{\tilde{\boldsymbol{\delta}}\in\Delta} \|\frac{1}{n}\sum_{t=1}^{n} m(\mathbf{q}_t; \tilde{\boldsymbol{\delta}}) - \mathbb{E}m(\mathbf{d}_t, \tilde{\boldsymbol{\delta}})\| \xrightarrow{p} 0$. This implies that

$$\sup_{\tilde{\boldsymbol{\delta}}\in\Delta} |\frac{1}{n}\sum_{t=1}^{n} \|m(\mathbf{d}_t, \tilde{\boldsymbol{\delta}})\|^2 - \|\mathbb{E}m(\mathbf{d}_t, \tilde{\boldsymbol{\delta}})\|^2| \xrightarrow{p} 0 \ ,$$

and since $\widehat{\boldsymbol{\delta}} = \operatorname{argmin}_{\tilde{\boldsymbol{\delta}}\in\Delta} \frac{1}{n}\sum_{t=1}^{n} \|m(\mathbf{d}_t, \tilde{\boldsymbol{\delta}})\|^2$ we may apply Newey and McFadden (1994, Theorem 2.1) to conclude $\widehat{\boldsymbol{\delta}} \xrightarrow{p} \boldsymbol{\delta}$.

Next for part 2. Assumption R-1 ensures that the ergodic theorem implies that $h_n(\tilde{\boldsymbol{\psi}}) \xrightarrow{p} \mathbb{E}h(\mathbf{d}_t, \tilde{\boldsymbol{\psi}})$. Further, $\boldsymbol{\gamma}_n \xrightarrow{p} \boldsymbol{\gamma}$ implies that $\mathbf{W}_n \xrightarrow{p} \mathbf{W}$ where

$$\mathbf{W}_n = \begin{pmatrix} 1 & \boldsymbol{\gamma}_n' & 0 \\ \boldsymbol{\gamma}_n & \boldsymbol{\gamma}_n\boldsymbol{\gamma}_n' & 0 \\ 0 & 0 & \mathbf{I}_{d_v} \end{pmatrix} \qquad \text{and} \qquad \mathbf{W} = \begin{pmatrix} 1 & \boldsymbol{\gamma}' & 0 \\ \boldsymbol{\gamma} & \boldsymbol{\gamma}\boldsymbol{\gamma}' & 0 \\ 0 & 0 & \mathbf{I}_{d_v} \end{pmatrix} \ ,$$

which are positive semi-definite. From Lemma 2.3 in Newey and McFadden (1994) it follows that if (a) $\mathbb{E}h(\mathbf{d}_t, \boldsymbol{\psi}) = 0$ and (b) $\mathbf{W}\mathbb{E}h(\mathbf{d}_t, \tilde{\boldsymbol{\psi}}) \neq 0$ for all $\tilde{\boldsymbol{\psi}} \neq \boldsymbol{\psi}$, then $Q(\tilde{\boldsymbol{\psi}}) = \mathbb{E}h(\mathbf{d}_t, \tilde{\boldsymbol{\psi}})'\mathbf{W}\mathbb{E}h(\mathbf{d}_t, \tilde{\boldsymbol{\psi}})$ has a unique minimum at $\boldsymbol{\psi}$.

For (a) note that by the law of iterated expectations

$$\mathbb{E}h(\mathbf{d}_t, \boldsymbol{\psi}) = \mathbb{E}\begin{pmatrix} z_t^{\perp}\xi_{t+h}^{\perp}\mathbb{E}(s_t|\mathbf{q}_t, z_t^{\perp})/\kappa(\mathbf{q}_t;\boldsymbol{\delta}) \\ \mathbf{v}_t\xi_{t+h}^{\perp}(1-\mathbb{E}(s_t|\mathbf{q}_t, z_t^{\perp})/\kappa(\mathbf{q}_t;\boldsymbol{\delta})) \\ \frac{\boldsymbol{\kappa}^{(1)}(\mathbf{q}_t;\boldsymbol{\delta})(\mathbb{E}(s_t|\mathbf{q}_t, z_t^{\perp})-\kappa(\mathbf{q}_t;\boldsymbol{\delta}))}{\kappa(\mathbf{q}_t;\boldsymbol{\delta})(1-\kappa(\mathbf{q}_t;\boldsymbol{\delta}))} \end{pmatrix} = 0 \ , \tag{S2}$$

where the second equality follows as $\mathbb{E}(s_t|\mathbf{q}_t, z_t^{\perp}) = \mathbb{E}(s_t|\mathbf{q}_t) = \kappa(\mathbf{q}_t;\boldsymbol{\delta})$ by Assumption 2, making the second and third entries zero. Further, by Assumption 1 we have $\mathbb{E}(z_t^{\perp}\xi_{t+h}^{\perp}) = 0$ ensuring that the first entry is also zero.

For (b) note that this is implied when $\mathbb{E}g(\mathbf{d}_t, \tilde{\boldsymbol{\psi}};\boldsymbol{\gamma}) \neq 0$ and $\mathbb{E}m(\mathbf{q}_t;\tilde{\boldsymbol{\delta}}) \neq 0$ for $\tilde{\boldsymbol{\psi}} \neq \boldsymbol{\psi}$, where g and m are defined in (14) and (15). Assumption R-4-(i) imposes $\mathbb{E}m(\mathbf{q}_t, \tilde{\boldsymbol{\delta}}) \neq 0$ for $\tilde{\boldsymbol{\delta}} \neq \boldsymbol{\delta}$, which can be verified for any preferred binary response model. Hence, for $\mathbb{E}g(\mathbf{d}_t, \tilde{\boldsymbol{\psi}}, \boldsymbol{\gamma})$ we only need to consider the case where $\tilde{\boldsymbol{\delta}} = \boldsymbol{\delta}$, i.e.

$$\begin{aligned}\mathbb{E}g(\mathbf{d}_t, \tilde{\theta}_h, \boldsymbol{\delta};\boldsymbol{\gamma}) =& \mathbb{E}(z_t^{\perp}\tilde{\xi}_{t+h}^{\perp}s_t/\kappa(\mathbf{q}_t;\boldsymbol{\delta})) + \boldsymbol{\gamma}'\mathbb{E}(\mathbf{v}_t\tilde{\xi}_{t+h}^{\perp}(1-s_t/\kappa(\mathbf{q}_t;\boldsymbol{\delta}))) \\ =& \mathbb{E}(z_t^{\perp}\tilde{\xi}_{t+h}^{\perp}\mathbb{E}(s_t|\mathbf{q}_t, z_t^{\perp})/\kappa(\mathbf{q}_t;\boldsymbol{\delta})) + \boldsymbol{\gamma}'\mathbb{E}(\mathbf{v}_t\tilde{\xi}_{t+h}^{\perp}(1-\mathbb{E}(s_t|\mathbf{q}_t)/\kappa(\mathbf{q}_t;\boldsymbol{\delta}))) \\ =& \mathbb{E}(z_t^{\perp}\tilde{\xi}_{t+h}^{\perp}) \ ,\end{aligned}$$

where $\tilde{\xi}_{t+h}^{\perp} = y_{t+h}^{\perp} - \tilde{\theta}_h p_t^{\perp} = \xi_{t+h}^{\perp} + (\theta_h - \tilde{\theta}_h)p_t^{\perp}$ and $\mathbb{E}(z_t^{\perp}\tilde{\xi}_{t+h}^{\perp}) = \mathbb{E}(z_t^{\perp}\xi_{t+h}^{\perp}) + (\theta_h - \tilde{\theta}_h)\mathbb{E}(z_t^{\perp}p_t^{\perp}) = (\theta_h - \tilde{\theta}_h)\mathbb{E}(z_t^{\perp}p_t^{\perp}) \neq 0$ unless $\theta_h = \tilde{\theta}_h$, which follows from Assumptions 1 and I. We conclude that the limiting objective function $Q(\tilde{\boldsymbol{\psi}})$ is uniquely minimized at $\boldsymbol{\psi}$.

Next, the parameter space for $\boldsymbol{\psi}$ is compact by Assumption R-2. Finally, to verify that $Q(\tilde{\boldsymbol{\psi}})$ is continuous with respect to $\tilde{\boldsymbol{\psi}}$ and that $\sup_{\tilde{\boldsymbol{\psi}}\in\Psi}|h_n(\tilde{\boldsymbol{\psi}})'\mathbf{W}_n h_n(\tilde{\boldsymbol{\psi}}) - Q(\tilde{\boldsymbol{\psi}})| \xrightarrow{p} 0$ we first check the conditions of Lemma 2.4 in Newey and McFadden (1994)[S1] to verify that $\mathbb{E}h(\mathbf{d}_t, \tilde{\boldsymbol{\psi}})$ is continuous with respect to $\tilde{\boldsymbol{\psi}}$ and $\sup_{\tilde{\boldsymbol{\psi}}\in\Psi}|h_n(\tilde{\boldsymbol{\psi}}) - \mathbb{E}h(\mathbf{d}_t, \tilde{\boldsymbol{\psi}})| \xrightarrow{p} 0$.

First, $h(\mathbf{d}_t, \tilde{\boldsymbol{\psi}})$ is continuous at each $\tilde{\boldsymbol{\psi}} \in \Psi$ with probability one. This follows as the functions $z_t^{\perp}(y_{t+h}^{\perp} - \theta_h p_t^{\perp})s_t/\kappa(\mathbf{q}_t;\boldsymbol{\delta})$ and $\mathbf{v}_t(y_{t+h}^{\perp} - \theta_h p_t^{\perp})(1-s_t/\kappa(\mathbf{q}_t;\boldsymbol{\delta}))$ are continuous in any $\tilde{\theta}_h \in \Theta$ (linear) and in any $\tilde{\boldsymbol{\delta}} \in \Delta$ as κ is continuous by R-4-(ii). The function $m(\mathbf{q}_t, \tilde{\boldsymbol{\delta}})$ is continuous at each $\tilde{\boldsymbol{\delta}} \in \Delta$ as it is a continuous function of κ and $\kappa^{(1)}$ which are both continuous by R-4-(ii).

[S1] Note that their lemma applies after replacing the iid assumption for the data with Assumption R-1.

Second, we verify the dominance condition $\mathbb{E}(\sup_{\tilde{\boldsymbol{\psi}}\in\Psi}\|h(\mathbf{d}_t,\tilde{\boldsymbol{\psi}})\|)<\infty$. Note that

$$\mathbb{E}\left(\sup_{\tilde{\boldsymbol{\psi}}\in\Psi}\|h(\mathbf{d}_t,\tilde{\boldsymbol{\psi}})\|\right) \leq \mathbb{E}\left(\sup_{\tilde{\boldsymbol{\psi}}\in\Psi}|z_t^{\perp}(y_{t+h}^{\perp}-\tilde{\theta}_h p_t^{\perp})s_t/\kappa(\mathbf{q}_t;\tilde{\boldsymbol{\delta}})|\right) + \mathbb{E}\left(\sup_{\tilde{\boldsymbol{\psi}}\in\Psi}\|\mathbf{v}_t(y_{t+h}^{\perp}-\tilde{\theta}_h p_t^{\perp})(1-s_t/\kappa(\mathbf{q}_t;\tilde{\boldsymbol{\delta}}))\|\right) + \mathbb{E}\left(\sup_{\tilde{\boldsymbol{\delta}}\in\Delta}\|m(\mathbf{q}_t;\tilde{\boldsymbol{\delta}})\|\right)$$

Using Assumption R-4-(iii) the first term can be bounded as follows

$$\mathbb{E}\left(\sup_{\tilde{\boldsymbol{\psi}}\in\Psi}|z_t^{\perp}(y_{t+h}^{\perp}-\tilde{\theta}_h p_t^{\perp})s_t/\kappa(\mathbf{q}_t;\tilde{\boldsymbol{\delta}})|\right) \leq c_1^{-1}\mathbb{E}\left(\sup_{\tilde{\theta}_h\in\Theta}|z_t^{\perp}(y_{t+h}^{\perp}-\tilde{\theta}_h p_t^{\perp})|\right) \leq c_1^{-1}(\mathbb{E}|z_t^{\perp}\xi_{t+h}^{\perp}| + \sup_{\tilde{\theta}_h\in\Theta}|\tilde{\theta}_h-\theta_h|\mathbb{E}|z_t^{\perp}p_t^{\perp}|)<\infty \tag{S3}$$

where the last inequality follows from R-2 and R-3. Similarly, for the second term

$$\mathbb{E}\left(\sup_{\tilde{\boldsymbol{\psi}}\in\Psi}\|\mathbf{v}_t(y_{t+h}^{\perp}-\tilde{\theta}_h p_t^{\perp})(1-s_t/\kappa(\mathbf{q}_t;\tilde{\boldsymbol{\delta}}))\|\right) \leq (1+c_1^{-1})\mathbb{E}\left(\sup_{\tilde{\theta}_h\in\Theta}\|\mathbf{v}_t(y_{t+h}^{\perp}-\tilde{\theta}_h p_t^{\perp})\|\right) \leq (1+c_1^{-1})(\mathbb{E}\|\mathbf{v}_t\xi_{t+h}^{\perp}\| + \sup_{\tilde{\theta}_h\in\Theta}|\tilde{\theta}_h-\theta_h|\mathbb{E}\|\mathbf{v}_t p_t^{\perp}\|)<\infty\ . \tag{S4}$$

Finally, the last term follows from (S1).

This completes the verification of the conditions of Lemma 2.4. We use these to conclude that $Q(\tilde{\boldsymbol{\psi}})$ is continuous with respect to $\tilde{\boldsymbol{\psi}}$ and since

$$|h_n(\tilde{\boldsymbol{\psi}})'\mathbf{W}_n h_n(\tilde{\boldsymbol{\psi}})-Q(\tilde{\boldsymbol{\psi}})| \leq \|h_n(\tilde{\boldsymbol{\psi}})-\mathbb{E}h(\mathbf{d}_t,\tilde{\boldsymbol{\psi}})\|^2\|\mathbf{W}_n\| + 2\|\mathbb{E}h(\mathbf{d}_t,\tilde{\boldsymbol{\psi}})\|\|h_n(\tilde{\boldsymbol{\psi}})-\mathbb{E}h(\mathbf{d}_t,\tilde{\boldsymbol{\psi}})\|\|\mathbf{W}_n\|\|\mathbb{E}h(\mathbf{d}_t,\tilde{\boldsymbol{\psi}})\|^2\|\mathbf{W}_n-\mathbf{W}\|$$

the uniform convergence of $h_n(\tilde{\boldsymbol{\psi}})-\mathbb{E}h(\mathbf{d}_t,\tilde{\boldsymbol{\psi}})$, the dominance condition and $\mathbf{W}_n-\mathbf{W}\xrightarrow{p}0$ imply the uniform convergence of the objective function.

This completes the verification of the conditions of Theorem 2.1 in Newey and McFadden (1994) and we may conclude that $\widehat{\boldsymbol{\psi}}(\boldsymbol{\gamma}_n)\xrightarrow{p}\boldsymbol{\psi}$ which implies $\widehat{\boldsymbol{\delta}}\xrightarrow{p}\boldsymbol{\delta}$ and $\hat{\theta}_h(\boldsymbol{\gamma}_n)\xrightarrow{p}\theta_h$. □

Lemma S2. *Given Assumptions 1-2 and Assumption R, we have that*

(i) $\sup_{\tilde{\boldsymbol{\psi}}\in N}\|\widehat{\mathbf{G}}_{\delta z}(\tilde{\boldsymbol{\psi}})-\mathbb{E}\widehat{\mathbf{G}}_{\delta z}(\tilde{\boldsymbol{\psi}})\|\xrightarrow{p}0$ *and* $\mathbb{E}\widehat{\mathbf{G}}_{\delta z}(\tilde{\boldsymbol{\psi}})$ *is continuous in* $\tilde{\boldsymbol{\psi}}$.

(ii) $\sup_{\tilde{\boldsymbol{\psi}}\in N}\|\widehat{\mathbf{G}}_{\delta v}(\tilde{\boldsymbol{\psi}})-\mathbb{E}\widehat{\mathbf{G}}_{\delta v}(\tilde{\boldsymbol{\psi}})\|\xrightarrow{p}0$ *and* $\mathbb{E}\widehat{\mathbf{G}}_{\delta v}(\tilde{\boldsymbol{\psi}})$ *is continuous in* $\tilde{\boldsymbol{\psi}}$.

(iii) $\sup_{\tilde{\boldsymbol{\psi}}\in N} \|\widehat{\mathbf{M}}(\tilde{\boldsymbol{\psi}}) - \mathbb{E}\widehat{\mathbf{M}}(\tilde{\boldsymbol{\psi}})\| \xrightarrow{p} 0$ *and* $\mathbb{E}\widehat{\mathbf{M}}(\tilde{\boldsymbol{\psi}})$ *is continuous in* $\tilde{\boldsymbol{\psi}}$.

(iv) $\sup_{\tilde{\boldsymbol{\psi}}\in N} \|\mathbf{H}_n(\tilde{\boldsymbol{\psi}}) - \mathbb{E}\mathbf{H}(\mathbf{d}_t, \boldsymbol{\psi})\| \xrightarrow{p} 0$ *and* $\mathbb{E}\mathbf{H}(\mathbf{d}_t, \boldsymbol{\psi})$ *is continuous in* $\tilde{\boldsymbol{\psi}}$.

where N *is some neighborhood of* $\boldsymbol{\psi}$ *and*

$$
\begin{aligned}
\widehat{\mathbf{G}}_{\delta z}(\tilde{\boldsymbol{\psi}}) &= -\frac{1}{n}\sum_{t\in\mathcal{N}} s_t z_t^{\perp}(y_{t+h}^{\perp} - \tilde{\theta}_h p_t^{\perp})\kappa^{(1)}(\mathbf{d}_t^{\pi}; \tilde{\boldsymbol{\delta}})/\kappa^2(\mathbf{d}_t^{\pi}, \tilde{\boldsymbol{\delta}}) \ , \\
\widehat{\mathbf{G}}_{\delta v}(\tilde{\boldsymbol{\psi}})' &= \quad \frac{1}{n}\sum_{t\in\mathcal{N}} s_t \mathbf{v}_t(y_{t+h}^{\perp} - \tilde{\theta}_h p_t^{\perp})\kappa^{(1)}(\mathbf{d}_t^{\pi}; \tilde{\boldsymbol{\delta}})'/\kappa^2(\mathbf{d}_t^{\pi}, \tilde{\boldsymbol{\delta}}) \ , \\
\widehat{\mathbf{M}}(\tilde{\boldsymbol{\psi}}) &= -\frac{1}{n}\sum_{t\in\mathcal{N}} \boldsymbol{\kappa}^{(1)}(\mathbf{d}_t^{\pi}; \tilde{\boldsymbol{\delta}})\boldsymbol{\kappa}^{(1)}(\mathbf{d}_t^{\pi}; \tilde{\boldsymbol{\delta}})'/(\kappa(\mathbf{d}_t^{\pi}; \tilde{\boldsymbol{\delta}})(1 - \kappa(\mathbf{d}_t^{\pi}; \tilde{\boldsymbol{\delta}}))) \\
\mathbf{H}_n(\boldsymbol{\psi}) &= \frac{1}{n}\sum_{t=1}^{n} \mathbf{H}(\mathbf{d}_t, \boldsymbol{\psi})
\end{aligned}
$$

with $\mathbf{H}(\mathbf{d}_t, \boldsymbol{\psi}) = \frac{\partial h(\mathbf{d}_t, \boldsymbol{\psi})}{\partial \boldsymbol{\psi}'}$,

$$
\mathbf{H}(\mathbf{d}_t, \boldsymbol{\psi}) = \begin{pmatrix} -z_t^{\perp} p_t^{\perp} s_t/\kappa(\mathbf{q}_t; \boldsymbol{\delta}) & -z_t^{\perp}(y_{t+h}^{\perp} - \theta_h p_t^{\perp}) s_t \boldsymbol{\kappa}^{(1)}(\mathbf{q}_t; \boldsymbol{\delta})'/\kappa^2(\mathbf{q}_t; \boldsymbol{\delta}) \\ -\mathbf{v}_t p_t^{\perp}(1 - s_t/\kappa(\mathbf{q}_t; \boldsymbol{\delta})) & \mathbf{v}_t(y_{t+h}^{\perp} - \theta_h p_t^{\perp}) s_t \boldsymbol{\kappa}^{(1)}(\mathbf{q}_t; \boldsymbol{\delta})'/\kappa^2(\mathbf{q}_t; \boldsymbol{\delta}) \\ 0 & \partial m(\mathbf{q}_t, \boldsymbol{\delta})/\partial \boldsymbol{\delta}' \end{pmatrix}
$$

and

$$
\begin{aligned}
\partial m(\mathbf{q}_t, \boldsymbol{\delta})/\partial \boldsymbol{\delta}' =& \boldsymbol{\kappa}^{(2)}(\mathbf{q}_t, \boldsymbol{\delta})(s_t - \kappa(\mathbf{q}_t, \boldsymbol{\delta}))/(\kappa(\mathbf{q}_t, \boldsymbol{\delta})(1 - \kappa(\mathbf{q}_t, \boldsymbol{\delta}))) \\
& - \boldsymbol{\kappa}^{(1)}(\mathbf{q}_t, \boldsymbol{\delta})\boldsymbol{\kappa}^{(1)}(\mathbf{q}_t, \boldsymbol{\delta})'(1 - 2\boldsymbol{\kappa}(\mathbf{q}_t, \boldsymbol{\delta}))(s_t - \kappa(\mathbf{q}_t, \boldsymbol{\delta}))/(\kappa(\mathbf{q}_t, \boldsymbol{\delta})(1 - \kappa(\mathbf{q}_t, \boldsymbol{\delta})))^2 \\
& - \boldsymbol{\kappa}^{(1)}(\mathbf{q}_t, \boldsymbol{\delta})\boldsymbol{\kappa}^{(1)}(\mathbf{q}_t, \boldsymbol{\delta})'/(\kappa(\mathbf{q}_t, \boldsymbol{\delta})(1 - \kappa(\mathbf{q}_t, \boldsymbol{\delta}))) \ .
\end{aligned}
$$

Proof. To prove (i)-(iv) we repeatedly apply Lemma 2.4 from Newey and McFadden (1994) which requires checking that the argument of the sum is continuous with respect to $\tilde{\boldsymbol{\psi}}$ and that the argument is dominated in expectation. Part (i). Note that by Assumption R-1:

$$
\mathbb{E}\widehat{\mathbf{G}}_{\delta z}(\tilde{\boldsymbol{\psi}}) = -\mathbb{E} s_t z_t^{\perp}(y_{t+h}^{\perp} - \tilde{\theta}_h p_t^{\perp})\kappa^{(1)}(\mathbf{d}_t^{\pi}; \tilde{\boldsymbol{\delta}})/\kappa^2(\mathbf{d}_t^{\pi}, \tilde{\boldsymbol{\delta}}) \ ,
$$

and $s_t z_t^{\perp}(y_{t+h}^{\perp} - \tilde{\theta}_h p_t^{\perp})\kappa^{(1)}(\mathbf{d}_t^{\pi}; \tilde{\boldsymbol{\delta}})/\kappa^2(\mathbf{d}_t^{\pi}, \tilde{\boldsymbol{\delta}})$ is continuous with respect to $\tilde{\boldsymbol{\delta}}$ on Δ by R-4-(ii). Also,

$$
\begin{aligned}
&\mathbb{E}\left(\sup_{\tilde{\boldsymbol{\psi}}\in N} \|z_t^{\perp}(y_{t+h}^{\perp} - \tilde{\theta}_h p_t^{\perp}) s_t \boldsymbol{\kappa}^{(1)}(\mathbf{q}_t; \tilde{\boldsymbol{\delta}})'/\kappa^2(\mathbf{q}_t; \tilde{\boldsymbol{\delta}})\|\right) \leq \\
&c_1^{-2}\mathbb{E}(\sup_{\tilde{\boldsymbol{\delta}}\in N} \|\boldsymbol{\kappa}^{(1)}(\mathbf{q}_t; \tilde{\boldsymbol{\delta}})\|^2)^{1/2} \times \mathbb{E}(\sup_{\tilde{\theta}_h\in N} (|z_t^{\perp}\xi_{t+h}^{\perp}| + |\tilde{\theta}_h - \theta_h||z_t^{\perp} p_t^{\perp}|)^2)^{1/2} < \infty \qquad \text{(S5)}
\end{aligned}
$$

by R-2,3,4-(iii,v). Part (ii). We have

$$\mathbb{E}\widehat{\mathbf{G}}_{\delta v}(\tilde{\boldsymbol{\psi}})' = \mathbb{E}s_t\mathbf{v}_t(y^{\perp}_{t+h} - \tilde{\theta}_h p^{\perp}_t)\kappa^{(1)}(\mathbf{d}^{\pi}_t;\tilde{\boldsymbol{\delta}})'/\kappa^2(\mathbf{d}^{\pi}_t,\tilde{\boldsymbol{\delta}})\ .$$

Continuity follows from the same argument as above and

$$\begin{aligned}
&\mathbb{E}\left(\sup_{\tilde{\boldsymbol{\psi}}\in N}\|\mathbf{v}_t(y^{\perp}_{t+h} - \tilde{\theta}_h p^{\perp}_t)s_t\boldsymbol{\kappa}^{(1)}(\mathbf{q}_t;\tilde{\boldsymbol{\delta}})'/\kappa^2(\mathbf{q}_t;\tilde{\boldsymbol{\delta}})\|\right) \leq \\
&c_1^{-2}\mathbb{E}(\sup_{\tilde{\boldsymbol{\delta}}\in N}\|\boldsymbol{\kappa}^{(1)}(\mathbf{q}_t;\tilde{\boldsymbol{\delta}})\|^2)^{1/2} \times \mathbb{E}(\sup_{\tilde{\theta}_h\in N}(\|\mathbf{v}_t\xi^{\perp}_{t+h}\| + |\tilde{\theta}_h - \theta_h|\|\mathbf{v}_t p^{\perp}_t\|)^2)^{1/2} < \infty
\end{aligned} \tag{S6}$$

by R-2,3,4-(iii,v). Part (iii). We have

$$\mathbb{E}\widehat{\mathbf{M}}(\tilde{\boldsymbol{\psi}}) = \mathbb{E}\boldsymbol{\kappa}^{(1)}(\mathbf{d}^{\pi}_t;\tilde{\boldsymbol{\delta}})\boldsymbol{\kappa}^{(1)}(\mathbf{d}^{\pi}_t;\tilde{\boldsymbol{\delta}})'/(\kappa(\mathbf{d}^{\pi}_t;\tilde{\boldsymbol{\delta}})(1-\kappa(\mathbf{d}^{\pi}_t;\tilde{\boldsymbol{\delta}})))$$

and

$$\begin{aligned}
&\mathbb{E}\left(\sup_{\tilde{\boldsymbol{\psi}}\in N}\|\boldsymbol{\kappa}^{(1)}(\mathbf{d}^{\pi}_t;\tilde{\boldsymbol{\delta}})\boldsymbol{\kappa}^{(1)}(\mathbf{d}^{\pi}_t;\tilde{\boldsymbol{\delta}})'/(\kappa(\mathbf{d}^{\pi}_t;\tilde{\boldsymbol{\delta}})(1-\kappa(\mathbf{d}^{\pi}_t;\tilde{\boldsymbol{\delta}})))\|\right) \leq \\
&\mathbb{E}\left(\sup_{\tilde{\boldsymbol{\delta}}\in N}\|\boldsymbol{\kappa}^{(1)}(\mathbf{q}_t;\tilde{\boldsymbol{\delta}})\|^2\right)/(c_1(1-c_2)) < \infty
\end{aligned} \tag{S7}$$

by R-4-(iii,v). Finally, for part (iv) note that $\mathbf{H}(\mathbf{d}_t,\boldsymbol{\psi})$ is dominated in expectation if its sub-blocks are dominated. The top and middle right blocks are dominated by (S5) and (S6). For the remaining terms we have

$$\mathbb{E}\left(\sup_{\tilde{\boldsymbol{\psi}}\in N}|z^{\perp}_t p^{\perp}_t s_t/\kappa(\mathbf{q}_t;\tilde{\boldsymbol{\delta}})|\right) \leq c_1^{-1}\mathbb{E}(|z^{\perp}_t p^{\perp}_t|) < \infty \tag{S8}$$

$$\mathbb{E}\left(\sup_{\tilde{\boldsymbol{\psi}}\in N}\|\mathbf{v}_t p^{\perp}_t(1 - s_t/\kappa(\mathbf{q}_t;\tilde{\boldsymbol{\delta}}))\|\right) \leq (1+c_1^{-1})\mathbb{E}(\|\mathbf{v}_t p^{\perp}_t\|) < \infty \tag{S9}$$

$$\begin{aligned}
&\mathbb{E}\left(\sup_{\tilde{\boldsymbol{\psi}}\in N}\|\nabla_{\tilde{\boldsymbol{\delta}}'}m(\mathbf{q}_t,\tilde{\boldsymbol{\delta}})\|\right) \leq \mathbb{E}\left(\sup_{\tilde{\boldsymbol{\delta}}\in N}\|\boldsymbol{\kappa}^{(2)}(\mathbf{q}_t;\tilde{\boldsymbol{\delta}})\|\right)(1+c_2)/(c_1(1-c_2)) \\
&+\mathbb{E}\left(\sup_{\tilde{\boldsymbol{\delta}}\in N}\|\boldsymbol{\kappa}^{(1)}(\mathbf{q}_t;\tilde{\boldsymbol{\delta}})\|^2\right)(2+3c_2+2c_2^2)(c_1(1-c_2))^{-1} < \infty
\end{aligned} \tag{S10}$$

where we used R-2,3,4-(ii),(iii),(v). The continuity of $\mathbf{H}(\mathbf{d}_t,\boldsymbol{\psi})$ follows from R-4-(ii) and conventional arguments. This allows to apply Lemma 2.4 of Newey and McFadden (1994) to conclude that (iv) holds. □

Lemma S3. *Given Assumptions 1-2 and R we have that*

$$\sqrt{n}h_n(\boldsymbol{\psi}) \xrightarrow{d} N(0, \mathbf{S}) \ .$$

Proof. Consider any $\boldsymbol{\lambda} \in \mathbb{R}^{\dim(h)}$ with $\boldsymbol{\lambda}'\boldsymbol{\lambda}$. Under assumption R-1 have that $\boldsymbol{\lambda}'h(\mathbf{d}_t, \boldsymbol{\psi})$ is stationary by White (2000, Theorem 3.35) and α-mixing of size $-r/(r-2)$ White (2000, Theorem 3.49). Also, by (S2), $\mathbb{E}(\boldsymbol{\lambda}'h(\mathbf{d}_t, \boldsymbol{\psi})) = 0$. Consider the moment condition

$$\begin{aligned}
\mathbb{E}|\boldsymbol{\lambda}'h(\mathbf{d}_t, \boldsymbol{\psi})|^{2+v} &\leq \mathbb{E}|\lambda_1 z_t^{\perp}\xi_{t+h}^{\perp} s_t/k(\mathbf{q}_t; \boldsymbol{\delta})|^{2+v} + \mathbb{E}|\boldsymbol{\lambda}_2'\mathbf{v}_t\xi_{t+h}^{\perp}(1 - s_t/k(\mathbf{q}_t; \boldsymbol{\delta}))|^{2+v} \\
&\quad + \mathbb{E}|\boldsymbol{\lambda}_3'\boldsymbol{\kappa}^{(1)}(\mathbf{q}_t; \boldsymbol{\delta})(s_t - \kappa(\mathbf{q}_t; \boldsymbol{\delta}))/(\kappa(\mathbf{q}_t; \boldsymbol{\delta})(1 - \kappa(\mathbf{q}_t; \boldsymbol{\delta})))|^{2+v} \\
&\leq |\lambda_1|^{2+v} c_1^{-2-v}\mathbb{E}|z_t^{\perp}\xi_{t+h}^{\perp}|^{2+v} + (1 + c_1^{-2-v})\|\boldsymbol{\lambda}_2\|^{2+v}\mathbb{E}\|\mathbf{v}_t\xi_{t+h}^{\perp}\|^{2+v} \\
&\quad + \mathbb{E}\|\boldsymbol{\kappa}^{(1)}(\mathbf{q}_t; \boldsymbol{\delta})\|^{2+v}\|\boldsymbol{\lambda}_3\|^{2+v}(1 + c_2)^{2+v}(c_1(1 - c_2))^{-2-v} < \infty \ ,
\end{aligned}$$

by Assumptions R-3,4-(iii),(vi). Also, note that by R-5 we have $\lim_{n\to\infty} \text{var}(\sqrt{n}\boldsymbol{\lambda}'h_n(\boldsymbol{\psi})) > 0$ as $\mathbf{S}$ is positive definite. These conditions are sufficient to apply the central limit theorem for α-mixing processes as stated in White (2000, Theorem 5.20) to $\sqrt{n}\boldsymbol{\lambda}'h_n(\boldsymbol{\psi})$. Subsequently by the Cramer-Wold device (White, 2000, Proposition 2.1) we have $\sqrt{n}h_n(\boldsymbol{\psi}) \xrightarrow{d} N(0, \mathbf{S})$.

□

S2 Deterministic selection

Recall the selection indicator

$$s_t = \begin{cases} 1 & \text{if} \quad t \in \mathcal{G} = \{t \in \mathcal{N} : z_t^{\perp} \text{ credible}\} \\ 0 & \text{if} \quad t \in \mathcal{B} \end{cases} \ ,$$

which classifies whether period t was used for identifying the causal effect of interest. In the main text we treated s_t as a stationary random sequence that took on the value one with probability $\pi(\mathbf{q}_t)$. While this approach covers numerous applications there are others that cannot be covered by this assumption, yet innovation powering can still be used.

A prominent category arises when s_t is a deterministic series. As a concrete example we can think of Gertler and Karadi (2015), where the good period is post 1990 since high frequency monetary surprises can only be computed once a fed funds futures market exist. The period before 1990 is then the bad period. In this case $s_t = 0$ prior to 1990 and then shifts to $s_t = 1$ for all post-1990 periods and we can treat s_t as a deterministic sequence. Note that here the stationarity assumption of the main text (Assumption R-1) does not hold. The good news is that innovation powering can be applied without much modifications.

When s_t is deterministic the selection assumption 2 can be replaced by the following assumption.

Assumption S1. *The indicator* $\{s_t\}$ *that defines the sample* $\mathcal{G}$ *is deterministic and*

$$\frac{1}{n}\sum_{t\in\mathcal{N}}\mathrm{Cov}(\mathbf{v}_t,k_t)-\frac{1}{n_{\mathcal{G}}}\sum_{t\in\mathcal{G}}\mathrm{Cov}(\mathbf{v}_t,k_t)\to 0 \quad as \quad n\to\infty \quad with \quad n_{\mathcal{G}}/n\to\pi>0\ ,$$

for $k_t\in\{\xi^{\perp}_{t+h},p^{\perp}_t,y^{\perp}_{t+h}\}$.

This assumption is clearly satisfied for stationary time series $(\mathbf{v}_t,p^{\perp}_t,y^{\perp}_{t+h})$, but it can be also satisfied under weaker conditions.[S2] The key difference with the main text is that there, since s_t is random, we need to ensure that $\mathbb{E}(s_t\mathbf{v}_t\xi^{\perp}_{t+h})$ is equal to $\mathbb{E}(\mathbf{v}_t\xi^{\perp}_{t+h})$ which requires an assumption on $\mathbb{E}(s_t|\mathbf{v}_t,p^{\perp}_t,y^{\perp}_{t+h},z^{\perp}_t)$, i.e. Assumption 2. As discussed in the main text, a random s_t arises for example when the narrative accounts are more likely to detect an informative shock when the shock is large, or, more generally, whenever detecting an informative shock depends on the state of the economy. In the case where the sample is split into good and bad for data availability reasons no such considerations arise and s_t can be treated as deterministic.

The methodology of the main text simplifies under this assumption as no selection model needs to be estimated and the only moment condition of interest is as follows.

Lemma S4. *Given Assumptions 1 and S1 for any* $\boldsymbol{\gamma}\in\mathbb{R}^{d_v}$ *we have that*

$$\begin{aligned}&\mathbb{E}g(\mathbf{d}_t,\theta_h,\boldsymbol{\delta};\boldsymbol{\gamma})=0\ ,\\ &g(\mathbf{d}_t,\theta_h,\boldsymbol{\delta};\boldsymbol{\gamma})=z^{\perp}_t(y^{\perp}_{t+h}-\theta_h p^{\perp}_t)s_t/\pi+\boldsymbol{\gamma}'\mathbf{v}_t(y^{\perp}_{t+h}-\theta_h p^{\perp}_t)(1-s_t/\pi)\ .\end{aligned}\tag{S11}$$

Here π replaces the selection model. Since π can be estimated by $n_{\mathcal{G}}/n$ the IPIV estimator simplifies to

$$\hat{\theta}_h(\boldsymbol{\gamma})=\frac{\frac{1}{n}\sum_{t\in\mathcal{N}}\boldsymbol{\gamma}'\mathbf{v}_t y^{\perp}_{t+h}+\frac{1}{n_{\mathcal{G}}}\sum_{t\in\mathcal{G}}(z^{\perp}_t-\boldsymbol{\gamma}'\mathbf{v}_t)y^{\perp}_{t+h}}{\frac{1}{n}\sum_{t\in\mathcal{N}}\boldsymbol{\gamma}'\mathbf{v}_t p^{\perp}_t+\frac{1}{n_{\mathcal{G}}}\sum_{t\in\mathcal{G}}(z^{\perp}_t-\boldsymbol{\gamma}'\mathbf{v}_t)p^{\perp}_t}\ ,\tag{S12}$$

which is the same estimator that one would use when $P(s_t=1|\mathbf{q}_t,z^{\perp}_t)=\pi$, i.e. completely random selection as considered in the simple illustrative example of Section 2. In both cases the selection model drops out and the frequency $n_{\mathcal{G}}/n$ reweights for the fraction of good periods.

[S2] Note that the stationary assumption in Assumption R is only imposed for convenience in the main text — it is not a required structural assumption. E.g. alternative laws of large numbers and central limit theorems can be applied for heterogeneous processes.

Similarly, the IPAR statistic also simplifies to

$$\mathrm{AR}^*_{\mathrm{IP}}(\bar{\theta}_h) = \frac{\frac{1}{\sqrt{n}}\sum_{t\in\mathcal{N}}\widehat{\boldsymbol{\gamma}}(\bar{\theta}_h)'\mathbf{v}_t(y^{\perp}_{t+h} - \bar{\theta}_h p^{\perp}_t) + \frac{\sqrt{n}}{n_{\mathcal{G}}}\sum_{t\in\mathcal{G}}(z^{\perp}_t - \widehat{\boldsymbol{\gamma}}(\bar{\theta}_h)'\mathbf{v}_t)(y^{\perp}_{t+h} - \bar{\theta}_h p^{\perp}_t)}{\sqrt{(1,\widehat{\boldsymbol{\gamma}}(\bar{\theta}_h)')\widehat{\boldsymbol{\Sigma}}(\bar{\theta}_h)(1,\widehat{\boldsymbol{\gamma}}(\bar{\theta}_h)'}} \ , \tag{S13}$$

where $\widehat{\boldsymbol{\Sigma}}(\bar{\theta}_h)$ is any consistent estimate for

$$\boldsymbol{\Sigma}(\bar{\theta}_h) = \lim_{n\to\infty}\mathrm{var}(\sqrt{n}f_n(\bar{\theta}_h)) \qquad f_n(\bar{\theta}_h) = \frac{1}{n}\sum_{t=1}^{n} f(\mathbf{d}_t,\bar{\theta}_h)$$

with

$$f(\mathbf{d}_t,\bar{\theta}_h) = \begin{pmatrix} z^{\perp}_t(y^{\perp}_{t+h} - \bar{\theta}_h p^{\perp}_t)s_t/\hat{\pi} \\ \mathbf{v}_t(y^{\perp}_{t+h} - \bar{\theta}_h p^{\perp}_t)(1 - s_t/\hat{\pi}) \end{pmatrix}$$

and $\hat{\pi} = n_{\mathcal{G}}/n$. Estimators such as those proposed in Newey and West (1987), Andrews (1991) and de Jong and Davidson (2000), can be used.

Further, the estimate for the efficient weights also simplifies to

$$\widehat{\boldsymbol{\gamma}}(\bar{\theta}_h) = -\widehat{\boldsymbol{\Sigma}}_{vv}(\bar{\theta}_h)^{-1}\widehat{\boldsymbol{\Sigma}}_{vz}(\bar{\theta}_h) \ ,$$

where the sub-blocks conform to the partition in $f(\mathbf{d}_t,\bar{\theta}_h)$.

Under $H_0 : \theta_h = \bar{\theta}_h$ and given Assumptions 1, S1 and regularity conditions R we continue to have that $\mathrm{AR}^*_{\mathrm{IP}}(\bar{\theta}_h) \xrightarrow{d} N(0,1)$. Also, the construction of the confidence set remains the same.

As such for practical purposes we can now distinguish between two options:

1. s_t is a random sequence and Assumption 2 holds. In this case the methodology presented in the main text can be followed.

2. s_t is a deterministic sequence or Assumption 2 holds with $\pi(\mathbf{q}_t) = \pi$ (i.e. complete random selection); in case the expressions (S12) and (S13) can be used.

S3 Compatible underlying structural models

The main text did not specify an underlying structural model for p_t and y_t. In this section we show that multiple different structural models are compatible with the general framework and our methodology applies regardless which is true. Specifically, we will show that innovation powered inference improves over standard IV *regardless* of whether the structural shocks of interest actually do *or* do not exist on the $\mathcal{B}$ sample.

Narratively undetected structural shocks

To set the stage we first take a conventional perspective and postulate that the vector of macro variables $\mathbf{w}_t = (p_t, y_t, w_{3t}, \dots, w_{Kt})$ follows a structural vector autoregressive model (SVAR) model of order p.

$$\mathbf{w}_t = \sum_{j=1}^{p} \Phi_j \mathbf{w}_{t-j} + \mathbf{e}_t \qquad \mathbf{e}_t \sim \text{iid}(0, \Sigma)\ , \qquad \mathbf{e}_t = \Phi_0 \boldsymbol{\eta}_t\ , \qquad t \in \mathcal{N}\ , \tag{S14}$$

where $\mathbf{e}_t$ are the reduced form shocks with variance Σ and $\boldsymbol{\eta} = (\varepsilon_t, \eta_{2t}, \dots, \eta_{Kt})'$ is the vector of structural shocks that has unit variance such that $\Sigma = \Phi_0\Phi_0'$. We are interested in ε_t which is normalized to have unit effect on p_t. The other shocks do not necessarily have a structural interpretation.[S3] For this underlying model we can characterize narrative instruments similarly as in the illustrative example:

$$z_t^{\perp} = f(\varepsilon_t, \zeta_t) = \begin{cases} \varepsilon_t + \zeta_t & \text{if} \quad t \in \mathcal{G} \\ 0 & \text{if} \quad t \in \mathcal{B} \end{cases}\ ,$$

where ζ_t is required to be uncorrelated with $\xi_{t+h}^{\perp} = y_{t+h}^{\perp} - \theta_h p_t^{\perp}$ and for simplicity we assume that the variables that are projected out include exactly the p lags of the variables in the SVAR model. We note that as long as the zeros on $\mathcal{B}$ are determined as a function of ε_t and ζ_t we have that z_t is an exogenous instrument and Assumption 1 holds.

In this underlying model structural shocks ε_t arrive every period, but the narrative does not always record the shock. We do not need to take a stance on why this is, but a plausible reason is that the narrative only captures large exogenous events, i.e. large values of ε_t, and misses the smaller ones.

This perspective underlies the narrative SVAR literature (Ludvigson, Ma and Ng, 2017; Antolín-Díaz and Rubio-Ramírez, 2018; Giacomini, Kitagawa and Read, 2022), where the identified set of structural parameters in (S14) is refined using the available information in the narrative accounts. The narrative series of Romer and Romer (1989, 2023) is often used as an illustration: monetary policy shocks arrive every period but the narrative accounts only capture a few of such exogenous events. This justifies imposing an SVAR as the underlying model — which assumes that structural shocks arrive every period — and using the narrative to further inform the direction or magnitude of the structural shocks on specific dates.

[S3] We adopt an SVAR model for comparability with the existing literature. From the perspective of IP inference this section could also be written for any structural vector (autoregressive) moving average model.

Non-existent structural shocks

To motivate a second plausible underlying structural model, suppose that the narrative series z_t corresponds to the military defense spending news series of Ramey and Zubairy (2018). This series is equal to zero for about 80% of periods, but the reason for why it is often zero is quite straightforward: the US was not involved in any major war and there were simply no surprises to military defense spending. This leaves us with two options.

First, if we postulate that ε_t is the shock to government spending (arriving every period) and we assume that the components of this shock, including military defense spending, have homogeneous effects on the macro outcomes, then the conventional SVAR model (S14) remains appropriate. In this case we can view the narrative instrument as

$$z_t^{\perp} = f(\varepsilon_t, \zeta_t) = \begin{cases} \varepsilon_{kt} + \zeta_t & \text{if} \quad t \in \mathcal{G} \\ 0 & \text{if} \quad t \in \mathcal{B} \end{cases} ,$$

where $f()$ now isolates the military defense component ε_{kt} of the shock ε_t. Under this assumption $z_t^{\perp}$ remains a valid instrument (e.g. Stock and Watson, 2018), but the implicit homogeneous effects assumption is strong.

Second, we can instead postulate that ε_t corresponds specifically to the shock to military defense spending and suppose that this shock is often zero. Let $s_t^{\varepsilon} = \mathbf{1}(\varepsilon_t \text{ occurs})$ be the indicator for whether the shock materializes and π_t^{ε} the probability that $s_t^{\varepsilon} = 1$, which may depend on the state of the economy. To incorporate the non-random arrival of structural shocks we consider the extended SVAR model

$$\mathbf{w}_t = \sum_{j=1}^{p} \Phi_j \mathbf{w}_{t-j} + \mathbf{e}_t , \qquad \mathbf{e}_t = s_t^{\varepsilon} \Phi_0 \boldsymbol{\eta}_t + (1 - s_t^{\varepsilon}) \mathbf{e}_t , \qquad t \in \mathcal{N} , \tag{S15}$$

where the reduced form shocks are only mapped to the structural shocks when $s_t^{\varepsilon} = 1$, i.e. when the structural shocks occur. We note that the model remains stationary and the regularity conditions needed for IP inference can be verified to hold.[S4]

For simplicity, we impose that $s_t = s_t^{\varepsilon}$ which implies that the narrative is observed

[S4] In contrast, if (S15) is the true underlying structural model, the conventional SVAR model (S14) is misspecified and the narrative SVAR methods of Antolín-Díaz and Rubio-Ramírez (2018); Giacomini, Kitagawa and Read (2022) cannot be applied.

whenever the structural shock occurs, and we define[S5]

$$z_t^{\perp} = f(\varepsilon_t, \zeta_t) = \begin{cases} g(\varepsilon_t) + \zeta_t & \text{if} \quad t \in \mathcal{G} \\ 0 & \text{if} \quad t \in \mathcal{B} \end{cases} ,$$

which implies a subtle difference in interpretation: on $\mathcal{B}$ there were no shocks to defense spending and the narrative documents the *correct* value of zero, i.e. $0 = \varepsilon_t$ for $t \in \mathcal{B}$. This instrument series together with model (S15) fits in the general framework of the previous section, and hence innovation powered narrative inference can be applied in this setting.

Invariance of Innovation Powered Inference

Next, we clarify why innovation powered inference is invariant to the underlying structural model and we show that the gains in efficiency are the same regardless whether model (S14) or (S15) is true. In other words, the reductions in variance are the same regardless whether structural shocks exist on $\mathcal{B}$.

For simplicity suppose that a researcher uses innovations $\mathbf{v}_t = p_t^{\perp}$. Model (S14) implies – after projecting out the lags — that $p_t^{\perp} = \Phi_{1\bullet,0}\boldsymbol{\eta}_t = \Phi_{11,0}\varepsilon_t + \sum_{j=2}^{K} \Phi_{1j,0}\eta_{jt}$ and the innovations depend on the structural shock of interest for all time periods $t \in \mathcal{N}$. In contrast for model (S15) we have $p_t^{\perp} = \varepsilon_t + \sum_{j=2}^{K} \Phi_{1j,0}\eta_{jt}$ only on $\mathcal{G}$ and $p_t^{\perp} = \mathbf{e}_{1t}$ on $\mathcal{B}$.

Consider the main innovation powered sample moments for horizon $h = 0$

$$\underbrace{\frac{1}{n}\sum_{t\in\mathcal{N}} z_t^{\perp}(y_t^{\perp} - \theta_0 p_t^{\perp}) s_t/\kappa(\mathbf{q}_t;\boldsymbol{\delta})}_{\text{narrative part}} + \underbrace{\widehat{\gamma}}_{\text{weight}} \underbrace{\frac{1}{n}\sum_{t\in\mathcal{N}} p_t^{\perp}(y_t^{\perp} - \theta_0 p_t^{\perp})(1 - s_t/\kappa(\mathbf{q}_t;\boldsymbol{\delta}))}_{\text{innovations part}} ,$$

which determines the IPIV estimator and the IPAR test statistic.

The narrative part of the moment condition depends on ε_t via $z_t^{\perp}$ and hence this part exploits the correlation between the structural shock and the macro variables. However, this term is only nonzero for $t \in \mathcal{G}$, and therefore the existence of structural shocks on $\mathcal{B}$ does not matter — they never contribute —.

The innovations part is non-zero for all t, but in contrast to the narrative term it does not exploit any correlations among structural shocks and observables. Indeed, we can write

$$\frac{1}{n}\sum_{t\in\mathcal{N}} p_t^{\perp}(y_t^{\perp} - \theta_0 p_t^{\perp})(1 - s_t/\kappa(\mathbf{q}_t;\boldsymbol{\delta})) = \frac{1}{n}\sum_{t\in\mathcal{N}} \mathbf{e}_{1t}(\mathbf{e}_{2t} - \theta_0\mathbf{e}_{1t})(1 - s_t/\kappa(\mathbf{q}_t;\boldsymbol{\delta})) ,$$

[S5] Violation of this assumption can go two ways. First the narrative may miss certain structural shocks (e.g. because they were small) in which case the narrative instrument will remain exogenous (similar as in the scenario above). Second if the narrative documents a structural shock when there was none, we violate Assumption 1 and we fall outside the scope of innovation powered inference.

which only depends on the reduced form correlations among e_t. As both models have identical reduced form shocks these correlations are identical for (S14) and (S15). The fact that different structural shocks may have generated them is irrelevant. As long as the stationarity assumption continues to hold and the correlations among the macro variables are not different on $\mathcal{G}$ and $\mathcal{B}$ innovation powering will reduce the variance.

Finally, one may wonder whether the efficient weight estimate $\widehat{\gamma}$ change when structural shocks are missing. To show that this is not the case consider the expression in equation (28). There are two types of terms: those that depend on $z_t^{\perp}$ and those that do not. Similar as above the terms that depend on $z_t^{\perp}$ are only computed on periods $\mathcal{G}$, whereas the other terms only exploit reduced form shocks which are identical across the different SVAR models.

We conclude that each of the components of the innovation powered sample moments is invariant to whether or not structural shocks arrive on $\mathcal{B}$, hence the resulting estimators and tests are are as well.

S4 Simulation study

We discuss the details of the simulation study that we conducted to verify the finite sample properties of the IPIV estimates and IPAR tests.

S4.1 Simulation design

We simulate data from two different models: the static simultaneous equations model that was discussed in the illustrative example and an SVAR model calibrated to different empirical settings considered in the literature. The first model aims to explore the performance of our method in a stylized setting, whereas the second class of SVAR models aims to assess the workings of our method in realistic macro examples.

Within in each model class we consider different specifications that we detail below. For each specific simulation design we simulate $M = 5000$ datasets.

Linear simultaneous equations model We consider the model

$$y_t = \theta p_t + \sigma \xi_t \ , \qquad p_t = c\varepsilon_t + \nu\sigma\xi_t \ , \qquad \text{for } t \in \mathcal{N} \ ,$$

where $(\xi_t, \varepsilon_t)' \overset{iid}{\sim} \mathcal{N}(0, I_2)$. The scale σ is fixed at one, the degree of endogeneity is fixed at $\nu = 1.5$, instrument strength is governed by c which is varied between $c = 0.5$ and $c = 1.0$, and $\theta = 1$ is the parameter of interest. The narrative instrument is generated as $z_t = \varepsilon_t$ and the innovations are set as

$$v_t = \rho z_t + (1 - \rho)\xi_t \ ,$$

where $\rho \in (0.4, 0.8)$ captures the relevance of the innovations. Note that all these parameter choices interact for determining instrument strength. Therefore, for each combination of parameters, we summarize the model by reporting the average F statistic across simulations.

The narrative instrument is observed either completely at random or as a function of the state variables. The first case is included to explore the minimum $n_{\mathcal{G}}$ requirement for which our methodology has good finite sample properties. Specifically, to have exact control over the number of narratives we select a fixed number of $n_{\mathcal{G}}$ periods to be observed at random from n. We consider $n_{\mathcal{G}} = 25, 50, 100, 150$ and $n = 200, 500$.

For the second case we model $\pi(\mathbf{q}_t)$ using the logistic specification

$$\pi(\mathbf{q}_t) = \frac{\exp(\mathbf{q}_t' \boldsymbol{\delta})}{1 + \exp(\mathbf{q}_t' \boldsymbol{\delta})} \ ,$$

and we consider two parametrizations: $\mathbf{q}_t = (1, p_t, y_t)'$ and $\boldsymbol{\delta} = (c, 0.2, 0, 2)$ — indicated by $\pi = 1$ in the results — and $\mathbf{q}_t = (1, |\varepsilon_t|)'$ and $\boldsymbol{\delta} = (c, 0.2)$— indicated by $\pi = 2$ —. For each specification we only consider draws for s_t such that $50 \leq n_{\mathcal{G}} \leq 150$. To ensure this happens on average we pick the constant $c = -1$ for $n = 200$ and $c = -2$ for $n = 500$.

Structural VARs In our second design we simulate data from an SVAR model that is calibrated to monetary or fiscal policy data from the US. The general model is given by

$$\mathbf{w}_t = \sum_{j=1}^{p} \Phi_j \mathbf{w}_{t-j} + A_0 \boldsymbol{\eta}_t \ ,$$

where $\mathbf{w}_t$ is a vector of macro variables that includes y_t and p_t, and $\boldsymbol{\eta}_t$ are the structural shocks which include ε_t, which is the shock that generates exogenous variation in p_t. We are interested in estimating θ_h in model $y_{t+h} = \theta_h p_t + \boldsymbol{\beta}_h' \mathbf{x}_t + u_{t+h}$ where $\mathbf{x}_t$ includes a constant and the lags of y_t, p_t.

The narrative instrument is generated in a dynamic way

$$z_t = c_1 \varepsilon_t + c_2 \varepsilon_{t-1} + \zeta_t \ , \qquad \zeta_t \overset{iid}{\sim} \mathcal{N}(0, 1) \ ,$$

which allows for serial correlation in the instrument. The latter is important as most instrument series in macro are serially correlated (Alloza, Gonzalo and Sanx, 2019). We consider $(c_1, c_2) = (0.9, 0.5)$ or $(c_1, c_2) = (0.5, 0.2)$. Similar as for the simultaneous equations model we select the observed narrative periods $s_t = 1$ either completely at random or with probability $\pi(\mathbf{q}_t)$ taken as a logit with $\mathbf{q}_t = (1, p_t^{\perp}, y_{t+h}^{\perp})'$ and $\boldsymbol{\delta} = (c, 0.2, 0, 2)$ — indicated by $\pi = 1$ in the results — and $\mathbf{q}_t = (1, |\varepsilon_t|)'$ and $\boldsymbol{\delta} = (-1, 0.2)$— indicated by $\pi = 2$ —, and $c = -1$ for $n = 200$ and $c = -2$ for $n = 500$.

The variables and parameters are based on our empirical studies.

1. Monetary policy. We consider $\mathbf{w}_t = (w_{1t}, w_{2t}, w_{3t})'$ with unemployment $y_t = w_{1t}$, PCE inflation w_{2t} and the short term interest rate $p_t = w_{3t}$. Using $p = 12$ the model is fitted using monthly US data over the 1954M7-2016M12 period. The structural shocks are ordered as $\boldsymbol{\eta}_t = (\eta_{1t}, \eta_{2t}, \varepsilon_t)'$, where ε_t is the monetary policy shock. This shock is identified by imposing short run recursive restrictions on the impact matrix A_0. The parameters are estimated using OLS and we use these parameter to generate the data and the define the true impulse response. The results are shown for θ_h corresponding to the unemployment response at horizon $h = 30$.

   The innovations that we consider are: (i) $\mathbf{v}_t = w_{3t}^{\perp}$ and (ii) $\mathbf{v}_t = w_{1t}^{\perp}$, which are the interest rate or growth rate after projecting out the lags of p_t, y_t. Intuitively, (i) which uses interest rate innovations is more promising as the monetary policy shock is typically more pronounced in these series. In contrast, (ii) uses GDP innovations and variance decompositions show that monetary policy explains little variation in output. Note that neither choice makes use of the lower triangular ordering in the SVAR used for simulating the data.

2. Fiscal policy. We consider $\mathbf{w}_t = (w_{1t}, w_{2t}, w_{3t})$ with w_{1t} is taxes, $p_t = w_{2t}$ is spending and $y_t = w_{3t}$ is gdp growth. We fit the reduced form VAR to quarterly US data from 1947Q1 until 2015Q4. We use the short run restrictions from Blanchard and Perotti (2002) to identify the effects of spending. The innovations that we consider are: (i) $\mathbf{v}_t = w_{3t}^{\perp}$ and (ii) $\mathbf{v}_t = w_{1t}^{\perp}$, which are the interest rate or growth rate after projecting out the lags of p_t, y_t. The results are shown for θ_h corresponding to the gdp growth response at horizon $h = 30$.

S4.2 Evaluation criteria

We compare the accuracy of the $\hat{\theta}_h^{\text{cue}}$ IPIV estimator and the conventional IV estimator that is based on the instrument z_t that includes the zeros. We report the root mean squared error (RMSE) and the mean absolute error (MAE). We note that due to the relatively weak instruments in some simulation design some outliers in the parameter estimates occur (notably in the conventional IV estimator). This makes the comparison based on mean squared error somewhat distorted and the results for MAE are more reliable.

Further, we evaluate the size of the IPAR and AR tests and the length of the confidence interval based on inverting these tests. To evaluate the empirical size of the test we compare the empirical null rejection probability (ERP) the nominal level $\alpha = 0.05$. Finally, the length of the confidence set is capped at 50 for each draw to reduce the computational time. This implies that for very weak IV designs we will not see much differences in the confidence set length as both IPAR and AR are uninformative.

S4.3 Results

Linear simultaneous equations model The results for the simultaneous equations model are shown in Tables S1 and S2, where S1 considers random selection for good periods and S2 considers $\mathcal{G}$ selection via the discussed logit models.

The following patterns can be detected under random selection from Table S1.

(i) The length of the confidence set based on the IPAR test is virtually always smaller on average when compared to the length based on the AR test. The leading determinants for the magnitude in reduction are: (i) the correlation between the innovation and the instrument as captured by ρ — larger ρ leads to more gains — and (ii) $n_{\mathcal{G}}/n$ ratio – when $n_{\mathcal{G}}$ is small relative to n the gains are larger. The strength of the instrument as measured by the F-statistic is moderately important for the reduction: stronger instruments lead to slightly larger improvements for IPAR, but a part of this is driven by the maximum length of the confidence intervals being capped at 50.

   For example, when $n = 200, n_{\mathcal{G}} = 50$, $\rho = 0.4$ and $F = 7.6$, the gain in confidence set length reduction is 32%. When we increase ρ to 0.8, the gain becomes to around 80%. In the same scenario varying $n_{\mathcal{G}} = 25$ to $n_{\mathcal{G}} = 150$ modifies the gains between 30% and 0%.

(ii) The size of the IPAR test is generally good. Only when $n_{\mathcal{G}} = 25$ we find that the test starts over-rejecting. This is not surprising as our asymptotic approximation is based on $n_{\mathcal{G}} \to \infty$. Interestingly, when ρ increases — i.e. the innovations become more relevant — the size distortions also disappear for $n_{\mathcal{G}} = 25$. Nonetheless, we strongly recommend to only use the IPAR test when $n_{\mathcal{G}} \geq 50$. In future work we aim to explore alternative ways of approximating the distribution of IPAR.

(iii) The reductions in RMSE and MAE are very large for the IPIV continuous updating estimator (relative to the conventional IV). This holds for virtually all specifications considered. We note that due to the large outliers when the instruments are weak the patterns are not always clear. Further inspection shows that the reductions all come from the variance part of the MSE, the bias term is is virtually identical for IV and IPIV.

Table S2 shows the results for the setting where the selection probabilities depend on the state of the economy. The conclusions made above also hold for this design. We make the following additional observations.

(iv) The specification of the selection model does not affect the results much. The gains for the case where the selection probability depends on $|\varepsilon_t|$ — i.e. the size of the shock — are generally a bit smaller when compared to the case where the logit model is defined by observables p_t, y_t. Nonetheless, the reductions in confidence set length can still be large.

Structural VARs The estimation results for the structural VAR model are shown in Tables S3 and S4 for the monthly monetary SVAR and in Tables S5 and S6 for the quarterly fiscal SVAR. Again, the first table corresponds to the completely ignorable selection case where $\mathcal{G}$ is picked at random and the second tables correspond to the case where the logit selection model is included.

We make the following observations.

(v) The IP gains in RMSE, MAE and CS length reduction clearly depend on the choice of the innovation. While in the previous simulations this was tightly controlled by the parameter ρ, here it becomes clear when we compare using $p_t^\perp$ and $y_t^\perp$ as innovation. In each design the shock of interest, e.g. the monetary shock, is much more pronounced in $p_t^\perp$ and therefore using this as innovation leads to much improved results. In contrast when we use $v_t = y_t^\perp$ the gains are much more modest, and in some designs we even lose relative to conventional IV.

(vi) The size of the IPAR test is somewhat less well controlled overall — the ERP is slightly above the nominal level. We note that this is also observed for the conventional AR test and is most likely due to the estimation of the long run variance matrix. This is a well known difficult problem and our choice for relying on the Newey and West (1987) estimator can possibly be improved (e.g. Lazarus et al., 2018).

(vii) Comparing the monetary and fiscal VARs we find little differences. The gains are somewhat larger in the quarterly VAR, which is possibly due to the reduced serial correlation in the model.

S5 Additional empirical results

S5.1 Romer and Romer (1989) meets Romer and Romer (2004)

In this section, we revisit our empirical application on the effects of monetary policy by considering an alternative candidate for the innovations $\mathbf{v}_t$: the Romer and Romer (2004) monetary shocks, henceforth RR04. RR04 estimate a time series model of the fed funds rate (a policy rule), where they use internal forecasts from the Fed Board staff —the Greenbook forecasts— to control for the Fed's information set. This provides a more credible identification strategy than Sims (1980)'s original approach, as the Fed reacts to a lot more information than can be included in a macro time series model. By controlling for the staff forecasts for unemployment and inflation, RR04 can isolate movements in the policy rate that are orthogonal to the expected paths of unemployment and inflation.

Unlike the narrative approach of Romer and Romer (1989) (henceforth RR89) however, the Greenbook forecasts may not control for all the information that the Fed is reacting to, such that the RR04 innovations may be contaminated by residual endogeneity. As Romer and Romer (2004) note (p. 1066), "To the extent that policy makers employ useful information about the paths of output and inflation beyond what is in the Greenbooks,[...], our series could still include [some] anticipatory actions."

The comparison between RR89 and RR04 is a great illustration of the power-credibility in time series macro. On the one hand, RR89 is a very credible approach to isolate exogenous changes in the policy rate, but the method can only isolate a few episodes: credibility is very high but power is lower. On the other hand, RR04 identify monetary innovations at every single dates but some of these innovations may not be entirely exogenous: power is high but credibility is lower.

Figure S1 plots our IPIV estimates using RR04 as innovations. Compared to the Sims (1980)'s type innovations, the RR04 innovations are closer to the true underlying monetary shocks, and the correlation with RR89's narrative shock proxy is higher (0.43 vs 0.40), implying larger gains from innovation powering. Counterbalancing this effect, the RR04 series is shorter, covering only 1967-2007 (using Tenreyro and Thwaites (2016)'s extension) instead of the 1954-2023 sample underlying our baseline estimates in Figure 4, implying smaller expected efficiency gains from innovation powering.

Overall, the results are similar to those presented in the main text with substantial reductions in confidence bands for both inflation and unemployment. The reductions are a bit smaller for unemployment however, with reductions in the order of 20 percent at horizons above 10 months. This is likely due to the smaller sample period covered by the RR04 series.

S5.2 Romer and Romer (1989) with selection

On possible worry for our IPIV estimate of the effects of monetary policy is that the selection process for the Romer and Romer (1989) dates are not random. Indeed, Romer and Romer (1989) are careful to select times when output is at potential but inflation is suddenly deemed too high by the central bank. This indicates that the selection of RR89 dates is not random but depends on the underlying state of the economy.

To address this concern, we model the credible event selection probability by considering a linear logit model in $p_t^{\perp}, y_{t+h}^{\perp}$ with $y_{t+h}^{\perp}$ the unemployment and inflation rates. Figure S2 plots the results and show similar efficiency improvements, in fact slightly larger for the unemployment response while a bit smaller for the inflation response.

References

Alloza, Mario, Jesus Gonzalo, and Carlos Sanx. 2019. "Dynamic Effects of Persistent Shocks." working paper.

Andrews, Donald W. K. 1991. "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation." *Econometrica*, 59(3): 817–858.

Antolín-Díaz, Juan, and Juan F. Rubio-Ramírez. 2018. "Narrative Sign Restrictions for SVARs." *American Economic Review*, 108(10): 2802–29.

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

de Jong, Robert M., and James Davidson. 2000. "Consistency of Kernel Estimators of Heteroscedastic and Autocorrelated Covariance Matrices." *Econometrica*, 68(2): 407–423.

Gertler, Mark, and Peter Karadi. 2015. "Monetary Policy Surprises, Credit Costs, and Economic Activity." *AEJ: Macroeconomics*, 7: 44–76.

Giacomini, Raffaella, Toru Kitagawa, and Matthew Read. 2022. "Narrative Restrictions and Proxies." *Journal of Business & Economic Statistics*, 40(4): 1415–1425.

Lazarus, Eben, Daniel J. Lewis, James H. Stock, and Mark W. Watson. 2018. "HAR Inference: Recommendations for Practice." *Journal of Business & Economic Statistics*, 36(4): 541–559.

Ludvigson, Sydney C, Sai Ma, and Serena Ng. 2017. "Shock Restricted Structural Vector-Autoregressions." National Bureau of Economic Research Working Paper 23225.

Newey, Whitney K., and Daniel McFadden. 1994. "Chapter 36 Large sample estimation and hypothesis testing." In . Vol. 4 of *Handbook of Econometrics*, 2111 – 2245. Elsevier.

Newey, Whitney K., and Kenneth D. West. 1987. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, 55(3): 703–708.

Ramey, Valerie A., and Sarah Zubairy. 2018. "Government Spending Multipliers in Good Times and in Bad: Evidence from U.S. Historical Data." *Journal of Political Economy*, 126.

Romer, Christina D, and David H Romer. 1989. "Does monetary policy matter? A new test in the spirit of Friedman and Schwartz." *NBER macroeconomics annual*, 4: 121–170.

Romer, Christina D., and David H. Romer. 2004. "A New Measure of Monetary Shocks: Derivation and Implications." *American Economic Review*, 94: 1055–1084.

Romer, Christina D., and David H. Romer. 2023. "Presidential Address: Does Monetary Policy Matter? The Narrative Approach after 35 Years." *American Economic Review*, 113(6): 1395–1423.

Sims, Christopher A. 1980. "Macroeconomics and reality." *Econometrica*, 1–48.

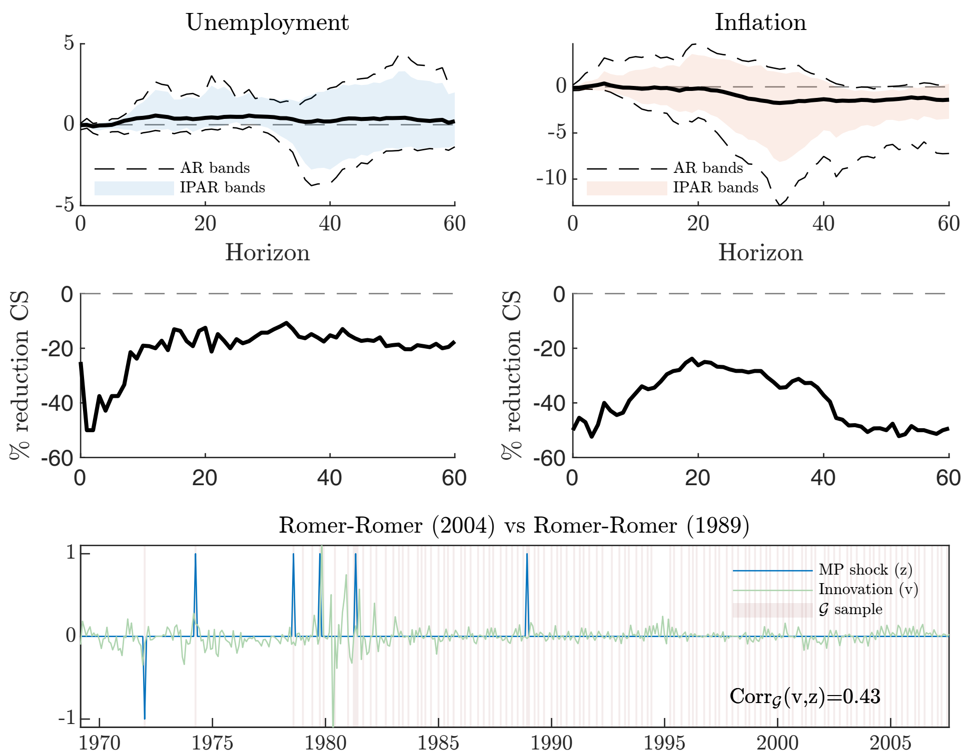
Stock, James H., and Mark W. Watson. 2018. "Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments." *The Economic Journal*, 128(610): 917–948.

Tenreyro, Silvana, and Gregory Thwaites. 2016. "Pushing on a String: US Monetary Policy Is Less Powerful in Recessions." *American Economic Journal: Macroeconomics*, 8(4): 43–74.

White, Halbert L. Jr. 2000. *Asymptotic Theory for Econometrians - 2nd edition.* San Diego, California:Academic Press.

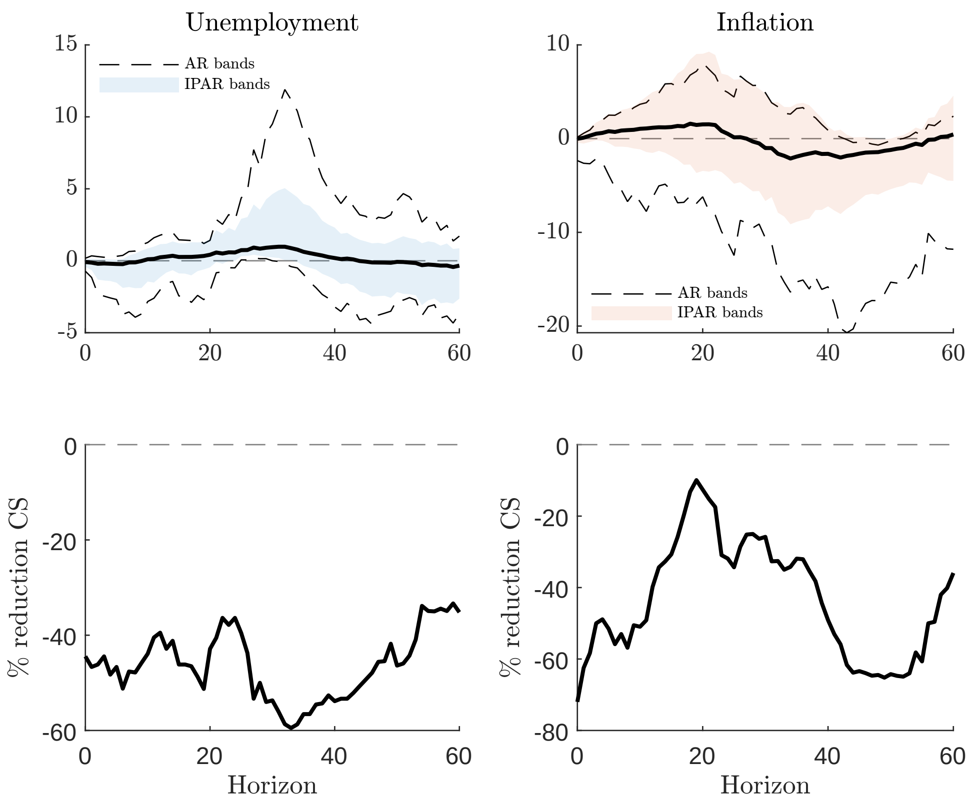
Appendix: Tables and Figures

Figure S1: Effects of Monetary Policy, Innovation-Powered IV, RR89 meets RR04



Top row: IPIV point estimates ($\hat{\theta}_h^{\text{cue}}$, thick black line) for unemployment and inflation, together with the innovation-powered AR bands (IPAR) and regular AR bands at 95% confidence levels. *Middle row:* differences in the lengths of the 95% bands for the regular IV estimator ("AR bands") and the Innovation-Powered IV estimator ("IPAR bands"). *Bottom row:* Time series of the Romer and Romer (1989) shocks and Romer and Romer (2004) innovations. The table reports the mean absolute error (MAE), Empirical Rejection Probability (ERP) and the width of the confidence interval (wCS) for the standard IV estimator with AR confidence set and the innovation powered versions.

Figure S2: Effects of Monetary Policy, Innovation-Powered IV with selection model



Notes: The top row shows the IPIV point estimate ($\hat{\theta}_h^{\text{cue}}$, thick black line) for unemployment and inflation, together with the innovation-powered AR bands (IPAR) and regular AR bands at 95% confidence levels. The bottom row reports the difference in the lengths of the 95% bands for the regular IV estimator ("AR bands") and the Innovation-Powered IV estimator ("IPAR bands")

Table S1: Simulation Evidence — Simultaneous Equations Model (i)

				RMSE		MAE		ERP		CS length			
n	c	ρ	$n_{\mathcal{G}}$	IV	IPIV	IV	IPIV	AR	IPAR	AR	IPAR	%Δ	F^{eff}
200	0.5	0.4	25	19.39	5.58	1.68	1.13	0.048	0.110	36.41	25.37	-0.30	5.44
200	0.5	0.4	50	11.09	5.60	1.01	0.64	0.044	0.075	22.75	15.21	-0.33	7.67
200	0.5	0.4	100	0.71	0.71	0.22	0.21	0.032	0.050	7.84	6.54	-0.16	12.79
200	0.5	0.4	150	0.26	0.26	0.15	0.15	0.055	0.051	2.48	2.48	0.00	18.62
200	0.5	0.8	25	5.91	0.38	1.22	0.19	0.046	0.049	37.11	6.87	-0.81	5.26
200	0.5	0.8	50	17.71	0.27	1.82	0.16	0.045	0.048	22.37	3.82	-0.82	7.67
200	0.5	0.8	100	0.41	0.21	0.21	0.14	0.051	0.052	7.85	1.77	-0.77	12.82
200	0.5	0.8	150	0.41	0.35	0.17	0.15	0.054	0.061	3.03	1.69	-0.44	18.30
200	1.0	0.4	25	17.61	1.33	0.87	0.29	0.055	0.100	13.74	6.09	-0.55	16.27
200	1.0	0.4	50	0.20	0.19	0.13	0.12	0.053	0.071	2.21	1.18	-0.46	27.07
200	1.0	0.4	100	0.11	0.11	0.08	0.08	0.055	0.052	0.46	0.41	-0.10	49.10
200	1.0	0.4	150	0.08	0.08	0.06	0.06	0.042	0.048	0.32	0.31	-0.02	70.24
200	1.0	0.8	25	1.60	0.14	0.27	0.08	0.048	0.049	14.40	0.52	-0.96	16.68
200	1.0	0.8	50	0.19	0.09	0.12	0.06	0.047	0.045	2.00	0.34	-0.82	27.78
200	1.0	0.8	100	0.11	0.08	0.08	0.06	0.048	0.054	0.46	0.29	-0.36	49.56
200	1.0	0.8	150	0.09	0.08	0.06	0.06	0.061	0.070	0.33	0.28	-0.15	68.69
500	0.5	0.4	25	6.95	5.64	1.48	1.03	0.040	0.116	36.44	24.47	-0.32	5.35
500	0.5	0.4	50	9.52	2.65	0.82	0.46	0.057	0.072	21.91	13.90	-0.36	8.05
500	0.5	0.4	100	0.55	0.43	0.22	0.20	0.038	0.065	7.31	4.54	-0.37	13.74
500	0.5	0.4	150	0.26	0.25	0.16	0.15	0.058	0.055	2.59	1.85	-0.28	18.25
500	0.5	0.8	25	8.22	0.64	1.34	0.15	0.035	0.069	37.25	3.38	-0.90	5.35
500	0.5	0.8	50	3.06	0.15	0.60	0.11	0.055	0.053	23.39	1.10	-0.95	7.56
500	0.5	0.8	100	1.02	0.13	0.25	0.09	0.049	0.039	7.97	0.57	-0.92	13.22
500	0.5	0.8	150	0.31	0.12	0.16	0.08	0.050	0.049	2.77	0.48	-0.82	18.31
500	1.0	0.4	25	3.94	1.22	0.37	0.27	0.050	0.115	15.54	5.95	-0.61	16.26
500	1.0	0.4	50	0.19	0.18	0.12	0.11	0.051	0.081	1.93	1.01	-0.47	27.59
500	1.0	0.4	100	0.11	0.10	0.08	0.07	0.055	0.059	0.46	0.38	-0.15	49.78
500	1.0	0.4	150	0.08	0.08	0.06	0.06	0.039	0.058	0.31	0.28	-0.11	72.62
500	1.0	0.8	25	0.40	0.08	0.20	0.06	0.040	0.064	13.22	0.34	-0.97	17.49
500	1.0	0.8	50	0.19	0.06	0.13	0.05	0.062	0.040	2.04	0.24	-0.88	27.10
500	1.0	0.8	100	0.11	0.05	0.08	0.04	0.054	0.041	0.46	0.19	-0.58	48.38
500	1.0	0.8	150	0.08	0.05	0.06	0.04	0.053	0.046	0.32	0.16	-0.48	71.59

*Notes:*The table reports the root mean squared error (RMSE), mean absolute error (MAE), empirical rejection probability (ERP) and the width of the confidence interval (CS length) for the standard IV estimator with AR confidence set and the innovation powered versions.

Table S2: Simulation Evidence — Simultaneous Equations Model (ii)

				RMSE		MAE		ERP		CS length			
n	c	ρ	π	IV	IPIV	IV	IPIV	AR	IPAR	AR	IPAR	$\%\Delta$	F^{eff}
200	0.5	0.4	1	1.98	0.29	0.34	0.17	0.048	0.059	12.56	2.43	-0.80	10.74
200	0.5	0.4	2	2.64	6.34	0.50	0.59	0.050	0.067	16.41	10.87	-0.33	9.73
200	0.5	0.8	1	0.86	0.19	0.26	0.13	0.056	0.068	12.52	1.23	-0.90	10.57
200	0.5	0.8	2	1.47	0.46	0.38	0.16	0.041	0.055	15.76	2.31	-0.85	9.75
200	1.0	0.4	1	0.13	0.11	0.10	0.08	0.055	0.076	0.51	0.40	-0.21	42.57
200	1.0	0.4	2	0.15	0.14	0.10	0.10	0.046	0.078	0.89	0.59	-0.33	34.99
200	1.0	0.8	1	0.14	0.07	0.11	0.06	0.051	0.070	0.69	0.25	-0.63	42.77
200	1.0	0.8	2	0.14	0.08	0.10	0.06	0.050	0.059	0.94	0.29	-0.68	34.42
500	0.5	0.4	1	1.25	0.13	0.25	0.10	0.053	0.042	8.18	0.56	-0.93	12.93
500	0.5	0.4	2	0.48	0.51	0.25	0.23	0.066	0.082	12.58	7.24	-0.42	10.72
500	0.5	0.8	1	0.64	0.09	0.22	0.07	0.050	0.058	8.82	0.35	-0.96	13.08
500	0.5	0.8	2	7.80	0.13	0.57	0.09	0.044	0.054	11.84	0.54	-0.95	11.25
500	1.0	0.4	1	0.13	0.09	0.11	0.07	0.058	0.062	0.35	0.31	-0.11	54.09
500	1.0	0.4	2	0.14	0.12	0.09	0.09	0.062	0.067	0.67	0.45	-0.32	40.97
500	1.0	0.8	1	0.13	0.04	0.11	0.03	0.047	0.049	0.36	0.13	-0.64	53.12
500	1.0	0.8	2	0.13	0.06	0.09	0.04	0.056	0.059	0.68	0.18	-0.72	38.87

Notes: The table reports the root mean squared error (RMSE), mean absolute error (MAE), empirical rejection probability (ERP) and the width of the confidence interval (CS length) for the standard IV estimator with AR confidence set and the innovation powered versions.

Table S3: Simulation Evidence — Monetary SVAR (i)

				RMSE		MAE		ERP		CS length			
v_t	n	(c_1, c_2)	$n_{\mathcal{G}}$	IV	IPIV	IV	IPIV	AR	IPAR	AR	IPAR	%Δ	F^{eff}
$p_t^\perp$	200	$(0.9, 0.5)$	25	14.41	1.50	1.43	0.80	0.085	0.061	17.47	10.13	-0.42	10.14
$p_t^\perp$	200	$(0.9, 0.5)$	50	0.82	2.65	0.52	0.55	0.080	0.064	5.72	4.08	-0.28	17.53
$p_t^\perp$	200	$(0.9, 0.5)$	100	0.44	0.42	0.33	0.32	0.095	0.090	1.62	1.39	-0.14	35.20
$p_t^\perp$	200	$(0.9, 0.5)$	150	0.35	0.35	0.27	0.27	0.102	0.085	1.09	1.00	-0.08	57.89
$p_t^\perp$	200	$(0.5, 0.2)$	25	12.58	7.46	2.48	2.12	0.062	0.062	28.09	21.92	-0.22	5.02
$p_t^\perp$	200	$(0.5, 0.2)$	50	10.60	29.21	1.33	2.29	0.091	0.079	16.76	14.00	-0.16	7.99
$p_t^\perp$	200	$(0.5, 0.2)$	100	0.74	0.73	0.53	0.53	0.094	0.108	5.41	4.72	-0.12	14.29
$p_t^\perp$	200	$(0.5, 0.2)$	150	0.54	0.55	0.40	0.41	0.107	0.084	2.45	2.24	-0.08	21.36
$p_t^\perp$	500	$(0.9, 0.5)$	25	1.35	20.68	0.83	1.32	0.057	0.047	15.18	7.03	-0.53	15.39
$p_t^\perp$	500	$(0.9, 0.5)$	50	0.65	0.58	0.50	0.40	0.054	0.050	3.75	2.32	-0.38	26.20
$p_t^\perp$	500	$(0.9, 0.5)$	100	0.42	0.36	0.33	0.28	0.063	0.051	1.65	1.34	-0.18	48.66
$p_t^\perp$	500	$(0.9, 0.5)$	150	0.33	0.29	0.26	0.22	0.060	0.048	1.24	1.07	-0.14	72.28
$p_t^\perp$	500	$(0.5, 0.2)$	25	6.80	8.67	1.70	1.77	0.051	0.048	26.07	18.81	-0.27	7.11
$p_t^\perp$	500	$(0.5, 0.2)$	50	9.25	1.37	1.10	0.76	0.066	0.055	12.83	9.75	-0.23	11.34
$p_t^\perp$	500	$(0.5, 0.2)$	100	0.67	0.63	0.50	0.46	0.062	0.050	4.05	3.23	-0.20	20.03
$p_t^\perp$	500	$(0.5, 0.2)$	150	0.52	0.49	0.41	0.38	0.066	0.066	2.23	2.02	-0.09	28.62
$y_t^\perp$	200	$(0.9, 0.5)$	25	14.41	2.75	1.43	1.00	0.085	0.087	17.47	20.33	0.16	10.14
$y_t^\perp$	200	$(0.9, 0.5)$	50	0.82	1.15	0.52	0.52	0.080	0.082	5.72	6.32	0.10	17.53
$y_t^\perp$	200	$(0.9, 0.5)$	100	0.44	0.43	0.33	0.33	0.095	0.105	1.62	1.58	-0.02	35.20
$y_t^\perp$	200	$(0.9, 0.5)$	150	0.35	0.34	0.27	0.26	0.102	0.098	1.09	1.01	-0.07	57.89
$y_t^\perp$	200	$(0.5, 0.2)$	25	12.58	9.57	2.48	2.14	0.062	0.082	28.09	29.27	0.04	5.02
$y_t^\perp$	200	$(0.5, 0.2)$	50	10.60	7.64	1.33	1.35	0.091	0.099	16.76	17.34	0.03	7.99
$y_t^\perp$	200	$(0.5, 0.2)$	100	0.74	0.76	0.53	0.53	0.094	0.107	5.41	5.25	-0.02	14.29
$y_t^\perp$	200	$(0.5, 0.2)$	150	0.54	0.54	0.40	0.40	0.107	0.080	2.45	2.30	-0.06	21.36
$y_t^\perp$	500	$(0.9, 0.5)$	25	1.35	1.33	0.83	0.83	0.057	0.081	15.18	17.17	0.13	15.39
$y_t^\perp$	500	$(0.9, 0.5)$	50	0.65	0.65	0.50	0.49	0.054	0.079	3.75	4.26	0.13	26.20
$y_t^\perp$	500	$(0.9, 0.5)$	100	0.42	0.42	0.33	0.32	0.063	0.065	1.65	1.63	-0.01	48.66
$y_t^\perp$	500	$(0.9, 0.5)$	150	0.33	0.32	0.26	0.25	0.060	0.058	1.24	1.25	0.00	72.28
$y_t^\perp$	500	$(0.5, 0.2)$	25	6.80	8.05	1.70	1.97	0.051	0.119	26.07	27.06	0.03	7.11
$y_t^\perp$	500	$(0.5, 0.2)$	50	9.25	1.65	1.10	0.85	0.066	0.091	12.83	14.00	0.09	11.34
$y_t^\perp$	500	$(0.5, 0.2)$	100	0.67	0.68	0.50	0.51	0.062	0.069	4.05	4.16	0.02	20.03
$y_t^\perp$	500	$(0.5, 0.2)$	150	0.52	0.52	0.41	0.40	0.066	0.070	2.23	2.21	-0.00	28.62

Notes: The table reports the root mean squared error (RMSE), mean absolute error (MAE), empirical rejection probability (ERP) and the width of the confidence interval (CS length) for the standard IV estimator with AR confidence set and the innovation powered versions.

Table S4: Simulation Evidence — Monetary SVAR (ii)

				RMSE		MAE		ERP		CS length			
v_t	n	(c_1, c_2)	π	IV	IPIV	IV	IPIV	AR	IPAR	AR	IPAR	%Δ	F^{eff}
$p_t^\perp$	200	$(0.9, 0.5)$	1	0.45	0.40	0.35	0.31	0.094	0.091	1.76	1.45	-0.17	33.97
$p_t^\perp$	200	$(0.9, 0.5)$	2	0.42	0.38	0.32	0.30	0.080	0.058	1.67	1.49	-0.11	35.35
$p_t^\perp$	200	$(0.5, 0.2)$	1	2.04	2.09	0.62	0.58	0.078	0.071	6.56	5.16	-0.21	13.93
$p_t^\perp$	200	$(0.5, 0.2)$	2	0.68	0.71	0.50	0.52	0.084	0.078	5.41	5.77	0.06	15.15
$p_t^\perp$	500	$(0.9, 0.5)$	1	0.45	0.37	0.35	0.28	0.055	0.047	1.87	1.45	-0.22	40.93
$p_t^\perp$	500	$(0.9, 0.5)$	2	0.43	0.38	0.33	0.29	0.056	0.042	1.70	1.50	-0.11	46.62
$p_t^\perp$	500	$(0.5, 0.2)$	1	0.72	0.67	0.52	0.48	0.058	0.060	5.12	4.00	-0.21	17.68
$p_t^\perp$	500	$(0.5, 0.2)$	2	4.54	7.49	0.65	0.75	0.060	0.057	4.30	4.56	0.05	19.16
$y_t^\perp$	200	$(0.9, 0.5)$	1	0.45	0.44	0.35	0.34	0.094	0.097	1.76	1.68	-0.04	33.97
$y_t^\perp$	200	$(0.9, 0.5)$	2	0.42	0.41	0.32	0.32	0.080	0.076	1.67	1.72	0.02	35.35
$y_t^\perp$	200	$(0.5, 0.2)$	1	2.04	0.84	0.62	0.54	0.078	0.091	6.56	6.73	0.02	13.93
$y_t^\perp$	200	$(0.5, 0.2)$	2	0.68	0.71	0.50	0.51	0.084	0.087	5.41	6.04	0.11	15.15
$y_t^\perp$	500	$(0.9, 0.5)$	1	0.45	0.44	0.35	0.34	0.055	0.062	1.87	1.92	0.03	40.93
$y_t^\perp$	500	$(0.9, 0.5)$	2	0.43	0.42	0.33	0.33	0.056	0.060	1.70	1.71	0.00	46.62
$y_t^\perp$	500	$(0.5, 0.2)$	1	0.72	0.72	0.52	0.52	0.058	0.071	5.12	5.40	0.05	17.68
$y_t^\perp$	500	$(0.5, 0.2)$	2	4.54	5.95	0.65	0.70	0.060	0.071	4.31	4.79	0.11	19.16

Notes: The table reports the root mean squared error (RMSE), mean absolute error (MAE), empirical rejection probability (ERP) and the width of the confidence interval (CS length) for the standard IV estimator with AR confidence set and the innovation powered versions.

Table S5: Simulation Evidence — Fiscal SVAR (i)

				RMSE		MAE		ERP		CS length			
v_t	n	(c_1, c_2)	π	IV	IPIV	IV	IPIV	AR	IPAR	AR	IPAR	%Δ	F^{eff}
$p_t^\perp$	200	(0.9, 0.5)	25	2.97	2.81	2.06	1.66	0.042	0.057	20.84	10.84	-0.48	18.05
$p_t^\perp$	200	(0.9, 0.5)	50	1.77	1.48	1.35	1.12	0.058	0.066	8.93	5.92	-0.33	30.95
$p_t^\perp$	200	(0.9, 0.5)	100	1.14	1.06	0.88	0.82	0.056	0.060	4.71	3.98	-0.15	61.82
$p_t^\perp$	200	(0.9, 0.5)	150	0.96	0.93	0.75	0.73	0.065	0.086	3.62	3.30	-0.08	97.17
$p_t^\perp$	200	(0.5, 0.2)	25	11.60	21.48	4.30	4.60	0.049	0.055	30.35	22.78	-0.24	7.95
$p_t^\perp$	200	(0.5, 0.2)	50	4.12	6.36	2.24	2.19	0.054	0.066	19.23	14.79	-0.23	12.73
$p_t^\perp$	200	(0.5, 0.2)	100	1.77	1.74	1.34	1.31	0.058	0.070	8.72	7.41	-0.15	22.82
$p_t^\perp$	200	(0.5, 0.2)	150	1.39	1.40	1.08	1.09	0.049	0.067	6.03	5.51	-0.08	33.66
$p_t^\perp$	500	(0.9, 0.5)	25	3.05	2.48	2.08	1.58	0.030	0.038	19.39	10.25	-0.47	21.28
$p_t^\perp$	500	(0.9, 0.5)	50	1.66	1.33	1.28	1.02	0.050	0.046	8.30	5.51	-0.33	35.98
$p_t^\perp$	500	(0.9, 0.5)	100	1.15	0.96	0.92	0.76	0.050	0.045	4.87	3.82	-0.21	66.04
$p_t^\perp$	500	(0.9, 0.5)	150	0.91	0.78	0.71	0.62	0.040	0.042	3.73	3.11	-0.16	97.18
$p_t^\perp$	500	(0.5, 0.2)	25	10.38	29.29	4.08	4.88	0.034	0.046	29.45	21.48	-0.27	8.85
$p_t^\perp$	500	(0.5, 0.2)	50	3.74	3.81	2.16	1.98	0.044	0.051	18.36	14.12	-0.23	13.76
$p_t^\perp$	500	(0.5, 0.2)	100	1.92	1.79	1.47	1.36	0.067	0.070	8.58	7.18	-0.16	25.23
$p_t^\perp$	500	(0.5, 0.2)	150	1.42	1.32	1.11	1.03	0.055	0.055	6.06	5.32	-0.12	35.90
$y_t^\perp$	200	(0.9, 0.5)	25	2.97	3.11	2.06	2.07	0.042	0.104	20.84	18.41	-0.11	18.05
$y_t^\perp$	200	(0.9, 0.5)	50	1.77	1.72	1.35	1.31	0.058	0.067	8.93	7.98	-0.10	30.95
$y_t^\perp$	200	(0.9, 0.5)	100	1.14	1.10	0.88	0.85	0.050	0.059	4.71	4.38	-0.07	61.82
$y_t^\perp$	200	(0.9, 0.5)	150	0.96	0.93	0.75	0.73	0.065	0.082	3.62	3.35	-0.07	97.17
$y_t^\perp$	200	(0.5, 0.2)	25	11.60	11.18	4.30	4.31	0.049	0.111	30.35	27.87	-0.08	7.95
$y_t^\perp$	200	(0.5, 0.2)	50	4.12	3.38	2.24	2.16	0.054	0.088	19.23	17.60	-0.08	12.73
$y_t^\perp$	200	(0.5, 0.2)	100	1.77	1.76	1.34	1.34	0.058	0.079	8.72	7.92	-0.09	22.82
$y_t^\perp$	200	(0.5, 0.2)	150	1.39	1.36	1.08	1.06	0.049	0.066	6.03	5.51	-0.08	33.66
$y_t^\perp$	500	(0.9, 0.5)	25	3.05	3.24	2.08	2.13	0.030	0.115	19.39	17.73	-0.08	21.28
$y_t^\perp$	500	(0.9, 0.5)	50	1.66	1.64	1.28	1.27	0.050	0.069	8.30	7.68	-0.07	35.98
$y_t^\perp$	500	(0.9, 0.5)	100	1.15	1.12	0.92	0.89	0.050	0.056	4.87	4.68	-0.03	66.04
$y_t^\perp$	500	(0.9, 0.5)	150	0.91	0.88	0.71	0.69	0.040	0.044	3.73	3.64	-0.02	97.18
$y_t^\perp$	500	(0.5, 0.2)	25	10.38	8.99	4.08	3.96	0.034	0.109	29.45	27.30	-0.07	8.85
$y_t^\perp$	500	(0.5, 0.2)	50	3.74	4.58	2.16	2.24	0.044	0.075	18.36	17.35	-0.05	13.76
$y_t^\perp$	500	(0.5, 0.2)	100	1.92	1.90	1.47	1.44	0.067	0.091	8.58	8.15	-0.05	25.23
$y_t^\perp$	500	(0.5, 0.2)	150	1.42	1.41	1.11	1.11	0.055	0.063	6.06	5.84	-0.03	35.90

Notes: The table reports the root mean squared error (RMSE), mean absolute error (MAE), empirical rejection probability (ERP) and the width of the confidence interval (CS length) for the standard IV estimator with AR confidence set and the innovation powered versions.

Table S6: Simulation Evidence — Fiscal SVAR (ii)

				RMSE		MAE		ERP		CS length			
v_t	n	(c_1, c_2)	π	IV	IPIV	IV	IPIV	AR	IPAR	AR	IPAR	$\%\Delta$	F^{eff}
$p_t^{\perp}$	200	$(0.9, 0.5)$	1	1.47	1.21	1.13	0.94	0.051	0.052	6.27	4.76	-0.23	44.77
$p_t^{\perp}$	200	$(0.9, 0.5)$	2	1.22	1.11	0.95	0.87	0.060	0.056	5.07	4.46	-0.12	57.31
$p_t^{\perp}$	200	$(0.5, 0.2)$	1	2.26	2.17	1.66	1.56	0.047	0.057	13.18	10.52	-0.20	17.60
$p_t^{\perp}$	200	$(0.5, 0.2)$	2	1.87	1.88	1.40	1.41	0.051	0.063	9.91	9.55	-0.03	22.25
$p_t^{\perp}$	500	$(0.9, 0.5)$	1	1.46	1.16	1.13	0.91	0.053	0.049	6.33	4.59	-0.27	47.39
$p_t^{\perp}$	500	$(0.9, 0.5)$	2	1.21	1.04	0.94	0.82	0.060	0.042	5.04	4.26	-0.15	63.00
$p_t^{\perp}$	500	$(0.5, 0.2)$	1	2.52	2.42	1.80	1.65	0.051	0.058	13.29	10.42	-0.21	18.64
$p_t^{\perp}$	500	$(0.5, 0.2)$	2	1.85	1.87	1.39	1.39	0.049	0.042	9.70	9.38	-0.03	22.94
$y_t^{\perp}$	200	$(0.9, 0.5)$	1	1.47	1.43	1.13	1.10	0.051	0.071	6.27	5.80	-0.07	44.77
$y_t^{\perp}$	200	$(0.9, 0.5)$	2	1.22	1.21	0.95	0.94	0.060	0.068	5.07	4.77	-0.05	57.31
$y_t^{\perp}$	200	$(0.5, 0.2)$	1	2.26	2.33	1.66	1.68	0.047	0.071	13.18	11.93	-0.09	17.60
$y_t^{\perp}$	200	$(0.5, 0.2)$	2	1.87	1.88	1.40	1.40	0.051	0.073	9.91	9.27	-0.06	22.25
$y_t^{\perp}$	500	$(0.9, 0.5)$	1	1.46	1.46	1.13	1.13	0.053	0.075	6.33	5.96	-0.05	47.39
$y_t^{\perp}$	500	$(0.9, 0.5)$	2	1.21	1.19	0.94	0.93	0.060	0.072	5.04	4.84	-0.03	63.00
$y_t^{\perp}$	500	$(0.5, 0.2)$	1	2.52	2.52	1.80	1.80	0.051	0.072	13.29	12.62	-0.05	18.64
$y_t^{\perp}$	500	$(0.5, 0.2)$	2	1.85	1.84	1.39	1.38	0.049	0.056	9.70	9.30	-0.04	22.94

Notes: The table reports the root mean squared error (RMSE), mean absolute error (MAE), empirical rejection probability (ERP) and the width of the confidence interval (CS length) for the standard IV estimator with AR confidence set and the innovation powered versions.