Dynamic Causal Effects for Time Series in Stata

David Schenck

Senior Econometrician
Stata

2025 Research Symposium in Econometrics and Finance
Dubai
April 24, 2025

Agenda

1. Identifying shocks
2. Local projections
3. Instrumental variables local projections
4. Instrumental variables vector autoregressions

Agenda

Introduction

- Goal: identify dynamic causal effects
- What is the effect of a tightening of monetary policy on output?
- What is the effect of a contraction in oil supply?
- Tax rates, government spending, productivity, ...
- These effects are often summarized in an impulse-response function

The challenge

- Most movements in economic variables are endogenous

$$y_t = \beta x_t + u_t$$
$$x_t = \phi y_t + e_t$$

- To disentangle casual effects, need exogenous variation
- Major research program in creating shock series
  - Narrative methods
  - High-frequency identification
- Once we have identified exogenous variation, we need to use it appropriately

Early attempts at shock identification

- Romer and Romer (1989); Ramey and Shapiro (1998)
- Isolate dates at which policy changed or a shock occurred for plausibly-exogenous reasons
- Similar theme: Hamilton (1983) identifies oil prices as exogenous to US before 1973
- Regress outcomes on these shock dates or exogenous series:

$$y_t = \sum_{j=1}^{p} \alpha_j y_{t-j} + \sum_{i=0}^{q} \beta_i d_{t-i} + u_t$$

- Compute the response function to a one-time shock to d_t

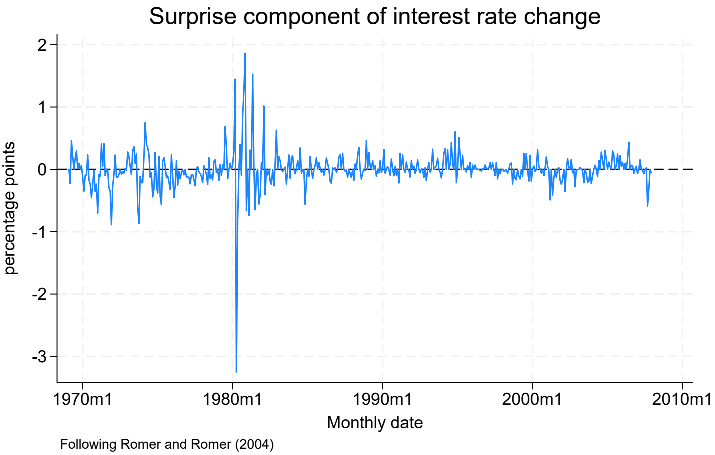
More sophisticated attempts at shock identification

- The key issue: a policy variable is changed for both endogenous and exogenous reasons

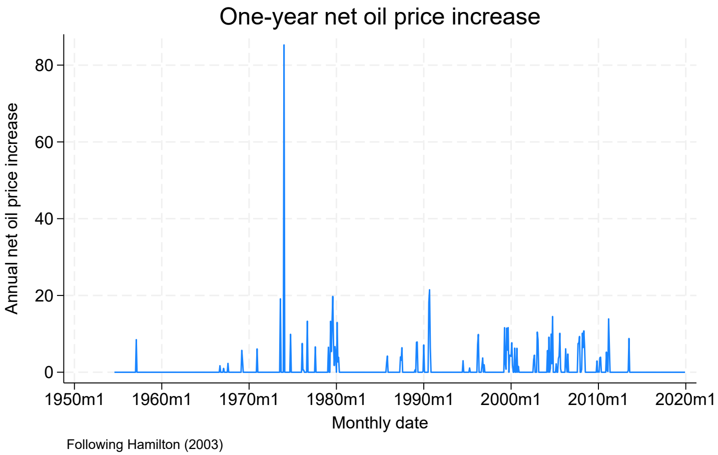
$$x_t = f(y_t, \pi_t, \dots) + e_t$$

- Extract the exogenous part e_t
- Many examples:
  - Romer and Romer (2004) monetary shock (Greenbook forecasts)
  - Swanson (2024) monetary shock (high-frequency)
  - Romer and Romer (2010) tax shocks (narrative)
  - Ramey (2011) defense buildups (narrative)
  - Hamilton (2003) oil price shock (net price increase)
  - Kilian (2008) oil supply and demand shocks
  - Useful summary: Ramey (2016 *Handbook of Macro*)

Example identified shock: The Romer monetary shocks



Example identified shock: The Hamilton oil shocks



Working with the identified shocks

- Once shocks have been identified, how to work wth them?
- Local projections (LP)
- Instrumental variables local projections (LP-IV)
- External instruments in a vector autoregression (IV-SVAR)

Agenda

Local projections

- Jorda (2005)
- For an outcome y_t and an identified shock z_t, regress the $t+h$ horizon outcome on the shock:

$$\begin{aligned} y_t &= \beta_0 z_t + \gamma' \mathbf{w}_t + u_t \\ y_{t+1} &= \beta_1 z_t + \gamma' \mathbf{w}_t + u_{t+1} \\ \vdots &= \vdots \\ y_{t+h} &= \beta_h z_t + \gamma' \mathbf{w}_t + u_{t+h} \end{aligned}$$

- The local projection estimator is the collection of $(\beta_0, \ldots, \beta_h)$ coefficients

Local projections in Stata

- Command `lpirf` (introduced in Stata 18)
- Syntax:

  `lpirf` *depvars* `[if] [in] [, options]`

- Useful options:
  - `lags(`*numlist*`)` – lags of the depvars included as controls
  - `exog()` – allows for exogenous variables
  - `step(#)` – number of impulse-response steps to compute

Local projections example

- Data: US CPI, US industrial production, Hamilton oil price shock
- Scaling: CPI and industrial production in 100 $\times$ log level
- Oil price shock scaled to represent a 10% increase in oil price

Local projections example: output

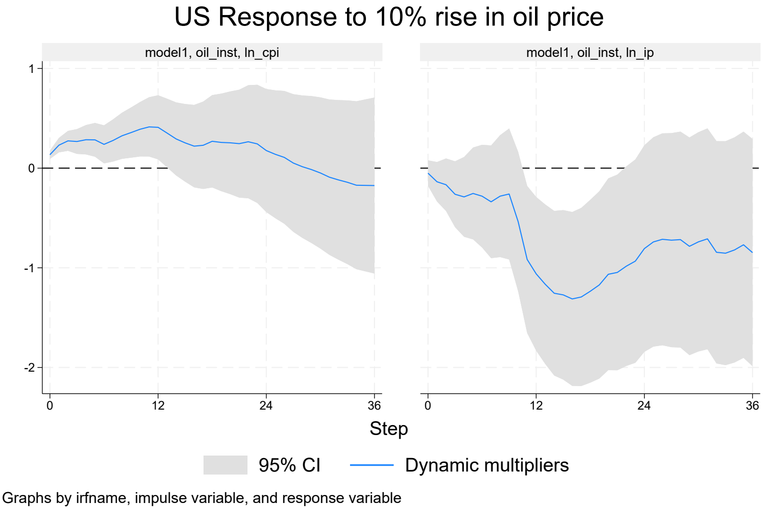
```
. lpirf ln_ip ln_cpi , exog(l(0/12).oil_inst) lag(1/12) step(6)

Local-projection impulse responses

Sample: 1960m1 thru 2015m4                        Number of obs        = 664
                                                  Number of impulses   =   3
                                                  Number of responses  =   2
                                                  Number of controls   =  34
```

	IRF coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
(output omitted)						
oil_inst						
ln_ip						
--.	-.0516982	.0639728	-0.81	0.419	-.1770826	.0736863
F1.	-.1363661	.098991	-1.38	0.168	-.330385	.0576528
F2.	-.1621691	.1316311	-1.23	0.218	-.4201612	.095823
F3.	-.2591914	.1652198	-1.57	0.117	-.5830163	.0646335
F4.	-.2829334	.2000399	-1.41	0.157	-.6750044	.1091376
F5.	-.247877	.2303465	-1.08	0.282	-.6993478	.2035938
ln_cpi						
--.	.1350983	.0214411	6.30	0.000	.0930746	.177122
F1.	.2326573	.0370489	6.28	0.000	.1600428	.3052718
F2.	.2788374	.0510551	5.46	0.000	.1787712	.3789035
F3.	.2749139	.0635589	4.33	0.000	.1503408	.3994871
F4.	.2936238	.0752978	3.90	0.000	.1460429	.4412048
F5.	.2949075	.0863967	3.41	0.001	.1255731	.4642419

Impulse responses from the local projections



Agenda

Using identified shocks as instruments

- So far I have treated the identified shocks like the true shocks:

$$z_t = e_{1t}$$

- More generally, identified shocks have the form

$$z_t = \gamma e_{it} + w_t$$

where $\gamma \neq 0$ is a bias term and w_t allows for measurement error

- Identified shocks retain two useful properties:

$$cov(z_t, e_{it}) \neq 0$$
$$cov(z_t, e_{jt}) = 0 \quad \text{for } j \neq i$$

so can be used as instruments

Using identified shocks as instruments II

- Let y_t be an outcome variable and let x_t be an impulse variable
- We wish to know how y_t is affected by x_t under a specific shock
- We have z_t, a noisy instrument for the shock
- Estimate the local projections

$$y_{t+h} = \beta_h x_t + u_{t+h}$$

  using z_t as an instrument for x_t
- The $(\beta_0, \dots, \beta_h)$ coefficients trace out an impulse response function
- Jorda and Taylor (2024)

IV local projections in Stata

- Command `ivlpirf` (introduced in Stata 19)
- Syntax:

  `ivlpirf` *depvars* `[if] [in] [, options]`

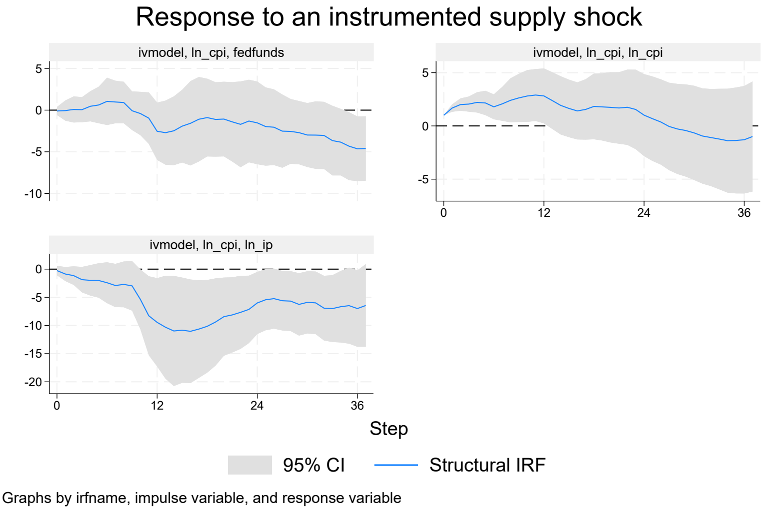
- Useful options:
  - `endog(`*endovar* `=` *instrument*`)` – specifies instrument and target shock
  - `step(#)` – number of impulse-response steps to compute

Instrumental variables local projections example

```
. ivlpirf ln_ip fedfunds, endog(ln_cpi = oil_inst) lag(1/12) nolog
Final GMM criterion Q(b) = 1.27e-32
note: model is exactly identified.
Instrumental-variables local-projection impulse responses
Sample: 1960m1 thru 2015m5                          Number of obs = 665
 ( 1)  [ln_cpi]ln_cpi = 1
```

	IRF coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
ln_ip						
--.	-.2592901	.414697	-0.63	0.532	-1.072081	.553501
F1.	-.8569246	.6341303	-1.35	0.177	-2.099797	.3859479
F2.	-1.131432	.8421365	-1.34	0.179	-2.781989	.5191249
F3.	-1.858664	1.15601	-1.61	0.108	-4.124401	.4070737
fedfunds						
--.	-.090686	.2470411	-0.37	0.714	-.5748776	.3935056
F1.	-.0327343	.6194848	-0.05	0.958	-1.246902	1.181434
F2.	.1211373	.8026353	0.15	0.880	-1.451999	1.694274
F3.	.1147244	.7732231	0.15	0.882	-1.400765	1.630214
ln_cpi						
--.	1	(constrained)				
F1.	1.690476	.1799686	9.39	0.000	1.337744	2.043208
F2.	2.048869	.2990583	6.85	0.000	1.462725	2.635012
F3.	2.113608	.3676475	5.75	0.000	1.393032	2.834184

Impulse responses from the IV local projections



Agenda

Vector autoregressions

- The setting:

$$\mathbf{y}_t = \mathbf{A}_1\mathbf{y}_{t-1} + \cdots + \mathbf{A}_p\mathbf{y}_{t-p} + \mathbf{u}_t$$
$$\mathbf{u}_t = \mathbf{B}\mathbf{e}_t$$

- $\mathbf{y}_t$ are observed variables
- $\mathbf{u}_t$ are VAR residuals
- $\mathbf{e}_t$ are unobserved shocks
- $\mathbf{B}$ is the impact matrix, from which we compute impulse responses
- Problem: $\mathbf{B}$ is not identified by data on $\mathbf{y}_t$
- Typical solution: restrict some values of $\mathbf{B}$ to zero

Instrumental variables in a VAR

- Consider again our two-equation example

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}$$

- This system would require one further restriction be identified
- The instrument behaved as follows:

$$z_t = \gamma e_{2t} + w_t$$

- Stack the instrument at the bottom of the VAR:

$$\begin{pmatrix} y_t \\ x_t \\ z_t \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & 0 \\ b_{21} & b_{22} & 0 \\ 0 & \gamma & \sigma \end{pmatrix} \begin{pmatrix} e_{1t} \\ e_{2t} \\ w_t \end{pmatrix}$$

- The 3-variable system requires 3 restrictions
- All of which are provided by the instrument

Estimation with multiple shocks I

- Angelini and Fanelli (2019) extend this logic to multiple instruments
- Consider a three-variable VAR; residuals are related to shocks via

$$u_{1t} = b_{11}e_{1t} + b_{12}e_{2t} + b_{13}e_{3t}$$
$$u_{2t} = b_{21}e_{1t} + b_{22}e_{2t} + b_{23}e_{3t}$$
$$u_{3t} = b_{31}e_{1t} + b_{32}e_{2t} + b_{33}e_{3t}$$

- And we have two measured instruments for two latent shocks

$$z_{1t} = \gamma_1 e_{1t} + w_{1t}$$
$$z_{2t} = \gamma_2 e_{2t} + w_{2t}$$

Estimation with multiple shocks II

- As before we write this system as a large VAR

$$\begin{pmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \\ z_{1t} \\ z_{2t} \end{pmatrix} = \begin{bmatrix} b_{11} & b_{12} & b_{13} & 0 & 0 \\ b_{21} & b_{22} & b_{23} & 0 & 0 \\ b_{31} & b_{32} & b_{33} & 0 & 0 \\ \gamma_{11} & \gamma_{12} & 0 & \sigma_1 & 0 \\ \gamma_{21} & \gamma_{22} & 0 & \sigma_{12} & \sigma_2 \end{bmatrix} \begin{pmatrix} e_{1t} \\ e_{2t} \\ e_{3t} \\ w_{1t} \\ w_{2t} \end{pmatrix}$$

- Compact notation:

$$\begin{pmatrix} \mathbf{u}_t \\ \mathbf{z}_t \end{pmatrix} = \begin{bmatrix} \mathbf{B}_1 & \mathbf{B}_2 & 0 \\ \mathbf{P} & 0 & \mathbf{\Sigma}_w^{1/2} \end{bmatrix} \begin{bmatrix} \mathbf{e}_t \\ \epsilon_t \\ \mathbf{w}_t \end{bmatrix}$$

- The minimum distance estimator recovers $(\mathbf{B}_1, \mathbf{P})$
- Instruments provide “credible zero restrictions”
- Method still requires $r(r-1)/2$ additional restrictions

Structural VARs in Stata

- `svar` – fully specified structural VARs
- `ivsvar gmm` – IV-GMM for one identified shock (Stata 19)
- `ivsvar mdist` – IV for multiple identified shocks (Stata 19)

ivsvar mdist setup

- Mapping the mathematical setup to Stata:

$$\begin{pmatrix} \mathbf{u}_t \\ \mathbf{z}_t \end{pmatrix} = \begin{bmatrix} \mathbf{B}_1 & \mathbf{B}_2 & 0 \\ \mathbf{P} & 0 & \boldsymbol{\Sigma}_w^{1/2} \end{bmatrix} \begin{bmatrix} \mathbf{e}_t \\ \epsilon_t \\ \mathbf{w}_t \end{bmatrix}$$

- Syntax:

  `ivsvar mdist` *depvars* `(`*endog* `=` *instr*`) [if] [in] [, options]`

- Useful options:
  - `beq(matrix)` – specify restrictions on $\mathbf{B}_1$
  - `peq(matrix)` – specify restrictions on $\mathbf{P}$

ivsvar mdist example

- Setting: three variables ip_growth, inflation, fedfunds
- Two identified shocks: oil price instrument and monetary surprise instrument
- Goals:
  - Identify impact effects of each shock
  - Assess any correlation between the shocks
  - Compute and graph impulse response functions
- Stata-speak:

```
. matrix P = (., 0  ., .)
. ivsvar mdist ip_growth (fedfunds infl = money_inst oil_inst), peq(P)
```

ivsvar mdist output

```
. matrix P = (.,0 \ .,.)
. ivsvar mdist ip_growth (fedfunds inflation = money_inst oil_inst), peq(P)
```

(*output omitted*)

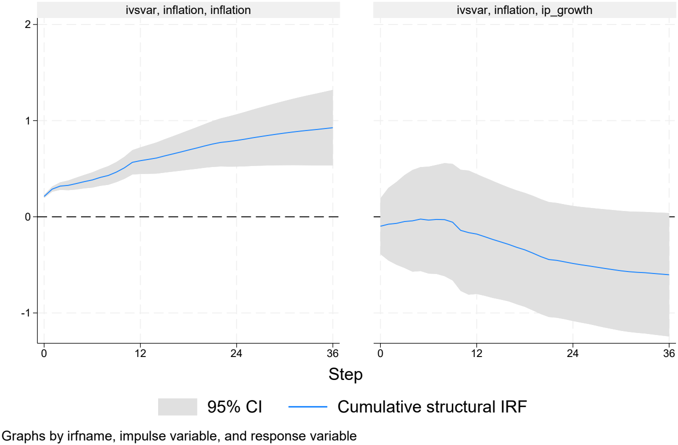
```
Instrumental-variables SVAR                          Number of obs = 468
Endogenous sample: 1954m10 thru 2019m12
Instrument sample:  1969m1 thru 2007m12

 ( 1)  [e.inflation]money_inst = 0
```

Effect	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
e.fedfunds						
ip_growth	.1597805	.0624209	2.56	0.010	.0374378	.2821232
fedfunds	.4485307	.01496	29.98	0.000	.4192097	.4778518
inflation	.0271413	.0182219	1.49	0.136	-.008573	.0628556
e.inflation						
ip_growth	-.1218286	.1342909	-0.91	0.364	-.3850338	.1413767
fedfunds	-.0222955	.0301973	-0.74	0.460	-.081481	.0368901
inflation	.2238954	.0086224	25.97	0.000	.2069959	.2407949
e.fedfunds						
money_inst	.1693461	.01252	13.53	0.000	.1448074	.1938847
oil_inst	.0378892	.2443338	0.16	0.877	-.4409963	.5167747
e.inflation						
money_inst	0	(constrained)				
oil_inst	1.298603	.2247333	5.78	0.000	.8581339	1.739072

```
Wald test of instrument relevance: chi2(6) = 243.5        Prob > chi2 = 0.000
```

Impulse responses from the IV-SVAR



Summary

- I described several methods and examples of constructing shock series
- I described three methods in Stata for estimating the dynamic effects of shocks – `lpirf`, `ivlpirf`, and `ivsvar`