Introduction to Local Projections

Centre for Central Banking Studies

Simon Lloyd

Bank of England and Centre for Macroeconomics

September 2024

The views expressed here do not necessarily reflect the position of the Bank of England.

These slides draw on material from Òscar Jordà and Karel Mertens' *Techniques of Empirical Macroeconomics* course, part of the AEA's 2023 Continuing Education Webcast (available here).

Common ‘Complaints’ About VARs

- VARs not robust to misspecification
  - $\Rightarrow$ If you have the ‘wrong’ VAR, you have the wrong IRFs
- Not easy to account for *non-linearities* within VARs
  - VARs are best, linear one-step ahead predictors
  - $\Rightarrow$ Quality of h-step-ahead responses from VARs depends on quality of 1-step-ahead responses, which are iterated upon
- Inference around VAR-implied impulse responses is complex
  - Standard errors around VAR IRFs are highly non-linear functions of estimated parameters
  - $\Rightarrow$ Practitioners often have to resort to complex bootstrap methods for inference

Enter...Local Projections

Òscar Jordà (SF Fed, UC Davis)

Key Advantages

- Estimated by single-equation OLS, with standard regression packages
- Simple, analytic, joint inference for impulse responses
- More robust to misspecification
- Experimentation with non-linearities straightforward

Enter...Local Projections

Òscar Jordà (SF Fed, UC Davis)

Key Advantages

- Estimated by single-equation OLS, with standard regression packages
- Simple, analytic, joint inference for impulse responses
- More robust to misspecification
- Experimentation with non-linearities straightforward

But not as different to VARs as you might think

[Plagborg-Møller & Wolf, 2021]

Basic Idea

Impulse Response Definition: A Comparison of Two Averages

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

where:

- y_{t+h}: outcome
- s_t: intervention
    - s: baseline, e.g., $s = 0$
    - δ: 'dose', e.g., $\delta = 1$, $\delta =$ var$(\varepsilon)^{1/2}$,...
- $\mathbf{x}_t$: vector of exogenous and predetermined variables

Estimation by Local Projection

For $h = 1, 2, ..., H$, can estimate:

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

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

$$\mathcal{R}(h) = \mathbb{E}[y_{t+h}|s_t = s_1, \mathbf{x}_t] - \mathbb{E}[y_{t+h}|s_t = s_0, \mathbf{x}_t] = \beta_h(s_1 - s_0)$$

Remarks

- **Single equation estimation**: generalises to panel; easy to extend to non-linear setup
- **Effects 'local' to each h**: no cross-period restrictions ▸ In Practice
- **Errors serially correlated**: needs fixing (more on this later)
- Can consider binary 'dose' or continuous policy 'treatment'

Remarks

- **Single equation estimation**: generalises to panel; easy to extend to non-linear setup
- **Effects 'local' to each h**: no cross-period restrictions

▸ In Practice

- **Errors serially correlated**: needs fixing (more on this later)
- Can consider binary 'dose' or continuous policy 'treatment'

Even so, still many assumptions implicit in *linear* formulation:

- **Symmetry**: increase in dose has same effects as decrease
- **Scale independence**: double the dose, double the effect
- **State independence**: the $\mathbf{x}_t$ do not affect $\mathcal{R}(h)$
- **Treatment does not affect covariate effects**: γ^h invariant to s_t
- $\delta|\mathbf{x}_t$ should be **randomly assigned**

Though, these assumptions can be generalised

A STATA Illustration

`LP_example.do`

- Simple illustration of different variable transformations:
  - Levels vs. Differences (e.g., price index vs. inflation)
  - Levels = Long-differences = Cumulation of differences

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

- Code shows a simple way to construct the loop and plot LPs

- Code can be built upon...as we will do later on!

Relation to VARs

Without Identification

Propagation in AR(1)

Suppose:

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

By recursive substitution, we have:

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

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

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

Remarks

- **Iterative approach** for AR(1): from $\hat{\psi}$ obtain $\hat{\psi}^h$
  - Here, inference on $\hat{\psi}^h$ would need to be based on delta method

$$H_0: \ \psi = 0 \implies H_0: \ ATE(h) = \mathcal{R}(h) = \psi^h = 0$$

- **Direct approach** with local projection:

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

  - where: $v_{t+h} = u_{t+h} + \psi u_{t+h-1} + ... + \psi^h u_t$

$\Rightarrow$ Hence $\mathbb{E}[y_{t-1} v_{t+h}] = 0 \implies \hat{\psi}_{h+1} \xrightarrow{p} \psi^{h+1}$

  - Here, inference should correct error serial correlation (we will see how):

$$H_0: \ ATE(h) = \mathcal{R}(h) = \psi_h = 0$$

Extends to Propagation in a VAR(2)

$$\underset{k\times 1}{\mathbf{y}_t} = \underset{k\times k}{\mathbf{A}_1}\mathbf{y}_{t-1} + \mathbf{A}_2\mathbf{y}_{t-2} + \mathbf{u}_t$$

By recursive substitution, we have:

$$\mathbf{y}_{t+1} = \left(\mathbf{A}_1^2 + \mathbf{A}_2\right)\mathbf{y}_{t-1} + \mathbf{A}_1\mathbf{A}_2\mathbf{y}_{t-2} + \mathbf{u}_{t+1} + \mathbf{A}_1\mathbf{u}_t$$

and, then:

$$\begin{aligned}\mathbf{y}_{t+2} = &\left(\mathbf{A}_1^3 + \mathbf{A}_2\mathbf{A}_1 + \mathbf{A}_1\mathbf{A}_2\right)\mathbf{y}_{t-1} + \left(\mathbf{A}_1^2\mathbf{A}_2 + \mathbf{A}_2^2\right)\mathbf{y}_{t-2} \\ &+ \mathbf{u}_{t+2} + \mathbf{A}_1\mathbf{u}_{t+1} + \left(\mathbf{A}_1^2 + \mathbf{A}_2\right)\mathbf{u}_t\end{aligned}$$

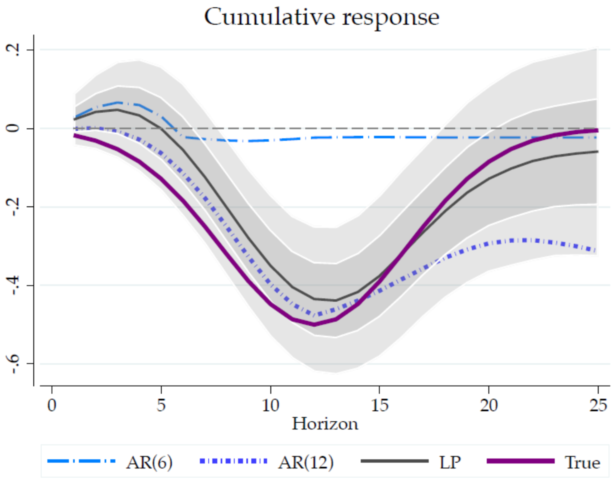
Takeaway: $\mathcal{R}(h)$ a complicated function of $\mathbf{A}_i$

A Note on Lag Lengths in VARs and LPs

- Iterated VAR-based forecasts need *correct specification*
  - If not, responses will be biased
  - $\mathcal{R}(h)$ will be consistent only if in VAR(p) s.t. $p \to h$ as $h \to \infty$
- Instead, local projections are approximations
  - No correct specification is assumed
  - Smaller lag lengths are fine for consistency under mild assumptions
  - However, lag-augmentation can be helpful for inference (more later)
- E.g.: Using VAR(p) to estimate VMA(∞), consistency of VAR(p) only guaranteed up to $h = p$
  - For $h \leq p$: VARs and LPs estimate same response [Plagborg-Møller & Wolf, 2021]
  - For $h > p$: VARs biased, but LPs are not (under certain conditions) [Jordà, Singh & Taylor, 2020]

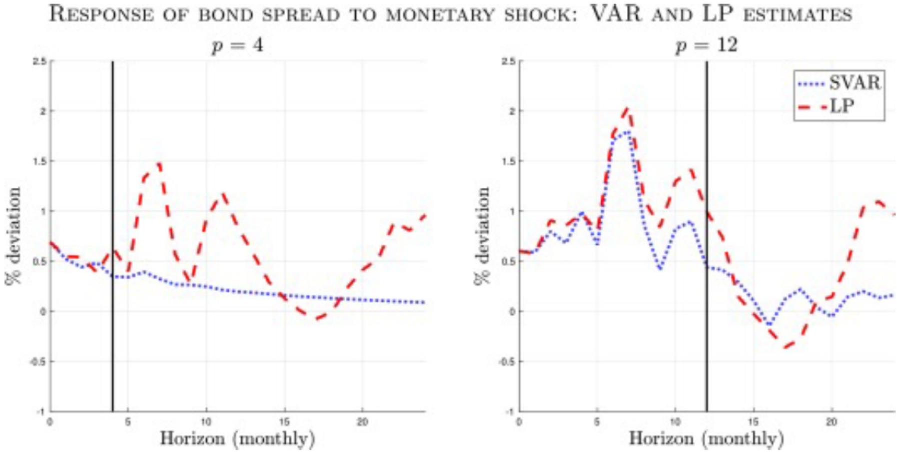
Example I of VAR vs. LP Bias

Based on MA(24) Model



Example II of VAR vs. LP Bias

Plagborg-Møller & Wolf (2021)



Local Projections in Practice

So You Want To Run a Local Projection? What Next?

$$\Delta^h y_{t+h} = \alpha_h + \beta_h \varepsilon_t + \boldsymbol{\gamma}_h \mathbf{x}_t + v_{t+h}$$

#1 Where do I get my shocks ε from?

#2 Once I have my shocks, how should I use them?

#3 What else should I control $\mathbf{x}$ for?

#4 What standard errors should I use?

#5 How can I assess the contribution of shocks to the variance of outcome variables?

#6 How can I test and account for non-linearities?

#7 How do things change if I decide to use panel data?

#8 What non-OLS estimators can I use?

#1. Fantastic Shocks and Where To Find Them

Where do I get my shocks ε from?

Unbiasedness of LP estimator β_h depends critically on 'exogeneity' of shock ε conditional on observables $\mathbf{x}$—i.e., $\varepsilon_t|\mathbf{x}_t$ must be randomly assigned.

Fortunately, a huge literature has developed a range of shocks (constructed in various ways):

- **Monetary Policy**:
  - US: Romer and Romer (1989, 2004) [Extended by Wieland and Yang (2020)]; Gürkaynak, Sack and Swanson (2005); Gertler and Karadi (2015); Miranda-Agrippino and Ricco (2021); Bauer and Swanson (2023)
  - UK: Cloyne and Hürtgen (2016); Cesa-Bianchi, Thwaites and Vicondoa (2020); Braun, Miranda-Agrippino and Saha (2023)
  - EA: Altavilla, Brugnolini, Gürkaynak, Motto and Ragusa (2019); Jaroncinski and Karadi (2020)
  - RoW: Choi, Willems and Yoo (2023)

  Amongst others

#1. Fantastic Shocks and Where To Find Them

Where do I get my shocks ε from?

Unbiasedness of LP estimator β_h depends critically on 'exogeneity' of shock ε conditional on observables $\mathbf{x}$—i.e., $\varepsilon_t|\mathbf{x}_t$ must be randomly assigned.

- **Tax**: Romer and Romer (2010); Mertens and Ravn (2013, 2014); Cloyne and Surico (2017); Cloyne, Martinez, Mumtaz and Surico (2022); Cloyne, Postel-Vinay and Dimsdale (2023)
- **Government Spending**: Ramey (2011); Owyang, Ramey and Zubairy (2014); Ramey and Zubairy (2018)
- **Macroprudential Policy**: Richter, Schularick and Shim (2019); Fernández-Gallardo, Lloyd and Manuel (2023); Bluwstein and Patozi (2023)
- **General Aggregate Demand or Supply**: Angeletos, Collard and Dellas (2020)
- **Oil Shocks**: Känzig (2021)
- **Credit Shocks**: Gilchrist and Zakrejsek (2012)
- **FX Shocks**: Bippus, Lloyd & Ostry (2023)

#2. Handle With Care: Shock-Handling Instructions

Once I have my shocks, how should I use them?

Once you have your shock ε, you can in principle use it in two ways:

#1 **Use Shock Directly in OLS**: Estimate the following OLS regression:

$$\Delta^h y_{t+h} = \alpha_h + \beta_h \varepsilon_t + \boldsymbol{\gamma}_h \mathbf{x}_t + v_{t+h}$$

Assuming: Shock observed without error!

#2 **Use Shock as Instrument in 2SLS**: Use shock ε as an instrument for z in following IV regression:

$$\Delta^h y_{t+h} = \alpha_h + \beta_h z_t + \boldsymbol{\gamma}_h \mathbf{x}_t + v_{t+h}$$

Requirement: No weak instruments (i.e., first-stage $F > 10$)

In practice, two should be comparable. Preference may be a function of interpretation.

#2. Handle With Care: Shock-Handling Instructions

Once I have my shocks, how should I use them?

Conditions for LP-IV

[Stock and Watson, 2018]

- Suppose ε_t a vector of instruments for z_t and denote $\varepsilon_t^p = \varepsilon_t - \mathcal{P}(\varepsilon_t | \mathbf{w}_t)$ where $\mathbf{w}_t$ collects all controls in the LP
  - #1 **Relevance**: $\mathbb{E}\left[\mathbf{z}_{i,t}^{p} \varepsilon_t^{p'}\right] = \boldsymbol{\alpha}' \neq 0$
  - #2 **Basic Exogeneity**: $\mathbb{E}\left[\mathbf{z}_{-i,t}^{p} \varepsilon_t^{p'}\right] = 0$, $j \neq 1$
  - #3 **Lead-Lag Exogeneity**: $\mathbb{E}\left[\mathbf{z}_{-i,t+h}^{p} \varepsilon_t^{p'}\right] = 0$, $\forall j, h \neq 0$
- Usual IV conditions, except lead-lag exogeneity due to dynamics
- A STATA Illustration: `LPIV_example.do`

#3. Controlling Behaviour

What else should I control $\mathbf{x}$ for?

- **Simple Answer**: control for anything that could be correlated by ε and y

#3. Controlling Behaviour

What else should I control x for?

- **Simple Answer**: control for anything that could be correlated by ε and y
- **But**: common to construct shock in a first-stage regression (e.g., of policy z on x)
  - See: Lloyd and Manuel (2024), “Controls, Not Shocks: Estimating Dynamic Causal Effects in Macroeconomics,” Bank of England Staff Working Paper No. 1079.

#3. Controlling Behaviour

What else should I control x for?

- **Simple Answer**: control for anything that could be correlated by ε and y
- **But**: common to construct shock in a first-stage regression (e.g., of policy z on x)
  - See: Lloyd and Manuel (2024), "Controls, Not Shocks: Estimating Dynamic Causal Effects in Macroeconomics," Bank of England Staff Working Paper No. 1079.
- So what's the story then?
  - If adding controls to improve 'exogeneity', you could be undermining the exogeneity of the shock ε you picked in the first place
  - If 'shock' constructed with controls, use same controls in LP (better still: do it in one step)
  - If you do include controls, you're more reasonably doing it to improve fit ($\uparrow R^2$) and $\downarrow$ s.e.

#4. Error-Free Error Bands

What standard errors should I use?

Why is inference different with LPs? The MA structure of the residuals

- Recall the AR(1) example: $y_t = \rho y_{t-1} + u_t$. By recursive substitution:

$$y_{t+h} = \rho^{h+1} y_{t-1} + u_{t+h} + \rho u_{t+h-1} + \dots + \rho^h u_t$$

- So, in a LP:

$$y_{t+h} = \beta_{h+1} y_{t-1} + v_{t+h}; \quad v_{t+h} = u_{t+h} + \rho u_{t+h-1} + \dots + \rho^h u_t$$

In general, we do not know the MA structure!

⇒ Jordà (2005) recommends HAC standard errors (e.g., Newey & West, 1987), but this is an active area of research (see, e.g., ‘significance bands’ by Inoue, Jordà and Kuersteiner, 2023)

#4. Error-Free Error Bands

What standard errors should I use?

Lag augmentation offers a simpler, more elegant solution [Montiel-Olea & Plagborg-Møller, 2021]

The Logic in a Simple Example

- *DGP*: $y_t = \rho y_{t-1} + u_t$, where u_t strictly stationary s.t. $\mathbb{E}[u_t | \{u_s\}_{s \neq t}] = 0$
- *LP*: $y_{t+h} = \beta_h y_t + v_{t+h}$ where $v_{t+h} \sim MA(h)$
- *Plug DGP into LP*: $y_{t+h} = \beta_h u_t + \gamma y_{t-1} + v_{t+h}$
- *FWL Logic*: obtain β_h by regressing $y_{t+h} - \gamma_h y_{t-1}$ on $y_t - \rho y_{t-1}$

$$
\begin{aligned}
\hat{\beta}_h =& \frac{\sum_{t=1}^{T-h} (y_{t+h} - \gamma_h y_{t-1})(y_t - \rho y_{t-1})}{\sum_{t=1}^{T-h} (y_t - \rho y_{t-1})^2} = \frac{\sum_{t=1}^{T-h} (\beta_h u_t + v_{t+h}) u_t}{\sum_{t=1}^{T-h} (y_t - \rho y_{t-1})^2} \\
=& \beta_h + \frac{\sum_{t=1}^{T-h} v_{t+h} u_t}{\sum_{t=1}^{T-h} u_t^2}
\end{aligned}
$$

#4. Error-Free Error Bands

What standard errors should I use?

Lag augmentation offers a simpler, more elegant solution [Montiel-Olea & Plagborg-Møller, 2021]

- Recall:

$$\hat{\beta}_h = \beta_h + \frac{\sum_{t=1}^{T-h} v_{t+h} u_t}{\sum_{t=1}^{T-h} u_t^2} \quad \rightarrow \quad \hat{\sigma}^2(\hat{\beta}_h) = \frac{\sum_{t=1}^{T-h} \hat{v}_{t+h}^2 \hat{u}_t^2}{\left(\sum_{t=1}^{T-h} \hat{u}_t^2\right)^2}$$

- Although $v_{t+h} \sim MA(h)$, note that $v_{t+h}u_t \sim MA(0)$ since for any $s < t$:

$$\begin{aligned}
\mathbb{E}[v_{t+h}u_t v_{s+h}u_s] =& \mathbb{E}[\mathbb{E}[v_{t+h}u_t v_{s+h}u_s | u_{s+1}, u_{s+2}, \ldots]] \\
=& \mathbb{E}[v_{t+h}u_t v_{s+h} \underbrace{\mathbb{E}[u_s | u_{s+1}, u_{s+2}, \ldots]]}_{=0}
\end{aligned}$$

$\Rightarrow$ Do lag-augmented LP with White-corrected errors (no need for Newey-West), *or* wild bootstrap procedure (see paper for more)

#5. Shocking Importance

How can I assess the contribution of shocks to the variance of outcome variables?

- Note that we can always write:

$$y_{t+h} = \hat{\mathbb{E}}_t[y_{t+h}] + \hat{v}_{t+h}$$

- Then R^2 from regression of $\hat{v}_{t+h}$ on $\varepsilon_{t+h}, ..., \varepsilon_t$ measures percent of forecast-error variance explained by shock
- If multiple shocks, can run this exercise shock by shock

[Gorodnichenko and Lee, 2020; Plagborg-Møller and Wolf, 2022]

#6. Impulse Response Heterogeneity

How can I test and account for non-linearities?

An Example From Applied Micro

- Think of observed y as coming from latent mixture:

$$y = (1-\varepsilon)y_0 + \varepsilon y_1 = y_0 + \varepsilon(y_1 - y_0), \quad \varepsilon = 0, 1$$

- Assumption:

$$y_i \sim f(\mu_j, \sigma_j), \quad j = 0, 1 \quad \text{unobserved random variables}$$

- We would like: $\mathbb{E}[y_1 - y_0]$, the average treatment effect ATE

#6. Impulse Response Heterogeneity

How can I test and account for non-linearities?

Assume linear model for latent variables: y_j, $j = 0, 1$

- Let $y_j = \mu_j + v_j$, where $\mathbb{E}[v_j] = 0$, $j = 0, 1$ and v_j captures heterogeneity
- Let $v_j = \mathbf{x} - \boldsymbol{\mu}_x)\boldsymbol{\gamma}_j + e_j$ with $\mathbb{E}[e_j] = 0$ and $\mathbb{E}[e_j|\mathbf{x}] = 0$, then:

$$\underbrace{\mathbb{E}_x\left[\mathbb{E}[y_1|\varepsilon = 1, \mathbf{x}] - \mathbb{E}[y_0|\varepsilon = 0, \mathbf{x}]\right]}_{ATE} = \left(\mu_1 + \mathbb{E}_x\left[\mathbb{E}[\mathbf{x} - \boldsymbol{\mu}_x|\varepsilon = 1]\right]\boldsymbol{\gamma}_1\right)$$

$$- \left(\mu_0 + \mathbb{E}_x\left[\mathbb{E}[\mathbf{x} - \boldsymbol{\mu}_x|\varepsilon = 0]\right]\boldsymbol{\gamma}_0\right)$$

- Add/subtract counterfactual: $\mathbb{E}_x\left[\mathbb{E}[\mathbf{x} - \boldsymbol{\mu}_x|\varepsilon = 1]\right]\boldsymbol{\gamma}_1$:

$$\begin{aligned} ATE =& (\mu_1 - \mu_0) + \mathbb{E}_x\left[\mathbb{E}[\mathbf{x} - \boldsymbol{\mu}_x|\varepsilon = 1]\right](\boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_0) \\ & + \mathbb{E}_x\left[\mathbb{E}[\mathbf{x} - \boldsymbol{\mu}_x|\varepsilon = 1] - \mathbb{E}[\mathbf{x} - \boldsymbol{\mu}_x|\varepsilon = 0]\right]\boldsymbol{\gamma}_0 \end{aligned}$$

#6. Impulse Response Heterogeneity

How can I test and account for non-linearities?

Kitagawa-Oaxaca-Blinder Decomposition Components

$$ATE = \underbrace{(\mu_1 - \mu_0)}_{\text{Direct}} + \underbrace{\mathbb{E}_x\left[\mathbb{E}[\mathbf{x} - \boldsymbol{\mu}_x | \varepsilon = 1]\right](\boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_0)}_{\text{Indirect}}$$
$$+ \underbrace{\mathbb{E}_x\left[\mathbb{E}[\mathbf{x} - \boldsymbol{\mu}_x | \varepsilon = 1] - \mathbb{E}[\mathbf{x} - \boldsymbol{\mu}_x | \varepsilon = 0]\right]\boldsymbol{\gamma}_0}_{\text{Composition}}$$

- **Direct**: ATE under random assignment
- **Indirect**: Treatment spillovers on covariates
- **Composition**: Failure of random assignment? Small-sample bias

#6. Impulse Response Heterogeneity

How can I test and account for non-linearities?

Null Hypotheses in Linear Case

$$y_i = \mu_0 + (\mathbf{x}_i - \overline{\mathbf{x}}_0)\,\boldsymbol{\gamma}_0 + \varepsilon_i\,[\beta + (\mathbf{x}_i - \overline{\mathbf{x}}_1)\,\boldsymbol{\theta}] + \omega_i$$

Noting: $\beta = \mu_1 - \mu_0$, $\boldsymbol{\theta} = \boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_0$ and $\omega_i = e_{0,i} + \varepsilon_i\,(e_{1,i} - e_{0,i})$

- H_0: $\beta = 0$ null of no *direct* treatment effect
- H_0: $\boldsymbol{\theta} = 0$ null of no *indirect* treatment effect
- H_0: $\mathbb{E}[\mathbf{x}|\varepsilon = 1) - \mathbb{E}[\mathbf{x}|\varepsilon = 0] = 0$ null of no *composition* effect
- H_0: $\boldsymbol{\gamma}_0 = 0$ null of no *random assignment* (hence no composition effect possible)

#6. Impulse Response Heterogeneity

How can I test and account for non-linearities?

What does this mean for local projections?

- Let $\mathbf{y}_t = [y_t, y_{t+1}, ..., y_{t+H}]$ and $\mathbf{y}$ denote associated random variable
- Assume conditional-mean independence, let $\mathbb{E}[y_s] = \mu_s$ for $s \in \{0, 1\}$, wlog $y_s = \mu_s + v_s$
- Under linearity: $v_s = (\mathbf{x} - \boldsymbol{\mu}_x)\mathbf{b}_s + e_s$, then:

$$\mathbb{E}[y_s|\mathbf{x}] = \mu_s, \quad \mathbb{E}[v_s] = 0, \quad \mathbb{E}[e_s|\mathbf{x}] = 0, \quad s \in \{0, 1\}$$

- Hence

$$y_{t+h} = \underbrace{\mu_0^h + (\mathbf{x}_t - \overline{\mathbf{x}})\gamma_0^h + \varepsilon_t \beta^h}_{\text{usual local projection}} + \underbrace{\varepsilon_t (\mathbf{x}_t - \overline{\mathbf{x}})\boldsymbol{\theta}^h}_{\text{Kitagawa term}} + \omega_{t+h}$$

for $h = 0, 1, ..., H$, $t = h, ..., T$

#6. Impulse Response Heterogeneity

How can I test and account for non-linearities?

Kitagawa-Oaxaca-Blinder Decomposition Components

$$y_{t+h} = \underbrace{\mu_0^h + (\mathbf{x}_t - \overline{\mathbf{x}})\gamma_0^h + \varepsilon_t \beta^h}_{\text{usual local projection}} + \underbrace{\varepsilon_t (\mathbf{x}_t - \overline{\mathbf{x}})\boldsymbol{\theta}^h}_{\text{Kitagawa term}} + \omega_{t+h}$$

for $h = 0, 1, ..., H$, $t = h, ..., T$

- **Direct Effect**: $\hat{\mu}_1^h - \hat{\mu}_0^h = \hat{\beta}^h$
- **Indirect Effect**: $(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}})(\hat{\gamma}_1^h - \hat{\gamma}_0^h) = (\overline{\mathbf{x}}_1 - \overline{\mathbf{x}})\hat{\boldsymbol{\theta}}^h$
- **Composition Effect**: $(\overline{\mathbf{x}}_1 - \overline{\mathbf{x}}_0)\hat{\gamma}_0^h$

Ergodicity: needed to ensure $\overline{\mathbf{x}} \to \boldsymbol{\mu}_x$

#6. Impulse Response Heterogeneity

How can I test and account for non-linearities?

Implications: State-Dependence Suppose $\mathbf{x} = \mathbf{x}^*$ then total response is:

$$
\begin{aligned}
\mathbb{E}[y_1|\mathbf{x}^*, \varepsilon = \delta] - \mathbb{E}[y_0|\mathbf{x}^*, \varepsilon = 0] \\
&= \delta\mu_1 + \delta\left[\mathbf{x}^* - \mathbb{E}[\mathbf{x}]\right]\boldsymbol{\gamma}_1 - \left[\mu_0 + (\mathbf{x}^* - \mathbb{E}[\mathbf{x}])\boldsymbol{\gamma}_0\right) \\
&= \delta\beta + \delta(\mathbf{x}^* - \mathbb{E}[\mathbf{x}])\boldsymbol{\theta}
\end{aligned}
$$

- Dependence on $\mathbf{x}^*$ is only partial equilibrium
- Need identification (instruments) for $\mathbf{x}$
- Usual single-variable stratification omits other terms in $\mathbf{x} \rightarrow$ bias

#6. Impulse Response Heterogeneity

How can I test and account for non-linearities?

State-Dependence Example: Monetary Policy

- How effective was monetary policy in...
  - #1 November 1987 (Post-Stock Market Crash)
    - Stocks 23% lower by end-October
    - Fed lowered funds rate by $50bo$
  - #2 February 1996 (Middle of a Long Expansion)
    - Middle of stable funds rate
- **Idea**: two different scenarios, but similar policy paths $\rightarrow$ differences not due to policy

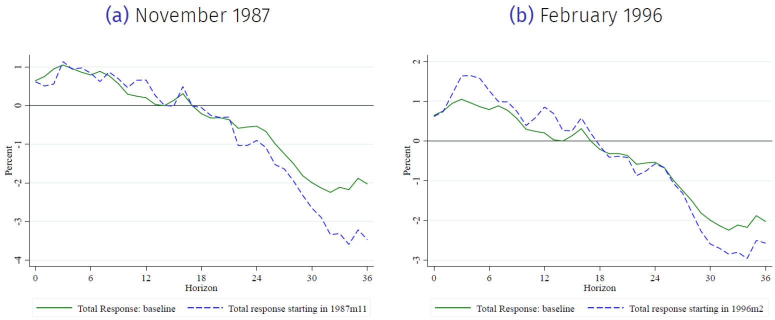
#6. Impulse Response Heterogeneity

How can I test and account for non-linearities?

State-Dependence Example: Monetary Policy

Funds rate paths nearly identical

Federal funds rate



Baseline = average over sample

#6. Impulse Response Heterogeneity

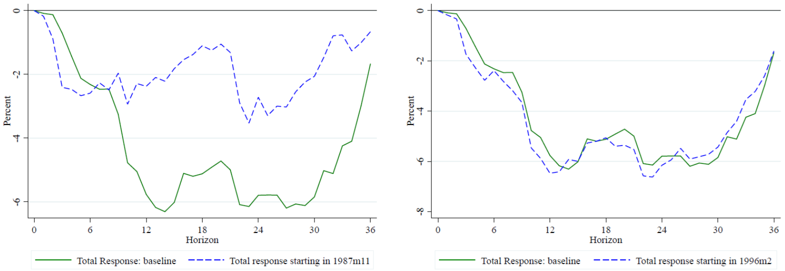
How can I test and account for non-linearities?

State-Dependence Example: Monetary Policy

Policy unable to boost activity post-1987 crash

Industrial production

(a) November 1987 (b) February 1996



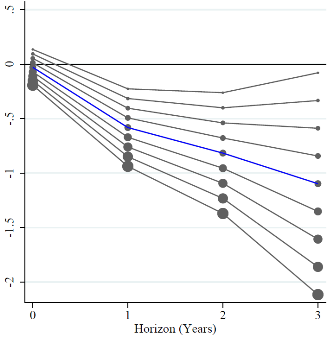
Baseline = average over sample

#6. Impulse Response Heterogeneity

How can I test and account for non-linearities?

State-Dependence Example: Fiscal Policy

GDP response to fiscal policy varies with monetary stance



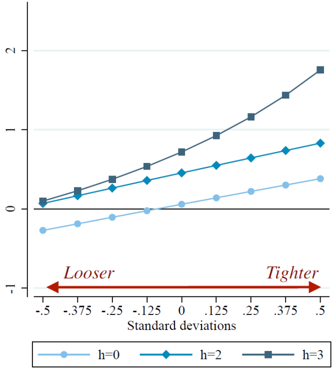
Cloyne, Jordà and Taylor (2023)

#6. Impulse Response Heterogeneity

How can I test and account for non-linearities?

State-Dependence Example: Fiscal Policy

Variation in the multiplier by horizon and stance



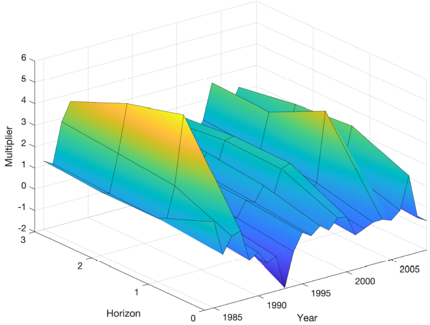
Cloyne, Jordà and Taylor (2023)

#6. Impulse Response Heterogeneity

How can I test and account for non-linearities?

State-Dependence Example: Fiscal Policy

Time-varying estimates of the multiplier



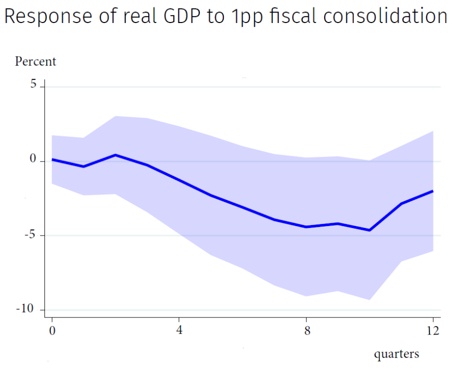
Cloyne, Jordà and Taylor (2023)

#6. Impulse Response Heterogeneity

How can I test and account for non-linearities?

A STATA Illustration: Usual LPIV

`kob_example.do`



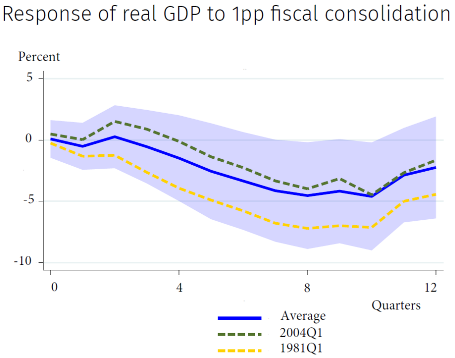
90% error bands

#6. Impulse Response Heterogeneity

How can I test and account for non-linearities?

A STATA Illustration: Choosing Two Dates

`kob_example.do`



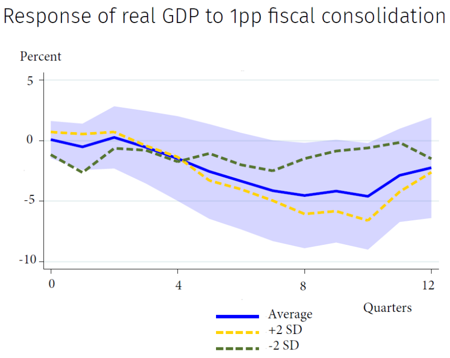
90% error bands

#6. Impulse Response Heterogeneity

How can I test and account for non-linearities?

A STATA Illustration: Monetary Offset

`kob_example.do`



90% error bands

#7. LPs in Panel Data

How do things change if I decide to use panel data?

$$y_{i,t+h} = \alpha_i + \delta_t + \varepsilon_{i,t}\beta_h + \mathbf{x}_{i,t}\boldsymbol{\gamma}_h + v_{i,t+h}, \quad i = 1, ..., N, \ t = 1, ..., T$$

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

A STATA Illustration: `LP_example_panel.do`

#7. LPs in Panel Data

How do things change if I decide to use panel data?

Remarks: Usual panel data issues, plus:

- LP is costly in short panels (lose time dimension)
- But cross-section brings more power
- Incidental parameter issues (fixed effects)
  - Beware of high autocorrelation and low T [Alvarez and Arellano, 2003]
  - May need Arellano-Bond or similar estimator
- Inference
  - N, T large: Two-way clustering helps MA(h) and heteroskedasticity
  - N large, T small: Cluster by unit helps with MA(h)
  - T large, N small: Cluster by time helps heteroskedasticity
  - Else Driscoll-Kraay s.e. are like Newey-West s.e. for panel data (`xtscc` in STATA)
  - When clustering with small N, T, may need bootstap (`summclust` and `boottest` in STATA)

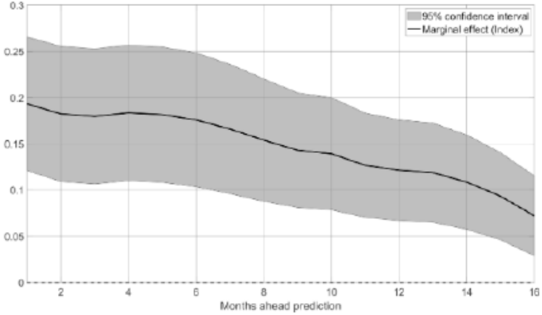
#8. Making Things Funky

What non-OLS estimators can I use?

$$\Delta^h y_{t+h} = \alpha_h + \beta_h \varepsilon_t + \boldsymbol{\gamma}_h \mathbf{x}_t + v_{t+h}$$

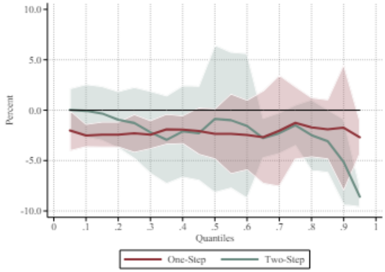
Don't *only* need to use OLS to estimate this!

Binary Events: Response of recession prob. to newspaper index



Quantiles: Response infl. distbn. to mon. pol.

(d) Quantile Response at 4-year Horizon from QR-LP



Relation to VARs*

With Identification

Short-Run Restrictions

Recall the Cholesky decomposition from VARs: $\mathbf{\Sigma} = \mathbf{P}\mathbf{P}'$ with $\mathbf{P}$ lower triangular

- This always exists, is unique, and is trivial to implement but:
  - Different ordering of variables implies different $\mathbf{P}$
  - Implied 0 restrictions need not be correct
  - Just-identification $\Rightarrow$ ordering can't be tested

- **Interpretation**:
  - $y_{(1),t}$ does not contemporaneously depend on others
  - $y_{(2),t}$ only depends on $y_{(1),t}$ contemporaneously
  - $y_{(3),t}$ only depends on $y_{(1),t}, y_{(2),t}$ contemporaneously
  - and so on...

Recursive Identification in LPs

- Let $\mathbf{y}_t$ be an $n \times 1$ vector, and decide causal ordering
- Include contemporaneous values of variables causally-ordered first:

$$y_{j,t+h} = \mu_j^h + \beta_{j,1}^h y_{1,t} + \ldots + \beta_{j,i-1}^h y_{i-1,t} + \beta_{j,i}^h y_{i,t} + \sum_{k=0}^{p} \mathbf{c}_{j,k}^h \mathbf{y}_{t-k} + v_{j,t+h}$$

- Structural LP estimate: $\hat{\mathcal{R}}_{ij}(h) = \hat{\beta}_{j,t}^h$, $h = 0, 1, \ldots, H$, $i, j \in \{1, \ldots, n\}$
- Good idea to order treatment variable $y_{i,t}$ last $\rightarrow$ variation cannot be explained by observables

Long-Run Restrictions

Two-Step Procedure in LPs

Blanchard and Quah (1989) Example: $\mathbf{y}_t = [x_t, u_t]'$ where x_t log real GDP and u_t unemp. rate

#1 Long-Run LP, where H is large

$$x_{t+H} - x_{t-1} = \alpha_H + \boldsymbol{\delta}_{x,H}\mathbf{y}_t + \sum_{k=1}^{p} \mathbf{c}_{x,k}^{H}\mathbf{y}_{t-k} + v_{x,t+H}$$

$\boldsymbol{\delta}_{x,H}$: linear combination that best explains long-run GDP (i.e., supply shock)

#2 Regress:

$$y_{j,t+h} = \mu_h + \beta_{j,h}\left(\hat{\boldsymbol{\delta}}_{x,H}\mathbf{y}_t\right) + \sum_{k=1}^{p} \mathbf{x}_{j,k}^{h}\mathbf{y}_{t-k} + v_{j,t+h}, \quad j = x, u,\ h = 0, 1, ..., H$$

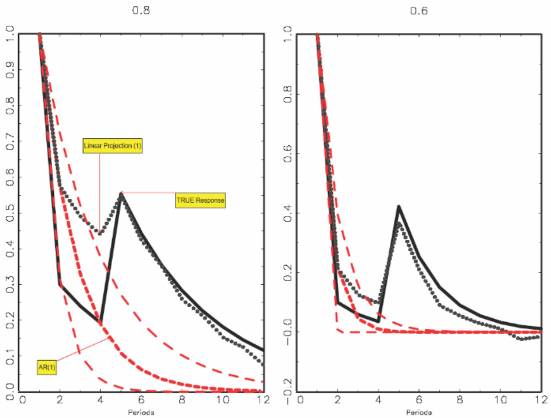
$\beta_{j,h}$ is response of jth variable to supply shock in period h

Little guidance on how to choose H, try different value

Could generalise idea in a number of ways (e.g., medium-run identification)

Appendix

Example: AR(1) vs. Local Projection



$y_t = \rho y_{t-1} + \varepsilon_t - 0.5\varepsilon_{t-1} + 0.4\varepsilon_{t-4}$, $T = 180$, 100 replications $\rho = 0.8, 0.6$

Back