Local Projections

Based on Jordà & Taylor (2025)

Yuxuan He   Linh Nguyen   Chenyang Pan

Department of Economics, UC3M

December 14, 2025

Outline

Part 1
Foundations & Intuition

(Sections 1-2: What & Why)

1. The Local Projection (LP) Approach

The Central Question: How does an exogenous shock s_t affect outcome y over time? (The **Impulse Response Function**).

The LP Definition: A sequence of separate regressions.

The LP Equation

$$y_{t+h} = \alpha_h + \beta_h s_t + \gamma_h' x_t + v_{t+h}, \quad h = 0, 1, ..., H$$

- β_h: Directly estimates the response at horizon h.
- **VS. VARs:** Instead of iterating a complex system forward, we estimate each horizon *separately*.

Why LPs? (Advantages)

- **Robust:** Single-equation method; no full system dynamics needed.
- **Flexible:** Handles nonlinearities and panel data easily.

Part 2
Estimation, Bias, Multipliers & Smoothing

(Sections 3-5: Best Practices)

2. Specification Choice: The Bias Problem

Large Samples: LPs and VARs are equivalent.
Small Samples: LPs suffer from bias if data is persistent ($\rho \approx 1$).

- **Level Specification:** Regressing y_{t+h} on y_{t-1}.
- **Visual Evidence (Fig 1):**
  - **Solid Green:** True Persistence.
  - **Dashed Blue:** LP in Levels.
- **Result:** Severe downward bias (underestimates persistence).

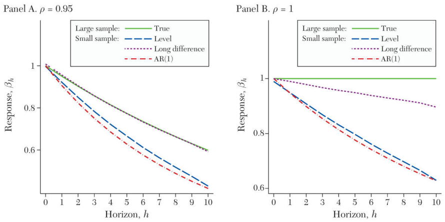

Figure: Small-Sample Bias ($\rho = 0.95, T = 100$)

Solving the Bias: The Long-Difference Specification

Best Practice: Use the **Long-Difference** specification.

Long-Difference LP

$$y_{t+h} - y_{t-1} = \alpha_h + \beta_h s_t + \gamma_h (y_{t-1} - y_{t-2}) + v_{t+h}$$

- Regress the **cumulative change** $(y_{t+h} - y_{t-1})$ on the shock.
- **Visual Evidence:** Look at the **Dotted Red Line**.
- **Result:** It hugs the Truth (Green) almost perfectly.

(Refer to Figure 1 again)

Takeaway

For persistent data, always prefer Long-Differences to eliminate bias.

3. Interpretation: Calculating Multipliers

The Context: Initial shocks often trigger cascades of policy changes (e.g., dynamic tax cuts).

- **Standard IRF:** Shows total effect, but ignores the cost.
- **The Multiplier $m(h)$:** The "Bang for the Buck".

$$m(h) = \frac{\text{Cumulative Outcome Response}}{\text{Cumulative Policy Response}} = \frac{\sum \mathcal{R}_{sy}(h)}{\sum \mathcal{R}_{ss}(h)}$$

Estimation Problem: Dividing two estimated coefficients is "messy for inference".

Solution (LP-IV)

$$y^c_{t,h} = m(h) s^c_{t,h} + \epsilon_{t+h}$$

- $y^c_{t,h}, s^c_{t,h}$: Cumulative outcome and policy variables.
- **Instrument:** Use the shock Z_t to instrument for the cumulative policy $s^c_{t,h}$.
- **Benefit:** Avoids the "messy inference" of dividing two coefficients.

Example: Fiscal Multipliers (Fig 2)

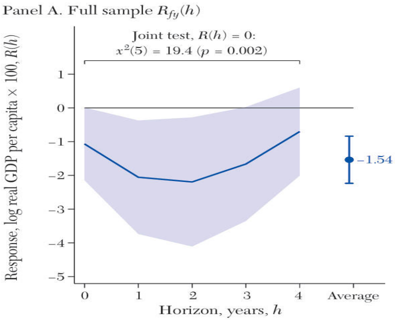

Figure: Cumulative Response vs. Multiplier

Left Panel (IRF)
Output drops, but magnitude is hard to interpret.

Right Panel (Multiplier)
Stable at ≈ -1.
"1% consolidation reduces output by 1%."

4. The Bias-Variance Trade-off & Smoothing

The Trade-off:

- **VARs:** Low Variance, potentially High Bias.
- **LPs:** Low Bias, but High Variance ("Choppy" or "Noisy" responses).

Solution: Impose smoothness using **Gaussian Basis Functions (GBF)**.

GBF Approximation

$$\phi(h;\Theta) = a \cdot \exp\left(-\frac{(h-b)^2}{c^2}\right)$$

The Efficiency Gain:

- Instead of estimating **20 separate coefficients**...
- We estimate just **3 parameters** (a, b, c).

Visualizing Gaussian Basis Functions (GBF)

3 Intuitive Parameters:

- $\mathbf{a}$: Height of the peak.
- $\mathbf{b}$: Timing of the peak.
- $\mathbf{c}$: Width / Persistence.

Result: fitting this curve to noisy LP estimates drastically reduces standard errors (Low Variance) while retaining robustness (Low Bias).

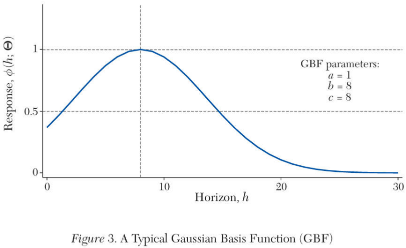

Figure: Typical Gaussian Basis Function

Summary: Best Practices (Sec 1-5)

To estimate Impulse Responses effectively with LPs:

1. **Method:** Use LPs for flexibility and robustness.
2. **Specification:** Use **Long Differences** to fix small-sample bias.
3. **Interpretation:** Use **Multipliers** for dynamic policy analysis.
4. **Refinement:** Use **Smoothing** (GBF) to gain efficiency.

Coming Next...

Now that we have the estimates, are they statistically significant? Is the shock endogenous?

Next: Inference and Identification.

Part 3
Inference, Robustness, Identification

("How to Use LP Safely")

Pointwise Inference

Residual Structure of LPs: Consider an AR(1) data-generating process

$$w_t = \phi w_{t-1} + u_t.$$

A local projection of horizon h implies:

$$w_{t+h} = \phi^{h+1} w_{t-1} + v_{t+h},$$

where

$$v_{t+h} = u_{t+h} + \phi u_{t+h-1} + \cdots + \phi^h u_t.$$

- The residual v_{t+h} is an MA(h) process.
- LP inference must account for this serial dependence.

Inference and Estimation Challenges

1. Serial Correlation & Standard Errors

- The error term $\xi_t(\phi, h)$ follows a **Moving Average (MA)** structure.
- This structure violates standard OLS assumptions regarding independent errors.
- **Implication:** Jordà (2005) recommends using **HAC-robust standard errors** to ensure valid hypothesis testing.

2. Near-Unit Root Asymptotics ($\phi \to 1$)

- The estimator $\hat{\beta}(\phi, h)$ follows a near-unit root distribution rather than a standard normal distribution.
- **Bias:** This results in a well-known **downward bias** in impulse response estimates (e.g., Pesavento & Rossi, 2006).
- **Validity:** Inference based on standard normal critical values is **not valid uniformly** over all $\phi \in [-1, 1]$, though it remains asymptotically normal strictly inside the stationary region.

Lag-Augmented Local Projections: Theory

Augmented regression:

$$y_{t+h} = \alpha_h + \beta_h s_t + \gamma_h' x_t + \delta_h y_{t-1} + \varepsilon_{t+h}.$$

Results (Montiel Olea & Plagborg-Møller, 2021):

- Adding y_{t-1} absorbs the VAR truncation error.
- Removes the first-order bias in $\hat{\beta}_h$.
- Ensures:
$$\sqrt{T}(\hat{\beta}_h - \beta_h) \xrightarrow{d} N(0, V_h) \quad \text{uniformly in } h.$$
- Works even when y_t is nonstationary.

Bias Under Model Misspecification

Montiel Olea et al. (2024) showed the robustness of LP to misspecification. VAR estimation solves:

$$y_t = By_{t-1} + \cdots + Cy_{t-p} + e_t.$$

If incorrect lag length or omitted MA structure:

$$\hat{\beta}_h^{VAR} = \psi_h + \text{bias}(h).$$

LPs instead approximate:

$$y_{t+h} = \psi_h s_t + \text{error}.$$

Properties:

- LP bias increases slowly with h.
- VAR bias increases sharply under DGP misspecification.
- LP confidence intervals exhibit uniform robustness.

Joint Inference

The reason for that is the dependence across LP horizons. Vector of IRFs:

$$\beta = (\beta_0, \beta_1, \ldots, \beta_H)'.$$

Sampling distribution:

$$\sqrt{T}(\hat{\beta} - \beta) \xrightarrow{d} N(0, \Omega),$$

with

$$\Omega_{hh'} = \mathsf{Cov}(\hat{\beta}_h, \hat{\beta}_{h'}).$$

Implication:

- Pointwise confidence intervals ignore cross-horizon dependence.
- Cannot answer questions such as "Is the entire IRF positive?".

1. System Setup & Identification

Goal: Joint Estimation

We estimate Local Projections as a system to handle hypothesis tests across multiple horizons ($h = 0, \ldots, H$).

System Variables:

- **Outcomes ($\boldsymbol{y}_t$):** Vector of future outcomes.
- **Intervention ($\boldsymbol{S}_t$):** The shock/treatment variable.
- **Instruments ($\boldsymbol{Z}_t$):** External instruments for identification.

Key Assumptions:

1. **Relevance:** Instruments must predict the intervention.
2. **Exogeneity:** Instruments must be uncorrelated with future error terms (Lead-lag exogeneity).

Population Moment Condition

$$E[\boldsymbol{Z}_t'(\boldsymbol{y}_t(H) - \boldsymbol{S}_t\boldsymbol{\beta})] = \mathbf{0}$$

The GMM Estimator

Minimizing the GMM objective function yields the closed-form estimator:

$$\hat{\beta} = \left(\sum \boldsymbol{S}'\boldsymbol{Z}\hat{\boldsymbol{\Lambda}}^{-1}\boldsymbol{Z}'\boldsymbol{S}\right)^{-1}\left(\sum \boldsymbol{S}'\boldsymbol{Z}\hat{\boldsymbol{\Lambda}}^{-1}\boldsymbol{Z}'\boldsymbol{y}\right)$$

HAC Correction ($\hat{\Lambda}$)

- We cannot use standard weights due to serial correlation.
- **Solution:** Use a HAC-robust weighting matrix (e.g., Newey-West Bartlett kernel).

Inference Properties

- The estimator is **consistent** and **asymptotically normal**.
- Standard errors are robust to heteroskedasticity and autocorrelation.

Simultaneous Inference: The Sup-t Method

1. The Problem: Joint Hypotheses

- Practitioners often test hypotheses over a *range* of horizons (e.g., "Is the effect significant between Year 1 and Year 3?").
- Standard point-wise confidence intervals (using 1.96) **do not** provide the correct coverage for these joint tests.
- **Result:** Point-wise bands are too narrow, leading to false positives when assessing shapes or paths.

2. The Solution: Sup-t Confidence Bands

- **Approach:** Construct simultaneous confidence bands that cover the entire impulse response path jointly.
- **Method:** Use the **Sup-t procedure** (Montiel Olea & Plagborg-Møller, 2019).
- **Key Difference:** Instead of a fixed critical value (like 1.96), we compute a specific value c based on the **maximum t-statistic** across all horizons.
- This ensures the bands are valid for the entire trajectory of the response.

Simultaneous Inference: The Sup-t Method

Objective: Construct simultaneous confidence bands for parameters β_h.

- **Step 1: Simulation**
  Draw M i.i.d. normal vectors from the estimated asymptotic distribution:
  $$\hat{V}^{(m)} \sim \mathcal{N}_H(\mathbf{0}_H, \hat{\Omega}_\beta), \quad m = 1, \dots, M.$$
- **Step 2: Critical Value**
  Define $\hat{q}_{1-\alpha}$ as the empirical $1-\alpha$ quantile of the maximum standardized deviation across m:
  $$\max_h \left| \hat{\sigma}_h^{-1} \hat{V}_h^{(m)} \right| \quad \text{across } m = 1, \dots, M.$$
- **Step 3: Confidence Band Construction**
  Construct the simultaneous band using the critical value $\hat{q}_{1-\alpha}$:
  $$\hat{\mathcal{B}}(\hat{q}_{1-\alpha}) = \bigcap_{h=0}^{H} \left[\hat{\beta}_h - \hat{\sigma}_h \hat{q}_{1-\alpha},\ \hat{\beta}_h + \hat{\sigma}_h \hat{q}_{1-\alpha} \right].$$

Note: Step 2 can be adapted for bootstrap or Bayesian methods (Montiel Olea and Plagborg-Møller, 2019).

Causality: Identification via Selection on Observables

1. The Identification Problem

- Local Projections (LP) by themselves only measure correlations.
- **Goal:** Uncover causal relations (counterfactual difference in mean outcomes).
- **Challenge:** The intervention s_t is rarely randomly assigned in macroeconomic data; it is often determined endogenously.

2. Solution: Selection on Observables (LP-OLS)

- **Assumption:** Conditional on a set of control variables $\mathbf{x}_t$, the variation in the shock s_t is "as good as random."
- **Implementation:** Simply include the controls $\mathbf{x}_t$ in the LP regression:

$$y_{t+h} = \alpha_h + \beta_h s_t + \boldsymbol{\gamma}_h' \mathbf{x}_t + v_{t+h}$$

- **Relation to VARs:** This approach is asymptotically equivalent to a Cholesky-identified VAR if $\mathbf{x}_t$ includes the appropriate contemporaneous variables (Plagborg-Møller & Wolf, 2021).

Inverse Propensity Scores: LP-IPW and LP-IPWRA

1. The Challenge: Non-linearity

- Including controls $\mathbf{x}_t$ linearly in OLS assumes a specific functional form.
- **Problem:** Covariates may affect the intervention s_t **non-linearly**.
- **Solution:** Use **propensity scores** to reweight sample averages, a method dating back to Horvitz and Thompson (1952).

2. LP-IPW Estimator

- Define propensity score: $p_t = \Pr(s_t = 1|\mathbf{x}_t)$.
- Reweight the outcomes to estimate the impulse response $\hat{\mathcal{R}}_{sy}(h)$:

$$\hat{\mathcal{R}}_{sy}(h) = \frac{\sum y_{t+h} s_t}{\sum p_t} - \frac{\sum y_{t+h}(1 - s_t)}{\sum (1 - p_t)}$$

(See Equation 41)

3. LP-IPWRA (Doubly Robust)

- **Refinement:** Combine weighting with regression adjustment.
- **Method:** Include controls $\mathbf{x}_t$ linearly *and* use Weighted Least Squares (WLS) based on propensity scores.
- **Benefit:** This is a **doubly robust** estimator—consistent if *either* the propensity model *or* the outcome regression model is correct.

Instrumental Variables: LP-IV Identification

1. The Method

- **Intuition:** When the intervention s_t is endogenous, use an external instrument z_t to isolate exogenous variation.
- Referred to as **LP-IV** (Jordà, Schularick, Taylor, 2015).

2. Assumptions

- **Relevance:** The instrument must be correlated with the intervention ($E[z_t s_t] \neq 0$).
- **Exogeneity (Lead-Lag):** The instrument must be uncorrelated with future residuals ($E[z_t v_{t+h}] = 0$).
- *Note: This differs slightly from static settings (Stock & Watson, 2018).*

3. Key Advantage over VARs

- **Invertibility:** VARs often fail when structural shocks cannot be recovered from current/lagged variables ("non-invertibility").
- **LP-IV Solution:** Identification is achieved **even in non-invertible settings** (Plagborg-Møller & Wolf, 2022).

Indirect Inference: Impulse Response Matching

1. The Concept: Linking Structure to Data

- **Goal:** Estimate structural parameters $\boldsymbol{\theta}$ (e.g., from a DSGE model) that are hard to estimate directly.
- **Approach:** Express them as functions of "auxiliary" parameters $\boldsymbol{\pi}$ which are easy to estimate (e.g., LP impulse responses).
- **Mapping:** $\boldsymbol{\theta} = g(\boldsymbol{\pi})$.

2. The Estimator: Minimum Distance

We estimate $\boldsymbol{\theta}$ by minimizing the distance between the structural parameters and the estimated auxiliary parameters $\hat{\boldsymbol{\pi}}$:

$$\min_{\boldsymbol{\theta}} (\boldsymbol{\theta} - g(\hat{\boldsymbol{\pi}}))' \mathbf{W}_T (\boldsymbol{\theta} - g(\hat{\boldsymbol{\pi}}))$$

- Using the optimal weighting matrix $\mathbf{W}_T = \boldsymbol{\Omega}_{\pi}^{-1}$, this yields asymptotically normal estimates.
- **Application:** Widely used to estimate DSGE models by matching impulse responses.

Part 4
Advanced Applications
(Sections 10-13: Counterfactuals, Decompositions, Nonlinearities & Panels)

1. Counterfactual Paths

The impulse response before reflects the effect of the initial shock on the unemployment rate, as well as the effect of the shock on the monetary policy variable (say, the federal funds rate) over time, and how it feeds back into the unemployment response. Thus, we may ask, what would happen to the unemployment rate response if the policy path itself were to deviate from its usual pattern?

Deriving counterfactual response of the funds rate

Denote with β_u and β_r the $H \times 1$ response of the unemployment rate and the funds rate, respectively, to a monetary shock. Hence, we may write

$$\begin{pmatrix} \hat{\beta}_u \\ \hat{\beta}_r \end{pmatrix} \to \mathcal{N}\left(\begin{pmatrix} \beta_u \\ \beta_r \end{pmatrix}; \begin{pmatrix} \Omega_{uu} & \Omega_{ur} \\ \Omega_{ru} & \Omega_{rr} \end{pmatrix} \right). \tag{1}$$

Denote by β_r^c a counterfactual response of the funds rate. Based on the rules of the multivariate normal distribution, we can then calculate the unemployment rate response conditional on β_r^c as follows,

$$\beta_u^c = \hat{\beta}_u + \Omega_{ur}\Omega_{rr}^{-1}(\beta_r^c - \hat{\beta}_r), \tag{2}$$

$$\Omega_{uu}^c = \Omega_{uu} - \Omega_{ur}\Omega_{rr}^{-1}\Omega_{ru}. \tag{3}$$

Example

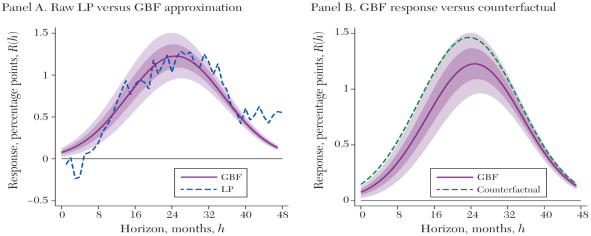

Figure 8. The Response of the Unemployment Rate to a Counterfactual Funds Rate Response

- Panel A shows the response of the unemployment rate estimated using LP-IV and using a GBF approximation. Panel B considers a counterfactual path for the response of the federal funds rate.
- The figure repeats the original GBF response estimate of the unemployment rate (as a solid line) along with its counterfactual response (as a dashed line). As expected, the counterfactual experiment results in the unemployment rate being higher earlier on, and peaking slightly sooner, before returning back to 0

.

2. State-Dependent Responses: A Decomposition

- **Why is State-Dependent Responses needed?**
  When impulse responses are state dependent, estimating a traditional local projection by conditioning on past information without also conditioning on the state will mix up the state-specific responses to yield only an overall average response. The correct approach is to condition on the state as well, and to estimate a state-dependent LP. In general, this will require the full set of interactions of the state variable with all controls for past information.
- **Construction** Conditional on today's state, LPs directly estimate the average response across all possible trajectories that the economy may follow in the future, including possible future shifts in the state, given today's state and conditional on controls.

Example

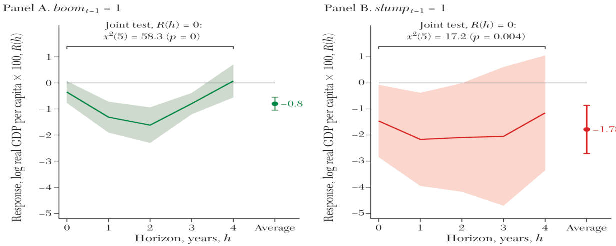

Figure 9. State-dependent Cumulative Fiscal Impulse Response $\mathcal{R}_{yf}(h)$

- The stratification variable D_t takes the value 1 in a boom (or 0 in a slump) respectively. Critically, a key assumption is that consolidations are not influenced by whether the economy is in a boom or a slump.
- The figure shows that fiscal consolidations are contractionary over horizon years 0 to 4, in both split samples. Tests of both the average response and the joint test of nonzero response indicate that the differences are statistically significant. However, the output response is much larger when fiscal consolidations are implemented during slumps, as compared to booms.

Heterogeneity

Abandoning linearity usually comes at a steep cost in complexity, at least when working with VARs. However, since LPs are a single-equation method, these costs tend to be lower. In fact, a great degree of heterogeneity can be achieved with specifications that remain linear in parameters and hence easy to estimate with standard methods. In this section we rely on recent work by Cloyne, Jordà, and Taylor (2023) to explain some of these extensions and their interpretation.

Construction

Consider a departure from the binary example discussed earlier, where the economy can be either in a boom or a slump. Instead, think of the economy as being in a continuum of states. For example, the state of the economy is determined by the vector x_t of controls. We may therefore be interested in comparing the responses resulting from moving from, say $x_t = x_0$ (such as, for example, $x_0 = \bar{x}$), where as before, x_t denotes lags of the outcome, the intervention, and other exogenous and predetermined variables. The state of interest is some deviation δ_x from this equilibrium state. As before, let s_t denote the policy variable that will be shifted from s_0 to $s_0 + \delta_s$.

Decompostion

The researcher is thus usually interested in evaluating the effectiveness of an intervention in a given state via

$$\mathcal{R}_{sy|x}(h) = E[y_{t+h} \mid s_t = s_0 + \delta_s; x_t = x_0 + \delta_x] - E[y_{t+h} \mid s_t = s_0; x_t = x_0 + \delta_x], \tag{4}$$

where δ_s is the only difference between these two expectations. This response can be further decomposed by adding and subtracting $E[y_{t+h} \mid s_t = s_0 + \delta_s; x_t = x_0]$ and $E[y_{t+h} \mid s_t = s_0; x_t = x_0]$. Simple manipulations allow us to decompose the equation into

$$\begin{aligned} \mathcal{R}_{sy|x}(h) = & \underbrace{E[y_{t+h} \mid s_t = s_0 + \delta_s; x_t = x_0 + \delta_x] - E[y_{t+h} \mid s_t = s_0 + \delta_s; x_t = x_0]}_{\mathcal{R}_{xy|s=s_0+\delta_s}(h)} \\ & - \underbrace{(E[y_{t+h} \mid s_t = s_0; x_t = x_0 + \delta_x] - E[y_{t+h} \mid s_t = s_0; x_t = x_0])}_{\mathcal{R}_{xy|s=s_0}(h)} \\ & + \underbrace{E[y_{t+h} \mid s_t = s_0 + \delta_s; x_t = x_0] - E[y_{t+h} \mid s_t = s_0; x_t = x_0]}_{\mathcal{R}_{sy|x=x_0}(h)}. \end{aligned} \tag{5}$$

Under an usual LP linear specification

- Based on these simple derivations Cloyne, Jordà, and Taylor (2023) propose the following extension to the usual LP linear specification,

$$y_{t+h} = \underbrace{\mu_{0h} + \beta_h(s_t - s_0) + \gamma_h(x_t - x_0)}_{\text{usual local projection}} + \underbrace{\theta_h(s_t - s_0)(x_t - x_0)}_{\text{extension}} \quad (6)$$

  with $h = 0, 1, \ldots, H; t = h, \ldots, T$.

- Going back to the decomposition above, note the terms involving a shift in the state:

$$\mathcal{R}_{xy|s=s_0+\delta_s}(h) - \mathcal{R}_{xy|s=s_0}(h) = \theta_h \delta_s \delta_x,$$

  whereas the term directly related to the policy intervention is the usual impulse response coefficient:

$$\mathcal{R}_{sy|x=x_0}(h) = \beta_h \delta_s,$$

  The sum of the two is the state-dependent response where now clearly the term $\theta_h \delta_s \delta_x$ will attenuate/amplify the original response $\beta_h \delta_s$ depending on the sign of θ_h.

Time-Varying Responses

As long as $\theta_h \neq 0$, then the impulse response will vary depending on the value that x_t takes in relation to x_0. That is, the impulse response is time varying. This article ties the time variation of the responses directly to the state of the economy characterized by the value of x_t at each point in the sample, which may be very useful.

Construction of Example

- For now, we provide a simple simulation exercise to illustrate the main features of the Cloyne, Jordà, and Taylor (2023) approach.
- Assume that there are two exogenous variables of interest and s_t will be the primary intervention of interest, whereas x_t will be a secondary exogenous variable. You can think of it as a secondary intervention, such as when one examines fiscal policy given monetary policy. The DGP is as follows:

$$\begin{cases} s_t = 0.75 s_{t-1} + v_{s,t}, \\ x_t = 0.75 x_{t-1} + v_{x,t}, \\ y_t = 0.75 y_{t-1} + \gamma x_{t-1} + \mathbf{1}\{|s_t| > 1\}(\beta s_t + \theta x_t s_t) + v_{y,t}; \end{cases} \tag{7}$$

  with $v_{i,t} \sim \mathcal{N}(0,1)$ for $i = y, s, x$.
- For the simulation, we set $\gamma = 0.75$ and $\beta = \theta = 0.5$ to keep the simulation simple. We initialize the data with 500 burn-in replications that we disregard and study instead the subsequent 500 observations. Then we estimate the LPs.

Results of the Example

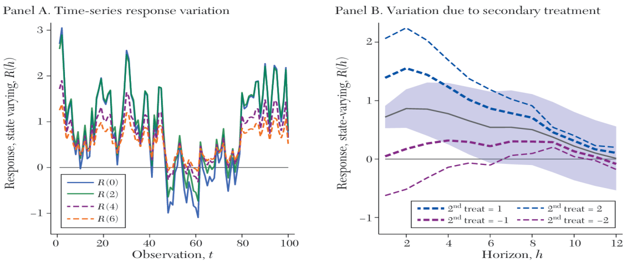

Figure 10. Variation in the Impulse Response Due to Secondary Treatment

- Panel A shows the response coefficients at horizons $0, 2, 4, 6$ for the first 100 observations. The impact is mostly positive for the first 50 observations and mostly negative for the next 25.
- Panel B shows the average impulse response along with the attenuation (in purple)/amplification (in blue) generated by x_t for $x_t = -2, -1, 1, 2$. The response on average begins around 0.75 on impact and by period 12 it has died off to zero. When $x_t = 2$ the response on impact can be as large as 2 whereas when $x_t = -2$ the response on impact can be as low as about -0.75.

3. Nonlinearities

- In this section we review a general observation about nonlinear LPs and highlight a few of the studies from the literature. Illustrating the main issues can be done with a simple motivating example. Hence consider the following nonlinear (separable) local projection

$$y_{t+h} = \mu_h(s_t; x_t; \theta) + v_{t+h}; \quad h = 0, 1, \ldots, H. \tag{8}$$

The corresponding impulse response will be

$$\mathcal{R}_{sy}(h, s_0, \delta; x_t) = \mu_h(s_t = s_0 + \delta; x_t; \theta) - \mu_h(s_t = s_0; x_t; \theta). \tag{9}$$

- With instruments s_t, the moment conditions that we want to exploit are

$$E[z_t'(y_{t+h} - \mu_h(s_t; x_t; \theta))] = 0; \quad h = 0, 1, \ldots, H. \tag{10}$$

Example

Considering the following local projection

$$y_{t+h} = \beta_{1h} s_t + \beta_{2h} s_t^2 + \beta_{3h} s_t x_t + \gamma_h x_t + v_{t+h}; \quad h = 0, 1, \ldots, H. \tag{11}$$

The corresponding response is $\mathcal{R}_{sy}(h, s_0, \delta; x_t) = \beta_1 + \beta_2(\delta^2 + 2s_0\delta) + \beta_3 \delta x_t$. This response is no longer symmetric (since δ^2 is always positive); it also varies with the size of the intervention, δ; it further depends on where the response is benchmarked, $s_t = s_0$; and lastly it will vary depending on the value of the control, x_t.

4. Panel Data

- LPs are well suited to handle this type of data. Estimating a single panel regression is far more convenient than estimating a system of panel regressions, as would be necessary with a VAR. A typical panel data local projection could be specified as:

$$y_{it+h} = \mu_{ih} + \delta_{th} + \beta_h s_{it} + \gamma_h' x_{it} + v_{it+h}, \tag{12}$$

  where the main differences versus earlier specifications are the presence of individual and time-fixed effects, and a sample of $i = 1, \ldots, N$ individual units observed over $t = 1, \ldots, T$ time periods.
- Specification, identification, and analysis using LPs along the lines discussed in previous sections remain largely the same and many of the same methods are directly applicable to panel data. There are, however, two areas worth discussing in more detail: inference and difference-in-differences identification.

Inference

At a basic level and taking a similar approach to that originally proposed in Jordà (2005), one could adjust for heteroskedasticity and autocorrelation using Driscoll–Kraay robust standard errors (Driscoll and Kraay 1998), that is, the direct analog of Newey–West standard errors for panels. The asymptotic justification for this method relies on $T \to \infty$ with N fixed, or N growing at a slower rate than T. A cluster-robust approach could be used in situations where $N \to \infty$ with T fixed to correct for autocorrelation. However, if T is relatively small, a recommended correction for heteroskedasticity is to use the wild cluster bootstrap (see Cameron, Gelbach, and Miller 2008; Canay, Santos, and Shaikh 2021; Roodman et al. 2019).

Difference-in-Differences Estimation

The literature has evolved from some simple settings (such as the well known two-way fixed effects (TWFE) estimator) to include more complex situations. This includes cases where more than one group of individuals receives treatment and this treatment is perhaps not administered at the same time (i.e., it is staggered). Moreover, treatment effects may vary across groups depending on when treatment is received (i.e., they are heterogeneous) and the effects may also change over time after treatment (i.e., they are dynamic). These extensions have generated an extensive new literature, well summarized in the surveys by Roth et al. (2023) and de Chaisemartin and D'Haultfoeuille (2023), for example.

Thank you!