Local Projection

국제금융연구팀 김민 과장

April 30, 2024

통계분석연구회 4회차

Introduction

- "What happens to KRW/USD when the BoK raise policy rates?"
- Easy in theory, hard in empirics
  - "How to identify policy rate (structural) shocks?"
  - "What would be the confounding factors that should be controlled for?"
  - "Would those effects be valid in other settings?"
- These are questions about **dynamic causal effects** often involving two steps
  i. Identifying exogenous variation in policy (the shocks)
  ii. Estimating an impulse response function given the shocks
- Conventional way: Structural VAR
  - identification with short-run, long-run, sign restrictions etc.
- Let us begin with comparison between SVAR and LP

Outline

1. SVAR vs. LP
2. On Inference
3. Counterfactual
4. Nonlinear and Panel LP
5. LP-DiD

SVAR vs. LP (in population)

- Suppose you are interested in dynamic responses of y_t to s_t
  - with observed data $w_t = (\overbrace{r_t}^{n_r \times 1}, \overbrace{s_t}^{1 \times 1}, \overbrace{y_t}^{1 \times 1}, \overbrace{q_t}^{n_q \times 1})'$
- Setting up VAR:

$$w_t = c + \sum_{\ell=1}^{p} A_\ell w_{t-\ell} + u_t, \quad \text{or} \quad A(L) w_t = c + B\eta_t$$

where B is lower triangular with positive diagonal entries

- IRF from SVAR is $\{\theta_h\}_{h \geq 0}$

$$\theta_h \equiv C_{n_r+2,\cdot,h} B_{\cdot,n_r+1}$$

where $A(L)^{-1} = C(L) = \sum_{\ell=0}^{p} C_\ell L^\ell$

SVAR vs. LP (in population)

- IRF from LP is $\{\beta_h\}_{h\geq 0}$ where

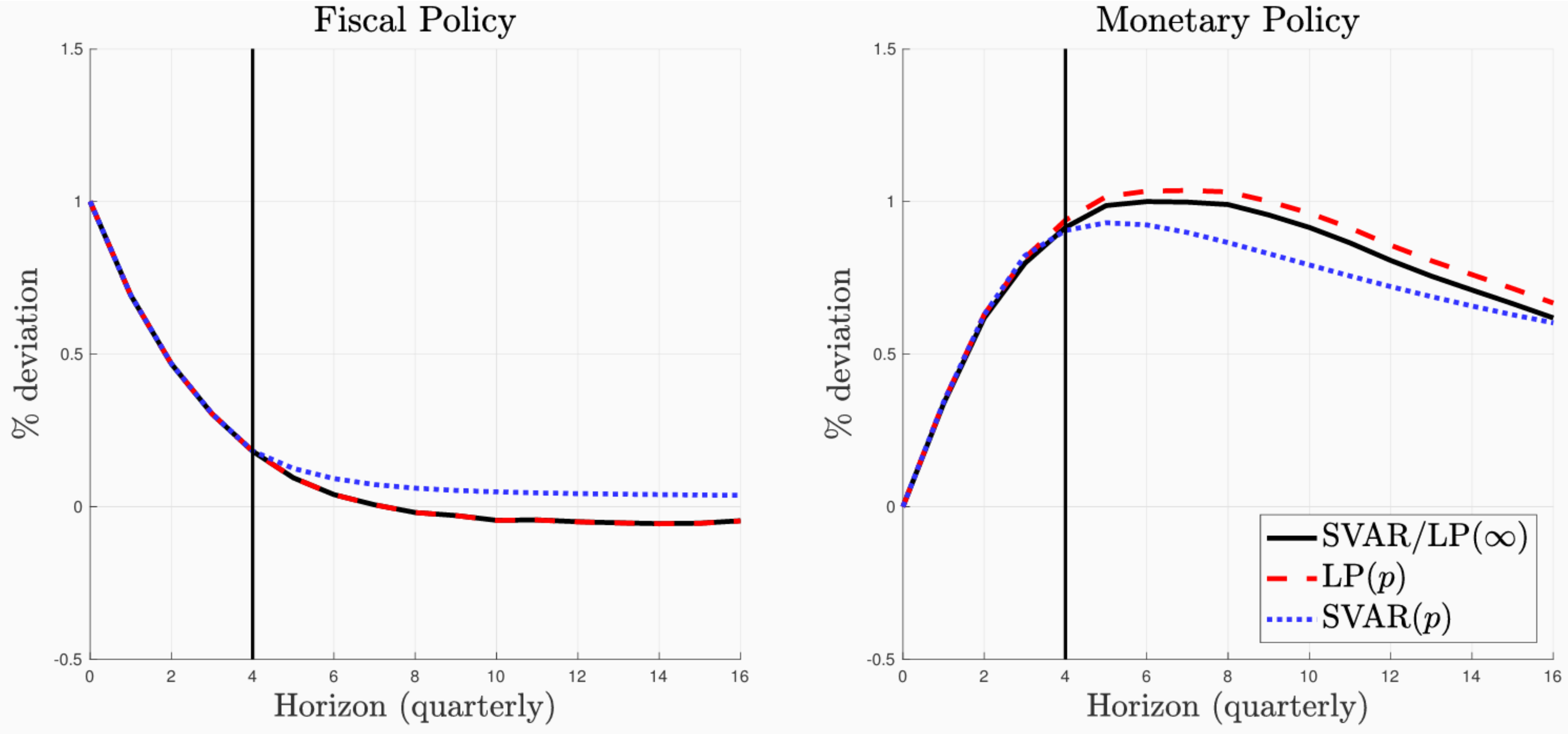
$$y_{t+h} = \mu_h + \beta_h s_t + \gamma_h' r_t + \sum_{\ell=1}^{p} \delta_{h,\ell}' w_{t-\ell} + \xi_{h,t},$$

Note that r_t is controlled while q_t is not

- Are $\{\beta_h\}_{h\geq 0}$ and $\{\theta_h\}_{h\geq 0}$ the same?
- Yes, for $h \leq p$, and more so as $p \to \infty$
  - This results also hold with other non-recursive SVAR identifications
  - Note that a general covariance stationary DGP is assumed

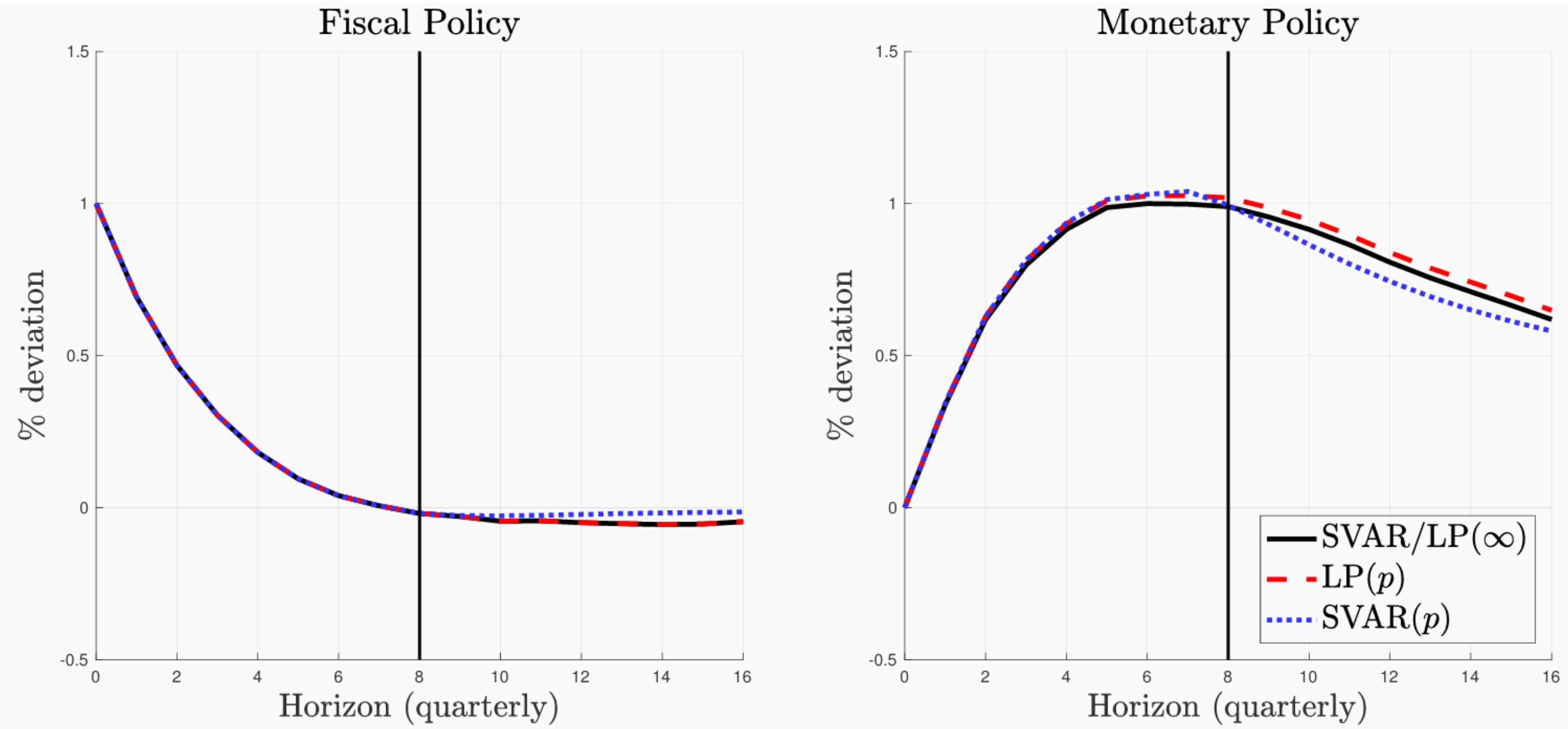
SVAR vs. LP (in population)

- Using Smets-Wouters (2007) simulation data, run VAR(p) and LP(p)
- Left panel: $w_t = (z_{gt}, y_t)$; Right panel: $w_t = (y_t, \pi_t, r_t)$



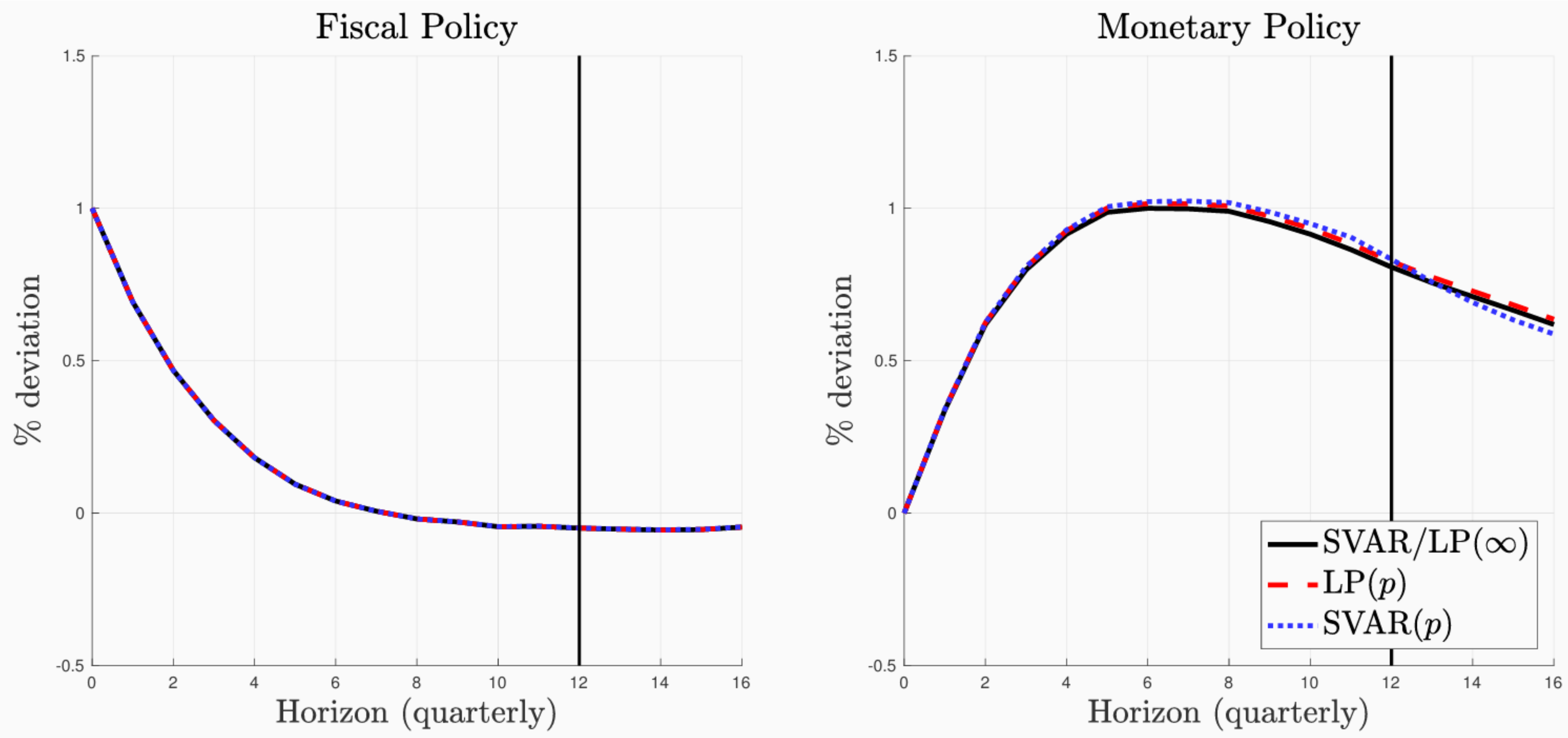
Note: $p =$ 4. Left panel: shock observed. Right panel: recursive ID.

SVAR vs. LP (in population)



Note: $p =$ 8. Left panel: shock observed. Right panel: recursive ID.

SVAR vs. LP (in population)



Note: $p = 12$. Left panel: shock observed. Right panel: recursive ID.

SVAR vs. LP (in sample)

- The results hold in sample too (under mild conditions)
- Assume that DGP is a SVMA

$$w_t = \mu + \Theta(L)\varepsilon_t, \quad \Theta(L) \equiv \sum_{\ell=0}^{\infty} \Theta_\ell L^\ell \quad \text{and} \quad \varepsilon_t \overset{\text{i.i.d.}}{\sim} N\left(0, I_{n_\varepsilon}\right).$$

  - The parameter of interest is $\Theta_{n_r+2,1,h}$ (i.e., IRF of y_t to $\varepsilon_{1,t}$)
- Assuming invertibility (i.e., $\varepsilon_{1,t}$ is a function of current and past observables),
  - any SVAR identification finding b s.t. $\varepsilon_{1,t} = b' u_t$ can be implemented by LP
- In the case of non-invertibility, one can do recursive SVAR with IV ordered first
  - which can also be implemented with LP-IV

LP-IV

- Often we find external instrument z_t for s_t (e.g. GK shock for FFR) s.t.

$$Cov(z_t, \varepsilon_{j,s} | \{z_\tau, w_\tau\}_{-\infty<\tau<t}) \neq 0 \quad \text{iff} \quad j = 1, s = t$$

  in words: (i) relavance, (ii) basic exogeneity, and *(iii) lead-lag exogeneity*

- Or equivalently,

$$z_t = c_z + \sum_{\ell=1}^{\infty} (\Psi_\ell z_{t-\ell} + \Lambda_\ell w_{t-\ell}) + \alpha \varepsilon_{1,t} + v_t$$

  where $\alpha \neq 0$ and v_t is independent measurement error

  - Ψ_ℓ and Λ_ℓ may possibly be zero

LP-IV

- $W_t \equiv (z_t, w_t')'$ and consider the "reduced-form" IV projection

$$y_{t+h} = \mu_{\mathrm{RF},h} + \beta_{\mathrm{RF},h} z_t + \sum_{\ell=1}^{\infty} \delta'_{\mathrm{RF},h,\ell} W_{t-\ell} + \xi_{\mathrm{RF},h,t}$$

- Consider also the "first-stage" IV projection

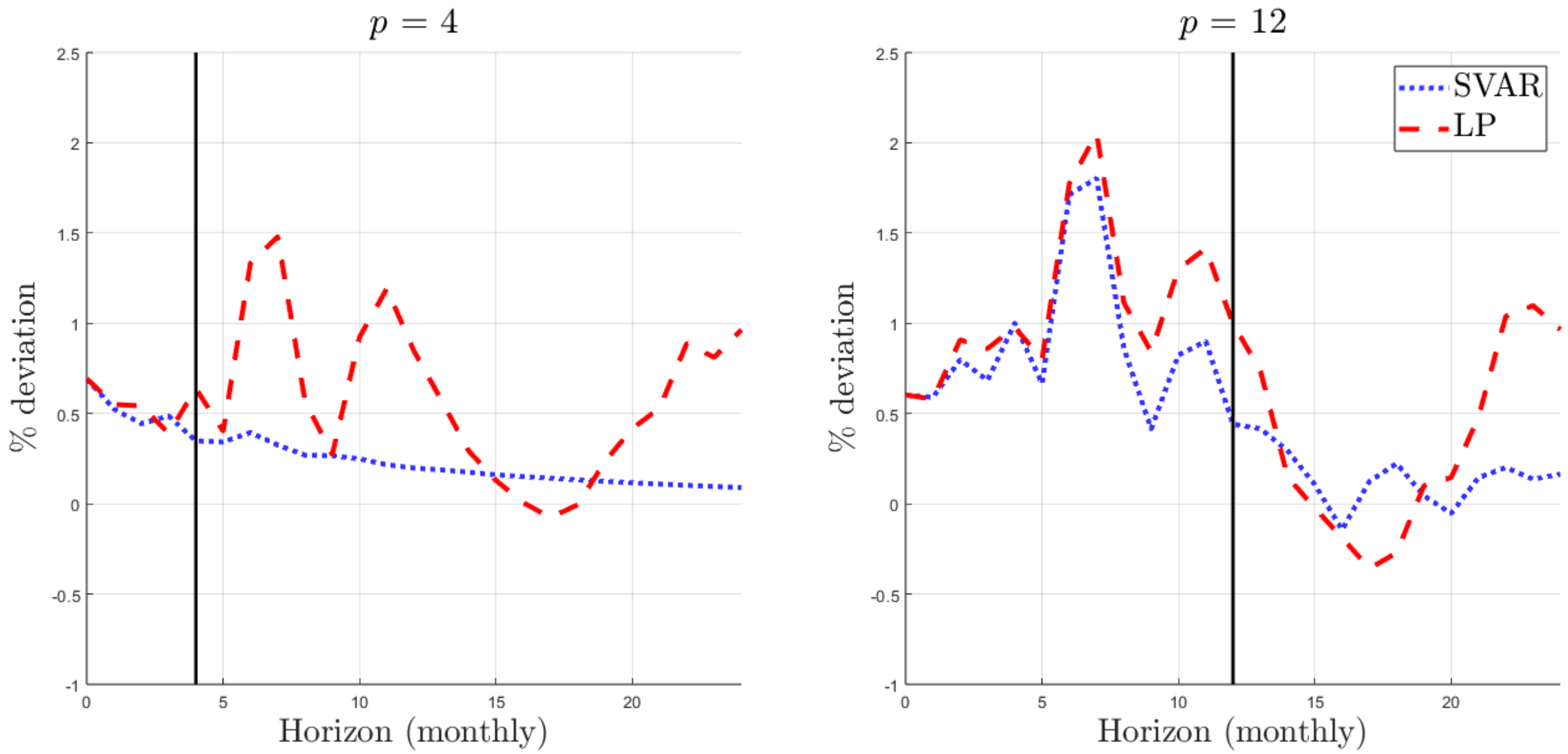
$$s_t = \mu_{\mathrm{FS}} + \beta_{\mathrm{FS}} z_t + \sum_{\ell=1}^{\infty} \delta'_{\mathrm{FS},\ell} W_{t-\ell} + \xi_{\mathrm{FS},t}.$$

- The LP-IV estimand is $\beta_{\mathrm{LPIV},h} \equiv \beta_{\mathrm{RF},h} / \beta_{\mathrm{FS}}$
  - which correctly identifies relative IRF $\Theta_{n_r+2,1,h} / \Theta_{n_r+1,1,0}$

SVAR vs. LP (in sample)

- w_t: output growth, inflation, 1yr gov bond rate, excess bond premium
  - z_t is GK monetary shock

- IRFs from LP and SVAR of EBP to monetary policy are below



SVAR vs. LP

- LP has many advantages over VAR
  - simple univariate regression without strong assumption on the entire system
  - more flexible to accommodate different settings (e.g. nonlinear or panel data)
- Keep in mind there is bias-variance trade-off
  - whenever DGP is not well approximated by finite-order VAR
  - VAR typically has larger bias but smaller variance

LP specification

- LP on levels (y_{t+h}), long-differences $(\Delta_h y_{t+h} = y_{t+h} - y_{t-1})$, cumulative effect on first differences $(\Delta y_t = y_t - y_{t-1})$ are the same

$$\mathcal{R}_{sy}(h) = \mathcal{R}_{s\Delta_h y}(h) = \sum_{j=0}^{h} \mathcal{R}_{s\Delta y}(j)$$

(i) $\mathcal{R}_{sy}(h) = \mathcal{R}_{s\Delta_h y}(h)$ as long as y_{t-1} is included in the RHS

(ii) $\mathcal{R}_{s\Delta_h y}(h) = \sum_{j=0}^{h} \mathcal{R}_{s\Delta y}(j)$ since cumulative sum of first-diff is long-diff

LP specification

- Often the moment of interest is not a simple IRF but a transformed one
  - e.g. mutiplier = (overall gains in GDP / overall fiscal expenditure) over time

- Consider $y_t = \gamma s_t + u_t^y$ and $s_t = \rho s_{t-1} + u_t^s$, with $E\left(u_t^y, u_r^s\right) = 0$
- It is easy to see that $\mathcal{R}_{sy}(h) = \gamma\rho^h$ and $\mathcal{R}_{ss}(h) = \rho^h$. Define the multiplier as

$$m_h = \frac{\sum_{j=0}^{h} \mathcal{R}_{sy}(j)}{\sum_{j=0}^{h} \mathcal{R}_{ss}(j)} = \frac{\gamma \sum_{j=0}^{h} \rho^j}{\sum_{j=0}^{h} \rho^j} = \gamma$$

- In general, set up LP-IV as below and estimate with instrument z_t

$$\sum_{i=0}^{h} y_{t+i} = \alpha_h + m_h \sum_{i=0}^{h} s_{t+i} + Controls + Trend + \nu_{t+h}$$

On inference

- Let us begin with AR(1) example: $y_t = \rho y_{t-1} + u_t$
  - u_t is stationary & mean-independent relative to past and future innovations
  - the parameter of interest $\beta(\rho, h) \equiv \rho^h$
- LP in this setting:

$$y_{t+h} = \beta(\rho, h) y_t + \xi_t(\rho, h)$$

  where $\xi_t(\rho, h) \equiv \sum_{\ell=1}^{h} \rho^{h-\ell} u_{t+\ell}$ is serially-correlated even if u_t is iid
- Conventional approach:
  - heteroscedasticity and autocorrelation robust (HAR) standard errors (e.g. NW)
- Instead, lag-augmented LP: $y_{t+h} = \beta(\rho, h) u_t + \beta(\rho, h+1) y_{t-1} + \xi_t(\rho, h)$
  - and use heteroscedasticity robust $\hat{s}(h) \equiv \frac{\left(\sum_{t=1}^{T-h} \hat{\xi}_t(h)^2 \hat{u}_t(h)^2\right)^{1/2}}{\sum_{t=1}^{T-h} \hat{u}_t(h)^2}$

On inference

- To see why,

$$\hat{\beta}(h) \approx \frac{\sum_{t=1}^{T-h} \{y_{t+h} - \beta(\rho, h+1) y_{t-1}\} u_t}{\sum_{t=1}^{T-h} u_t^2}$$

$$= \beta(\rho, h) + \frac{\sum_{t=1}^{T-h} \xi_t(\rho, h) u_t}{\sum_{t=1}^{T-h} u_t^2}.$$

- Note that $\xi_t(\rho, h) u_t$ (regression score) is not serially correlated
  - under the assumption we made on u_t
- $100(1-\alpha)\%$ of confidence interval:

$$\hat{C}(h, \alpha) \equiv \left[\hat{\beta}(h) - z_{1-\alpha/2} \hat{s}(h), \hat{\beta}(h) + z_{1-\alpha/2} \hat{s}(h)\right]$$

  - Worried about small sample? do boostrap!

Counterfactual

- Suppose

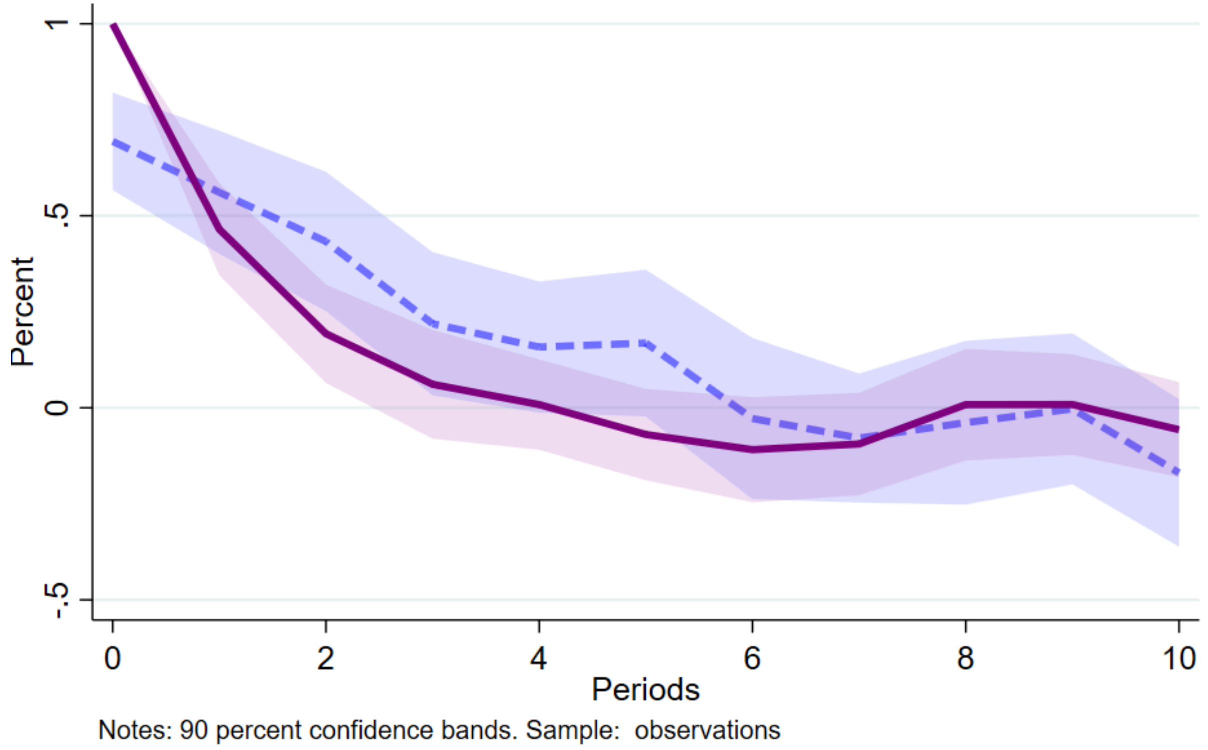
$$\begin{aligned}\Delta y_t &= \beta \Delta s_t + \rho \Delta y_{t-1} + u_t^y \\ \Delta s_t &= \theta \Delta s_{t-1} + u_t^s\end{aligned}$$

- The IRFs are $\mathcal{R}_{ss}(h) = \theta^h$ and

$$\mathcal{R}_{\mathrm{ys}}(h) = \underbrace{\beta\theta^h + \beta\theta^{h-1} + \ldots + \beta\theta\rho^{h-1}}_{\text{due to policy persistence}} + \underbrace{\beta\rho^h}_{\text{internal propagation}}$$

Counterfactual

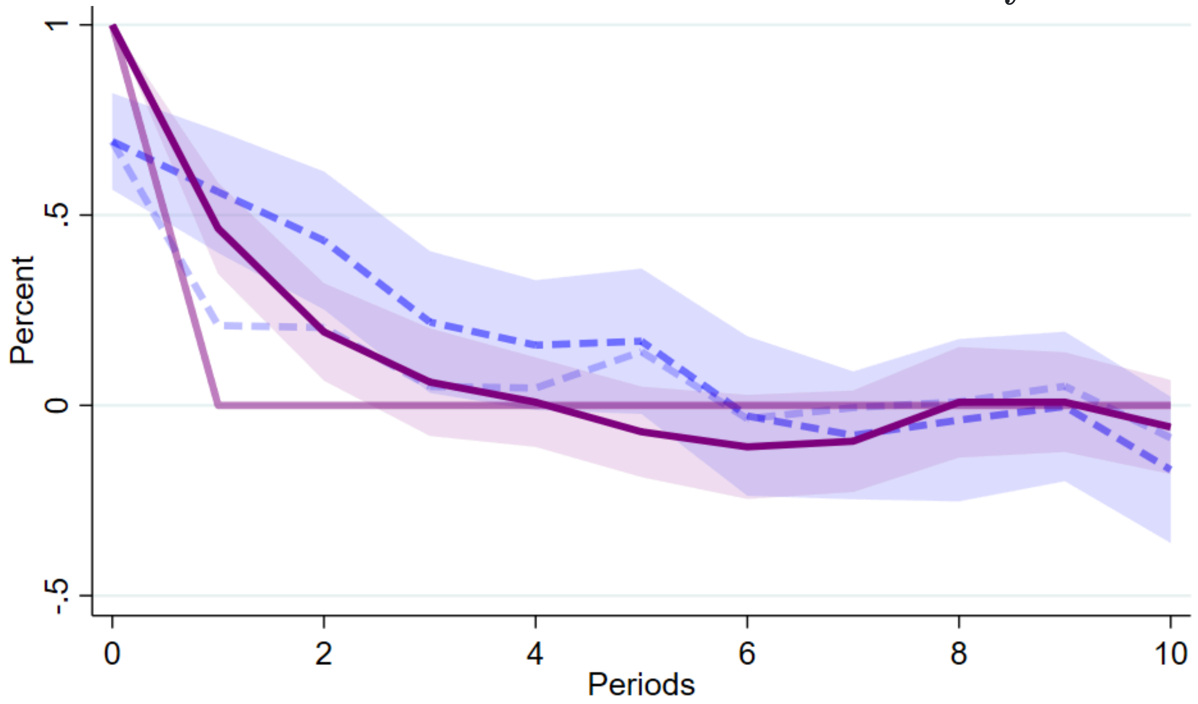
- Here is $\mathcal{R}_{\text{ys}}(h)$ (blue) and $\mathcal{R}_{\text{ss}}(h)$ (purple)



- Setting $\theta = 0$, $\mathcal{R}^*_{\text{ys}}(h) = \beta\rho^h$ (only the internal propagation)

Counterfactual

- Compare $\mathcal{R}_{\text{ys}}(h)$ (blue) and $\mathcal{R}_{\text{ss}}(h)$ (purple) with $\mathcal{R}^*_{\text{ys}}(h)$ and $\mathcal{R}^*_{\text{ss}}(h)$



Notes: 90 percent confidence bands. Sample: 200 observations

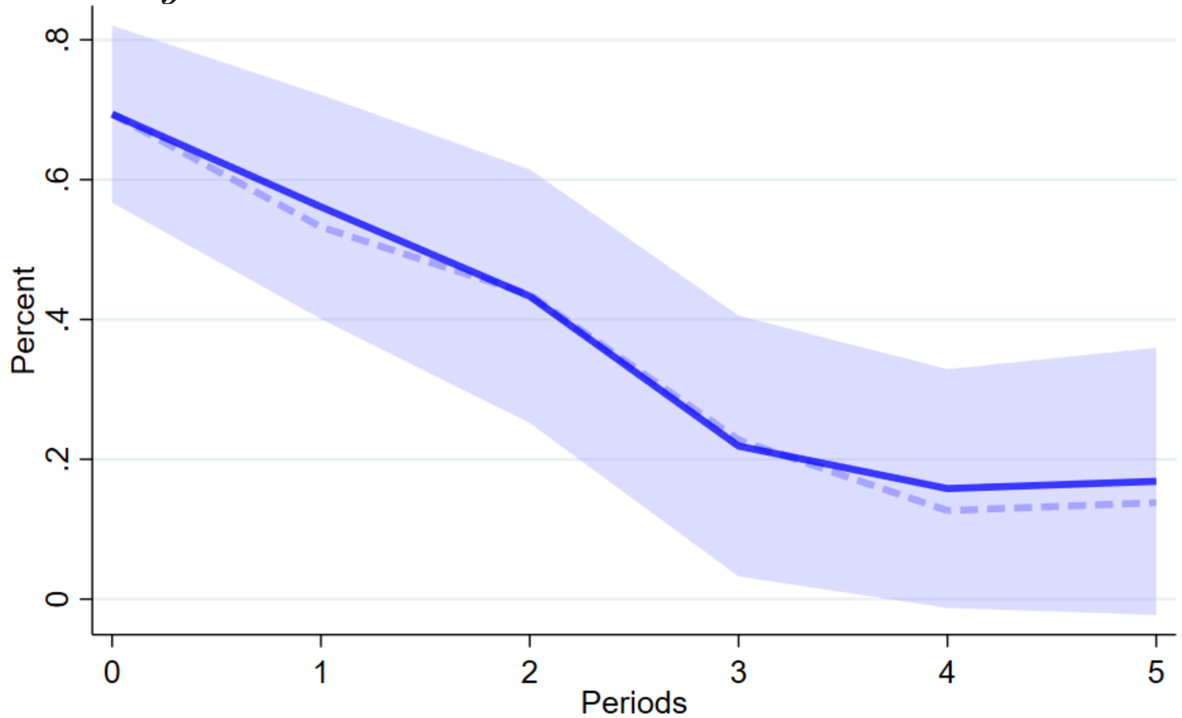
Counterfactual

- Can we recover $\mathcal{R}_{\text{ys}}(h)$ with $\mathcal{R}^{*}_{\text{ys}}(h)$ and $\mathcal{R}_{\text{ss}}(h)$?
- Letting $\mathcal{R}^{c}_{\text{ss}}(h) = \mathcal{R}_{\text{ss}}(h)$ follow the procedure below

$$
\begin{aligned}
\mathcal{R}^{c}_{ys}(0) &= \mathcal{R}^{*}_{ys}(0)\mathcal{R}^{c}_{ss}(0) \\
\mathcal{R}^{c}_{ys}(1) &= \mathcal{R}^{*}_{ys}(0)\mathcal{R}^{c}_{ss}(1) + \mathcal{R}^{*}_{ys}(1)\mathcal{R}^{c}_{ss}(0) \\
\mathcal{R}^{c}_{ys}(2) &= \mathcal{R}^{*}_{ys}(0)\mathcal{R}^{c}_{ss}(2) + \mathcal{R}^{*}_{ys}(1)\mathcal{R}^{c}_{ss}(1) + \mathcal{R}^{*}_{ys}(2)\mathcal{R}^{c}_{ss}(0) \\
\mathcal{R}^{c}_{ys}(3) &= \mathcal{R}^{*}_{ys}(0)\mathcal{R}^{c}_{ss}(3) + \mathcal{R}^{*}_{ys}(1)\mathcal{R}^{c}_{ss}(2) + \mathcal{R}^{*}_{ys}(2)\mathcal{R}^{c}_{ss}(1) + \mathcal{R}^{*}_{ys}(3)\mathcal{R}^{c}_{ss}(0) \\
\vdots &= \vdots
\end{aligned}
$$

Counterfactual

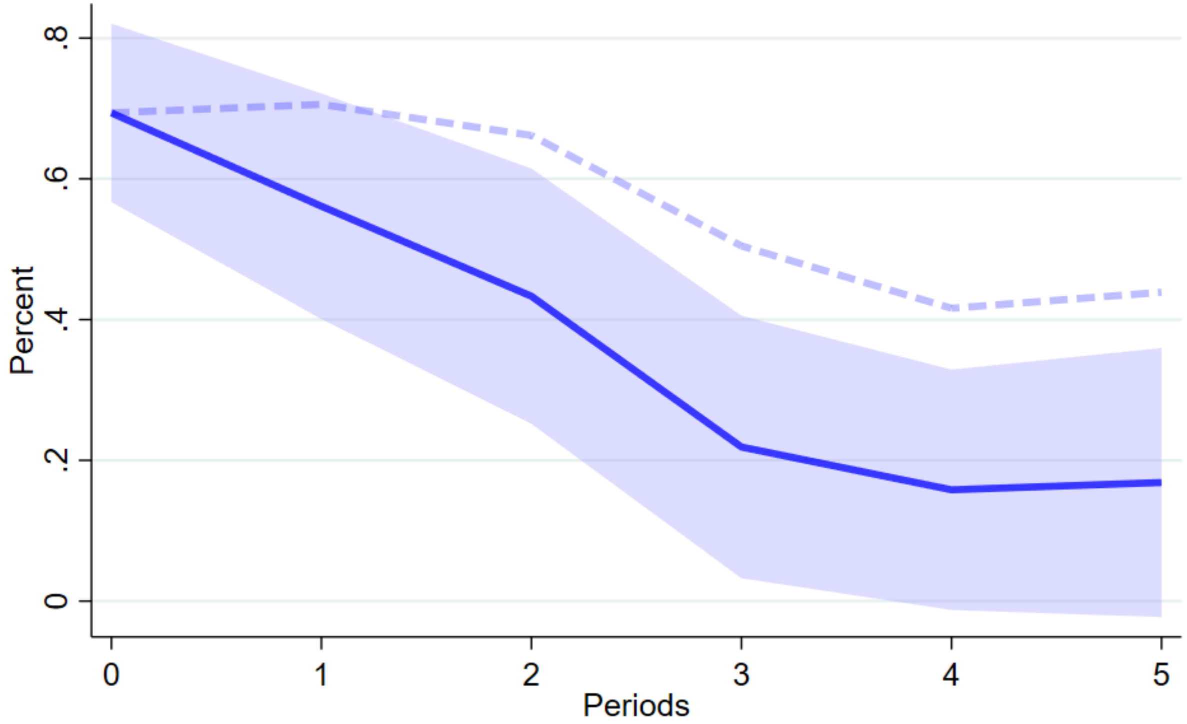
- Now we compare $\mathcal{R}^c_{ys}(h)$ (light blue) with $\mathcal{R}_{ys}(h)$ (blue)



Notes: 90 percent confidence bands. Sample: 200 observations

Counterfactual

- Now as a counterfactual: $\mathcal{R}^c_{\text{ss}}(h) = \mathcal{R}_{\text{ss}}(h) + 0.25$
- Then $\mathcal{R}^c_{ys}(h)$ becomes



Notes: 90 percent confidence bands. Sample: 200 observations

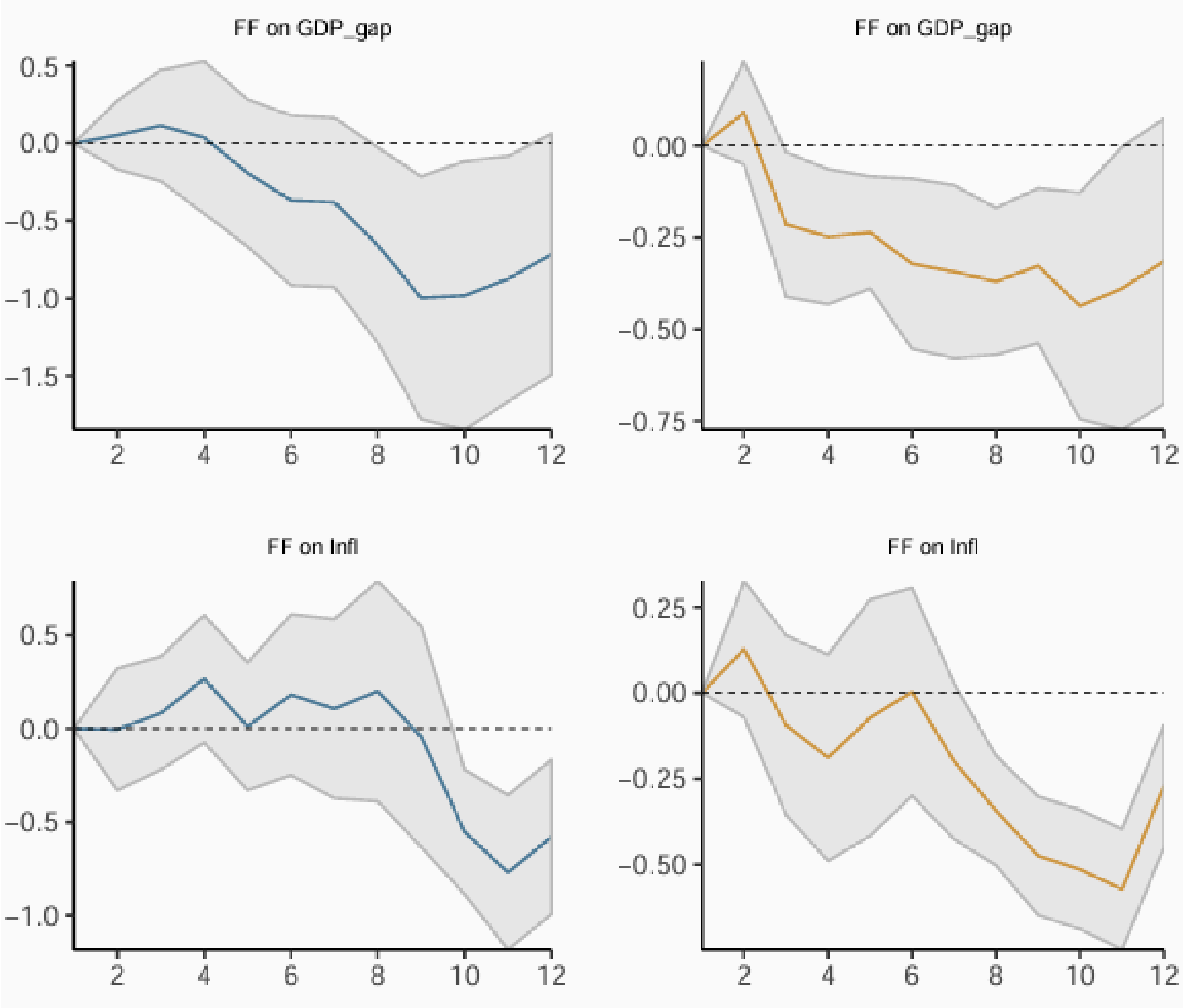
Nonlinear LP

- State-dependency: separate data into two regime using dummy variable I_t

$$
\begin{aligned}
y_{t+h} = & I_t[\mu_{A,h} + \beta_{A,h} s_t + \gamma'_{A,h} r_t + \sum_{\ell=1}^{p} \delta'_{A,h,\ell} w_{t-\ell}] \\
& + (1 - I_t)[\mu_{B,h} + \beta_{B,h} s_t + \gamma'_{B,h} r_t + \sum_{\ell=1}^{p} \delta'_{B,h,\ell} w_{t-\ell}] + \xi_{h,t}
\end{aligned}
$$

- For example, a threshold of 4.75% for π_{t-3}
  - to show Y_t and π_t are more responsive to i_t in the low-inflation regime

Nonlinear LP



Panel LP

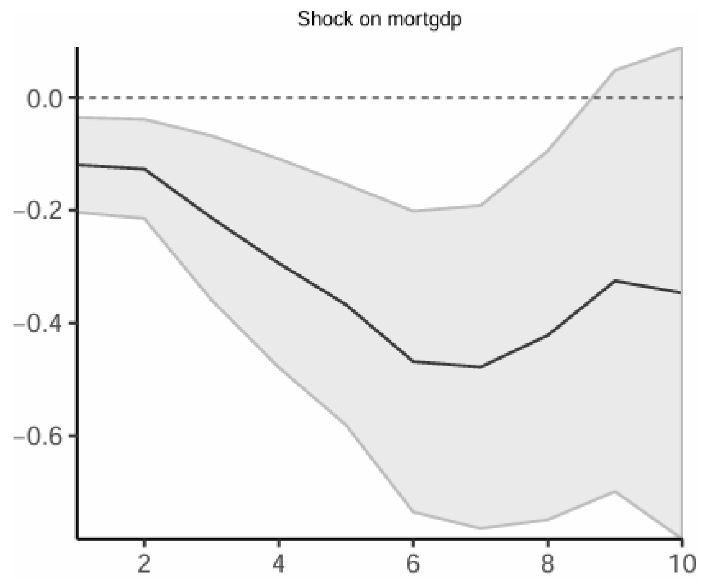
- Consider a panel LP for $i = 1, \ldots, n$ and $t = 1, \ldots, T$,

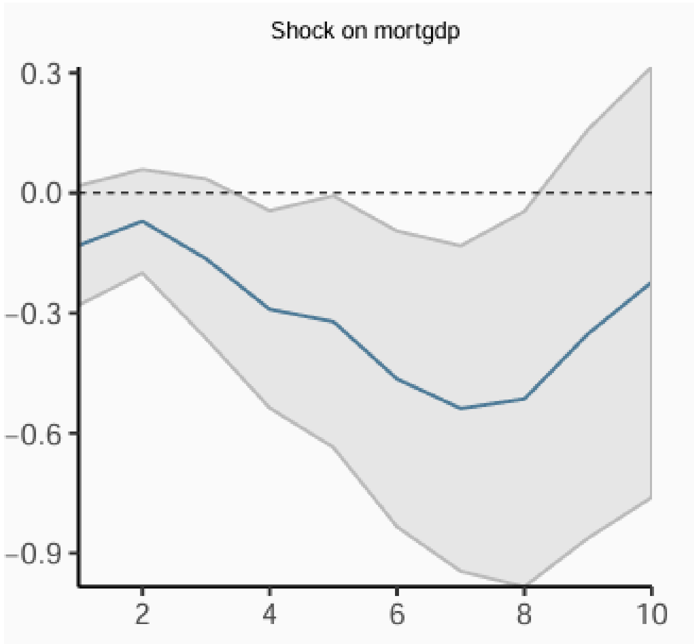
$$y_{i,t+h} = \alpha_i + \delta_t + \beta_h \boldsymbol{s}_{it} + \gamma_h \boldsymbol{x}_{it} + v_{i,t+h}$$

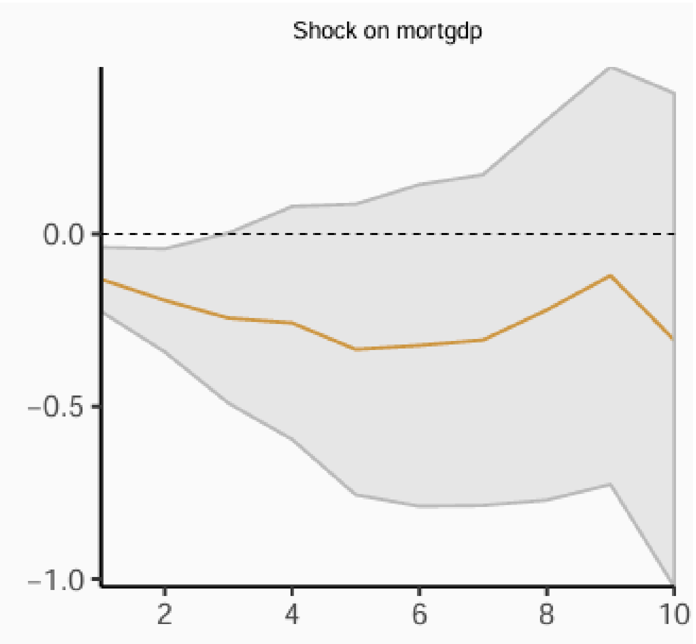
  - where $\boldsymbol{x}_{it}$ is control including lagged endogenous variables

- Given country panel data, you ask
  - "What happens to mortgage lending relative to GDP when you increase interest rates?"
  - "Would that effect be stronger in periods of economic expansions?"

Panel LP


Shock on mortgdp
0.0
−0.2
−0.4
−0.6
2
4
6
8
10


Shock on mortgdp
0.3
0.0
−0.3
−0.6
−0.9
2
4
6
8
10


Shock on mortgdp
0.0
−0.5
−1.0
2
4
6
8
10

LP-DiD

- Consider a policy evaluation with a typical DiD setting with
  - $P_t = 1$ for post, 0 for pre; $A_i = 1$ for treated, 0 for control

- Standard approaches: under (i) parallel trend and (ii) no anticipation
  - Static TWFE

$$y_{it} = \alpha_i + \delta_t + \beta^{\text{TWFE}} D_{it} + \epsilon_{it}; \quad D_{it} = P_t \times A_i$$

  - Event-study (distributed lags) TWFE

$$y_{it} = \alpha_i + \delta_t + \sum_{m=-Q}^{M} \beta_m^{\text{TWFE}} D_{it-m} + \epsilon_{it}$$

- TWFE is okay in the 2×2 setting
  - or when treatment occurs at the same time

LP-DiD

- TWFE is biased even under parallel trends with staggered treatment
  - if treatment effects are dynamic and heterogeneous
  - Problem: you compare newly treated with earlier treated

- LP-DiD

$$
\begin{aligned}
y_{i,t+h} - y_{i,t-1} = & \beta^{h,LP-DiD} \Delta D_{it} && \} \text{ treatment indicator} \\
& + \delta_t^h && \} \text{ time effects} \\
& + e_{it}^h
\end{aligned}
$$

where

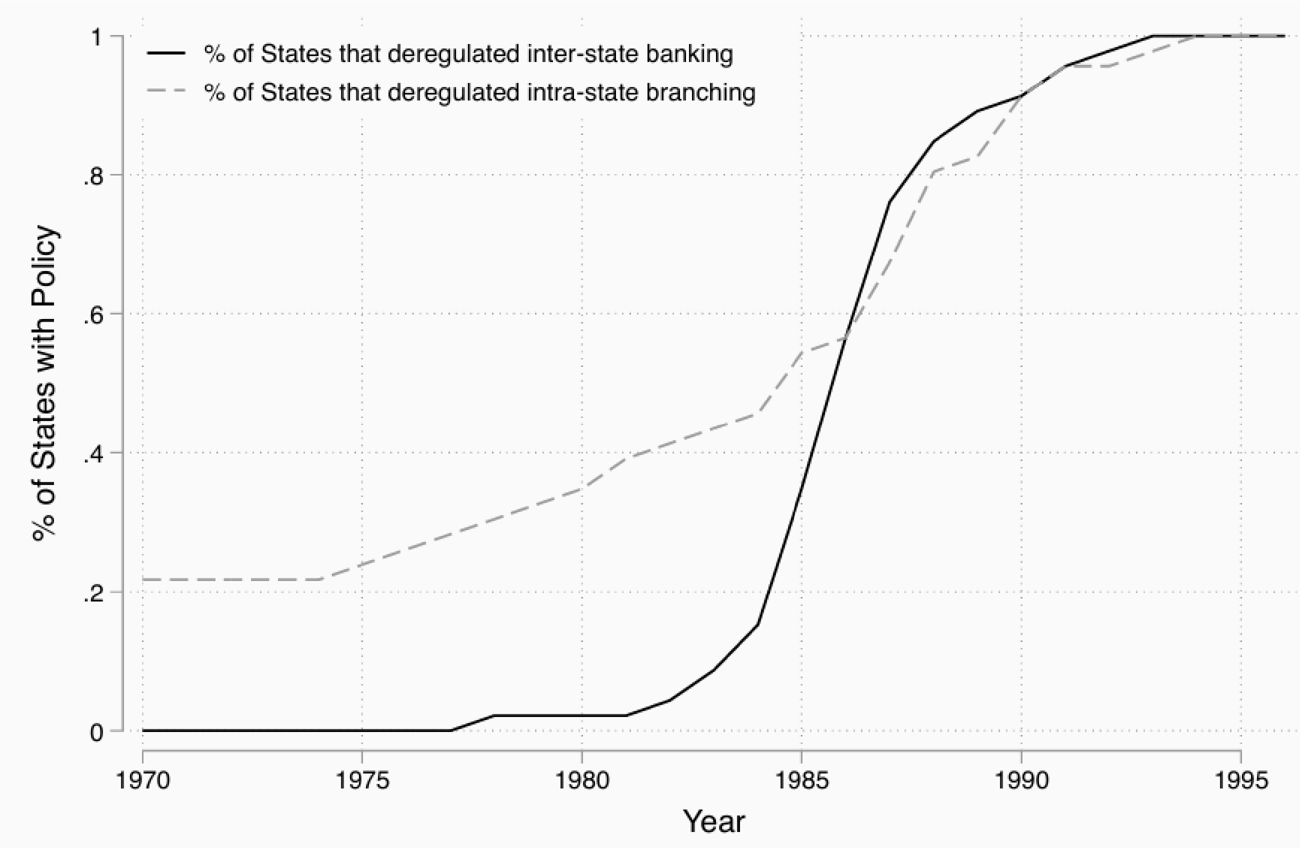
$$
\begin{cases}
\text{newly treated} & \Delta D_{it} = 1, \\
\text{or clean control} & \Delta D_{i,t+h} = 0
\end{cases}
$$

  - You can also add lagged outcomes and exogenous covariates as controls

LP-DiD

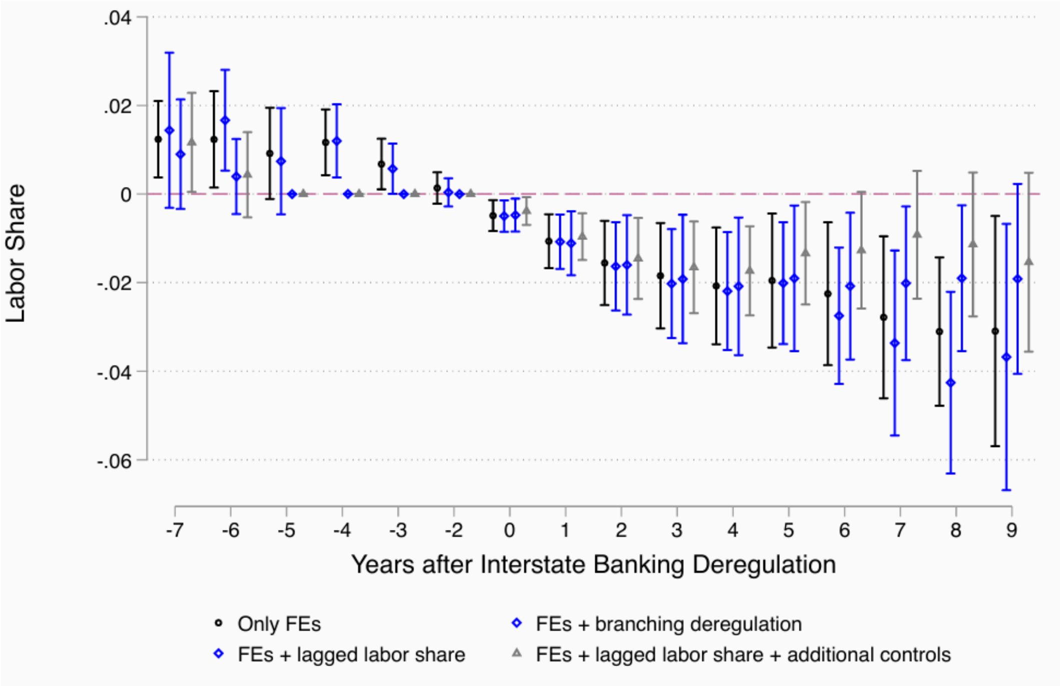
- Example: the effects of banking sector deregulation in late '70s on labor share
  - financial development has direct consequences on how firms finance inputs

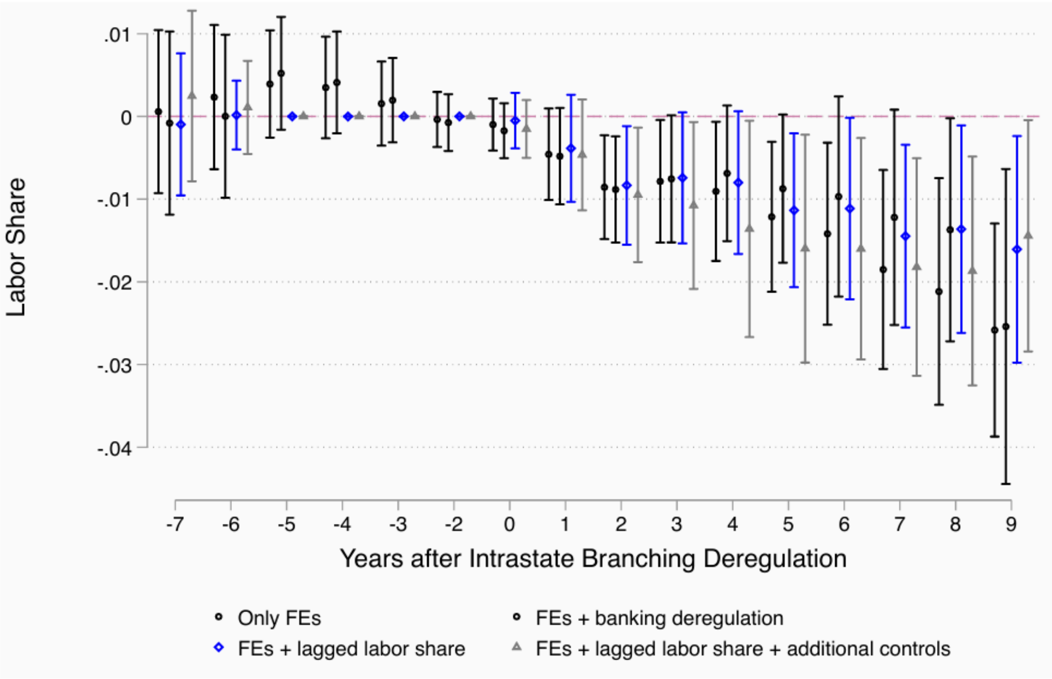
- The policy was implemented in a staggered way



LP-DiD

- You can find negative effects on labor share





References

Jordà (2023) "Local Projections for Applied Economics," *Annual Review of Economics*

Jordà (2023) "Techniques for Empirical Macroeconomics," *AEA Continuing Education 2023*

Nakmura-Steinsson (2018) "Identification in Macroeconomics," *Journal of Economic Perspectives*

Stock-Watson (2018) "Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments," *The Economic Journal*

Ramey (2016) "Macroeconomic Shocks and Their Propagation," *Handbook of Macro, Chapter 2*

Plagborg-Møller & Wolf (2021) "Local Projections and VARs Estimate the Same Impulse Responses," *ECMA*

Li, Plagborg-Møller & Wolf (2024) "Local projections vs. VARs: Lessons from thousands of DGPs," *Journal of Econometrics*

Dube et al. (2023) "A local projections approach to difference-in-differences event studies." NBER WP

Adämmer (2019) "lpirfs: An R Package to Estimate Impulse Response Functions by Local Projections," The R Journal