Introduction to Local Projections

Òscar Jordà and Karel Mertens

AEA Continuing education 2023

Last updated: December 27, 2022

See also:
https://sites.google.com/site/oscarjorda/home/local-projections

Basic ideas

Borrowing from applied micro to draw a parallel

Back to index

Impulse responses: a comparison of two averages

$$\mathcal{R}(h) = E(E[y_{t+h}|s_t = s + \delta, \boldsymbol{x}_t] - E[y_{t+h}|s_t = s, \boldsymbol{x}_t]))$$

- y_{t+h}: outcome
- s_t: intervention
- s: baseline, e.g., $s = 0$
- δ: dose, e.g., $\delta = 1$; $\delta = var(\epsilon)^{1/2}; \ldots$
- $\boldsymbol{x}_t$: vector of exogenous and predetermined variables

Main issues to be solved

- Identification: next section
- Estimation of $E[y_{t+h}|s_t; \boldsymbol{x}_t]$
- Interpretation: multipliers
- Inference: discussed later

A trivial example

Suppose $s_t \in \{0, 1\}$ is *randomly assigned*, then:

$$\begin{aligned} \mathcal{R}(h) &= \frac{1}{N_1} \sum_{t=1}^{T-h} y_{t+h} s_t - \frac{1}{N_0} \sum_{t=1}^{T-h} y_{t+h} (1 - s_t) \\ N_1 &= \sum_{t=1}^{T-h} s_t; \quad T - h = N_1 + N_0 \end{aligned}$$

Remarks:

- inefficient (not using $\boldsymbol{x}_t$), but *consistent*
- could control for $\boldsymbol{x}_t$ with Inverse Propensity score Weighting (IPW)
- feels like the potential outcomes paradigm used in micro
- could have regressed y_{t+h} on s_t, same thing (could add $\boldsymbol{x}_t$ easily)

Estimation by Local projections

Linear case:

$$y_{t+h} = \alpha_h + \beta_h s_t + \boldsymbol{\gamma}_h \boldsymbol{x}_t + v_{t+h}; \quad \underbrace{v_{t+h} = u_{t+h} + \psi_1 u_{t+h-1} + \ldots + \psi_h u_t}_{\text{will see later why this residual MA(h)}}$$

As long as $s_t, \boldsymbol{x}_t$ exogenous w.r.t. v_t, then $\hat{\beta}_h \to \beta_h$ (identification) and then:

$$\mathcal{R}_{sy}(h) = E[y_{t+h}|s_t = s_1; \boldsymbol{x}_t] - E[y_{t+h}|s_t = s_0; \boldsymbol{x}_t] = \beta_h(s_1 - s_0)$$

General case:

$$y_{t+h} = m(s_t, \boldsymbol{x}_t; \boldsymbol{\theta}_h) + v_{t+h} \to \mathcal{R}_{sy}(h) = m(s_1, \boldsymbol{x}_t; \boldsymbol{\theta}_h) - m(s_0, \boldsymbol{x}_t; \boldsymbol{\theta}_h)$$

i.e. $m(s_t, \boldsymbol{x}_t; \boldsymbol{\theta}_h)$ can be a nonlinear function

Remarks

- **single equation estimation**: easily scales to panel, easy to extend to nonlinear specifications
- **effects 'local' to each h**: no cross-period restrictions
- **errors serially correlated**: needs fixing
- from binary to continuous treatment (dose)

Many assumptions implicit in **linear** formulation:

- **symmetry**: increase in dose same as decrease
- **scale independence**: double dose, double the effect
- **state independence**: the $\mathbf{x}_t$ don't affect $\mathcal{R}(h)$
- **treatment does not affect covariate effects**: $\boldsymbol{\gamma}_h^0 = \boldsymbol{\gamma}_h^1$
- $\delta|\mathbf{x}$ **randomly assigned**

We will analyze/generalize each of these assumptions

A STATA illustration

`LP_example.do`

- simple illustration of different variable transformations:
  - *levels vs. differences* (e.g. price index vs inflation)
  - *levels = long-differences = cumulative of differences*

$$\begin{aligned} \Delta y_{t+h} + \ldots + \Delta y_t &= y_{t+h} - y_{t+h-1} + y_{t+h-1} - y_{t+h-2} + \ldots y_t - y_{t-1} \\ &= y_{t+h} - y_{t-1} \end{aligned}$$

- shows a simple way to construct the loop and plot LPs
- maybe useful to build upon. Much left undone. Will come back to it

Relation to VARs reminder

Set aside identification discussion for now

Back to index

Propagation in an AR(1)

suppose:

$$(y_t - \mu) = \psi(y_{t-1} - \mu) + u_t$$

by recursive substitution:

$$(y_{t+h} - \mu) = \psi^{h+1}(y_{t-1} - \mu) + \underbrace{u_{t+h} + \psi u_{t+h-1} + \ldots + \psi^h u_t}_{\text{intrinsic MA residuals}}$$

suppose the intervention is $u_t = \delta$; $(u_{t+1} = \ldots = u_{t+h} = 0)$; $y_{t-1} = y^*$

$$\begin{aligned}
\mathcal{R}(h) &= E\left(E[y_{t+h}|u_t = \delta; y_{t-1} = y^*] - E[y_{t+h}|u_t = 0; y_{t-1} = y^*]\right) \\
&= E\left(\{\psi^{h+1}(y^* - \mu) + \psi^h \delta\} - \psi^{h+1}(y^* - \mu)\right) \\
&= E(\psi^h \delta) = \psi^h \delta
\end{aligned}$$

Remarks

- **iterative approach** with AR(1): from $\hat{\psi}$ obtain $\hat{\psi}^h$
- inference based on *delta method*:
  $H_0 : \psi = 0 \implies H_0 : ATE(h) = \mathcal{R}(h) = \psi^h = 0$
- **direct approach** with local projections:

$$y_{t+h} = \alpha_{h+1} + \psi_{h+1} y_{t-1} + v_{t+h}; \quad h = 0, 1, \ldots$$

- note: $v_{t+h} = u_{t+h} + \psi u_{t+h-1} + \ldots + \psi^h u_t$
- hence $E[y_{t-1}, v_{t+h}] = 0 \implies \hat{\psi}_{h+1} \xrightarrow{p} \psi^{h+1}$
- inference: correct error serial correlation (we will see how)
- $H_0 : ATE(h) = \mathcal{R}(h) = \psi_h = 0$

propagation in a VAR(2)

just to see the details

$$\underset{k\times 1}{\mathbf{y}_t} = \underset{k\times k}{A_1}\mathbf{y}_{t-1} + A_2\mathbf{y}_{t-2} + \mathbf{u}_t$$

by recursive substitution:

$$\mathbf{y}_{t+1} = (A_1^2 + A_2)\mathbf{y}_{t-1} + A_1A_2\mathbf{y}_{t-2} + \mathbf{u}_{t+1} + A_1\mathbf{u}_t$$

one more time:

$$\begin{aligned}\mathbf{y}_{t+2} =& (A_1^3 + A_2A_1 + A_1A_2)\mathbf{y}_{t-1} + (A_1^2A_2 + A_2^2)\mathbf{y}_{t-2} + \\ & \mathbf{u}_{t+2} + A_1\mathbf{u}_{t+1} + (A_1^2 + A_2)\mathbf{u}_t\end{aligned}$$

takeaway: $\mathcal{R}(h)$ a complicated function of A_1, A_2
(more on this later, an issue also raised in recent Plagborg-Møller papers)

Further exploration of the VAR—LP nexus

Back to index

A note on lag lengths

- iterated VAR-based forecasts need *correct specification*
- if not, responses will be biased
- consistency of $\mathcal{R}(h)$ only if in VAR(p) s. t. $p \to h$ as $h \to \infty$
- local projections are approximations
- no correct specification assumed
- smaller lag lengths ok for consistency under mild assumptions
- however, lag-augmentation can be very helpful for inference (later)

Some results derived more formally later

Using a VAR to construct $E[y_{t+h}|s_t, \boldsymbol{x}_t]$

Reduced-form only to explain VAR(p) vs. VAR(∞) issues

consider a VAR(p): (assume s_t and $\boldsymbol{x}_t$ in $\boldsymbol{y}_t$)

$$\underset{k\times 1}{\boldsymbol{y}_t} = \underset{k\times k}{A_1}\boldsymbol{y}_{t-1} + \ldots + A_p\boldsymbol{y}_{t-p} + \boldsymbol{u}_t; \quad E(\boldsymbol{u}_t\boldsymbol{u}_t') = \Sigma_u$$

by recursive substitution, $VMA(\infty)$:

$$\boldsymbol{y}_t = \boldsymbol{u}_t + B_1\boldsymbol{u}_{t-1} \ldots + B(\infty)\boldsymbol{y}_0;$$

$$B(\infty)\boldsymbol{y}_0 \to 0 \text{ if } |A(z)| \neq 0 \text{ for } |z| \leq 1 \quad \text{MA invertibility}$$

$B(\infty) = B(A_1, \ldots, A_p)$, e.g., see Slide 13

$\boldsymbol{y}_0$ is distant initial condition. MA invertibility $\implies B(\infty) \to 0$

Relation between $VAR(p)$ and $VMA(\infty)$

Recall the impulse response representation

$$
\begin{aligned}
B_1 &= A_1 \\
B_2 &= A_1 B_1 + A_2 \\
\vdots &= \vdots \\
B_i &= A_1 B_{i-1} + A_2 A_{i-2} + \ldots + A_p B_{i-p}; \quad i \geq p
\end{aligned}
$$

or compactly

$$
\boxed{B_i = \sum_{j=1}^{i} B_{i-j} A_j; \quad i = 1, 2, \ldots; \quad B_0 = I_k}
$$

Constructing $E[y_{t+h}|s_t, \boldsymbol{x}_t]$ using $VMA(\infty)$

from:

$$\boldsymbol{y}_{t+h} = \boldsymbol{u}_{t+h} + \ldots + B_{h-1}\boldsymbol{u}_{t+1} + B_h\boldsymbol{u}_t + B_{h+1}\boldsymbol{u}_{t-1} + \ldots$$

then:

$$E[y_{i,t+h}|u_{j,t} = 1, \boldsymbol{u}_{t-1}, \ldots] = B_h(i,j)$$

where, $s_t = u_{j,t}$ and $\boldsymbol{x}_t = \boldsymbol{u}_{t-1}, \boldsymbol{u}_{t-2}, \ldots$ hence

$$\mathcal{R}(h) = B_h(i,j); \quad \hat{B}_h = \sum_{j=1}^{h} \hat{B}_{h-j}\hat{A}_j; \quad \hat{A}_j \text{ from VAR(p)}$$

Important: in reduced form, $E(u_{i,t}u_{l,t}) \neq 0$ for $i \neq l$, usually

hence, this is not yet a well defined experiment

Fitting a finite $VAR(p)$ to a $VAR(\infty)$ (1 of 2)

A good assumption if true DGP is VARMA (e.g. many DSGE models)

Suppose the DGP is:

$$\boldsymbol{y}_t = \sum_{i=1}^{\infty} A_i \boldsymbol{y}_{t-i} + \boldsymbol{u}_t \quad with \quad \sum_{i=1}^{\infty} ||A_i|| < \infty$$

hence:

$$\boldsymbol{y}_t = \sum_{i=0}^{\infty} B_i \boldsymbol{u}_{t-i}; \quad B_0 = I; \quad det\left(\sum_{i=0}^{\infty} B_i z_i\right) \neq 0$$

$$for \quad |z_i| \leq 1 \quad and \quad \sum_{i=0}^{\infty} i^{1/2} ||B_i|| < \infty$$

Fitting a finite $VAR(p)$ to a $VAR(\infty)$ (2 of 2)

Results from Lewis and Reinsel (1985), a key paper in this literature

Let p_T denote the order of the $VAR(p_T)$. If:

$$p_T \to \infty; \quad \frac{p_T^3}{T} \to 0; \quad \sqrt{T} \sum_{i=p_T+1}^{\infty} ||A_i|| \to 0 \quad as \quad T \to \infty$$

then:

$$\sqrt{T}[vec(\hat{A}_1' \ldots \hat{A}_{p_T}') - vec(A_1 \ldots A_{p_T})] \xrightarrow{d} N(\mathbf{0}, \Sigma_a^*); \quad \Sigma_a^* \neq \Sigma_a$$

where Σ_a refers to finite $VAR(p)$, and

$$\sqrt{T}[vec(\hat{B}_h') - vec(B_h)] \xrightarrow{p} N\left(\mathbf{0}, \Sigma_u \otimes \sum_{j=0}^{h-1} B_j \Sigma_u B_j'\right); h \leq p_T$$

Note: consistency not guaranteed for $h > p_T$

Takeaways and references

- $VAR(\infty)$ results in, e.g., Lütkepohl (2005, Chapter 15)
- many DSGE have VARMA reduced form or $VAR(\infty)$
- note p_T grows with T but at a slower rate
- consistency of B_h only guaranteed up to $h = p_T$
- unlike $VAR(p)$, response S.E.s $\to 0$ as $h \to \infty$
- Plagborg-Møller and Wolf (2021): for $h \leq p_T$ VARs and LPs estimate the same response
- Jordà, Singh, and Taylor (2020): for $h > p_T$ VAR responses are biased, but LPs are not (under certain conditions)

VAR vs. LP Bias in infinite lag processes

Or why LPs can be more reliable for long-horizon responses

Intuition:

- suppose D.G.P. is:

$$\boldsymbol{y}_t = \sum_{j=0}^{\infty} A_j \boldsymbol{y}_{t-j} + \boldsymbol{u}_t; \quad \sum_{j=1}^{\infty} ||A_j|| < \infty$$

- fit VAR(1)
- true vs. VAR(1) IRFs

$VAR(\infty)$	$VAR(1)$
$B_1 = A_1$	$B_1^* = A_1$
$B_2 = A_1^2 + A_2$	$B_2^* = A_1^2$
$B_3 = A_1^3 + 2A_1A_2 + A_3$	$B_3^* = A_1^3$
$B_4 = A_1^4 + 3A_1^2A_2 + 2A_1A_3 + A_4$	$B_4^* = A_1^4$

VAR bias

Consistency guaranteed up to p only for $VAR(\infty)$

objective: truncate $VAR(\infty)$ so that remaining lags are "small"

$$\frac{1}{T^{1/2}} \sum_{j=p+1}^{\infty} ||A_j|| \to 0; \quad p, T \to \infty$$

however, from the usual $VAR \to VMA$ recursion, these terms are missing for $h > p$:

$$BIAS: \; A_{p+1}B_{h-(p+1)} + \ldots + A_{h-1}B_1 + A_h; \quad h > p$$

problem: in practice VARs are truncated too early

LP bias
or lack thereof

when is the LP consistent? i,e, when is this condition met:

$$||\hat{A}_{h,1} - B_h|| \xrightarrow{p} 0; \quad p, T \rightarrow \infty$$

in the LP:

$$\boldsymbol{y}_{t+h} = A_{h,1}\boldsymbol{y}_{t-1} + \ldots + A_{h,p}\boldsymbol{y}_{t-p} + \boldsymbol{u}_{t+h}$$

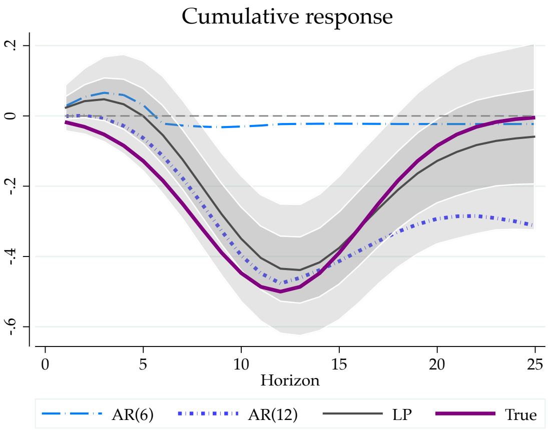
turns out same as consistency of $VAR(p)$, i.e.

$$p^{1/2}\sum_{j=0}^{\infty}||A_{k+j}|| \rightarrow 0$$

see proof in Jordà, Singh, Taylor (2020)

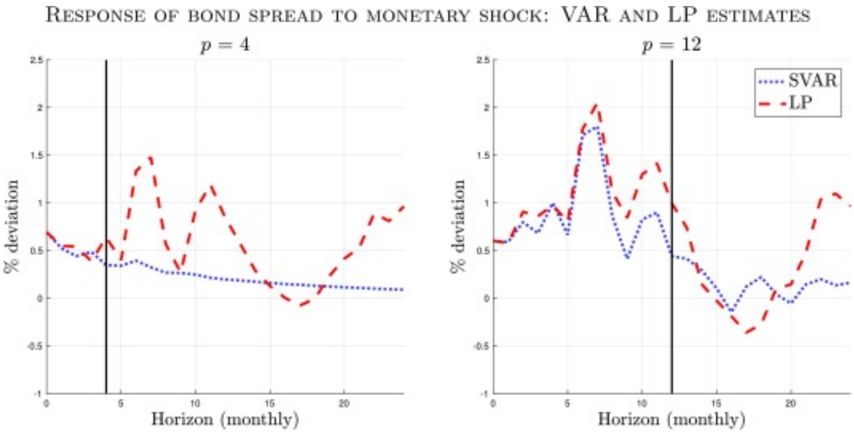
Illustration of VAR vs. LP bias

Based on MA(24) model



Another example

Figure 2 in Palgborg-Møller and Wolf (2021, ECTA)



Multipliers and Counterfactuals

Back to index

Two models, same response, different conclusions

Alloza, Gonzalo, Sanz (2020)

$$(a)\begin{cases}\Delta y_t &= \beta \Delta s_t + u_t^y \\ \Delta s_t &= \rho \Delta s_{t-1} + u_t^s\end{cases};\quad (b)\begin{cases}\Delta y_t &= \beta \Delta s_t + \rho \Delta y_{t-1} + u_t^y \\ \Delta s_t &= u_t^s\end{cases};\ \boldsymbol{u}_t \sim D(\mathbf{0}, I)$$

Note: $\mathcal{R}_{sy}^a(h) = \beta\rho^h = \mathcal{R}_{sy}^b(h)$. Both can be estimated with the LP:

$$\Delta y_{t+h} = \gamma_h \Delta s_t + \psi_h \Delta y_{t-1} + v_{t+h}$$

Propagation in (a), due to correlated treatment, in (b) correlated outcome. Consider augmenting LP with treatment leads:

$$\Delta y_{t+h} = \gamma_h \Delta s_t + \psi_h \Delta y_{t-1} + \sum_{i=1}^{h} \phi_i \Delta s_{t+i} + v_{t+h};$$

$\tilde{\mathcal{R}}_{sy}^a(h) = \beta$; $\quad \tilde{\mathcal{R}}_{sy}^b(h) = \beta\rho^h$

What is going on?

- in both cases, Δs_t is strictly exogenous. Leads are allowed in the LP
- in model (a), including leads removes the effect from future potential treatments (due to treatment serial correlation)
- in model (b), on average, there is no expectation of additional treatment. The leads do not matter
- what is the effect of a single treatment? In (a) β, in (b) $\beta\rho^h$
- think of the LP MA(h) residual structure. In general, the MA would have terms in u^y_{t+i} and u^s_{t+i}. But in model (b) coeffs on u^s_{t+i} are all zero
- another way to think about these effects is using multipliers

From previous example

Consider the following model (model (a) earlier):

$$\begin{cases} \Delta y_t & = \beta \Delta s_t + u_t^y \\ \Delta s_t & = \rho \Delta s_{t-1} + u_t^s \end{cases} ; \quad \boldsymbol{u}_t \sim D\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} ; \begin{pmatrix} \sigma_y & 0 \\ 0 & \sigma_s \end{pmatrix} \right)$$

Trivially: $\mathcal{R}_{sy}(h) = \beta \rho^h$; $\mathcal{R}_{ss}(h) = \rho^h$

The cumulative impact, $\mathcal{C}_{ij}(h) = \sum_{k=0}^{h} \mathcal{R}_{ij}(k)$ can be directly estimated from:

$$\begin{aligned} y_{t+h} - y_{t-1} = \Delta_h y_{t+h} &= \theta_h \Delta s_t + v_{t+h}^y; \quad v_{t+h}^y \sim MA(h) \\ s_{t+h} - s_{t-1} = \Delta_h s_{t+h} &= \psi_h \Delta s_t + v_{t+h}^s; \quad v_{t+h}^s \sim MA(h) \end{aligned}$$

with $\mathcal{C}_{sy}(h) = \theta_h = \beta \sum_{k=0}^{h} \rho^k$; $\mathcal{C}_{ss}(h) = \psi_h = \sum_{k=0}^{h} \rho^k$

Calculating the multiplier

Define:

$$m_h = \frac{\mathcal{C}_{sy}(h)}{\mathcal{C}_{ss}(h)} = \frac{\beta \sum_{k=0}^{h} \rho^k}{\sum_{k=0}^{h} \rho^k} = \beta; \text{cum. change in } y \text{ due to cum. change in } s$$

Suppose Δz_t is a valid instrument for Δs_t then:

$$\begin{aligned} E(\Delta_h y_{t+h}, \Delta z_t) &= \theta_h E(\Delta s_t \Delta z_t) \\ E(\Delta_h s_{t+h}, \Delta z_t) &= \psi_h E(\Delta s_t \Delta z_t) \end{aligned}$$

hence m_h can be directly estimated from the IV projection:

$$\Delta_h y_{t+h} = m_h \Delta_h s_{t+h} + \eta_{t+h}; \quad \text{instrumented with } \Delta z_t$$

References

- Ramey, Valerie A. 2016. Macroeconomic shocks and their propagation. In *Handbook of Macroeconomics*, Vol. 2, ed. John Taylor and Harald Uhlig. Elsevier, 71–162. Chapter 2.
- Ramey, Valerie A. and Sarah Zubairy. 2018. Government spending multipliers in good times and in bad: Evidence from U.S. historical data. *Journal of Political Economy*, 126(2):850–901.
- Stock, James H. and Mark Watson. 2018. Identification and estimation of dynamic causal effects in macroeconomics using external instruments. *Economic Journal*, 128(610): 917–948.

Panel data applications

Back to index

LPs in panels

The set-up

$$y_{i,t+h} = \alpha_i + \delta_t + s_{i,t}\beta_h + \boldsymbol{x}_{i,t}\boldsymbol{\gamma}_h + v_{i,t+h}; \quad i = 1, \ldots, n; \quad t = 1, \ldots, T$$

- α_i unit-fixed effects
- δ_t time-fixed effects
- $\boldsymbol{x}_{i,t}$ exogenous and pre-determined variables
- $s_{i,t}$ treatment variable
- β_h response coefficient of interest

Sample code: LP_example_panel.do

Panel-LPs

Remarks: usual panel data issues appear here too

- LP is costly in short-panels (lost time dimension cross-sections)
- but cross-section brings more power
- incidental parameter issues (fixed effects):
  - beware of high autocorr and low T (Alvarez and Arellano, 2003 ECTA)
  - will need Arellano-Bond or similar estimator
- inference
  - n, T large $\rightarrow$ two-way clustering helps MA(h) and heteroscedasticity
  - n large, T small $\rightarrow$ cluster by unit helps with MA(h)
  - T large, n small $\rightarrow$ cluster by time helps heteroscedasticity
  - else, Driscoll-Kraay is like Newey-West for panel data
  - when clustering with small n, T, may need bootstrap.
    See papers here and here.
    See also `summclust` and `boottest` STATA ado files

Cointegration

A brief detour

Back to index

What is cointegration?

Idea: two variables can be I(1) but their linear combination is I(0). Example:

$$\begin{cases} y_{1,t} & = \gamma y_{2,t} + u_{1,t} \\ y_{2,t} & = y_{2,t} + u_{2,t} \end{cases} ; \quad y_{1,t}, y_{2,t} \sim I(1) \quad but \quad z_t = y_{1,t} - \gamma y_{2,t} \sim I(0)$$

In general:

$$\boldsymbol{y}_t = \alpha + \Phi_1 \boldsymbol{y}_{t-1} + \ldots + \Phi_p \boldsymbol{y}_{t-p} + \boldsymbol{u}_t$$

cointegration means:

$$\Phi(1) \equiv I - \Phi_1 - \ldots - \Phi_p \quad then \quad rank(\Phi(1)) = g < n$$

that is, the system has $n - g$ unit roots and g cointegrating vectors, s.t. $\Phi(1) = BA'$ with A, B $n \times g$ matrices, and $A'\boldsymbol{y}_t = z_t$ cointegrating vectors

The VECM representation

Using general representation of a VAR(p)

$$\begin{aligned} \boldsymbol{y}_{t+1} &= \Phi_1 \boldsymbol{y}_t + \ldots + \Phi_{p+1} \boldsymbol{y}_{t-p} + \boldsymbol{\alpha} + \boldsymbol{u}_{t+1} \\ \boldsymbol{y}_{t+1} &= \Psi_1 \Delta \boldsymbol{y}_t + \ldots + \Psi_p \Delta \boldsymbol{y}_{t-p+1} + \Pi \boldsymbol{y}_t + \boldsymbol{\alpha} + \boldsymbol{u}_{t+1} \end{aligned}$$

with $\Psi_j = -[\Phi_{j+1} + \ldots + \Phi_{p+1}]$; for $j = 1, \ldots, p$ and $\Pi = \sum_{j=1}^{p+1} \Phi_j$
subtracting $\boldsymbol{y}_t$ on both sides:

$$\Delta \boldsymbol{y}_{t+1} = \Psi_1 \Delta \boldsymbol{y}_t + \ldots + \Psi_p \Delta \boldsymbol{y}_{t-p+1} + \Psi_0 \boldsymbol{y}_t + \boldsymbol{\alpha} + \boldsymbol{u}_{t+1}$$

Note: $\Psi_0 = -\Phi(1) = BA'$ when there is cointegration, and $\boldsymbol{z}_t = A' \boldsymbol{y}_t$

VECM

$$\Delta \boldsymbol{y}_{t+1} = \Psi_1 \Delta \boldsymbol{y}_t + \ldots + \Psi_p \Delta \boldsymbol{y}_{t-p+1} - B \boldsymbol{z}_t + \boldsymbol{\alpha} + \boldsymbol{u}_{t+1}$$

How does cointegration affect impulse responses?

Remarks

- responses from levels VAR **always** correct
- responses from differenced VAR **only** correct if no cointegration
- cointegration improves efficiency …
- … but estimation and inference more dificult
- responses often not used to investigate LR equilibrium relationships but should
- useful to impose LR exclusion identification restrictions

Cointegrated systems in state-space form

notice:

$$\Psi_0 = \Pi - I = -\Phi(1);$$

if rank $(\Psi_0) < n \rightarrow \Phi(1) = BA'$; cointegrating vector: $\boldsymbol{z}_t = A'\boldsymbol{y}_t$

$$\begin{bmatrix} \boldsymbol{z}_{t+1} \\ \Delta\boldsymbol{y}_{t+1} \\ \Delta\boldsymbol{y}_t \\ \vdots \\ \Delta\boldsymbol{y}_{t-p+1} \end{bmatrix} = \begin{bmatrix} A'\Pi & A'\Psi_1 & \dots & A'\Psi_{p-1} & A'\Psi_p \\ -B & \Psi_1 & \dots & \Psi_{p-1} & \Psi_p \\ 0 & I & \dots & 0 & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & I & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{z}_t \\ \Delta\boldsymbol{y}_t \\ \Delta\boldsymbol{y}_{t-1} \\ \vdots \\ \Delta\boldsymbol{y}_{t-p} \end{bmatrix} + \begin{bmatrix} A'\boldsymbol{u}_{t+1} \\ \boldsymbol{u}_{t+1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$Z_{t+1} = \boldsymbol{\Psi} Z_t + V_{t+1}$$

Usefulness of state-space representation

Calculating impulse responses through recursive substitution

long-run dynamics:

$$\boldsymbol{z}_{t+h} = \boldsymbol{\Psi}^h_{[1,1]}\boldsymbol{z}_t + \boldsymbol{\Psi}^h_{[1,2]}\Delta y_t + \sum_{j=3}^{p-2} \boldsymbol{\Psi}^h_{[1,j]}\Delta y_{t-j+2} + \boldsymbol{\nu}_{t+h}$$

$$\boldsymbol{\nu}_{t+h} = A'\boldsymbol{u}_{t+h} + A'(I+\Gamma_1)U_{t+h-1} + \ldots + A'(I+\Gamma_1+\ldots+\Gamma_{h-1})U_{t+1}$$

short-run dynamics:

$$\Delta \boldsymbol{y}_{t+h} = \boldsymbol{\Psi}^h_{[2,1]}\boldsymbol{z}_t + \boldsymbol{\Psi}^h_{[2,2]}\Delta \boldsymbol{y}_t + \sum_{j=3}^{p-2} \boldsymbol{\Psi}^h_{[2,j]}\Delta \boldsymbol{y}_{t-j+2} + \boldsymbol{\upsilon}_{t+h}$$

$$\boldsymbol{\upsilon}_{t+h} = \boldsymbol{u}_{t+h} + \Gamma_1\boldsymbol{u}_{t+h-1} + \ldots + \Gamma_{h-1}\boldsymbol{u}_{t+1}$$

where

$$\Delta \boldsymbol{y}_t = \sum_{j=0}^{\infty} \Gamma_j \boldsymbol{u}_{t-j}$$

Responses to equilibrium shocks

equilibrium dynamics, short- vs. long-run effects:

$$\mathcal{R}_z(h; A'u_{t+1} = 1) = (I + \Gamma_1 + \ldots + \Gamma_h)A = \underbrace{\boldsymbol{\Psi}^h_{[1,1]}}_{LR} + \underbrace{\boldsymbol{\Psi}^h_{[1,2]}A}_{SR}$$

short-run dynamics, short- vs long-run effects:

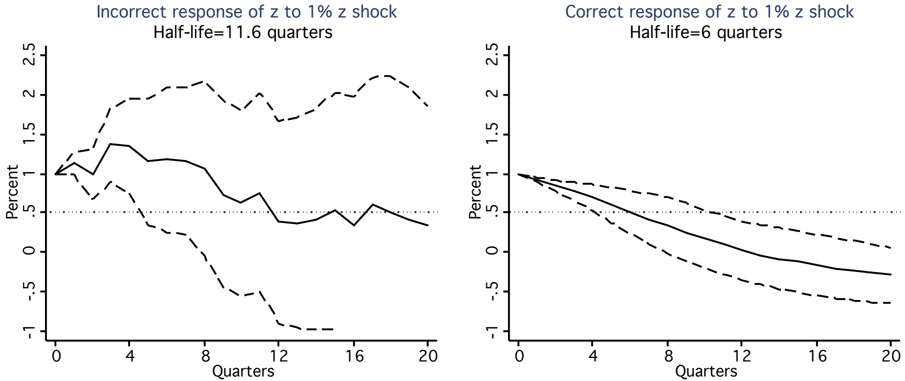
$$\mathcal{R}_{\Delta y}(h; A'u_{t+1} = 1) = \Gamma_h A = \underbrace{\boldsymbol{\Psi}^h_{[2,1]}}_{LR} + \underbrace{\boldsymbol{\Psi}^h_{[2,2]}A}_{SR}$$

remarks:

- note shock cointegrating vector, $\boldsymbol{z}$, not a variable
- each response, 2 parts:
  1. return to equilibrium (LR)
  2. short-run frictions (SR)

Application

Chong, Yanping, Òscar Jordà, and Alan M. Taylor. 2012. The Harrod-Balassa-Samuelson Hypothesis: Real Exchange Rates and their Long-Run Equilibrium. *International Economic Review*, 53(2): 609—634



Variance decompositions

Back to index

intuition

two important recent references:

- Gorodnichenko, Yuriy and Byoungchan Lee. 2020. Forecast error variance decompositions with local projections. Journal of Business and Economics Statistics
- Plagborg Møller, Mikkel and Christian K. Wolf. 2022. Instrumental variable identification of dynamic variance decompositions. Journal of Political Economy.

can always write $y_{t+h} = \hat{E}_t(y_{t+h}) + \hat{v}_{t+h}$

then R^2 of regression of $\hat{v}_{t+h}$ on $\epsilon_{j,t+h}, \ldots, \epsilon_{j,t}$ measures percent of FEV explained by j-shock

assumes structural shock $\epsilon_{j,t}$ available

Smooth local projections

Back to index

Smoothing

relevant references:

- Barnichon, Regis and Christian Brownlees. 2018. Impulse response estimation by smooth local projections. Available at: https://sites.google.com/site/regisbarnichon/research
- Barnichon, Regis and Christian Matthes. 2018. Functional approximations of impulse responses (FAIR). *Journal of Monetary Economics*, forthcoming.

Many solutions. A simple one: Gaussian Basis Functions

Intuition: impose some cross-horizon discipline to smooth LP wiggles. Can improve efficiency

Other options: bayesian shrinkage
see, e.g. Miranda-Agrippino and Rico. 2018. Bayesian Local Projections

A general approach to smoothing

GMM provides local projection estimates of the response $\hat{\mathcal{R}}$ given by $\hat{\gamma}$ and $\hat{\Sigma}_{\hat{\gamma}}$

a natural solution is minimum distance

let $\boldsymbol{\psi}(\hat{\boldsymbol{\gamma}}, \boldsymbol{\theta})$ be a function that returns a smoothed estimate of $\hat{\boldsymbol{\gamma}}$ based on auxiliary parameters $\boldsymbol{\theta}$, then:

$$\min_{\boldsymbol{\theta}} \left[\hat{\boldsymbol{\gamma}} - \boldsymbol{\psi}(\boldsymbol{\theta})\right]' \hat{\Sigma}_{\gamma} \left[\hat{\boldsymbol{\gamma}} - \boldsymbol{\psi}(\boldsymbol{\theta})\right]$$

delivering, $\hat{\boldsymbol{\theta}}, \hat{\Sigma}_{\hat{\theta}}$ and if $dim(\boldsymbol{\gamma}) > dim(\boldsymbol{\theta})$, a test of overidentifying restrictions for $\boldsymbol{\psi}(\boldsymbol{\theta})$

Smoothing with Gaussian Basis Functions

suppose no controls to simplify

$$\mathcal{R}(h; a, b, c) = \psi(h) = ae^{-\left(\frac{h-b}{c}\right)^2}$$

Using GMM set-up, two estimators: direct v. 2-step
Direct estimator:

$$\min_{a,b,c} \left[\sum_{t=1}^{T} Z_t'(\boldsymbol{y}_{t,H} - S_t\psi(h))\right]' \hat{W} \left[\sum_{t=1}^{T} Z_t'(\boldsymbol{y}_{t,H} - S_t\psi(h))\right]$$

2-step: Step-1 is usual LP, get $\hat{\gamma}, \hat{\Sigma}_\gamma$, then min. distance

$$\min_{a,b,c} \left[\hat{\gamma} - \phi(h)\right]' \hat{\Sigma}_\gamma \left[\hat{\gamma} - \phi(h)\right]$$

GBF–GMM

Remarks

- direct method requires NL estimation techniques
- however, problem is reasonably well behaved
- 2-step method provides useful intuition
- note H-period LP, but 3 parameters so $(H+1)-3$ overidentifying restrictions
- regardless of method, J-test natural specification test
- considerable gain in parsimony $\implies$ efficiency
- GBF approximation works well with ”single humps”
- multiple ”humps” require more basis functions $\implies$ GBF approach no longer as practical

approximation using gaussian basis functions

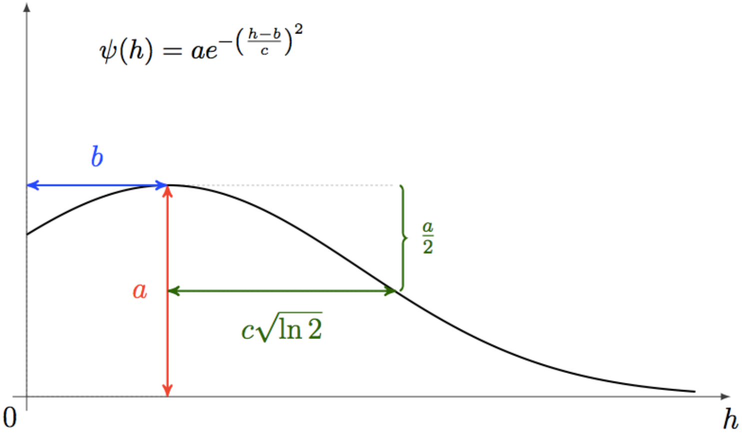
recall:

$$\mathcal{R}(h) = \psi(h) = ae^{-\left(\frac{h-b}{c}\right)^2}$$

what does each parameter do?

- a scales the entire response
- b dates the peak effect
- c measures the half-life

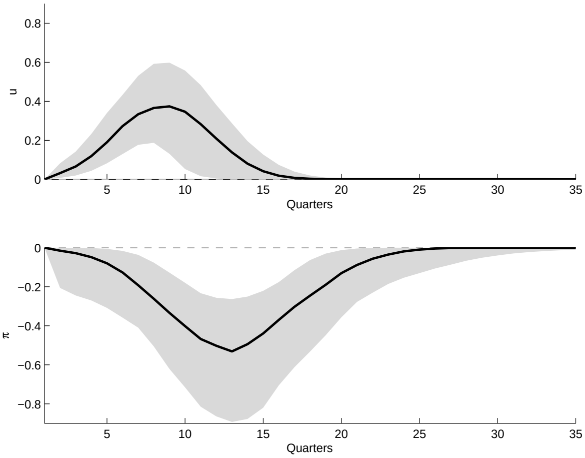
gaussian basis functions
the picture



Sample code: LP_GBF.do

GBF–GMM example

unemployment v. inflation response to monetary policy shock



Nonlinearities and other potential extensions

Back to index

The principle

What we are after:

$$\mathcal{R}_{sy}(h) = E[y_{t+h}|s_t = s_0 + \delta; \mathbf{x}_t] - E[y_{t+h}|s_t = s_0; \mathbf{x}_t]$$

No reason to assume the conditional expectation is linear

Example:

$$\begin{aligned} y_{t+h} &= \gamma_{1h} s_t + \gamma_{2h} s_t^2 + \boldsymbol{\gamma} \mathbf{x}_t + v_{t+h} \quad \rightarrow \\ \mathcal{R}_{sy}(h) &= \gamma_{1h}(s_0 + \delta) + \gamma_{2h}(s_0 + \delta)^2 + \boldsymbol{\gamma} \mathbf{x}_t - (\gamma_{1h} s_0 + \gamma_{2h} s_0^2 + \boldsymbol{\gamma} \mathbf{x}_t) \\ &= \gamma_{1h} + \gamma_{2h}(\delta^2 + 2 s_0 \delta) \end{aligned}$$

Hence, $\mathcal{R}_{sy}(h)$ depends on δ and s_0, just like NL regression

Binary dependent variable

Example: response probability of financial crisis to today's credit shock

$$\mathcal{R}_{sy}(h) = P(y_{t+h} = 1 | s_t = s_0 + \delta; \boldsymbol{x}_t) - P(y_{t+h} = 1 | s_t = s_0; \boldsymbol{x}_t)$$

Remarks:

- logit/probit $\rightarrow \mathcal{R}_{sy}(h)$ depends on s_0, δ and $\boldsymbol{x}_t$
- can estimate a linear probability model. But crises are tail events

Another example: Text-based recession probabilities

Ferrari Minesso, M., Lebastard, L. & Le Mezo, H. Text-Based Recession Probabilities. IMF Econ Rev (2022).

Response of recession probability

Marginal effect of 1% increase in newspaper-based index

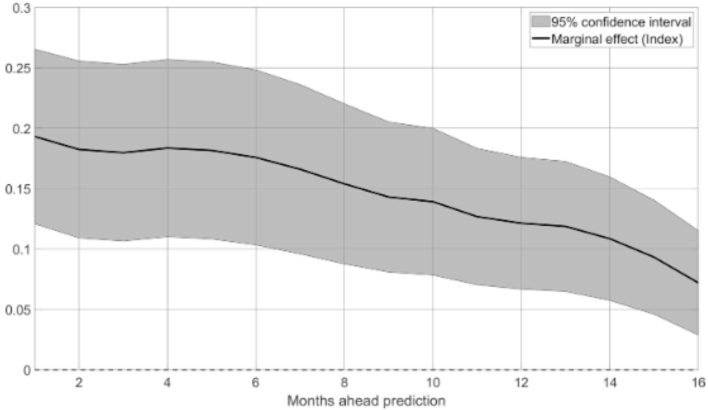


Fig. 2 Marginal effects from Eq. (4.1). Notes: Marginal effects $\left(\frac{\partial P(Recession_{t+k}=1|t)}{\partial Index_t}\right)$ from the probit regression for a 1% increase in the newspaper-based index (i.e., a 1% increase in the share of newspaper articles discussing a recession in the USA). Grey shaded areas report 95% confidence intervals

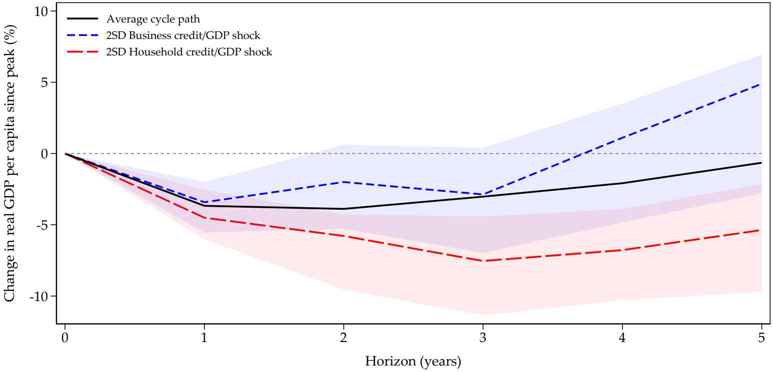
Quantile LPs

Example: does high corporate debt increase risk of left tail GDP draws? Does it depend on legal bankruptcy framework?

$$
\begin{aligned}
\hat{\gamma}_{h,\tau} \;=\; & \underset{\gamma_{h,\tau}}{\text{argmin}} \sum_{1}^{t(P)} \Big(\tau\, \mathbf{1}(\Delta_h y_{it(p)+h} \geq s_{it(p)}\gamma_{h,\tau}) \big| \Delta_h y_{it(p)+h} - s_{it(p)}\gamma_{h,\tau} \big| \\
+ \; & (1-\tau)\, \mathbf{1}(\Delta_h y_{it(p)+h} < s_{it(p)}\gamma_{h,\tau}) \big| \Delta_h y_{it(p)+h} - s_{it(p)}\gamma_{h,\tau} \big| \Big)
\end{aligned}
$$

Jordà, Kornejew, Schularick, and Taylor. 2022. Zombies at large? Corporate debt overhang and the macroeconomy. Review of Financial Studies

Figure A.4: *Business and household debt, responses at 20th percentile of real GDP per capita growth*



Notes: Figures show the predictive effects on growth of a two-SD business/household debt buildup in the five years preceding the recession based on a LP series of quantile regressions. Business credit booms shown in the left-hand side panel and household debt booms shown in the right-hand side panel. Shaded areas denote the 95% confidence interval based on bootstrap replications. See text.

Factor models

Idea: control for many covariates using factor model. Suppose:

$$\begin{cases} \underset{k\times 1}{\boldsymbol{x}_t} & = \lambda(L) \underset{q\times 1}{\boldsymbol{f}_t} + \boldsymbol{e}_t \\ \boldsymbol{f}_t & = \pi(L)\boldsymbol{f}_{t-1} + \boldsymbol{\eta}_t \end{cases} ; \quad k >> q;\ E(\boldsymbol{e}_t) = E(\boldsymbol{\eta}_t) = 0;\ E(\boldsymbol{e}_t\boldsymbol{\eta}'_{t-j}) = 0\ \forall j$$

Then LP can be specified as:

$$y_{t+h} = \beta_h s_t + \sum_{j=0}^{p} \boldsymbol{\gamma}_j \boldsymbol{f}_{t-j} + v_{t+h}$$