Inference for Local Projections

Atsushi Inoue[1] Òscar Jordà[2]
Guido Kuerseiner [3]

Royal Economic Society Meetings, Belfast, March 26, 2024

Last updated: March 24, 2024

1 Vanderbilt University
2 Federal Reserve Bank of San Francisco; University of California, Davis; and CEPR
3 University of Maryland

The views expressed herein do not necessarily represent those of any of the institutions in the Federal Reserve System

What are local projections?

Description of the problem

Objective

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

- y_{t+h} outcome
- s_t treatment, intervention, etc
- s_0 control level for the intervention (only relevant in NL settings)
- δ size of intervention
- $\boldsymbol{x}_t$ controls. May include lags of y, s, and other variables.
- $h = 0, 1, \ldots, H-1$

Intuition: suppose $s_t \in \{0, 1\}$ and randomly assigned

Difference in means

$$\hat{\mathcal{R}}_{sy}(h) = \frac{\sum_{t=h}^{T} y_{t+h}\, s_t}{\sum_{t=h}^{T} s_t} - \frac{\sum_{t=h}^{T} y_{t+h}\,(1 - s_t)}{\sum_{t=h}^{T}(1 - s_t)} = \overline{y}_1^h - \overline{y}_0^h$$

or in regression form:

$$y_{t+h} = \alpha_h + \beta_h s_t + u_{t+h} \quad \rightarrow \quad \hat{\beta}_h = \overline{y}_1^h - \overline{y}_0^h$$

Local projection

Idea: approximate $E(y_{t+h}|s_t, \boldsymbol{x}_t)$ with $y_{t+h} = m(s_t, \boldsymbol{x}_t) + u_{t+h}$, for example:

$$y_{t+h} = \alpha_h + \beta_h s_t + \boldsymbol{\gamma}_h \boldsymbol{x}_t + u_{t+h}$$

recall:

$$\begin{aligned}\mathcal{R}_{sy}(h, s_0, \delta) &\equiv E(y_{t+h}|s_t = s_0 + \delta; \boldsymbol{x}_t) - E(y_{t+h}|s_t = s_0, \boldsymbol{x}_t) \\ &= m(s_0 + \delta, \boldsymbol{x}_t) - m(s_0, \boldsymbol{x}_t)\end{aligned}$$

hence:

$$\mathcal{R}_{sy}(h, s_0, \delta) = (\alpha_h + \beta_h(s_0 + \delta) + \boldsymbol{\gamma}_h \boldsymbol{x}_t) - (\alpha_h + \beta_h s_0 + \boldsymbol{\gamma}_h \boldsymbol{x}_t) = \beta_h \delta$$

Suppose DGP is a VAR(1)

Consider $\boldsymbol{w}_t = \Phi \boldsymbol{w}_{t-1} + \boldsymbol{u}_t$ (e.g. $\boldsymbol{w}_t = (y_t, s_t, \boldsymbol{x}_t')'$). By recursive substitution:

$$\boldsymbol{w}_{t+h} = \Phi^{h+1} \boldsymbol{w}_{t-1} + \underbrace{\boldsymbol{u}_{t+h} + \Phi \boldsymbol{u}_{t+h-1} + \ldots + \Phi^h \boldsymbol{u}_t}_{\boldsymbol{v}_{t+h} \sim MA(h)}$$

Clearly:

$$\frac{\partial \boldsymbol{w}_{t+h}}{\partial \boldsymbol{u}_t} = \Phi^h = \mathcal{R}(h);$$

Two ways to estimate $\mathcal{R}(h)$:

$$\underbrace{\boldsymbol{w}_t = \Phi \boldsymbol{w}_{t-1} + \boldsymbol{u}_t}_{\text{VAR}} \rightarrow (\hat{\Phi})^h; \text{ or } \underbrace{\boldsymbol{w}_{t+h-1} = B_h \boldsymbol{w}_{t-1} + v_{t+h}}_{\text{local projection}} \rightarrow \hat{B}_h$$

Remarks on local projection inference

- response s.e. direct from regression (no delta method approx.)
- $MA(h)$ residuals (requires correction, no effect on consistency)
- responses correlated across horizons (matters for joint inference)
- truncation of infinite order DGPs (matters for consistency)
- unit or near to unit roots small sample inference distortions
- local projections and VARs imply a bias-variance trade-off (just like IV and OLS). Or do they?

What is the goal of impulse response inference?

1. A natural hypothesis is absence of a treatment effect, i.e.:

$$H_0 : \beta_0 = \ldots = \beta_H = 0$$

But this is not usually reported, instead...

2. report confidence bands: uncertainty about individual coefficients:

$$H_0 : \beta = 0; \ldots; H_0 : \beta_H = 0$$

3. could also care about a subset of coefficients, but very rarely done, e.g.:

$$H_0 : \beta_j = \ldots = \beta_{j+k} > 0 \quad \text{fot} \quad j \geq k, \quad j, k \in \{0, 1, \ldots, H\}$$

Clearly $1 \neq 2$ but 2 often used to say something about 1

Some results in the literature

Not comprehensive

- Jordà (2005): fix MA(h) error structure using Newey-West correction
- Montiel-Olea and Plagborg-Møller (2021): use lag augmentation
  - Then only need heteroscedasticity robust standard errors
  - Uniformly valid for $\phi \in [-1, 1]$
- Lusompa (2023): use GLS to parametrically fix MA(h) error structure
- Bruns and Lütkepohl (2022): lag augmentation + GLS
- Xu (2023): uniform asymptotic theory model's lag order unknown, possibly ∞
  - LPs are semiparametrically efficient as lag $\to \infty$
  - improved inference for unknown heteroscedasticity
- Jordà (2009)/Montiel-Olea and Plagborg-Møller (2019): on joint inference bounds based on Scheffé/sup-t procedure

Simulation-based inference and panels

Bootstrap:

- non-parametric: block wild bootstrap, e.g. Lusompa (2023)
- parametric: Gadea-Rivas and Jordà (2024) based on Paparoditis (1996)

Bayesian:

- Tanaka (2020)
- Ferreira, Miranda-Agrippino and Ricco (2024)

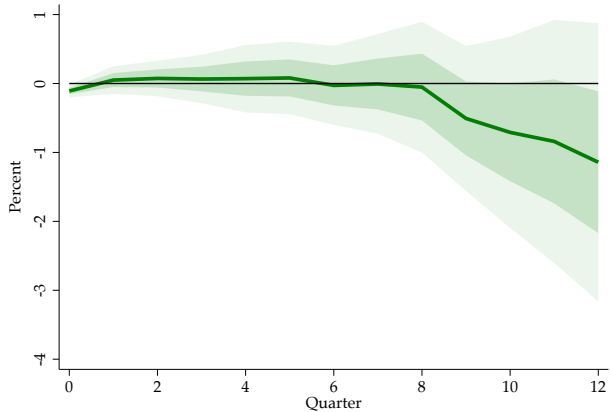
Panels:

- If $T \to \infty$, for N fixed or growing slowly: Driscoll-Kraay (1998)
- If $N \to \infty$, with T fixed, cluster-robust
- If $N \to \infty$, with T small, boostrap cluster-robust

Significance bands

Motivation: typical confidence bands

Response of the CPI price level to a Romer shock: 1969Q1–2007Q4



What is the correct interpretation?

Remarks

- Traditional confidence bands obtained from t-statistic inversion
- Danger: visual seems to indicate response is not significant ...
- ... but joint hypothesis test rejects significance null
- Formal joint test requires estimation of system of horizons... inconvenient?
- Is there a simpler way to assess statistical significance?
- Of course, always important to assess economic significance

An alternative/complement: significance bands

Idea:

- use the LM principle
- no need to estimate the model (under the null, there is no response)
- similar to ACF significance bands: $\pm 1.96/\sqrt{T}$
- bootstrap procedure easy to implement
- significance bands are conservative (Bonferroni bound)

Estimation using local projections

Omitting x_t for simplicity, and/or appeal Frisch-Waugh-Lovell

$$y_{t+h} = s_t\beta_h + u_{t+h} \quad \text{for } h = 0, 1, \ldots, H-1; \quad t = 1, \ldots, T$$

z_t instrumental variable (could be s_t if exogenous)

Stock and Watson (2018) assumptions:

- *Relevance:* $E(s_t z_t) \neq 0$.
- *Lead-lag exogeneity:* $E(u_{t+h} z_t) = 0 \, \forall h$.
- *Exclusion restriction:* $E(y_{t+h} z_t | s_t) = 0$.

Focusing on the IV estimator

$$\sqrt{T-h}(\hat{\beta}_h - \beta_h) = \frac{(T-h)^{-1/2} \sum_1^n z_t \, y_{t+h}}{(T-h)^{-1} \sum_1^n z_t \, s_t}$$

As usual:

$$\frac{1}{T-h} \sum_1^n z_t \, s_t = \hat{\gamma}_{sz} \xrightarrow{p} \gamma_{zs} = E(z_t \, s_t) \qquad \text{(denominator)}$$

At $H_0 : \beta_h = 0$:

$$\frac{1}{(T-h)^{1/2}} \sum_1^n z_t \, y_{t+h} \xrightarrow{d} N(0, \omega) \qquad \text{(numerator)}$$

Putting the LM principle to work

Under $H_0 : \beta_h = 0$ and lead-lag exogeneity

$$\omega = Var\left(\frac{1}{(T-h)^{1/2}}\sum_{1}^{n} z_t\, y_{t+h}\right) \approx \sum_{j=-\infty}^{\infty} E(z_t\, y_{t+h}\, z_{t-j}\, y_{t+h-j})$$

$$H_0 : \beta_h = 0 \;\; \& \;\; \text{lead-lag exogeneity} \quad = \sum_{j=-\infty}^{\infty} E(z_t\, z_{t-j}) E(y_{t+h}\, y_{t+h-j})$$

$$= \sum_{j=-\infty}^{\infty} \gamma_{z,j}\, \gamma_{y,j}$$

$$\sqrt{T-h}(\hat{\beta}_h - 0) \xrightarrow{d} N(0, \sigma^2); \quad \sigma^2 = \frac{\sum_{j=-\infty}^{\infty} \gamma_{z,j}\gamma_{y,j}}{\gamma_{zs}^2} = \frac{\omega}{\gamma_{zs}^2}; \quad \forall h$$

Theoretical significance bands

Based on Dunn (1961) and using a Bonferroni bound, the bands are:

$$\left[\zeta_{\alpha/2H}\frac{\sigma}{\sqrt{T-h}},\ \zeta_{1-\alpha/2H}\frac{\sigma}{\sqrt{T-h}}\right].$$

since:

$$P\left(\bigcap_{h=0}^{H-1}\left\{\zeta_{\alpha/2H}\frac{\sigma}{\sqrt{T-h}} < \hat{\beta}_h < \zeta_{(1-\alpha/2H)}\frac{\sigma}{\sqrt{T-h}}\right\}\right) \geq 1-\alpha$$

When $y = s = z$, under the null $\gamma_{y,0} = \gamma_{s,0} \implies \sigma^2 = 1$
The LP is an estimate of the ACF:

$$\sqrt{T}(\hat{\rho} - 0) \xrightarrow{d} N(0,1) \qquad \Longrightarrow \qquad \text{band: } \pm 1.96/\sqrt{T}$$

Analytical computation in small samples

1. Calculate the sample average of the product $s_t z_t$. Call this $\hat{\gamma}_{sz}$.
2. Construct the auxiliary variable $\eta_t = y_t z_t$ and regress η_t on a constant. The Newey-West estimate of the standard error of the intercept coefficient is an estimate of $s_{\hat{\eta}}$.
3. An estimate of $\sigma/\sqrt{T-h}$, call it $\hat{s}_{\beta_h}$, is therefore:

$$\hat{s}_{\beta_h} = \frac{\hat{s}_{\hat{\eta}}}{\hat{\gamma}_{sz}}$$

4. Construct the significance bands as:

$$\left[\zeta_{\alpha/2H}\hat{s}_{\beta_h},\ \zeta_{1-\alpha/2H}\hat{s}_{\beta_h}\right] \quad \text{where} \quad \int_{\zeta_{\alpha/2}}^{\zeta_{1-\alpha/2}} \phi(x)dx = 1-\alpha$$

Significance bands using the Wild-Block Bootstrap

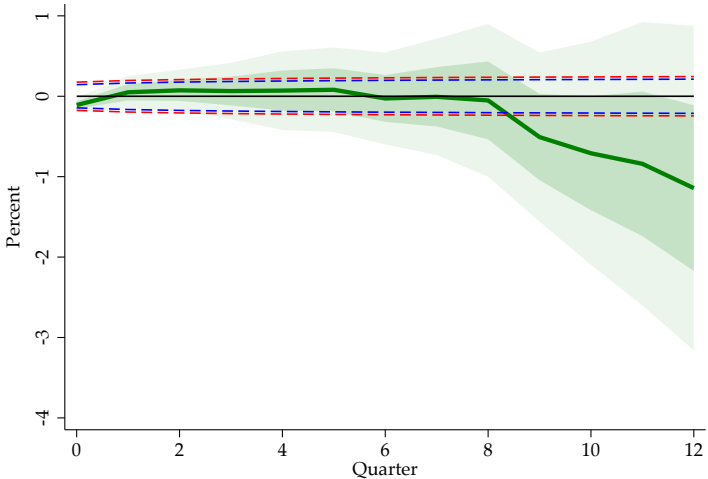
1. Calculate the sample average of $s_t z_t$. Call this $\hat{\gamma}_{sz}$.
2. Construct the auxiliary variable $\eta_t = y_t z_t$ and regress η_t on a constant. The Wild Block bootstrap estimate of the standard error of the intercept coefficient is an estimate of $s_{\hat{\eta}}$.
3. An estimate of $\sigma/\sqrt{T-h}$, call it $\hat{s}^b_{\beta_h}$, is therefore:

$$\hat{s}^b_{\beta_h} = \frac{\hat{s}^b_{\hat{\eta}}}{\hat{\gamma}_{sz}}$$

4. Construct the significance bands as:

$$\left[\hat{\zeta}_{\alpha/2H}\hat{s}^b_{\beta_h},\ \hat{\zeta}_{1-\alpha/2H}\hat{s}^b_{\beta_h}\right] \quad \hat{\zeta}_{\alpha/2H}, \zeta_{1-\alpha/2H} \quad \text{empirical quantiles}$$

Response of the CPI price level to a Romer shock



Red: bootstrap bands; Blue: analytical bands

Takeaways

Why use significance bands?

- s-bands are very easy to obtain. Avoids system estimation
- s-bands are a natural complement to confidence bands
- system estimation needed for formal multiple hypothesis testing

Summary

- Appropriate inference depends on the context
- Many tools now available
- Bias-variance trade-off increasingly tilting toward LPs
- Inference depends on the question posed

Always remember:

$$\text{statistical significance} \not\Longleftrightarrow \text{economic significance}$$