Techniques of Empirical Econometrics

Oscar Jordà
UC Davis, FRB San Francisco

Karel Mertens
FRB Dallas, CEPR

AEA Continuing Education, January 8-10, 2022

The views expressed do not necessarily reflect the views of the Federal Reserve Bank of Dallas, San Francisco or the Federal Reserve System.

Overview

3. More Inference for Impulse Responses

3.1 Weak Identification

3.2 Detecting Weak Instruments

3.3 Robust Inference

3.4 Proxy SVAR/SVAR-IV Inference

Inference with Instrumental Variables

IV methods are commonly used in empirical macro

- LP-IV
- Proxy SVAR/SVAR-IV
- But really, anytime you scale impulse responses to a fixed impact on an endogenous outcome variable at some horizon

IV estimates of impulse responses are quotients of sample moments

The numerator may be non-zero (identified), but small relative to the sampling error (weakly identified)

Division by deterministic zeros is not allowed

Division by statistics hard to distinguish from zero also creates problems

See Andrews, Stock, and Sun (2019), Keane and Neal (2022) for recent general discussions

General Model

Model with N endogenous regressors, $K \geq N$ instrumental variables

$$\begin{aligned} y &= Y\beta + \upsilon, \\ Y &= X\Pi + v, \end{aligned}$$

y :	$T \times 1$	outcome variable of interest
Y :	$T \times N$	endogenous regressors
X :	$T \times K$	instrumental variables, $K \geq N$
β :	$N \times 1$	parameters of interest
Π :	$K \times N$	'first stage' parameters
$\Gamma = \Pi\beta$:	$K \times 1$	'reduced form' parameters

All data is demeaned, wlg no exogenous regressors (Frisch Waugh), and $X'X/T = I_K$

Reduced form: $y = X\Gamma + w, \quad w = v\beta + \upsilon$

2SLS and $\hat{\mathbf{W}}$

Two-Stage-Least-Squares:

$$\hat{\beta}_{2SLS} = (Y'P_X Y)^{-1} Y'P_X y$$

where $P_X = XX'/T$

Let $\hat{\mathbf{W}}$ denote a robust estimator of the covariance matrix of the reduced-form and first-stage parameter estimates $[\hat{\Gamma}' \;\; \text{vec}(\hat{\Pi})']'$ with $\hat{\mathbf{W}} \xrightarrow{p} \mathbf{W}$

Use Newey-West, Huber-White, clustered, etc estimate as suited to the application

Partition $\mathbf{W} = \begin{bmatrix} \mathbf{W}_1 & \mathbf{W}_{12} \\ \mathbf{W}_{12}' & \mathbf{W}_2 \end{bmatrix} \quad (N+1)K \times (N+1)K$

So $\hat{\mathbf{W}}_2$ is the robust covariance of the first-stage parameter estimates

Weak Instruments Recap

Given the $X'X/T = I_K$ normalization, $\hat{\Pi} = X'Y/T$ and $\hat{\Gamma} = X'y/T$

Consider the $N = 1$ case:

$$\hat{\beta}_{2SLS} = \frac{Y'P_X y/T}{Y'P_X Y/T} = \frac{\hat{\Pi}'\hat{\Gamma}}{\hat{\Pi}'\hat{\Pi}}$$

Since $\Gamma = \Pi\beta$, the hope is that $\hat{\beta}_{2SLS} \approx \beta$ as $\Pi'\Pi$ cancels out in population

However, this never happens exactly with finite T and $\hat{\beta}_{2SLS}$ is biased

When $\hat{\Pi}'\hat{\Pi}$ is statistically small (the first stage is 'weak'), division-by-zero problems can cause the bias to be large

Since $\Pi'\Pi > 0$ by assumption, $Y'P_X Y/T = \hat{\Pi}'\hat{\Pi}$ being small is a small sample problem

Consequences of Weak Identification: Bias

Let's model the small sample distribution of $\hat{\beta}_{2SLS}$ using a weak instrument asymptotic approximation

Local-to-Zero Assumption

$Y = X\Pi + v$ with $\Pi = C/\sqrt{T}$ where C is a fixed full rank $K \times N$ matrix

First-stage relationship is local-to-zero ($\hat{\Pi}'\hat{\Pi}$ stays random even as $T \to \infty$)

$\hat{\beta}_{2SLS}$ **Asymptotics** Lewis and Mertens (2022)

Under **Local-to-Zero** and otherwise standard assumptions

$$\hat{\beta}_{2SLS} - \beta \xrightarrow{d} \beta^*_{2SLS} = \left(R'_{N,K}(\eta_2\eta_2' \otimes \mathcal{I}_K)R_{N,K}\right)^{-1} R'_{N,K}\,\mathrm{vec}(\eta_1\eta_2')$$

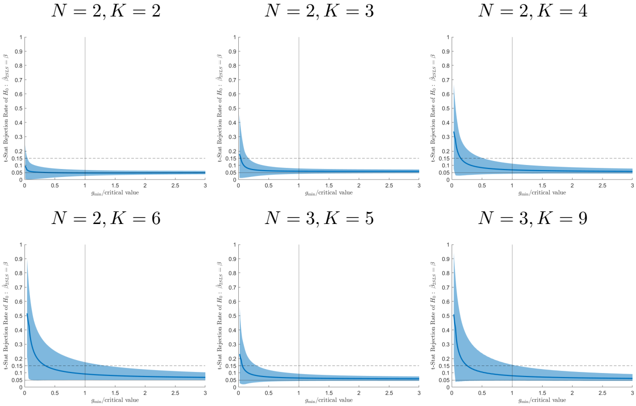
where $R_{N,K} = \mathcal{I}_N \otimes \mathrm{vec}(\mathcal{I}_K)$ and $[\eta_1' \ \eta_2'] \sim \mathcal{N}\left(\begin{pmatrix} \mathbf{0}_K \\ \mathrm{vec}(C') \end{pmatrix}, \mathbf{S}\right)$

$\mathbf{S}$ is the covariance of $T^{-\frac{1}{2}}[X'\upsilon \ \ \mathrm{vec}(X'v)']'$ as $T \to \infty$, and depends on $\mathbf{W}$ and β

β^*_{2SLS} is a (complicated) random variable and the asymptotic bias $E[\beta^*_{2SLS}]$ (when it exists) is not zero in general.

Consequences of Weak Identification: Size Distortions

Empirical rejection rates for nominal 5% two-sided t-statistics for the null that $\hat{\beta}_{2SLS} = \beta$ (y-axis) as a function of a measure of instrument strength (x-axis) across 5 million random DGPs



Weak instruments can cause serious size distortions for regular inference methods. Same for $N = 1$ models.

Detecting Weak Instruments

How do we know whether instruments are weak or not?

- **When are instruments considered weak or strong?**
- **What statistic is an indicator of instrument strength?**
- **What is the limiting distribution of this test statistic?**
- **Are critical values fast to obtain?**

Detecting Weak Instruments

How do we know whether instruments are weak or not?

Stock and Yogo (2005): Homoskedasticity, any N

- **When are instruments considered weak or strong?**

  Bias criterion, Size criterion

- **What statistic is an indicator of instrument strength?**

  Cragg and Donald (1993) statistic
  $=$ non-robust F-statistic when $N = 1$

- **What is the limiting distribution of this test statistic?**

  For $N = 1$: non-central χ^2
  For $N > 1$: unknown, use bounding non-central χ^2

- **Are critical values fast to obtain?**

  Yes, only depend on N and K. Look up in Stock and Yogo (2005) tables.

Detecting Weak Instruments

How do we know whether instruments are weak or not?

Montiel Olea and Pflueger (2013): Non-Homoskedasticity $N = 1$

- **When are instruments considered weak or strong?**

  Bias criterion

- **What statistic is an indicator of instrument strength?**

  'Effective F-statistic'
  = equals robust F-statistic when $K = 1$

- **What is the limiting distribution of this test statistic?**

  Unknown, weighted avg. of non-central χ^2's, use approx. distribution matching first two cumulants

- **Are critical values fast to obtain?**

  Yes, obtained numerically in each application using a second order approximation to the bias (`weakivtest.ado`)

Detecting Weak Instruments

How do we know whether instruments are weak or not?

Lewis and Mertens (2022): Non-Homoskedasticity, any N

- **When are instruments considered weak or strong?**

  Bias criterion

- **What statistic is an indicator of instrument strength?**

  Generalized First-Stage Statistic

- **What is the limiting distribution of this test statistic?**

  For $N = 1$, weighted avg. of non-central χ^2's, use approx. distribution matching first three cumulants

  For $N > 1$: unknown, mineval of matrix of traces, use approx. distribution matching first three cumulants of a bounding distribution

- **Are critical values fast to obtain?**

  Yes, obtained numerically in each application using a second order approximation to the bias (`gweakivtest.m`)

Defining Weak Instruments

Bias Criterion

$$B = \sqrt{E\left[\beta^*_{2SLS}\right]' \Phi E\left[\beta^*_{2SLS}\right]} / \sqrt{\mathrm{Tr}(\mathbf{S}_1)}$$

$\mathbf{S}_1$ is the covariance of $T^{-\frac{1}{2}}[X'v]'$ as $T \to \infty$

$B = 1$ in a worst-case scenario when the instruments are completely uninformative ($\Lambda = 0$) and v is a perfect linear combination of second-stage regressors $X\hat{\Pi}$

Bias criterion nests that in Stock and Yogo (2005) (under homoskedasticity) and Montiel Olea and Pflueger (2013) (when $N = 1$)

Fraction of the OLS bias under homoskedasticity, but not in general

Weak Instruments Definition

Instruments are weak for β and C such that $B \geq \tau$, where τ is a tolerance level

Nagar Bias

β^*_{2SLS} has intractable distribution

$E\left[\beta^*_{2SLS}\right]$ does not exist for $K = N$ (and $K = N + 1$ depending on assumptions)

$E\left[\beta^*_{2SLS}\right]$ has no known analytical form (except for $N = 1$ under homoskedasticity)

Two options:

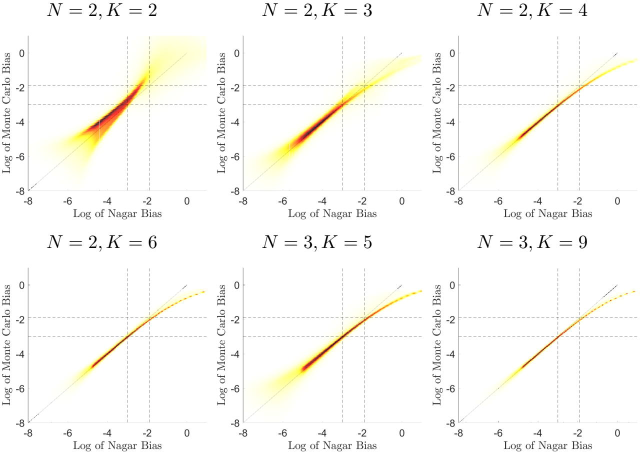
(1) obtain B by Monte Carlo simulations

(2) Nagar (1959) approximation B_n (second-order Taylor around $\eta_2 = \text{vec}(C'))$)

A function of **W**, so `weakivtest.ado` and `gweakivtest.m` use (2)

Nagar approximation reasonable for $K > N + 1$, but $N \leq K \leq N + 1$ can be problematic

Nagar Bias versus Monte Carlo Bias



Bounding the Bias

Nagar bias is a complicated function of β ($N \times 1$), C ($K \times N$) and $\mathbf{W}$

We can estimate $\mathbf{W}$ consistently, but not β and C

Nagar Bias Bounds Lewis and Mertens (2022)

- B_n has a sharp upper bound $B_n \leq \lambda_{\min}^{-1} \mathcal{B}(\mathbf{W})$ (**Worst-case Nagar bias**)
- B_n has an upper bound $B_n < \lambda_{\min}^{-1} \Psi(\mathbf{W})$

where $\lambda_{\min} = \text{mineval}\{\Lambda\}$ and $\Lambda = \Phi^{-\frac{1}{2}} C' C \Phi^{-\frac{1}{2}}$ is the **concentration matrix**

Both bounds depend only on a single unknown parameter $\lambda_{\min}$

In simulations (not a proof):

The worst-case Nagar bias is an effective upper bound on the Monte Carlo bias when $K > N + 1$

The non-sharp bound is an effective upper bound on the Monte Carlo bias when $K \leq N + 1$ when $N > 1$ and when $N = 1$ and $K = 2$

When $N = K = 1$, use the size-based Stock-Yogo test (OK, even under heteroskedasticity in this case)

Measuring Instrument Strength

We need a test statistic that is informative about λ_{min}

Generalized First-Stage Statistic Lewis and Mertens (2022)

For arbitrary N and $\hat{\mathbf{W}}$, the test statistic

$$g_{\text{min}} \quad = \quad \text{mineval}\{\hat{\Phi}^{-\frac{1}{2}} Y' P_X Y \hat{\Phi}^{-\frac{1}{2}}\},$$

measures instrument strength, $\hat{\Phi} = (\mathcal{I}_N \otimes \text{vec}(\mathcal{I}_K))'(\hat{\mathbf{W}}_2 \otimes I_K)(\mathcal{I}_N \otimes \text{vec}(\mathcal{I}_K))$

$\hat{\Phi}$ is the $N \times N$ matrix consisting of traces of the $K \times K$ partitions of $\hat{\mathbf{W}}_2$

The key property of the test statistic is that $E[g_{\text{min}}] = 1 + \lambda_{\text{min}}$

Test the null of weak instruments by testing whether $g_{\text{min}} \leq 1 + \mathcal{B}(\mathbf{W})/\tau$ (or $g_{\text{min}} \leq 1 + \Psi(\mathbf{W})/\tau$ for $K \leq N + 1$)

Measuring Instrument Strength

$$g_{\min} = \text{mineval}\{\hat{\Phi}^{-\frac{1}{2}} Y' P_X Y \hat{\Phi}^{-\frac{1}{2}}\}$$

Special cases:

- Under homoskedasticity and $N = 1$, $\hat{\mathbf{W}}_2 = \hat{\sigma}_v^2 \otimes I_K$, $\hat{\Phi} = K\hat{\sigma}_v^2$

  $g_{min} = Y' P_X Y / (K\hat{\sigma}_v^2)$ is the **non-robust F-statistic**

- Under homoskedasticity and $N \geq 1$, $\hat{\mathbf{W}}_2 = \hat{\Sigma}_v \otimes I_K$, $\hat{\Phi} = K\hat{\Sigma}_v$

  $g_{min} = K^{-1}\,\text{mineval}\{\hat{\Sigma}_v^{-\frac{1}{2}} Y' P_X Y \hat{\Sigma}_v^{-\frac{1}{2}}\}$ is the **Cragg and Donald (1993) test statistic** of the Stock and Yogo (2005) tests

- Under non-homoskedasticity and $N = K = 1$, $\hat{\mathbf{W}}_2 = (\hat{\sigma}_v^{rob})^2$, $\hat{\Phi} = (\hat{\sigma}_v^{rob})^2$

  $g_{min} = Y' P_X Y / (K(\hat{\sigma}_v^{rob})^2)$ is the **robust F-statistic**

- Under non-homoskedasticity and $N = 1$ and $K \geq 1$, $\hat{\Phi} = \text{Tr}(\hat{\mathbf{W}}_2)$

  $g_{min} = Y' P_X Y / \,\text{Tr}(\hat{\mathbf{W}}_2)$ is the **the effective F-statistic** of Montiel Olea and Pflueger (2013)

So $g_{\min}$ always gets it right (do not use Kleibergen and Paap (2006))

Critical Values

Under the null, g_{min} is the minimum eigenvalue of a matrix consisting of traces of the $K \times K$ partitions of $\mathbf{W}_2$

Lewis and Mertens (2022) derive analytical expressions for all the cumulants, and upper bounds on the cumulants that depend only on λ_{min}

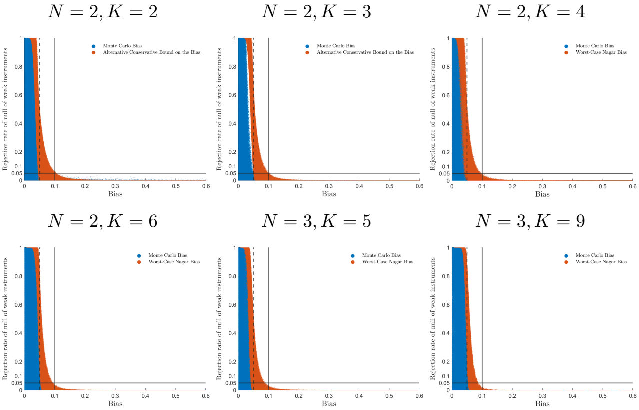
Critical values are obtained from a three-parameter approximating Imhof (1961) distribution that match the first three cumulants of the bounding limiting distribution

As in Stock and Yogo (2005), this makes the test conservative (false rejections with at most α probability, e.g. $\alpha = 5\%$)

The generalized test nevertheless has power

Size and Power

$\alpha = 5\%$, $\tau = 0.10$



Example: Ramey and Zubairy (2018) LP-IV

Ramey and Zubairy (2018) estimate state-dependent government spending multipliers using

$$
\begin{aligned}
\sum_{j=0}^{h} y_{t+j} =& I_{t-1}\left[\gamma_{A,h} + \phi_{A,h}(L) z_{t-1} + m_{A,h} \sum_{j=0}^{h} g_{t+j}\right] \\
&+ (1 - I_{t-1})\left[\gamma_{B,h} + \phi_{B,h}(L) z_{t-1} + m_{B,h} \sum_{j=0}^{h} g_{t+j}\right] + \omega_{t+h},
\end{aligned}
$$

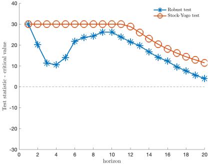
where I_t is an indicator of recession/boom or binding/non-binding ZLB

$N = 2,\ K = 4$

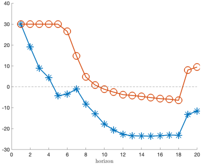
Example: Ramey and Zubairy (2018) LP-IV

(a) *Government Spending Interacted with Indicator of Slack*

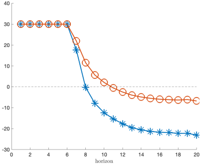
Full sample



Post-WWII

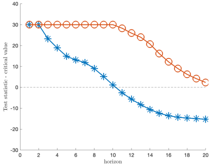


Excluding WWII

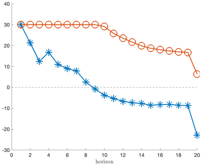


(b) *Government Spending Interacted with ZLB Indicator*

Full sample



Excluding WWII



Comments

- Download by clicking here: `gweakivtest.m`
- Same test statistic as Stock and Yogo (2005) and Montiel Olea and Pflueger (2013)
- Similar critical values to bias based test in Stock and Yogo (2005) under homoskedasticity (Nagar vs Monte Carlo)
- Virtually the same critical values as Montiel Olea and Pflueger (2013) when $N = 1$

  Except when $K = 2$:

  A more conservative bias bound is required when $K = 2$

  Example: Under homoskedasticity, the worst case nagar bias is zero when $K = 2$!
- Selection on first-stage statistics creates additional size distortions, e.g. Andrews, Stock, and Sun (2019)

Robust Inference Methods

When instruments are weak, several other robust inference methods are available

All work by test inversion

Suppose you have a test that correctly rejects at α % under the null that $\beta = \beta_0$

Perform the test over a set of possible values for β_0, the acceptance region contains β with $1 - \alpha$ % probability

Options under Non-Homoskedasticity:

- AR statistic, Anderson and Rubin (1949), Chernozhukov and Hansen (2008)
- KLM statistic, Kleibergen (2007), Chernozhukov and Hansen (2008)
- CLR statistic See overview in Andrews, Stock, and Sun (2019)

Not widely used, and an area of ongoing research

Inference based on Anderson and Rubin (1949)

Define $\hat{u}(\beta_0) = y - Y\beta_0$

Run the regression $\hat{u}(\beta_0) = X\gamma + w$ yielding $\hat{\gamma}(\beta_0) = X'\hat{u}(\beta_0)/T$

The AR statistic is the (robust) Wald-statistic

$$AR(\beta_0) = \hat{\gamma}(\beta_0)' Var(\hat{\gamma}(\beta_0))^{-1}\hat{\gamma}(\beta_0) \qquad AR(\beta_0)|_{H_0:\beta=\beta_0} \xrightarrow{d} \chi^2_K$$

Perform the test over a grid β_0, the $(1-\alpha)$ % confidence set is all grid points for which $AR(\beta_0) < \chi^2_{K,\alpha}$

Valid regardless of the value of Π, even $\Pi = 0$.

- Optimal for $N = K = 1$, inefficient when $N = 1$ and $K > 1$
- For $N > 1$, subvector confidence sets for β_j given by $\min_{\beta_0^{-j}} AR(\beta_0^j, \beta_0^{-j})$ are very conservative
- Refinements available, but not under non-homoskedasticity
- Can reject for all β_0 if the model is misspecified and overidentified

Example: Mertens and Montiel Olea (2018)

LP-IV estimates of elasticity of taxable income to $1 - AMTR$

AMTR is the average marginal tax rate

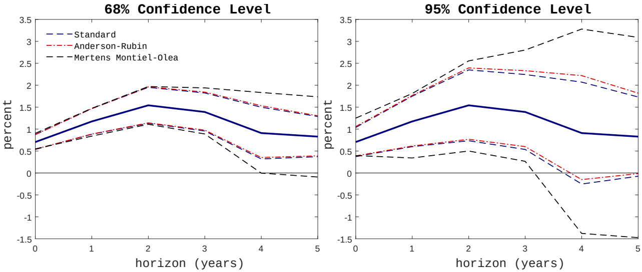


FIGURE B.II Confidence Intervals for LP-IV Estimates of Aggregate Tax Elasticities of Income

$N = K = 1$, $g_{\min}$ = 'Effective' F-statistic = 229.25

Well above critical value for $\alpha = 0.05$ and $\tau = .1$

Download the code here

Proxy SVAR/SVAR-IV Inference

Impulse estimation by LP-IV is an application of the general IV model discussed earlier

Very similar weak instrument problems also apply to impulse response estimation by Proxy SVARs/SVAR-IV

SVAR Inference with strong instruments: See overview in Kilian and Lütkepohl (2017)

- Delta Method
- Bootstrap Methods*
  - Parametric Bootstrap Montiel Olea, Stock, and Watson (2021)
  - Moving Block Bootstrap Jentsch and Lunsford (2019)

SVAR Inference with weak instruments: Montiel Olea, Stock, and Watson (2021)

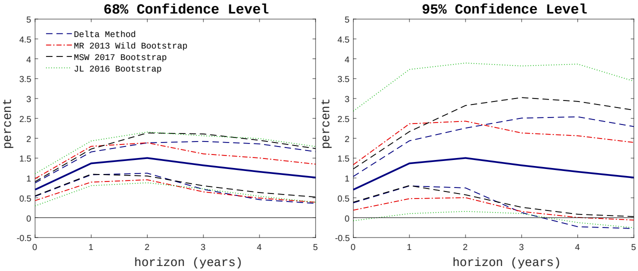
* Avoid Mertens and Ravn (2013) wild bootstrap for impulse responses in SVAR applications (regardless of identification scheme), see Jentsch and Lunsford (2019)

Example: Mertens and Montiel Olea (2018)

LP-IV estimates of elasticity of taxable income to $1 - AMTR$

AMTR is the average marginal tax rate

Inference methods with strong instruments



Download the code here

Proxy SVAR/SVAR-IV Recap

SVAR $B(L)z_t = \mathcal{D}\epsilon_t$ with

$$\mathcal{D} = \begin{bmatrix} \underset{1\times 1}{d_{11}} & \underset{1\times(N_z-1)}{d_{12}} \\ \underset{(N_z-1)\times 1}{d_{21}} & \underset{(N_z-1)\times(N_z-1)}{d_{22}} \end{bmatrix}, \; u_t = \begin{bmatrix} \underset{1\times 1}{u_{1,t}} \\ \underset{(N_z-1)\times 1}{u_{2,t}} \end{bmatrix}$$

Under fundamentalness, the following conditions on a scalar m_t

$E[m_t\epsilon_{1,t}] = \phi \neq 0$ (Relevance)

$E[m_t\epsilon_{-1,t}] = 0$ (Contemporaneous Exogeneity)

provide $N_z - 1$ covariance restrictions that suffice to identify $\mathcal{D}_1 = [d_{11} \; d_{21}']'$

The impulse response coefficients at horizon h are in $G_h\mathcal{D}_1 = M_h^1$ where G_h are the coefficients on L^h in $G(L) = B(L)^{-1}$, or

$$G_0 = \mathcal{I} \; ; \; G_h = \sum_{i=1}^{h} G_{h-i}B_i \; , \; h > 0$$

Weak Instruments in Proxy SVAR/SVAR-IV

Define $D_1 = \mathcal{D}_1 / d_{11} = [1 \ d'_{21}/d_{11}]'$

D_1 contains the impact coefficients scaled such that the impact on the first variable is normalized to unity

$$D_1 = \frac{E[m_t u_t]}{E[m_t u_{1,t}]}$$

The associated impulse response of variable n at horizon h is

$$\lambda_{n,h} = e'_n G_h D_1 = e'_n G_h \frac{E[m_t u_t]}{E[m_t u_{1,t}]}$$

where e_n is the n-th column of the identity matrix

The natural estimator is

$$\hat{\lambda}_{n,h} = e'_n \hat{G}_h \hat{D}_1 = e'_n \hat{G}_h \frac{um'/T}{u_1 m'/T} \quad , \quad \hat{G}_h = \sum_{i=1}^{h} \hat{G}_{h-i} \hat{B}_i \ , \ h > 0$$

where u is $N_z \times T$, u and m are $1 \times T$, and $\hat{B}_i$ are the OLS estimates of the VAR coefficients

Weak Instruments in Proxy SVAR/SVAR-IV

Local-to-Zero

$E[m_t \epsilon_{1,t}] = \phi$ with $\phi = c/\sqrt{T}$

Asymptotic Normality Regularity Condition

$$\sqrt{T}\left[\begin{array}{c} \text{vec}(\hat{B}) \\ um'/T \\ \text{vech}(uu'/T) \end{array}\right] \xrightarrow{d} \left[\begin{array}{c} \zeta \\ \xi \\ \varphi \end{array}\right] \sim \mathcal{N}(0, \Omega)$$

LP-IV (i.e. IV) assumes the IV reduced form parameters are (approximately) normal

Here, the assumption is instead that that VAR coefficients are approximately normal

Weak IV Asymptotic Representation of the Impulse Response Estimator

Montiel Olea, Stock, and Watson (2021)

Under **Local-to-Zero** and the **Regularity Condition**,

$$(\hat{\lambda}_{n,h} - \lambda_{n,h}) \to \lambda^*_{n,h} \xrightarrow{d} \frac{(e_n' G_h - \lambda_{n,h} e_1')' \xi}{e_1' \xi + c}$$

The impulse response bias is asymptotically distributed as a ratio of correlated normal variables.

The bias is decreasing in the concentration parameter c

Weak Instruments in Proxy SVAR/SVAR-IV

All of this is analogous to the standard IV framework with $N = K = 1$ under homoskedasticity (e.g. Staiger and Stock (1997))

Consider

$$z_{2,t} = \beta z_{1,t} + \sum_{i=1}^{\infty} B_i z_{t-i} + w_{2,t}$$

The SVAR IV estimator of $\beta = d_{21}/d_{11}$ is the equation-by-equation 2SLS estimator of β using m_t as the instrument

Weak instrument bias towards to OLS estimate of β, which is also the estimate from the lower triangular (Cholesky) factorization with $u_{1,t}$ ordered first

If Cholesky and SVAR-IV impulse response look similar, that is not evidence for the recursivity restrictions unless the instrument is strong.

Existing tests are also suited for detecting weak instruments in the equation above

Robust Inference in Proxy SVAR/SVAR-IV

Montiel Olea, Stock, and Watson (2021) propose Anderson and Rubin (1949) confidence sets

Under the null hypothesis $\lambda_{n,h} = \lambda^0_{n,h}$: $E[m_t u_{1,t}] - \lambda^0_{n,h} e_n' G_h E[m_t u_t] = 0$

Define $\hat{e}(\lambda^0_{n,h}) = \frac{u_1 m'}{T} - \lambda^0_{n,h} e_n' \hat{G}_h \frac{um'}{T}$ and construct the Wald-statistic

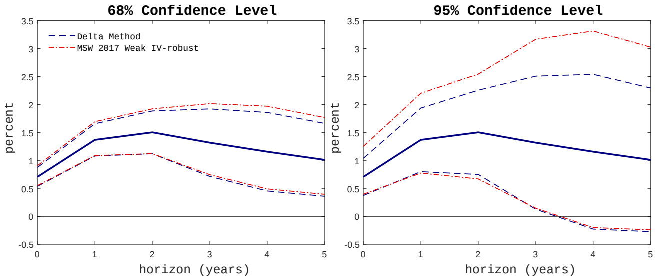
$$AR(\lambda^0_{n,h}) = \frac{\hat{e}(\lambda^0_{n,h})^2}{\hat{\omega}_{11} - 2\lambda^0_{n,h}\hat{\omega}_{12} + (\lambda^0_{n,h})^2\hat{\omega}_{22}}$$

where $\begin{bmatrix} \hat{\omega}_{11} & \hat{\omega}_{12} \\ \hat{\omega}_{12} & \hat{\omega}_{22} \end{bmatrix}$ is consistent for the asymptotic covariance of $[u_1 m'/T \;\; e_n' \hat{G}_h (um'/T)]'$.

Example: Mertens and Montiel Olea (2018)

SVAR-IV estimates of elasticity of taxable income to $1 - AMTR$

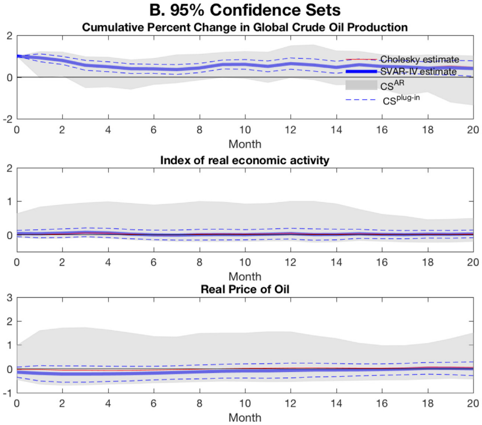
AMTR is the average marginal tax rate



Download the code here

Example: Montiel Olea, Stock, and Watson (2021)

Kilian (2008) proxy for oil shocks in a 3-variable SVAR



Oil supply have shocks have relatively little effect on oil prices under Cholesky identification Kilian (2009).

Conclusion appears robust to using Proxy SVAR and robust inference

Download the code here

Anderson, T. W. and Herman Rubin (1949). "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations". In: *The Annals of Mathematical Statistics* 20.1, pp. 46 –63.

Andrews, Isaiah, James H. Stock, and Liyang Sun (2019). "Weak Instruments in Instrumental Variables Regression: Theory and Practice". In: *Annual Review of Economics* 11.1, pp. 727–753. DOI: 10.1146/annurev-economics-080218-025643.

Chernozhukov, Victor and Christian Hansen (2008). "The reduced form: A simple approach to inference with weak instruments". In: *Economics Letters* 100.1, pp. 68–71. ISSN: 0165-1765. DOI: 10.1016/j.econlet.2007.11.012.

Cragg, John G. and Stephen G. Donald (1993). "Testing Identifiability and Specification in Instrumental Variable Models". In: *Econometric Theory* 9.2, pp. 222–240.

Imhof, J. P. (Dec. 1961). "Computing the distribution of quadratic forms in normal variables". In: *Biometrika* 48.3-4, pp. 419–426.

Jentsch, Carsten and Kurt G. Lunsford (2019). "The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States: Comment". In: *American Economic Review* 109.7, pp. 2655–78. DOI: 10.1257/aer.20162011.

Keane, Michael and Timothy Neal (2022). *A Practical Guide to Weak Instruments*. Discussion Papers 2021-05d. School of Economics, The University of New South Wales.

Kilian, Lutz (May 2008). "Exogenous Oil Supply Shocks: How Big Are They and How Much Do They Matter for the U.S. Economy?" In: *The Review of Economics and Statistics* 90.2, pp. 216–240. DOI: 10.1162/rest.90.2.216.

— (2009). "Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market". In: *American Economic Review* 99.3, pp. 1053–69. DOI: 10.1257/aer.99.3.1053.

Kilian, Lutz and Helmut Lütkepohl (2017). *Structural Vector Autoregressive Analysis*. Themes in Modern Econometrics. Cambridge University Press. DOI: 10.1017/9781108164818.

Kleibergen, Frank (2007). "Generalizing weak instrument robust IV statistics towards multiple parameters, unrestricted covariance matrices and identification statistics". In: *Journal of Econometrics* 139.1. Endogeneity, instruments and identification, pp. 181–216. ISSN: 0304-4076. DOI: https://doi.org/10.1016/j.jeconom.2006.06.010.

Kleibergen, Frank and Richard Paap (2006). "Generalized reduced rank tests using the singular value decomposition". In: *Journal of Econometrics* 133.1, pp. 97–126.

Lewis, Daniel and Karel Mertens (2022). *A Robust Test for Weak Instruments with Multiple Endogenous Variables*. Working Paper 2208. Federal Reserve Bank of Dallas. DOI: 10.24149/wp2208.

Mertens, Karel and José Luis Montiel Olea (2018). "Marginal Tax Rates and Income: New Time Series Evidence". In: *The Quarterly Journal of Economics* 133.4, pp. 1803–1884. DOI: 10.1093/qje/qjy008.

Mertens, Karel and Morten O. Ravn (2013). "The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States". In: *American Economic Review* 103.4, pp. 1212–47. DOI: 10.1257/aer.103.4.1212.

Montiel Olea, José Luis and Carolin Pflueger (2013). "A Robust Test for Weak Instruments". In: *Journal of Business & Economic Statistics* 31.3, pp. 358–369. DOI: 10.1080/00401706.2013.806694.

Montiel Olea, José L., James H. Stock, and Mark W. Watson (2021). "Inference in Structural Vector Autoregressions identified with an external instrument". In: *Journal of Econometrics* 225.1, pp. 74–87. DOI: 10.1016/j.jeconom.2020.05.014.

Nagar, A. L. (1959). "The Bias and Moment Matrix of the General k-Class Estimators of the Parameters in Simultaneous Equations". In: *Econometrica* 27.4, pp. 575–595.

Ramey, Valerie A. and Sarah Zubairy (2018). "Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data". In: *Journal of Political Economy* 126.2, pp. 850–901. DOI: doi.org/10.1086/696277.

Staiger, Douglas and James H. Stock (1997). "Instrumental Variables Regression with Weak Instruments". In: *Econometrica* 65.3, pp. 557–586.

Stock, James and Motohiro Yogo (2005). "Testing for Weak Instruments in Linear IV Regression". In: *Identification and Inference for Econometric Models*. Ed. by Donald W.K. Andrews. New York: Cambridge University Press, pp. 80–108.