Dynamic Identification Using System Projections and Instrumental Variables

Daniel Lewis Karel Mertens

ONLINE APPENDIX

A Testing the Null Hypothesis of Weak Instruments

This section describes the weak instruments test in the SP-IV model discussed in Section 2.2 of the main text. The test nests the popular bias-based test of Stock and Yogo (2005) when $H = 1$. The development of the latter test is analogous to that of the weak-instruments test in Lewis and Mertens (2022), which extends the Stock and Yogo (2005) test to be robust to autocorrelation and heteroskedasticity. Mathematically, the extension of the Stock and Yogo (2005) test in Lewis and Mertens (2022) closely resembles the extension required for SP-IV to allow $H > 1$.

We first establish some specific notation: $||U||_2$ is the spectral norm of U (the positive square root of the maximum eigenvalue of UU', also the ℓ_2-norm if U is a vector), $\mathbb{P}^n$ is the set of positive definite $n \times n$ matrices, $\mathbb{O}^{n\times m}$ is the set of $n \times m$ orthogonal real matrices U such that $UU' = I_n$, $\mathcal{K}_{n,m}$ denotes the $n \times m$ commutation matrix such that $\mathcal{K}_{n,m} \operatorname{vec}(U) = \operatorname{vec}(U')$ where $U \in \mathbb{R}^{n\times m}$. We also define the special matrix $R_{n,m} = I_n \otimes \operatorname{vec}(I_m)$. The dimension of $R_{n,m}$ is $nm^2 \times n$. For $U \in \mathbb{R}^{nm\times nm}$, the (i, j)-th element of $V = R'_{n,m}(U \otimes I_m)R_{n,m} \in \mathbb{R}^{n\times n}$ is $\operatorname{Tr}(U_{ij})$ where $U_{ij} \in \mathbb{R}^{m\times m}$ is (i, j)-th block of U and $\operatorname{Tr}(\cdot)$ is the trace. For $U \in \mathbb{R}^{nm\times m}$, the i-th element of $V = R'_{n,m} \operatorname{vec}(U') \in \mathbb{R}^n$ is equal to $\operatorname{Tr}(U_i)$ where $U_i \in \mathbb{R}^{m\times m}$ is the i-th row block of U. Note that $R'_{n,m}R_{n,m} = mI_N$.

A.1 Weak IV Representation of the SP-IV Estimator

Using the more general notation for the restriction matrix R defined above, the SP-IV estimator is

$$\text{(A.1)}\hat{\beta} \quad = \quad \left(R'_{K,H}(Y_H^{\perp}P_{Z^{\perp}}Y_H^{\perp\prime}\otimes I_H)R_{K,H}\right)^{-1}R'_{K,H}\,\text{vec}(y_H^{\perp}P_{Z^{\perp}}Y_H^{\perp\prime})\ ,$$

where $P_{Z^{\perp}} = Z^{\perp\prime}(Z^{\perp}Z^{\perp\prime})^{-1}Z^{\perp}$. As is standard in the literature, (see, e.g., Staiger and Stock (1997)), we assume identification but first-stage parameters that are local-to-zero.

Assumption 4. *$\Theta_Y = C/\sqrt{T}$ where $C \in \mathbb{R}^{HK\times N_z}$ is a fixed matrix and $R_{K,H}(CC'\otimes I_H)R_{K,H}$ is of full rank.*

This assumption implies that the instruments are weak under the null hypothesis. The following replace Assumptions 2 and 3 to allow the characterization of the weak instrument asymptotic distribution of $\hat{\beta}$.

Assumption 5. *The following limits hold as $T\to\infty$:*

$$\text{(5.a)}\qquad u_H^{\perp}u_H^{\perp\prime}/T \xrightarrow{p} \Sigma_{u_H^{\perp}} \ \in \mathbb{P}^{H},$$

$$u_H^{\perp}v_H^{\perp\prime}/T \xrightarrow{p} \Sigma_{u_H^{\perp}v_H^{\perp}} \ \in \mathbb{R}^{H\times HK},$$

$$v_H^{\perp}v_H^{\perp\prime}/T \xrightarrow{p} \Sigma_{v_H^{\perp}} \in \mathbb{P}^{HK}\ ,$$

$$\text{(5.b)}\qquad T^{-\frac{1}{2}}\begin{bmatrix}\text{vec}((Z^{\perp}Z^{\perp\prime})^{-\frac{1}{2}}Z^{\perp}w_H^{\perp\prime})\\ \text{vec}((Z^{\perp}Z^{\perp\prime})^{-\frac{1}{2}}Z^{\perp}v_H^{\perp\prime})\end{bmatrix}\xrightarrow{d}\mathcal{N}(0,\mathbf{W}\otimes I_{N_z})\ ,$$

$$\text{(5.c)}\qquad \textit{and } \hat{\mathbf{W}}\xrightarrow{p}\mathbf{W}$$

where $\mathbf{W} = \begin{bmatrix}\mathbf{W}_1 & \mathbf{W}_{12}\\ \mathbf{W}'_{12} & \mathbf{W}_2\end{bmatrix}\in\mathbb{P}^{(K+1)H}$.

$w_H^{\perp} = y_H^{\perp} - (\beta'\otimes I_H)\Theta_Y Q^{-\frac{1}{2}}Z^{\perp}$ are the reduced-form errors with covariance matrix $\mathbf{W}_1$, $v_H^{\perp} = Y_H^{\perp} - \Theta_Y Q^{-\frac{1}{2}}Z^{\perp}$ are first-stage error terms with covariance matrix $\mathbf{W}_2$, and $\mathbf{W}$ is the joint covariance of the reduced-form and first-stage errors.

The SP-IV estimator can be rewritten as

$$\text{(A.2)}\qquad \hat{\beta} \quad = \quad \left(R'_{K,H}(s_{ZY}s'_{ZY}\otimes I_H)R_{K,H}\right)^{-1}R'_{K,H}\,\text{vec}(s_{Zy}s'_{ZY}).$$

where $s_{Zy} = y_H^{\perp} Z^{\perp\prime}(Z^{\perp}Z^{\perp\prime})^{-\frac{1}{2}}$ and $s_{ZY} = Y_H^{\perp} Z^{\perp\prime}(Z^{\perp}Z^{\perp\prime})^{-\frac{1}{2}}$. This alternative expression reformulates $\hat{\beta}$ in terms of random vectors with asymptotic distributions given in Assumption 5. Define the random variables η_1 and η_2 ($H \times N_z$ and $HK \times N_z$ respectively) as

$$\begin{bmatrix} \text{vec}(\eta_1) \\ \text{vec}(\eta_2) \end{bmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mathbf{0}_{HN_z} \\ \text{vec}(C) \end{pmatrix}, \mathbf{S} \otimes I_{N_z} \right) \tag{A.3}$$

where $\mathbf{S} \in \mathbb{P}^{(K+1)H}$, partitioned as $\mathbf{W}$ with

(A.4)

$$\mathbf{S}_1 = \mathbf{W}_1 + (\beta' \otimes I_H)\mathbf{W}_2(\beta \otimes I_H) - (\beta' \otimes I_H)\mathbf{W}_{12}' - \mathbf{W}_{12}(\beta \otimes I_H),$$
$$\mathbf{S}_{12} = \mathbf{W}_{12} - (\beta' \otimes I_H)\mathbf{W}_2 \; , \; \mathbf{S}_2 = \mathbf{W}_2,$$

such that $\mathbf{S} \otimes I_{N_z}$ is the asymptotic covariance of $T^{-\frac{1}{2}} \left[\; \text{vec}(u_H^{\perp} Z^{\perp\prime}(Z^{\perp}Z^{\perp\prime})^{-\frac{1}{2}})' \quad \text{vec}(Y_H^{\perp} Z^{\perp\prime}(Z^{\perp}Z^{\perp\prime})^{-\frac{1}{2}})' \; \right)'$. Proposition 6 then characterizes the distribution of the random variable $\beta^* = \hat{\beta} - \beta$.

Proposition 6. *Under Assumptions 4 and 5, $s_{ZY} \xrightarrow{d} \eta_2$ and $s_{Zy} \xrightarrow{d} (\beta' \otimes I_H)\eta_2 + \eta_1$, and thus*

$$\hat{\beta} - \beta \xrightarrow{d} \beta^* = \left(R_{K,H}'(\eta_2\eta_2' \otimes I_H)R_{K,H}\right)^{-1} R_{K,H}' \,\text{vec}(\eta_1\eta_2').$$

Proof. The results follow directly from the stated assumptions, the expression for $\hat{\beta}$ in (A.2), and the continuous mapping theorem. □

Since β^* converges to a quotient of quadratic forms in normal random variables, $\hat{\beta}$ is not a consistent estimator of β. The asymptotic bias of the SP-IV estimator is the expected value $E[\beta^*]$. Before introducing the weak instruments set, we define the concentration matrix for the model.

Definition 1. *The concentration matrix is $\Lambda = \frac{1}{N_z}\Phi^{-\frac{1}{2}}R_{K,H}(CC' \otimes I_H)R_{K,H}\Phi^{-\frac{1}{2}}$ where $\Phi = R_{K,H}'(\mathbf{S}_2 \otimes I_H)R_{K,H}$.*

A.2 Definition of Weak Instruments

We consider instruments weak when a weighted ℓ_2-norm of the asymptotic bias $E[\beta^*]$ is large relative to a worst-case benchmark.

Definition 2. *The bias criterion is* $B = \text{Tr}(\mathbf{S}_1)^{-\frac{1}{2}}||E\left[\beta^*\right]' \Phi^{\frac{1}{2}}||_2$.

Following Stock and Yogo (2005), the ℓ_2-norm in the bias criterion aggregates the K elements of the bias through a quadratic loss function, such that B is weakly positive and penalizes larger biases more heavily. The criterion applies a weighting matrix, Φ, to put the elements of $E\left[\beta^*\right]$ on a comparable scale. The weighting matrix Φ effectively standardizes the regressors in the second stage, so that they have unit standard deviation and are orthogonal. The bias criterion also scales by $\text{Tr}(\mathbf{S}_1)$, which is the probability limit of $T^{-1}u_H^{\perp} P_{Z^{\perp}} u_H^{\perp\prime}$. This scaling expresses B as a ratio, relative to the same worst-case bias as in Montiel-Olea and Pflueger (2013), and Lewis and Mertens (2022). The intuition for the worst-case bias is given by the ad-hoc approximation of $E[\beta^*]$ in terms of a ratio of expectations as in Staiger and Stock (1997):

$$\text{(A.5)} \qquad E[\beta^*] \approx \frac{\text{vec}(\mathbf{S}_{12})' R_{K,H} \Phi^{-\frac{1}{2}}}{\text{Tr}(\mathbf{S}_1)^{\frac{1}{2}}} (I_K + \Lambda)^{-1} \Phi^{-\frac{1}{2}} \, \text{Tr}(\mathbf{S}_1)^{\frac{1}{2}}$$

Using this approximation, the bias criterion in (2) reaches a maximum of unity when the errors $u_H^{\perp}$ are perfect linear combinations of the second-stage regressors, $v_H^{\perp} P_Z^{\perp}$, such that the first term in (A.5) is a $K \times 1$ unit vector, and when the instruments are completely uninformative so the concentration matrix, Λ, is zero.

Definition 3. *The weak instrument set is*

$$\text{(A.6)} \qquad \mathbb{B}_\tau(\mathbf{W}) = \{C \in \mathbb{R}^{N \times K}, \beta \in \mathbb{R}^N : B \geq \tau\}.$$

The weak instrument set is the set of values for β and the first-stage parameters C such that bias B exceeds a tolerance level τ. This set depends on $\mathbf{W}$, which can be consistently estimated, but also on C, and the K unknown parameters in β.

A.3 Characterizing the Boundary of the Weak Instrument Set

Under Assumptions 4 & 5, the bias criterion in Definition 2 can be decomposed as $B = ||\mathbf{h}\rho||_2$, where

$$\mathbf{h} = HE\left[\left(R'_{K,H}(\mathcal{S}(l+\psi)(l+\psi)'\mathcal{S}' \otimes I_H)R_{K,H}\right)^{-1} R'_{K,H}\left(\mathcal{S}(l+\psi)\psi'\mathcal{S}^{-1} \otimes I_H\right)\right],$$
$$\rho = \left(\Phi^{-\frac{1}{2}} \otimes I_{H^2}\right) \mathrm{vec}\left(\mathbf{S}_{12}\right) / \sqrt{\mathrm{Tr}(\mathbf{S}_1)}\ ,$$

$l = \mathbf{S}_2^{-\frac{1}{2}} C$, $\psi = \mathbf{S}_2^{-\frac{1}{2}}(\eta_2 - C)$, $\mathrm{vec}(\psi) \sim \mathcal{N}(0, I_{KHN_z})$, and $\mathcal{S} = ((\Phi/H)^{-\frac{1}{2}} \otimes I_H)\mathbf{S}_2^{\frac{1}{2}}$. This decomposition is analogous to that of Lemma 1 in Lewis and Mertens (2022). The matrix $\mathbf{h}$ is the expected value of a random matrix that is a function of ψ, a matrix with i.i.d standard normal variables as elements. This expected value – when it exists – also depends on location parameters C and on $\mathbf{W}_2$. The vector ρ depends on $\mathbf{W}$ and β. In general, there is no tractable analytical expression for the integral underlying the expectation in $\mathbf{h}$, which is required to evaluate the bias. Following Montiel-Olea and Pflueger (2013) and Lewis and Mertens (2022), we adopt a Nagar (1959) approximation to $\mathbf{h}$ around $\psi = 0$, which we denote by $\mathbf{h}_n$. The Nagar bias is defined as $B_n = ||\mathbf{h}_n\rho||_2$. Using the eigenvalue decomposition $\Lambda = Q_\Lambda \mathcal{D}_\Lambda Q'_\Lambda$, the Nagar approximation of $\mathbf{h}$ around $\psi = 0$ is given by

(A.7)
$$\mathbf{h}_n = N_z^{-1} Q_\Lambda \mathcal{D}_\Lambda^{-\frac{1}{2}} M_1 (\mathcal{D}_\Lambda^{-\frac{1}{2}} Q_\Lambda \otimes L_0 \otimes I_K)(I_{KH} \otimes (I_{N_z} \otimes L_0)\mathcal{K}_{N_z,HN_z} R_{H,N_z}) M_2$$

with $L_0 = HN_z{}^{-\frac{1}{2}} Q'_\Lambda \Lambda^{-\frac{1}{2}} R'_{K,HN_z}(\mathcal{S}\,\mathrm{vec}(l) \otimes I_{HN_z}) \in \mathbb{O}^{K\times HN_z}$, $M_1 = R'_{K,K}\left(I_{K^3} + (\mathcal{K}_{K,K} \otimes I_K)\right)$ and $M_2 = R_{K,H}R'_{K,H}/(K+1) - I_{KH^2}$.

Analogous to Lewis and Mertens (2022), we base our test on

$$B_n \leq \lambda_{\min}^{-1} \mathcal{B}(\mathbf{W})\ , \tag{A.8}$$

where $\lambda_{\min} = \text{mineval}\{\Lambda\}$ and

(A.9)
$$\mathcal{B}(\mathbf{W}) = (N_z\sqrt{H})^{-1}\sup_{L_0}\{||M_1(I_K \otimes L_0 \otimes I_K)(I_{KH} \otimes (I_{N_z} \otimes L_0)\mathcal{K}_{N_z,HN_z}R_{H,N_z})M_2\Psi||_2\} \ ,$$

(A.10)
$$\Psi = \big(\mathcal{S}\mathbf{W}_2^{-\frac{1}{2}}[\mathbf{W}_{12} : \mathbf{W}_2]' \otimes I_H)R_{K+1,H}(R'_{K+1,H}(\mathbf{W} \otimes I_H)R_{K+1,H})^{-\frac{1}{2}} \ .$$

A.4 Null Hypothesis

Given a bias tolerance level τ, the test of the null hypothesis of weak instruments is based on a test of whether the minimum eigenvalue of Λ is less than or equal to a threshold value $\lambda^*_{\min}(\tau)$. More formally, the null and alternative hypotheses for the test are

(A.11)
$$H_0 : \lambda_{\min} \in \mathcal{H}(\mathbf{W}) \ \text{ vs. } \ H_1 : \lambda_{\min} \notin \mathcal{H}(\mathbf{W}), \quad \text{where } \ \mathcal{H}(\mathbf{W}) = \{\lambda_{\min} \in \mathbb{R}_+ : \lambda_{\min} \leq \lambda^*_{\min}(\tau)\},$$

where $\lambda^*_{\min}(\tau) = \mathcal{B}(\mathbf{W})/\tau$. The null hypothesis is that the minimum eigenvalue of the concentration matrix is in the set of values for which the worst-case Nagar bias is greater than the tolerance level τ. Under the alternative, the minimum eigenvalue is not in that set of values.

A.5 Test Statistic and Critical Values

The following proposition presents our statistic to test the null hypothesis.

Proposition 7. *Define the test statistic*

$$g = N_z^{-1}\,\text{mineval}\{\hat{\Phi}^{-\frac{1}{2}}(Y_H^{\perp}P_{Z^{\perp}}Y_H^{\perp\prime})\hat{\Phi}^{-\frac{1}{2}}\},$$

where $\hat{\Phi} = R'_{K,H}(\hat{\mathbf{W}}_2 \otimes I_H)R_{K,H}$. *Then, under Assumptions 4 and 5,*

$$g \xrightarrow{d} \text{mineval}\{R'_{K,H}(\zeta \otimes I_K)R_{K,H}/(HN_z)\},$$

where the $KH \times KH$ *random matrix* $\zeta = \mathcal{S}(l+\psi)(l+\psi)'\mathcal{S}'$ *has a noncentral Wishart distribution,* $\zeta \sim \mathcal{W}(N_z, \Sigma, \Omega)$, *with* N_z *degrees of freedom,* *covariance matrix* $\Sigma = \mathcal{S}\mathcal{S}' \in \mathbb{P}^{KH}$*, and a matrix of noncentrality parameters* $\Omega = \Sigma^{-1}\mathcal{S}ll'\mathcal{S}'$.[15]

Proof. The proposition follows from Slutsky's theorem, the continuous mapping theorem, and $Y_H^{\perp}P_{Z^{\perp}}Y_H^{\perp\prime} \xrightarrow{d} R'_{K,H}\left(\mathbf{S}_2^{\frac{1}{2}}(l+\psi)(l+\psi)'\mathbf{S}_2^{\frac{1}{2}\prime} \otimes I_K\right)R_{K,H}$, which implies the stated distribution of ζ. □

While ζ has a noncentral Wishart distribution, critical values for the test statistic g require the distribution of $\text{mineval}\{R'_{K,H}(\zeta \otimes I_H)R_{K,H}\}$, which is the minimum eigenvalue of the $K \times K$ matrix consisting of the traces of the $H \times H$ partitions of ζ. To the best of our knowledge, the distribution of this function of ζ is unknown. Moreover, the limiting distribution of g depends in general on all parameters in Σ and Ω, not just on the threshold for $\lambda_{\min}$.

To address both these challenges, we follow Stock and Yogo (2005) and Lewis and Mertens (2022) and obtain critical values from a bounding limiting distribution of g. Specifically, we consider the distribution of $\gamma' R'_{K,H}(\zeta \otimes I_H)R_{K,H}\gamma \geq \text{mineval}\{R'_{K,H}(\zeta \otimes I_H)R_{K,H}\}$ as a bounding distribution, where γ is the eigenvector associated with the minimum eigenvalue of Λ and $\gamma'\gamma = 1$. The following theorem is a straightforward extension of Theorem 2 in Lewis and Mertens (2022).

Theorem 1. *For* $\zeta \sim \mathcal{W}(N_z, \Sigma, \Omega)$,

(i) The n*-th cumulant of* $\gamma' R'_{K,H}(\zeta \otimes I_H)R_{K,H}\gamma$ *is*

$$\kappa_n = \; 2^{n-1}(n-1)!\Big(N_z \operatorname{Tr}\left(((\gamma\gamma' \otimes I_H)\Sigma)^n\right) + n \operatorname{Tr}\left(((\gamma\gamma' \otimes I_H)\Sigma)^n\Omega\right)\Big).$$

(ii) The n*-th cumulant* κ_n *with* $n > 1$ *is bounded by*

$$\begin{aligned}\kappa_n \leq & 2^{n-1}(n-1)!\Big(N_z \,\text{maxeval}\{R'_{K,H}(\Sigma^n \otimes I_H)R_{K,H}\} \\ & + nHN_z\lambda_{\min}\,\text{maxeval}\{\Sigma\}^{n-1}\Big).\end{aligned}$$

Proof. See Lewis and Mertens (2022). □

[15]We adopt the notational convention of Muirhead (1982) for the noncentral Wishart distribution.

As in Lewis and Mertens (2022), we consider the class of approximating distributions proposed by Imhof (1961), which match the first three cumulants of an unknown target distribution. We select the Imhof distribution with the largest critical value at significance level α subject to the constraints that the first cumulant, $\kappa_1 = HN_z(1 + \lambda_{\min})$, matches that of the target distribution, and that the second and third cumulants respect the analytical upper bounds on the cumulants of the limiting distribution of g. The resulting critical value is guaranteed to be conservative relative to the unknown critical value from the true limiting distribution of the test statistic, g.

B Additional Simulation Results

B.1 IRF Estimates in the Simulations

Figures B.1 and B.2 show the mean IRF estimates, together with 2.5% and 97.5% percentiles, across 5000 simulations from the Smets and Wouters (2007) model discussed in Section 3. The figures show IRFs estimated using a distributed lag specification, local projections with the set of predetermined control variables described in the main text, and a VAR in the variables of the control set. For brevity, we only show the IRFs associated with the monetary policy shock for $H = 20$ and $T = 250, 5000$. Results for the other specifications are available on request.

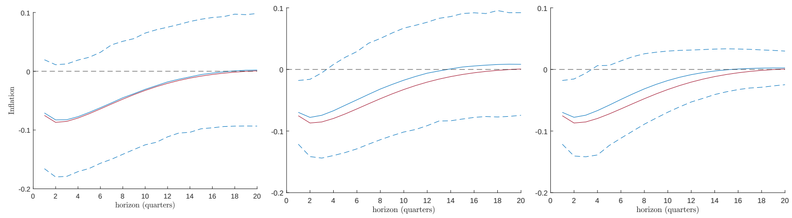
B.2 Simulation Results Using Three Instruments ($N_z = 3$)

This section presents the simulation results for specifications using three instruments: the monetary policy shock, the government spending shock, and the risk premium shock from the Smets and Wouters (2007) model. Panel a. of Table B.1 reports the mean estimates across 5000 Monte Carlo samples, Panel b. shows the standard deviations, and Table B.2 shows the empirical rejection rates.

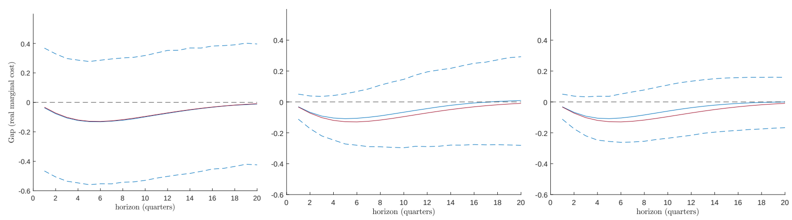
As mentioned in the main text, the results are qualitatively similar to those for a single instrument. SP-IV LP (without controls) and single-equation 2SLS (without controls) remain very close substitutes in terms of bias and variance. The addition of predetermined controls induces

FIGURE B.1: True and Estimated IRFs in Simulations, Small Sample ($T = 250$)

(a) Distributed Lag (b) LP with controls (c) VAR



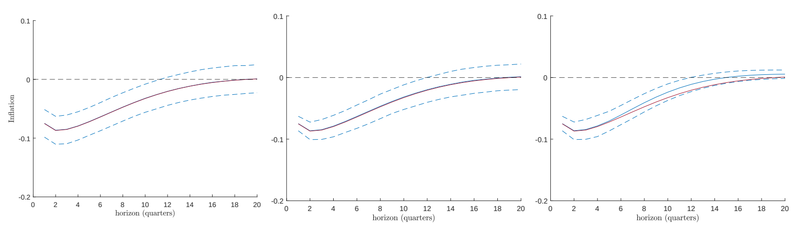
(d) Distributed Lag (e) LP with controls (f) VAR



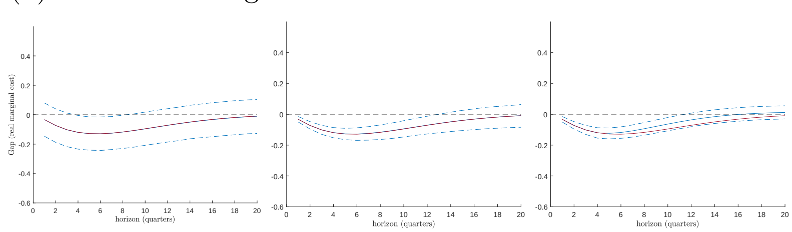
Notes: Figures show IRFs to a one s.t.d. contractionary monetary policy shock in data generated by the Smets and Wouters (2007) model. Red lines show the true IRFs. Blue lines show the mean and 2.5% and 97.5% percentiles of the estimated IRFs across 5000 samples.

FIGURE B.2: True and Estimated IRFs in Simulations, Large Sample ($T = 5000$)

(a) Distributed Lag (b) LP with controls (c) VAR



(d) Distributed Lag (e) LP with controls (f) VAR



Notes: See Figure B.1

Table B.1: Mean and Variance of parameter estimates, $N_z = 3$

a. Mean Parameter Estimates

	$T = 250$			$T = 500$			$T = 5000$		
Estimator	γ_b	γ_f	λ	γ_b	γ_f	λ	γ_b	γ_f	λ
True Value	0.15	0.85	0.05	0.15	0.85	0.05	0.15	0.85	0.05
OLS	0.47	0.47	0.00	0.48	0.48	0.00	0.48	0.48	0.00
$H = 8$									
2SLS	0.40	0.55	0.01	0.36	0.63	0.01	0.23	0.82	0.02
2SLS-C	0.36	0.59	-0.04	0.33	0.64	-0.02	0.22	0.83	-0.01
SP-IV LP	0.39	0.55	0.01	0.36	0.63	0.01	0.22	0.82	0.02
SP-IV LP-C	0.40	0.55	0.02	0.36	0.63	0.03	0.20	0.81	0.04
SP-IV VAR	0.34	0.69	0.01	0.29	0.75	0.02	0.20	0.83	0.04
$H = 20$									
2SLS	0.45	0.51	0.00	0.43	0.55	0.00	0.28	0.76	0.01
2SLS-C	0.40	0.52	-0.07	0.38	0.56	-0.05	0.25	0.77	0.00
SP-IV LP	0.44	0.51	0.00	0.42	0.56	0.00	0.28	0.76	0.01
SP-IV LP-C	0.44	0.50	0.01	0.43	0.55	0.01	0.27	0.76	0.02
SP-IV VAR	0.35	0.69	0.01	0.31	0.75	0.01	0.23	0.82	0.02

b. Standard Deviation of Parameter Estimates

	$T = 250$			$T = 500$			$T = 5000$		
Estimator	γ_b	γ_f	λ	γ_b	γ_f	λ	γ_b	γ_f	λ
$H = 8$									
2SLS	0.09	0.09	0.03	0.08	0.09	0.03	0.06	0.05	0.02
SP-IV LP	0.10	0.10	0.04	0.09	0.09	0.03	0.06	0.05	0.02
SP-IV LP-C	0.09	0.10	0.06	0.09	0.09	0.05	0.06	0.05	0.03
SP-IV VAR	0.11	0.13	0.05	0.11	0.11	0.05	0.06	0.05	0.03
$H = 20$									
2SLS	0.04	0.04	0.01	0.04	0.04	0.01	0.04	0.04	0.01
SP-IV LP	0.05	0.05	0.02	0.04	0.04	0.01	0.04	0.04	0.01
SP-IV LP-C	0.04	0.05	0.02	0.04	0.04	0.02	0.04	0.04	0.01
SP-IV VAR	0.09	0.10	0.02	0.09	0.10	0.02	0.05	0.04	0.02

Notes: Panel a. reports mean parameter estimates and Panel b. reports the standard deviation of estimates. The first row in each of the top panels contains the true parameter values $\beta = [\gamma_b, \gamma_f, \lambda]'$ of (2) in the Smets and Wouters (2007) model. The other rows show the mean or standard deviation of estimates across 5000 Monte Carlo samples of size T and with $h = 0, \ldots, H-1$. All IV estimators use the true monetary policy shock, government spending shock, and risk premium shock in the model as instruments. 2SLS-Almon is the estimator proposed in Barnichon and Mesters (2020); "-C" is appended to indicate the inclusion of the predetermined controls. SP-IV is the estimator in (9). LP and LP-C denote implementations based on local projections without and with controls, respectively, while VAR denotes a vector autoregression.

Table B.2: Empirical size of nominal 5% tests, $N_z = 3$

	$H=8$			$H=20$		
	$T=250$	$T=500$	$T=5000$	$T=250$	$T=500$	$T=5000$
WALD 2SLS	81.5	76.8	54.8	100.0	99.8	92.5
WALD 2SLS-C	71.9	62.3	93.6	99.9	99.1	96.9
WALD SP-IV LP	82.2	78.2	56.3	99.9	99.9	91.7
WALD SP-IV LP-C	73.2	59.6	17.7	100.0	99.7	76.6
WALD SP-IV VAR	38.8	28.2	13.2	86.1	76.5	53.8
AR 2SLS	13.7	8.9	4.0	55.3	25.7	3.9
AR 2SLS-C	9.8	9.9	74.7	32.4	16.0	44.2
AR SP-IV LP	7.4	6.5	5.1	14.6	8.7	4.9
AR SP-IV LP-C	7.2	6.7	5.2	17.3	9.9	5.2
AR SP-IV VAR	3.8	4.7	4.8	6.3	5.9	4.7
KLM 2SLS	4.9	5.3	4.2	0.5	12.7	4.9
KLM 2SLS-C	3.3	4.8	33.9	1.2	3.1	7.3
KLM SP-IV LP	6.0	5.5	4.9	8.2	6.3	5.3
KLM SP-IV LP-C	7.2	6.3	4.9	11.3	7.9	5.1
KLM SP-IV VAR	6.5	6.8	4.8	10.7	8.6	5.3

Notes: The table shows empirical rejection rates of nominal 5% tests of the true values of $\beta = [\gamma_b, \gamma_f, \lambda]'$ in 5000 Monte Carlo samples from the Smets and Wouters (2007) model using the monetary policy shock, government spending shock and risk premium shock in the model as instruments. All IV estimators are based on $h = 0, ..., H-1$ and use the true monetary policy shock, government spending shock, and risk premium shock in the model as instruments. SP-IV LP and LP-C denote implementations based on local projections without and with X_{t-1} (described in the text) as controls, respectively. SP-IV VAR denotes implementation with a vector autoregression for X_t with four lags. Robust tests for 2SLS use a HAR Newey-West variance matrix with Sun (2014) fixed-b critical values; inference procedures for SP-IV are described in Section 2.

considerable additional bias in the 2SLS specifications, whereas it leads to lower bias in the SP-IV specifications. Overall, the size distortions for $T = 250$ and $H = 20$ are larger with three instruments, due to the many instruments problem. The robust SP-IV inference procedures continue to perform better in general than the 2SLS AR or KLM tests.

B.3 Simulation Results for Generalized SP-IV estimators

This section presents simulation results for the (feasible) generalized SP-IV estimators based on a 2-step procedure. First, we estimate our baseline SP-IV estimators and estimate the covariance matrix $\hat{\Sigma}_u^{\perp}$ using (20).

Table B.3: Standard deviation of parameter estimates, GSP-IV

	$T = 250$			$T = 500$			$T = 5000$		
Estimator	γ_b	γ_f	λ	γ_b	γ_f	λ	γ_b	γ_f	λ
$H = 8, N_z = 1$									
GSP-IV LP	0.33	0.45	0.26	0.28	0.41	0.23	0.12	0.09	0.09
GSP-IV LP-C	0.35	0.35	0.31	0.31	0.24	0.29	0.12	0.06	0.08
GSP-IV VAR	0.36	0.43	0.34	0.33	0.27	0.30	0.13	0.06	0.09
$H = 20, N_z = 1$									
GSP-IV LP	0.15	0.18	0.07	0.12	0.15	0.06	0.07	0.05	0.03
GSP-IV LP-C	0.11	0.12	0.07	0.10	0.09	0.06	0.08	0.05	0.03
GSP-IV VAR	0.24	0.29	0.14	0.22	0.20	0.12	0.12	0.06	0.06
$H = 8, N_z = 3$									
GSP-IV LP	0.11	0.12	0.04	0.09	0.11	0.03	0.06	0.05	0.02
GSP-IV LP-C	0.10	0.10	0.05	0.09	0.08	0.05	0.06	0.04	0.03
GSP-IV VAR	0.11	0.12	0.05	0.10	0.11	0.05	0.06	0.05	0.03
$H = 20, N_z = 3$									
GSP-IV LP	0.03	0.04	0.01	0.03	0.03	0.01	0.04	0.04	0.01
GSP-IV LP-C	0.03	0.03	0.01	0.03	0.03	0.01	0.04	0.03	0.01
GSP-IV VAR	0.07	0.09	0.02	0.08	0.09	0.02	0.05	0.04	0.02

Notes: Rows show standard deviations across 5000 Monte Carlo samples of size T with $h = 0, \ldots, H-1$. $N_z = 1$ estimators use the monetary policy shock as an instrument; $N_z = 3$ add the government spending and risk premium shocks as instruments. GSP-IV is the (feasible) generalized estimator in (B.1), obtained in a two-step procedure using (20).

Then, we use the latter to obtain the generalized SP-IV estimators as in (B.1). The generalized SP-IV estimators are also the feasible 2-step efficient GMM estimators.

Table B.3 reports the standard deviations of the estimates in the simulations. The generalized SP-IV, or "GSP-IV", estimators are in theory asymptotically more efficient than our baseline estimators. However, the feasible versions do not generally improve performance in practice, at least not in realistic sample sizes and for our data-generating process. For $N_z = 1$, all GSP-IV variances slightly exceed those of their SP-IV counterparts in Table 3 in the main text. With more instruments ($N_z = 3$), there is some sporadic evidence of (small) efficiency gains of GSP-IV relative to their SP-IV counterparts. The fact that GSP-IV does not consistently provide efficiency gains (and often fares slightly worse) in small samples likely results from estimation error in the $H \times H$ weighting matrix, which itself depends on the estimate $\hat{\beta}$, which is only weakly identified.

For brevity, we do not report the simulation results for the bias and empirical rejection rates, but they are available on request. The results are comparable overall to the regular SP-IV estimators discussed in the main text. The GSP-IV estimators consistently show somewhat greater bias than their SP-IV counterparts when additional instruments are included. The size distortions are also generally worse for the GSP-IV estimators than their SP-IV counterparts. In sum, at least in our setting, the simulation results offer little motivation to prefer GSP-IV over SP-IV in practice.

Online Appendix References

Barnichon, R., & Mesters, G. (2020). Identifying Modern Macro Equations with Old Shocks. *The Quarterly Journal of Economics*, *135*(4), 2255–2298.

Imhof, J. P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika*, *48*(3-4), 419–426.

Lewis, D. J., & Mertens, K. (2022). *A Robust Test for Weak Instruments with Multiple Endogenous Regressors* (Staff Reports No. 1020). Federal Reserve Bank of New York.

Montiel-Olea, J. L., & Pflueger, C. (2013). A robust test for weak instruments. *Journal of Business & Economic Statistics*, *31*(3), 358–369.

Muirhead, R. J. (1982). *Aspects of multivariate statistical theory*. John Wiley; Sons Ltd.

Nagar, A. L. (1959). The bias and moment matrix of the general k-class estimators of the parameters in simultaneous equations. *Econometrica*, *27*(4), 575–595.

Smets, F., & Wouters, R. (2007). Shocks and frictions in us business cycles: A bayesian dsge approach. *American Economic Review*, *97*(3), 586–606.

Staiger, D., & Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, *65*(3), 557–586.

Stock, J., & Yogo, M. (2005). Testing for weak instruments in linear iv regression. In D. W. Andrews (Ed.), *Identification and inference for econometric models* (pp. 80–108). Cambridge University Press.

Sun, Y. (2014). Let's fix it: Fixed-b asymptotics versus small-b asymptotics in heteroskedasticity and autocorrelation robust inference. *Journal of Econometrics*, *178*(P3), 659–677.