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

1 **Time Series Representations of Dynamic Macro Models**
   Structural State Space Models, MA, VARMA and VAR representations; Estimating Dynamic Causal Effects; Misspecification: Nonfundamentalness, Nonlinearities, and Time Aggregation

2 **State-Space Models and the Kalman Filter**
   State Space Models, Kalman Filter, Forecasting, Maximum Likelihood Estimation

3 **Local Projections**
   Impulse Responses as Dynamic Treatment Effects, LP Estimation and Basic Inference, VAR-LP Impulse Response Equivalence

4 **Identification of Dynamic Causal Effects**
   Identification with Covariance Restrictions or Higher Moments. Proxy SVAR/SVAR-IV, Internal instrument SVAR

5 **Inference for Impulse Responses**
   Inference methods for VAR/LP impulse responses. Detecting weak instruments; Robust Inference Methods; Joint inference for VAR and LP impulse responses

6 **Impulse Response Heterogeneity**
   Kitagawa Decomposition, Time Varying Impulse Responses

7 **Other Uses of Impulse Responses**
   Impulse Response Matching and Indirect Inference; Estimating Structural Single Equations using Impulse Responses, SP-IV; Counterfactuals with Impulse Responses, Optimal Policy Perturbations
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

\[ y = Y\beta + \nu, \]
\[ Y = X\Pi + \nu, \]

$y$: $T \times 1$ outcome variable of interest

$Y$: $T \times N$ endogenous regressors

$X$: $T \times K$ instrumental variables, $K \geq N$

$\beta$: $N \times 1$ parameters of interest

$\Pi$: $K \times N$ ‘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: $y = X\Gamma + w$, $w = v\beta + \nu$
Two-Stage-Least-Squares:

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

where \( P_X = XX'/T \)

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

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

Partition \( W = \begin{bmatrix} W_1 & W_{12} \\ W'_1 & W_2 \end{bmatrix} \) \((N + 1)K \times (N + 1)K\)

So \( \hat{W}_2 \) 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**

$Y = X\Pi + \nu$ 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**

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

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

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

$S$ is the covariance of $T^{-\frac{1}{2}} [X'\nu \ \text{vec}(X'\nu)']'$ as $T \to \infty$, and depends on $W$ and $\beta$

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

<table>
<thead>
<tr>
<th>$N$ = 2, $K$ = 2</th>
<th>$N$ = 2, $K$ = 3</th>
<th>$N$ = 2, $K$ = 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$ = 2, $K$ = 6</td>
<td>$N$ = 3, $K$ = 5</td>
<td>$N$ = 3, $K$ = 9</td>
</tr>
</tbody>
</table>

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 \( \chi^2 \)
  For \( N > 1 \): unknown, use bounding non-central \( \chi^2 \)

- **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’
  
  $= \text{equals robust F-statistic when } K = 1$

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

  Unknown, weighted avg. of non-central $\chi^2$'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 $\chi^2$'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**

\[
B = \sqrt{E \left[ \beta_{2SLS}^* \right]' \Phi E \left[ \beta_{2SLS}^* \right] / \sqrt{\text{Tr}(S_1)}
\]

\(S_1\) is the covariance of \(T^{-\frac{1}{2}} [X'\nu]'\) as \(T \to \infty\)

\(B = 1\) in a worst-case scenario when the instruments are completely uninformative \((\Lambda = 0)\) and \(\nu\) 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 \(\beta\) and \(C\) such that \(B \geq \tau\), where \(\tau\) is a tolerance level
Nagar Bias

$\beta^*_{2SLS}$ has intractable distribution

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

$E[\beta^*_{2SLS}]$ 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 $\mathbf{W}$, 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

For each specification, we consider five million DGPs as described in the main text. For each DGP we take 1000 samples, and compute the Monte Carlo bias, $\hat{B}$, by numerical integration. The figure plots the (log of) Monte Carlo bias against the Nagar bias. The heatmap indicates the density of DGPs with a particular combination of Nagar and Monte Carlo biases. The dashed horizontal and vertical lines indicate bias levels of 0.05 and 0.15 to demarcate the typically relevant region for first-stage tests.

A.2 The Importance of Using the More Conservative Bound on the Bias in Models with $K \leq N + 1$.

Figures A.2 and A.3 illustrate the need to use the more conservative upper bound in the first stage test in models with degrees of overidentification less than two, $K \leq N + 1$. All panels in both figures plot rejection frequencies against the Monte Carlo bias in blue. Panel (a) in figure A.2 repeats the first two panels in Figure 2 for ease of comparison, and shows rejection frequencies based on the more conservative bound in red. The rejection frequencies shown in red in Panel (b) are instead based on the worst-case Nagar bias, i.e. as in the models with $K > N + 1$. The results clearly indicate that the worst-case Nagar bias is not a good criterion for a bias-based first-stage test in the models with $N = 2, K = 2$ and $N = 2, K = 3$. In both cases, there are a significant number of DGPs for which the Monte Carlo bias exceeds the bias tolerance but the rejection rates lies above the nominal level of 0.05. These positive size distortions are much more frequent in the just identified case with $N = K = 2$ than in the model with $N = 2, K = 3$. 

44
Bounding the Bias

Nagar bias is a complicated function of $\beta$ ($N \times 1$), $C$ ($K \times N$) and $W$

We can estimate $W$ consistently, but not $\beta$ and $C$

**Nagar Bias Bounds**

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

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

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)
We need a test statistic that is informative about $\lambda_{\text{min}}$.

**Generalized First-Stage Statistic**  
Lewis and Mertens (2022)

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

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

measures instrument strength, $\hat{\Phi} = (I_N \otimes \text{vec}(I_K))' (\hat{W}_2 \otimes I_K)(I_N \otimes \text{vec}(I_K))$

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

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 $g_{\text{min}} \leq 1 + B(\hat{W})/\tau$ (or $g_{\text{min}} \leq 1 + \Psi(\hat{W})/\tau$ for $K \leq N + 1$)
Measuring Instrument Strength

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

Special cases:

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

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

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

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

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

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

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

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

So \( g_{\text{min}} \) always gets it right (do not use Kleibergen and Paap (2006))
Under the null, $g_{\min}$ is the minimum eigenvalue of a matrix consisting of traces of the $K \times K$ partitions of $W_2$

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

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 $\alpha$ probability, e.g. $\alpha = 5\%$)

The generalized test nevertheless has power
Size and Power

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

$N = 2, K = 2$

$N = 2, K = 3$

$N = 2, K = 4$

$N = 2, K = 6$

$N = 3, K = 5$

$N = 3, K = 9$
Ramey and Zubairy (2018) estimate state-dependent government spending multipliers using

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

where \( l_t \) is an indicator of recession/boom or binding/non-binding ZLB

\( N = 2, \ K = 4 \)
(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)
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
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 $\alpha \%$ under the null that $\beta = \beta_0$.
Perform the test over a set of possible values for $\beta_0$, the acceptance region contains $\beta$ with $1 - \alpha \%$ probability.

Options under Non-Homoskedasticity:
- AR statistic, \cite{AndersonRubin1949,ChernozhukovHansen2008}
- KLM statistic, \cite{Kleibergen2007,ChernozhukovHansen2008}
- CLR statistic

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

$$AR(\beta_0) = \hat{\gamma}(\beta_0)' Var(\hat{\gamma}(\beta_0))^{-1}\hat{\gamma}(\beta_0)$$

$$AR(\beta_0)|_{H_0: \beta=\beta_0} \xrightarrow{d} \chi^2_K$$

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

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

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

![Diagram showing confidence intervals for LP-IV estimates of aggregate tax elasticities of income.](image)

**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](#).
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:

- Delta Method

- Bootstrap Methods*
  - Parametric Bootstrap
  - Moving Block Bootstrap

SVAR Inference with weak instruments:

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

\[ \text{SVAR } B(L)z_t = D \epsilon_t \text{ with} \]

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

Under fundamentalness, the following conditions on a scalar \( m_t \)

\[ E[m_t \epsilon_{1,t}] = \phi \neq 0 \text{ (Relevance)} \]
\[ E[m_t \epsilon_{-1,t}] = 0 \text{ (Contemporaneous Exogeneity)} \]

provide \( N_z - 1 \) covariance restrictions that suffice to identify \( D_1 = [d_{11} \ d_{21}]' \)

The impulse response coefficients at horizon \( h \) are in \( G_h D_1 = M_h^1 \) where \( G_h \) are the coefficients on \( L^h \) in \( G(L) = B(L)^{-1} \), or

\[ G_0 = I; \ G_h = \sum_{i=1}^{h} G_{h-i} B_i, \ h > 0 \]
Define $D_1 = D_1/d_{11} = [1 \ d_{21}'/d_{11}]'$

$D_1$ 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

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

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

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

Asymptotic Normality Regularity Condition

\[
\sqrt{T} \begin{bmatrix}
\text{vec}(\hat{B}) \\
(\hat{u}m')/T \\
\text{vech}(uu'/T)
\end{bmatrix} \xrightarrow{d} \begin{bmatrix}
\zeta \\
\xi \\
\varphi
\end{bmatrix} \sim \mathcal{N}(0, \Omega)
\]

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,

\[
(\hat{\lambda}_{n,h} - \lambda_{n,h}) \xrightarrow{d} \lambda_{n,h}^* \xrightarrow{} \lambda_{n,h} \xrightarrow{d} \frac{(e'_n G_h - \lambda_{n,h} e'_1)\xi}{e'_1 \xi + c}
\]

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 $\beta$ using $m_t$ as the instrument

Weak instrument bias towards OLS estimate of $\beta$, 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
Montiel Olea, Stock, and Watson (2021) propose Anderson and Rubin (1949) confidence sets

Under the null hypothesis \( \lambda_{n,h} = \lambda_{n,h}^0: E[m_t u_{1,t}] - \lambda_{n,h}^0 e'_n G_h E[m_t u_t] = 0 \)

Define \( \hat{e}(\lambda_{n,h}^0) = \frac{u_{1m}'}{T} - \lambda_{n,h}^0 e'_n \hat{G}_h \frac{u_{m}'}{T} \) and construct the Wald-statistic

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

where \( \left[ \begin{array}{cc} \hat{\omega}_{11} & \hat{\omega}_{12} \\ \hat{\omega}_{12} & \hat{\omega}_{22} \end{array} \right] \) is consistent for the asymptotic covariance of \( [u_{1m}/T \ e'_n \hat{G}_h(u_{m}/T)]' \).
SVAR-IV estimates of elasticity of taxable income to \(1 - AMTR\)

AMTR is the average marginal tax rate

Download the code here
Kilian (2008) proxy for oil shocks in a 3-variable SVAR

Oil supply 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


