Identification of Dynamic Causal Effects in Macroeconomics

James H. Stock, Harvard University

Joint work with Mark Watson, Princeton University

Sargan Lecture
Royal Economic Society
Bristol, U.K.

April 11, 2017

Figure on left (and idea of simultaneity bias) appeared in P.G. Wright (*QJE*, 1915)
Supply equation:
\[ O = eP + S \]

where:
- \( O \) = output
- \( P \) = price
- \( S \) = supply disturbance
- \( e \) = supply elasticity

Suppose this multiplication to be performed for every pair of price-output deviations and the results added, then:
\[ e\Sigma A.P = \Sigma A.O - \Sigma A.S_1 \quad \text{or} \quad e = \frac{\Sigma A.O - \Sigma A.S_1}{\Sigma A.P} \]

But \( A \) was a factor which did not affect supply conditions; hence it is uncorrelated with \( S_1 \); hence \( \Sigma A.S_1 = 0 \); and hence \( e = \frac{\Sigma A.O}{\Sigma A.P} \).

Similarly if \( B \) is a factor, say, yield per acre, which does not affect demand conditions we shall have:
\[ \eta = \frac{FH}{FG} = \frac{O-D_1}{P}; \quad \eta P = O-D_1; \quad \eta \Sigma B.P = \Sigma B.O - \Sigma B.D_1; \]
\[ \eta = \frac{\Sigma B.O - \Sigma B.D_1}{B.P} \]

But \( \Sigma B.D_1 = 0 \) Hence \( \eta = \frac{\Sigma B.O}{\Sigma B.P} \)

Success with this method depends on success in discovering factors of the type \( A \) and \( B \). Several such factors of each type should be used if possible. Because of the slow adjustment of price to marginal cost five-year (or four-year or six-year) averages should be used for \( P', O' \),
Philip Wright (1861-1934)
Economist, teacher, poet
MA Harvard, Econ, 1887
Lecturer, Harvard, 1913-1917

Sewall Wright (1889-1988)
genetic statistician
ScD Harvard, Biology, 1915
Prof., U. Chicago, 1930-1954
The Wrights’ letters, December 1925 - March 1926

March 4, 1926.

Dear Bureau:

It may interest you to see a very simple geometric demonstration which I have worked out for your use without reference to the theory of path coefficients.

...
Supply equation:

\[ O = eP + S \]

where:

- \( O \) = output
- \( P \) = price
- \( S \) = supply disturbance
- \( e \) = supply elasticity
March 15, 1926

Dear Devere:

... 

<table>
<thead>
<tr>
<th>Year</th>
<th>Real prices</th>
<th>Output</th>
<th>Average</th>
<th>Rainfall</th>
<th>Ratio value</th>
<th>Rain fall again</th>
<th>Rain fall again</th>
</tr>
</thead>
<tbody>
<tr>
<td>1903</td>
<td>126</td>
<td>27.3</td>
<td>6.1</td>
<td>8.4</td>
<td>3.40</td>
<td>9.3</td>
<td>12.8</td>
</tr>
<tr>
<td>4</td>
<td>153</td>
<td>23.1</td>
<td>2.26</td>
<td>10.3</td>
<td>2.19</td>
<td>7.5</td>
<td>14.0</td>
</tr>
<tr>
<td>5</td>
<td>123</td>
<td>28.5</td>
<td>2.53</td>
<td>11.2</td>
<td>4.17</td>
<td>9.5</td>
<td>18.6</td>
</tr>
<tr>
<td>6</td>
<td>146</td>
<td>25.6</td>
<td>2.51</td>
<td>12.2</td>
<td>3.30</td>
<td>9.3</td>
<td>18</td>
</tr>
<tr>
<td>7</td>
<td>133</td>
<td>25.9</td>
<td>2.86</td>
<td>9.0</td>
<td>2.66</td>
<td>1.9</td>
<td>18.7</td>
</tr>
<tr>
<td>8</td>
<td>157</td>
<td>25.8</td>
<td>2.67</td>
<td>7.6</td>
<td>3.30</td>
<td>7.6</td>
<td>17.5</td>
</tr>
<tr>
<td>9</td>
<td>204</td>
<td>19.7</td>
<td>2.08</td>
<td>9.5</td>
<td>3.10</td>
<td>9.5</td>
<td>21.3</td>
</tr>
</tbody>
</table>

... 

"Aver. for crop year beginning Sep. 1. The Minneapolis price was divided by wholesale price of all commodities to get "Real price."

Figure are for calendar year.

Figures are a simple average for rainfall (May, June, and July) for Duluth, Minn., Bismarck, N.D., Pierre, S.D.

The ratios of the values of planted area to spring wheat per acre suggest 1 year is

The ratios for the year shown in the table are nearly the ratios for the preceding year.
Modern (nonstructural) micro approach

Find a plausibly exogenous source of variation to identify the effect of interest (experiment, natural experiment):

\[ Y_{2i} = \theta Y_{1i} + \gamma W_i + u_i \]

Instrument \( z \):

(i) Relevance: \( \text{cov}(Y_{1}^\perp, z^\perp) \neq 0 \), where \( Y_{1}^\perp = Y_1 - \text{Proj}(Y_1 | W) \)

(ii) Exogeneity: \( E(u|W, z) = E(u| W) \)
Modern (nonstructural) micro approach

Find a plausibly exogenous source of variation to identify the effect of interest (experiment, natural experiment):

\[ Y_{2i} = \theta Y_{1i} + \gamma W_i + u_i \]

Instrument \( z \):

(i) Relevance: \( \text{cov}(Y_{1i}^\perp, z^\perp) \neq 0 \), where \( Y_{1i}^\perp = Y_{1i} - \text{Proj}(Y_{1i} | W) \)

(ii) Exogeneity: \( E(u|W, z) = E(u| W) \)

Modern (nonstructural) macro approach

Obtain impulse response function from a structural vector autoregression (SVAR).

\[
A(L)Y_t = \nu_t, \quad \nu_t | \nu_{t-1}, \nu_{t-2}, ... \sim (0, \Sigma_{\nu})
\]

\[
\nu_t = H \varepsilon_t, \quad \varepsilon_t \text{ structural shocks}
\]

\[
y_t = A(L)^{-1}H \varepsilon_t \quad \text{(IRFs from SVAR)}
\]
Modern (nonstructural) micro approach
Find a plausibly exogenous source of variation to identify the effect of interest (experiment, natural experiment):
\[ Y_{2i} = \theta Y_{1i} + \gamma W_i + u_i \]
Instrument \( z \):
(i) Relevance: \( \text{cov}(Y_1^\perp, z^\perp) \neq 0 \), where \( Y_1^\perp = Y_1 - \text{Proj}(Y_1 | W) \)
(ii) Exogeneity: \( E(u|W, z) = E(u|W) \)

Modern (nonstructural) macro approach
Obtain impulse response function from a structural vector autoregression (SVAR).
\[ A(L)Y_t = \nu_t, \quad \nu_t | \nu_{t-1}, \nu_{t-2}, \ldots \sim (0, \Sigma_{\nu}) \]
\[ \nu_t = H\varepsilon_t, \quad \varepsilon_t \text{ structural shocks} \]
\[ y_t = A(L)^{-1}H\varepsilon_t \quad \text{(IRFs from SVAR)} \]

This lecture
- Pull together IV approach to macro shocks
  - Conditions on \( z \) for identification of \( H \)
  - Conditions on \( z \) for identification of dynamic causal effects without a SVAR
- Follow-on: tests of SVAR validity, IV odds & ends, time series odds & ends
- [Are there reasons to prefer local projections over SVARs?]
Setup

**Structural MA:** \( Y_t = \Theta(L)\varepsilon_t \)

Structural shock: Define \( \varepsilon_{1t} = \) autonomous, unexpected change in \( Y_{1t} \)

All disturbances: \( \varepsilon_t = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{*t} \end{pmatrix} \), \( \varepsilon_t | \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots \sim (0, \Sigma_\varepsilon) \) (\( \bullet \) = “everything else”)

The structural IRF is the dynamic causal effect of an autonomous change in \( Y_{1t} \) on \( Y_{2t+h} \):
\[
\Theta_{h,21} = E \left( Y_{2t+h} | \varepsilon_{1t} = 1, \varepsilon_{*t}, \varepsilon_s, s \neq t \right) - E \left( Y_{2t+h} | \varepsilon_{1t} = 0, \varepsilon_{*t}, \varepsilon_s, s \neq t \right)
\]

**SVAR MA**

Wold representation: \( Y_t = C(L)\nu_t \), where \( \nu_t = Y_t - Y_t_{t-1}, \nu_t | \nu_{t-1}, \nu_{t-2}, \ldots \sim (0, \Sigma_\nu) \)

MA implied by SVAR: \( Y_t = C(L)H\varepsilon_t \)

**SVAR MA = structural MA if:** \( C(L)H = \Theta(L) \Leftrightarrow H = C(L)^{-1}\Theta(L) \)
Interpreting the condition $H = C(L)^{-1} \Theta(L)$

\[ H = C(L)^{-1} \Theta(L) = (I + C_1L + \ldots)( \Theta_0 + \Theta_1L + \ldots) = \Theta_0 + \text{terms in } L, L^2, \ldots \]

(1) **Impact effect:** $H = \Theta_0$. *Typically called the SVAR identification condition.*

- Timing restrictions (Cholesky, etc.), long-run restrictions
- Heteroskedasticity
- Sign restrictions
- Direct measurement of shock of interest
- Method of external instruments
Interpreting the condition $H = C(L)^{-1}\Theta(L)$

$$H = C(L)^{-1}\Theta(L) = (I + A_1L + \ldots)(\Theta_0 + \Theta_1L + \ldots) = \Theta_0 + \text{terms in } L, L^2, \ldots$$

(1) **Impact effect:** $H = \Theta_0$. *Typically called the SVAR identification condition.*

- Timing restrictions (Cholesky, etc.), long-run restrictions
- Heteroskedasticity
- Sign restrictions
- Direct measurement of shock of interest
- Method of external instruments

(2) **No lagged terms.** $Y_t = C(L)v_t$ and $Y_t = \Theta(L)\varepsilon_t$, so $v_t = C(L)^{-1}\Theta(L)\varepsilon_t$

"No lags": $E(v_t \mid \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) = 0 \iff E(Y_t \mid Y_{t-1}, Y_{t-2}, \ldots, \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) = E(Y_t \mid Y_{t-1}, Y_{t-2}, \ldots)$

$$\iff \text{span}(v_t) = \text{span}(\varepsilon_t)$$

$$\iff \text{Structural MA is invertible so } \varepsilon_t = \Theta_0^{-1}v_t$$

- Interpretation: “no omitted variables”
- Called the “invertibility” or “nonfundamentalness” problem
- There are two main solutions to OVB:
  - Include OVs (large SBVARs, SDFMs, FAVARs, etc.); or
  - IV estimation
The method of external instruments in SVARs ("SVAR-IV")

• Under the invertibility assumption, \( v_t = \Theta_0 \varepsilon_t \). The challenge is identifying \( \Theta_0 \).
• Suppose you have an instrument satisfying:

**Condition A**

(i) \( E \varepsilon_{1t} z_t = \alpha \neq 0 \) (relevance)
(ii) \( E \varepsilon_{1t} z_t = 0 \) (exogeneity w.r.t. other current shocks)

Then
\[
E v_{1t} z_t = \Theta_0 E \varepsilon_{1t} z_t = \Theta_0 E \begin{pmatrix} \varepsilon_{1t} z_t \\ \varepsilon_{1t} z_t \end{pmatrix} = \Theta_0 \begin{pmatrix} \alpha \\ 0 \end{pmatrix} = \begin{pmatrix} \Theta_{0,11} \alpha \\ \Theta_{0,11} \alpha \end{pmatrix}
\]

(1)

Adopt the:

**Unit effect normalization:** \( \Theta_{0,11} = 1 \)

Then, from (1),
\[
\frac{E v_{1t} z_t}{E v_{1t} z_t} = \frac{\Theta_{0,21}}{\Theta_{0,11}} = \Theta_{0,21}
\]

\( \iff \) IV estimator of \( \Theta_{0,21} \) in : \( v_{2t} = \Theta_{0,21} v_{1t} + u_t \) with IV \( z_t \)

**Unit effect vs. unit standard deviation normalization:** \( \Theta_{0,11} = 1 \) or \( \text{var}(\varepsilon_{1t}) = 1? \)
The method of external instruments (SVAR-IV), ctd.

1. Estimate VAR: \[ A(L)Y_t = \nu_t \]

2. Estimate \( \Theta_{0,21} \) by IV: \[ \hat{\nu}_{2t} = \Theta_{0,21}\hat{\nu}_{1t} + u_t \text{ using IV } z_t \]

3. Estimate structural MA as \( \hat{C}(L) \begin{pmatrix} 1 \\ \hat{\Theta}_{0,1} \end{pmatrix} \), where \( \hat{C}(L) = \hat{A}(L)^{-1} \)

4. SEs by parametric bootstrap (or another method)

References

Example: Gertler-Karadi (2015)

\[ Y_t = (\Delta \ln I P_t, \Delta \ln C P I_t, 1Yr \text{ Treasury rate}_t, E BP_t) \]

\[ E BP_t = \text{Gilchrist-Zakrajšek (2012) Excess Bond Premium} \]

\[ z_t = "\text{Announcement surprise}" = \text{change in 4-week Fed Funds Futures around FOMC announcement windows} \]

Sample period: 1990m1-2012m6 (monthly)

SVAR-IV

GK specification: 12 lag VAR

LP-IV

\[ W_t = Y_{t-1}, \ldots, Y_{t-4}, z_{t-1}, \ldots, z_{t-4} \]
Gertler-Karadi example, ctd.

Cumulative IRFs: **SVAR-IV** with ±1 SE bands

![Log(IP)](image1)

![Log(CPI)](image2)

![1-Year Bond Rate](image3)

![Excess Bond Premium](image4)
Identification of structural MA without SVAR step

Structural MA: \( Y_t = \Theta(L)\varepsilon_t \)

Focus on variables 1 and 2:

\[
Y_{1t} = \Theta_{0,11}\varepsilon_{1t} + \{\varepsilon_{t}, \varepsilon_{t-j}\} \quad (2)
\]

\[
Y_{2t+h} = \Theta_{h,21}\varepsilon_{1t} + \{\varepsilon_{t}, \varepsilon_{t+j}, \varepsilon_{t-j}\} \quad (3)
\]

Notation:

\( \{\varepsilon_{t}, \varepsilon_{t-j}\} \) = linear combination of \( \varepsilon_{t} \) and lags of \( \varepsilon \)

\( \{\varepsilon_{t}, \varepsilon_{t+j}, \varepsilon_{t-j}\} \) = linear combination of \( \varepsilon_{t} \), lags of \( \varepsilon \), and leads of \( \varepsilon \)

Again use the:

**Unit effect normalization:** \( \Theta_{0,11} = 1 \)

Use (2) with the unit effect normalization to substitute \( \varepsilon_{1t} = Y_{1t} - \{\varepsilon_{t}, \varepsilon_{t-j}\} \) into (3):

\[
Y_{2t+h} = \Theta_{h,21}Y_{1t} + \{\varepsilon_{t}, \varepsilon_{t+j}, \varepsilon_{t-j}\} \quad (4)
\]

OLS estimation of (4) suffers from simultaneity and OVB bias.
Local Projections-IV

\[ Y_{2t+h} = \Theta_{h,21} Y_{1t} + u_{t+h}^{(h)}, \text{ where } u_{t+h}^{(h)} = \{ \varepsilon_{t}, \varepsilon_{t+j}, \varepsilon_{t-j} \} \]  \hspace{1cm} (3)

Suppose the IV z satisfies:

**Condition B**

(i) \( E\varepsilon_{1t} z_{t} = \alpha \neq 0 \) \hspace{1cm} (relevance)

(ii) \( E\varepsilon_{i t} z_{t} = 0 \) \hspace{1cm} (exogeneity, other current shocks)

(iii) \( E\varepsilon_{t+j} z_{t} = 0, j \geq 1 \) \hspace{1cm} (\text{ shocks are mds wrt past } z, \varepsilon)

(iv) \( E\varepsilon_{t-j} z_{t} = 0, j \geq 1 \) \hspace{1cm} (\text{ } z_{t} \text{ is mds wrt past shocks})

Conditions (ii) – (iv) imply that \( Eu_{t+h} z_{t} = 0 \), so with condition (i),

\[ E( Y_{2t+h} z_{t} ) = \Theta_{h,11} E( Y_{1t} z_{t} ) \Rightarrow \Theta_{h,11} = \frac{E( Y_{2t+h} z_{t} )}{E( Y_{1t} z_{t} )} \]

- \( \Theta_{h,11} \) can be estimated by IV regression of \( Y_{2t+h} \) on \( Y_{1t} \) using \( z_{t} \) as an instrument

- Including control variables might reduce SEs, but isn’t necessary for identification under condition A.
LP-IV with control variables $W$ to relax condition (iv)

$$Y_{2t+h} = \Theta_{h,21} Y_{1t} + \gamma W_t + u_{t+h}^{(h)}$$

(5)

where $W_t$ contains past variables (some past $Y$’s, $z$’s). Suppose $z$ satisfies:

**Condition C**

(i) Relevance: $\text{cov}(Y_{2t}^\perp, z_t^\perp) = \alpha \neq 0$, where $Y_{2t}^\perp = Y_{2t} - \text{Proj}(Y_{2t} | W)$

(ii) Exogeneity: $E(u_{t+h} | W_t, z_t) = E(u_{t+h} | W_t)$

Then $\Theta_{h,21}$ can be estimated in (5) by IV using instrument $z$ and control variables $W$.

References:

Local projections (LP)
Jordà (2005) for LP terminology

Local projections-IV (LP-IV)
Owyang, Ramey, and Zubairy (2013), Mertens (2016), Barnichon and Brownless (2017),
Jordà, Schularick, and Taylor (2015), Ramey (2016), Ramey-Zubairy (forthcoming);
System estimation of structural MA without SVAR step
Plagborg-Møller (2016)
Gertler-Karadi example, ctd.

Cumulative IRFs: **LP-IV with ±1 SE bands**

$W = 4$ lags of $Y, z$
SVAR-IV and LP-IV: Open questions and reminders

LP-IV: \[ Y_{2t+h} = \Theta_{h,21} Y_{1t} + \gamma W_t + u_{t+h} \] using IV \( z_t \)

Condition C

(i) Relevance: \( \text{cov}(Y_{2t}^\perp, z_t^\perp) \neq 0 \)

(ii) Exogeneity: \( E(u_{t+h} \mid W_t, z_t) = E(u_{t+h} \mid W_t) \)

Open questions

1. If condition B(iv) fails, what are suitable control variables?
2. Is LP-IV IV robust to non-invertibility?
3. Can LP-IV and SVAR-IV be used to test for invertibility?
4. HAR inference – anything noteworthy?
5. LP-IV specification: Levels or first differences?
6. What if the instrument is weak?
7. How to handle news shocks?

Reminders

1. IV estimation of distributed lag, AR-distributed lag specifications yields correct impact effect but incorrect dynamics in general
2. SVAR-IV is more efficient than LP-IV, if correctly specified
3. Potentially can improve LP-IV efficiency by imposing smoothness
Q1. If condition A(iv) fails, what are suitable control variables?

\[ Y_{2t+h} = \Theta_{h,21} Y_{1t} + \gamma W_t + u_{t+h}^{(h)} \]

**Condition C**

(i) Relevance: \( \text{cov}(Y_{2t}^\perp, z_t^\perp) = \alpha \neq 0 \)

(ii) Exogeneity: \( E(u_{t+h} | W_t, z_t) = E(u_{t+h} | W_t) \)

A sufficient condition for C(ii) is that Conditions B(ii) and B(iii) hold and that \( W_t \) spans \( \{\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots\} \). Then

\[
E\left(u_{t+h}^{(h)} | W_t, z_t\right) = E\left(\{\varepsilon_{t+h}, \ldots, \varepsilon_{t+1}, \varepsilon_t, \varepsilon_{t-1}, \ldots\} | \varepsilon_{t-1}, ..., z_t\right) \\
= E\left(\{\varepsilon_{t+h}, \ldots, \varepsilon_{t+1}\} | \varepsilon_{t-1}, ..., z_t\right) + E\left(\{\varepsilon_t\} | \varepsilon_{t-1}, ..., z_t\right) + E\left(\{\varepsilon_{t-1}, \ldots\} | \varepsilon_{t-1}, ..., z_t\right) \\
= E\left(\{\varepsilon_{t-1}, \ldots\} | \varepsilon_{t-1}, \ldots\right) = E\left(u_{t+h}^{(h)} | W_t\right)
\]

**Remarks**

1. This (perhaps) suggests using generic instruments – e.g. factors from a DFM
   - But assuming condition C(ii) is satisfied using \( W_t = Y_{t-1}, \ldots \) is equivalent to assuming \( \text{span}(\varepsilon_t) = \text{span}(\nu_t) \) – that is, the SVAR is invertible.
     - If invertibility fails, then LP-IV using \( W_t = Y_{t-1}, \ldots \) will be inconsistent.
     - And if you *can* span \( \varepsilon_t \), you might as well use SVAR-IV!
Q1. If condition A(iv) fails, what are suitable control variables? (ctd)

Remarks, ctd.
2. In some cases, it should be possible to construct valid control variables using application-specific knowledge.
   • Announcement-day monetary shocks
   • Political disruptions (wars) as oil supply shocks
   • Legislation on fiscal policy

Toy example (shock that drags out over two periods)

Observe $z_t = \zeta_t + b\zeta_{t-1}$, where $\zeta_t$ satisfies condition $B$.

Then $z_t$ violates condition B(iv):

$$E\left(\varepsilon_{1t-1}z_t\right) = E\left[\varepsilon_{1t-1}(\zeta_t + b\zeta_{t-1})\right] = bE\left(\varepsilon_{1t-1}\zeta_{t-1}\right) = b\alpha$$

But if $(1 + bL)$ is invertible, then Condition C(ii) holds with $W_t = (1+bL)^{-1}z_{t-1}$

Implications:
1. Looking for generic instruments only leads you back to SVAR-IV
2. The instrument mds condition – or something close – is critical to valid inference
Q2. Is LP-IV robust to non-invertibility?

Yes, under Conditions B or C.

Under Condition A:

\[ \hat{\Theta}_{SVAR-IV}^h = \hat{C}_h \frac{\sum y_t^\perp z_t^\perp}{\sum \hat{v}_t^\perp z_t^\perp} \overset{p}{\rightarrow} C_h \Theta_{0,1}, \text{ where } C(L) = A(L)^{-1} \]

whereas under Condition B or C,

\[ \hat{\Theta}_{LP-IV}^h \overset{p}{\rightarrow} \Theta_{h,1} = \frac{\sum Y_{t+h}^\perp z_t^\perp}{\sum Y_{1t}^\perp z_t^\perp} \]

In general \( C_h \Theta_{0,1} \neq \Theta_{h,1} \) if \( \Theta(L) \) is not invertible.
Q3. Can LP-IV and SVAR-IV be used to test for invertibility?

Yes, under condition B or C.

Consider a **near-invertible local alternative**:  
\[ C(L)^{-1} \Theta(L) = \Theta_0 + T^{-1/2} \delta(L)L \]

so

\[ \nu_t = \Theta_0 \varepsilon_t + T^{-1/2} \delta(L) \varepsilon_{t-1} \text{ and } \Theta_h = C_h \Theta_0 + d_h / \sqrt{T} . \]

Then

\[
\Psi_T = \sqrt{T} \left( \hat{\Theta}_{h,1}^{SVAR-IV} - \hat{\Theta}_{h,1}^{LP-IV} \right) = \frac{1}{\sqrt{T}} \sum \left[ Y_{t+h}^\perp - \hat{C}_h Y_t^\perp \right] z_t^\perp \nrightarrow \frac{1}{T} \sum \nu_t z_t^\perp 
\]

* Test for mis-specification of VAR, in the spirit of a Hausman test
* This test based on \( \Psi_T \) differs from other invertibility tests in the literature, which test predictability of VAR forecast errors.
* This tests both predictability and multistep v. direct forecast coefficients, and does not require an invertible SVAR to exists (just a structural MA).
Gertler-Karadi example, ctd.

Cumulative IRFs: SVAR-IV and LP-IV and ±1 SE bands (parametric bootstrap)
Gertler-Karadi example, ctd.

Test statistics by horizon by variable: entries are $t$-statistics $\Psi_T / \sqrt{\hat{V}_h}$
Q4. HAR inference – anything noteworthy?

- HAR inference is needed for LP-OLS (standard LP method) – standard direct multiperiod ahead regression problem.

- But HAR isn’t needed for SVAR-IV under condition B (mds property of $z_t$)
Q5. LP-IV specification: Levels or first differences?

Consider estimation of cumulative causal effect:

In levels: \[ Y_{2t+h} = \Theta_{h,21} Y_{1t} + \gamma W_t + u_{t+h}^{(h)} \]

In first differences: \[ \Delta Y_{2t+h} + \ldots + \Delta Y_{2t+1} = \Theta_{h,21} Y_{1t} + \gamma W_t + u_{t+h}^{(h)} \]

Suppose \( Y_{1t} \) and \( Y_{2t} \) are persistent (e.g. local to unit root) and Condition B holds:

- If there is no \( W_t \), then for both the levels and first differences specifications:
  - Nonstandard distributions at all horizons
  - Not resolved by including linear time trend

- If \( W_t \) includes \( Y_{i-1}, \ldots \):
  - Levels and cumulated differences specifications of \( Y_{2t} \) are equivalent
  - For \( h \) s.t. \( h/T \to \lambda > 0 \), distribution of LP-IV is mixture of normals, with mean zero (heavy tailed)
Q6. What if the instrument is weak?

\[ Y_{2t+h} = \Theta_{h,21} Y_{1t} + \gamma W_t + u^{(h)}_{t+h} \]

We have a rich set of tools to handle weak instruments.

- Weak IV biases towards OLS – which here is bias towards Cholesky with shock 1 ordered first!
  - This true in both SVAR-IV and LP-IV (Montiel-Olea, Stock, and Watson)

- Single horizon weak-instrument robust inference
  - Single instrument: Anderson-Rubin (efficient if homoskedastic)
  - Multiple instruments: CLR (nearly efficient if homoskedastic)

The literature is aware of the weak IV possibility
Stock and Watson (2012), Gertler-Karadi (2015); Ramey (2016)

Gertler-Karadi example
- First stage \( F = 15.9 \) (SVAR-IV) and \( F = 23.7 \) (LP-IV)
- Anderson-Rubin confidence intervals…
Gertler-Karadi example, ctd.

LP-IV 68% bands: ±1 SE and Anderson-Rubin Confidence Interval
Q7. How to handle news shocks?

Essentially this just requires a change to the unit effect normalization.

Example

- $\varepsilon_{1t}$ is a productivity shock (invention)
- $z_t$ is news about that invention
- $\varepsilon_{1t}$ affects observed TFP with a lag
- $\varepsilon_{1t}$ affects consumption today via present value of future output

\[
Y_{1t} = \Delta \ln TFP_t = \Theta_{1,12} \varepsilon_{1t-1} + \text{lags and other shocks}
\]
\[
Y_{2t} = \Delta \ln Consumption_t = \Theta_{0,12} \varepsilon_{1t} + \Theta_{1,12} \varepsilon_{1t-1} + \text{lags and other shocks}
\]

The unit effect normalization fails (impact effect on TFP growth is 0), and $z_t$ is an irrelevant (weak) instrument for $\eta_{1t}$.

A 1-lag unit effect normalization succeeds: \( \Theta_{1,12} = 1 \)

- A unit shock to $\varepsilon_{1t}$ increases TFP next period by 1 unit.
- All parts of conditions B and C still hold.
- The scaling for the IV regression is $E\eta_{1t+1}z_t$
- The MA need not be invertible (news shock literature)
Reminders

1. IV estimation of distributed lag, AR-distributed lag specifications generally yields correct impact effect but incorrect dynamics.

   Distributed lag: \[ Y_{2t} = \Theta_{21}(L)Y_{1t} + \{\varepsilon_t, \varepsilon_{t-j}\} \]

   ADL: \[ Y_{2t} = \Theta_{21}(L)Y_{1t} + \rho(L)Y_{2t-1} + \{\varepsilon_t, \varepsilon_{t-j}\} \]

   • Even under condition B, \( z_{t-j} \) is correlated with \( \varepsilon_{t-j} \), so \( E u_t z_{t-j} \neq 0 \).

2. SVAR-IV is more efficient than LP-IV, if correctly specified

   Reference: Kim and Kilian (2011) for simulations; standard IV and VAR results for first-order asymptotics (e.g, Lütkepohl (2005))

3. Potentially can improve LP-IV efficiency by imposing smoothness

   References: Barnichon and Brownless (2017), Plagborg-Møller (2016)
Gertler-Karadi example, ctd.

Cumulative IRFs: \textit{SVAR-IV} and \textit{LP-IV} and ±1 SE bands (parametric bootstrap)
Microeconometric IV methods carry over to macro
   - arguably yielding more credible inference on (dynamic) causal effects;

The “dynamic” part requires some additional restrictions (e.g. $z$, mds);

Well-known lessons about IVs from microeconometrics also carry over; and

These lessons aren’t new…
The first IV regression (March 15, 1926)

<table>
<thead>
<tr>
<th>Year</th>
<th>Real prices</th>
<th>Other factors</th>
<th>B</th>
<th>B</th>
<th>B</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>1903</td>
<td>1.26</td>
<td>2.73</td>
<td>3.23</td>
<td>3.46</td>
<td>93</td>
<td>128</td>
</tr>
<tr>
<td>1904</td>
<td>1.53</td>
<td>2.34</td>
<td>2.66</td>
<td>1.30</td>
<td>15</td>
<td>110</td>
</tr>
<tr>
<td>1905</td>
<td>1.23</td>
<td>2.85</td>
<td>2.53</td>
<td>1.27</td>
<td>95</td>
<td>186</td>
</tr>
<tr>
<td>1906</td>
<td>1.26</td>
<td>2.56</td>
<td>2.57</td>
<td>1.92</td>
<td>93</td>
<td>181</td>
</tr>
<tr>
<td>1907</td>
<td>1.33</td>
<td>2.59</td>
<td>2.56</td>
<td>1.60</td>
<td>119</td>
<td>187</td>
</tr>
<tr>
<td>1908</td>
<td>1.57</td>
<td>2.58</td>
<td>2.68</td>
<td>1.65</td>
<td>3.88</td>
<td>76</td>
</tr>
<tr>
<td>1909</td>
<td>2.04</td>
<td>1.97</td>
<td>2.08</td>
<td>0.95</td>
<td>3.10</td>
<td>95</td>
</tr>
</tbody>
</table>

...
P.G. Wright’s flaxseed price and output data

- Prices are Minneapolis fall prices; annual data, 1904-1923, % deviation from trend
- $z$ = building permits on East coast

Estimated supply elasticity = -0.76

First stage $F = 1.25$
PG Wright to Sewall Wright, March 15, 1925

The economic "indices".
The problem, therefore, boils down to this: In the case of any specific commodity, is it possible to find factors which have such distinct causal relations with output or demand conditions, that the values of $x$ and $y$ computed from them can be accepted with any confidence as having any relation with actuality. Such factors, I fear, especially in the case of demand conditions, it is not easy to find. I have been experimenting with various and so far have arrived at no results over which I can place much confidence.

The most likely data which I have been able to secure.
The IV regression he never computed…

Wright 1925 data: demand estimation using rainfall in upper Midwest

\[ z = \text{rainfall in Minnesota, Wisconsin, North Dakota} \]

IV estimate of demand elasticity = -0.52 (SE = 0.15)

First stage \( F = 12.8 \)