Heterogeneity in impulse response functions

Martín Almuzara[1] Víctor Sancibrián[2]

[1] Federal Reserve Bank of New York

[2] CEMFI

October 10, 2024

Motivation

- Heterogeneity in transmission of shocks at individual level is key empirical object:
  - What is the effect of monetary policy shocks on household consumption and income?
  - What role does household heterogeneity play in the transmission mechanism?
- Large empirical literature estimates impulse responses using panel local projections:
  - Informative about observable heterogeneity.
- But what about unobservable heterogeneity?
  - Some policies (e.g., tax rebates) may have bigger impacts on certain units depending on dimensions we do not observe directly/very well (e.g., credit constraints).
- Presence of aggregate shocks poses additional challenges.

This paper

- We propose an estimator for the cross-sectional distribution of impulse responses:
  - Unit-by-unit estimation of individual responses + aggregation step.
  - Aggregation step corrects for noise in individual estimates and does dimension reduction.
  - Computationally easy to implement using local projections.

- We study its large-sample properties in a general class of DGPs:
  - Macro and micro shocks + unrestricted, heterogeneous dynamics.
    - ➡ $Y_{it} = \mu_i + \sum_{\ell=0}^{\infty} \theta_{i\ell} X_{t-\ell} + v_{it}$

- Key. Heterogeneity is low-dimensional linear combination of basis shapes:
  - ➡ $\theta_{ih} = \beta_0(h) + \beta(h)' \eta_i$ with η_i low-dim nonparametric object.

- We then study the impact of monetary policy on workers' labor income and firm's turnover using Spanish admin data.

Selected literature

- Ingredients related to extensive econometric literatures:
  - Unit-by-unit estimation: Pesaran, Smith (1995)
  - Bias reduction: Jochmans, Weidner (2021)
  - Heterogeneity and factor models: Alan, Browning, Ejrnæs (2018)

- Recent papers look at heterogeneity in impulse responses from different angles:
  - Huang (2022): discretized (group) heterogeneity
  - Chen, Chang, Schorfheide (2022): responses of cross-sectional distributions (no panel data)

➡ This paper: recovering heterogeneous distribution + panel data + aggregate shocks

Unit-by-unit estimation and challenges

Bias in empirical distributions

- Simple approach is to estimate heterogeneous IRs unit-by-unit and then construct empirical distribution,

$$\{\hat{\theta}_{h,i}\}_{i=1}^{N} \quad \mapsto \quad \hat{F}_h(\theta) = \frac{1}{N}\sum_{i=1}^{N} 1\left[\hat{\theta}_{h,i} \leq \theta\right] \tag{1}$$

- In small samples $\hat{F}_h$ tends to be wider than the actual distribution F

- Recently, Jochmans–Weidner (2021) considered the general problem in (1) and showed that the bias is proportional to T^{-1},

$$E\left[\hat{F}_{N,T}(\theta)\right] - F(\theta) = \frac{\boldsymbol{b(\theta)}}{\boldsymbol{T}} + O\left(T^{-3/2}\right).$$

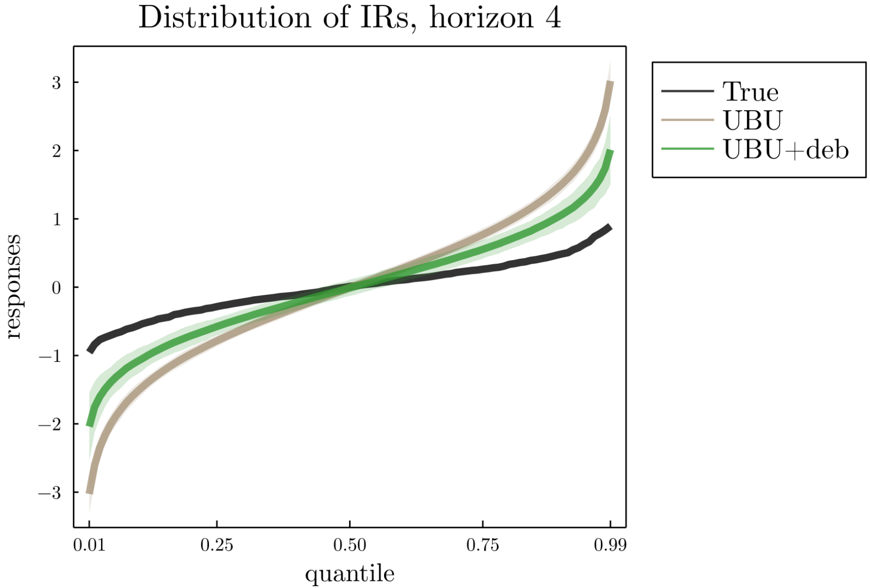
Illustration

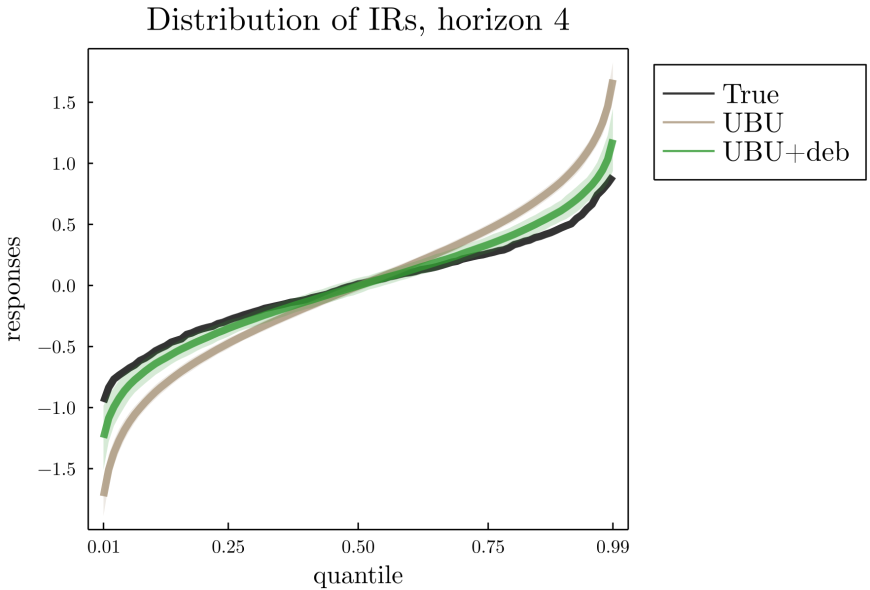
- Jochmans–Weidner (2021) also propose jackknife bias corrections for the distribution and quantile function

- The following simulations illustrate how bias correction performs in the context of impulse responses. Consider the following DGP:

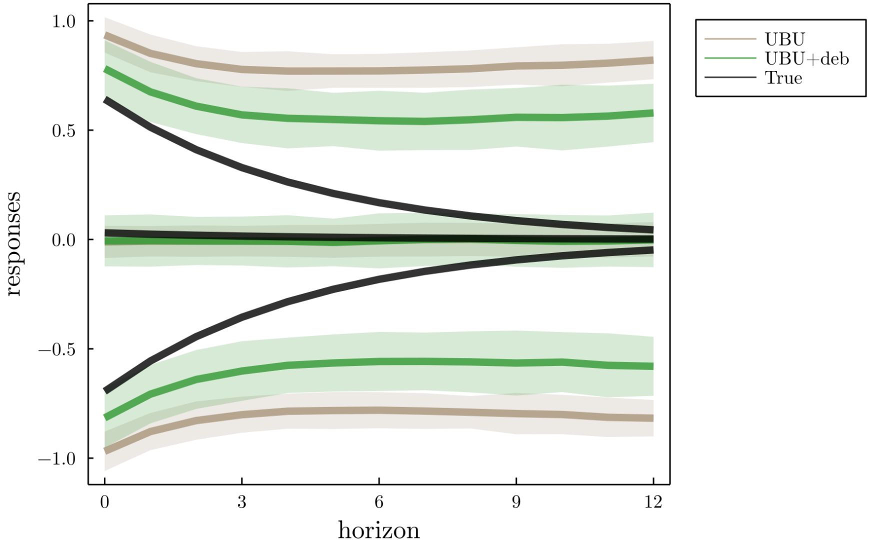
$$y_{i,t} = \mu + \theta_{0,i} x_{i,t} + \rho_i y_{i,t-1} + \varepsilon_{i,t}$$

  - $x_{i,t} \sim N(0,1)$, $\varepsilon_{i,t} \sim N(0,1)$
  - $\rho_i \sim U(0,1)$, $\theta_{i,0} = g(\eta_i)$, $\eta_i \sim N(0,1)$, g is the logistic function
  - IRs given by $\theta_{h,i} = \rho_i^h \theta_{0,i}$ for $h \geq 0$
  - Monte Carlo: 200 samples, $N = 1000$, $T = 20$

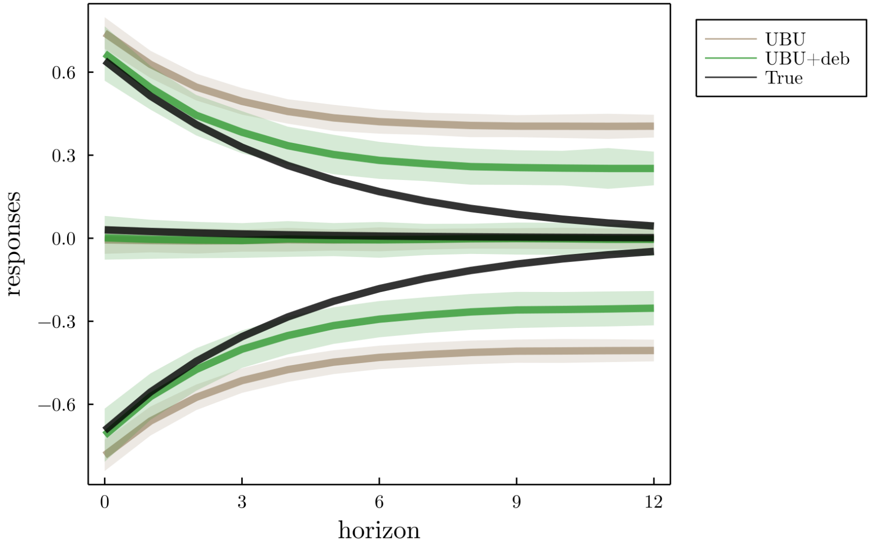

Distribution of IRs, horizon 4
responses
quantile
3
2
1
0
−1
−2
−3
0.01
0.25
0.50
0.75
0.99
True
UBU
UBU+deb


Distribution of IRs, horizon 4
responses
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
0.01
0.25
0.50
0.75
0.99
quantile
True
UBU
UBU+deb

Distribution of IRs, quantiles 0.25, 0.5 and 0.75



Distribution of IRs, quantiles 0.25, 0.5 and 0.75



Setup and method

Setup

- Micro data y_{it} and an aggregate x_t modeled as

$$\begin{pmatrix} y_{it} \\ x_t \end{pmatrix} = \begin{pmatrix} \mu_{y,i} \\ \mu_x \end{pmatrix} + \sum_{\ell=0}^{\infty} \begin{pmatrix} \Psi_{yy,i,\ell} & \Psi_{yx,i,\ell} \\ 0 & \Psi_{xx,\ell} \end{pmatrix} \begin{pmatrix} \varepsilon_{i,t-\ell} \\ \upsilon_{t-\ell} \end{pmatrix}$$

- In many applications x_t is measure of a "shock" (or interaction)
  - $x_t = \upsilon_t$
  - Interest in $\theta_i = (\Psi_{xx,i,0}, \Psi_{xx,i,1}, \ldots, \Psi_{xx,i,H})$
  - Today $\implies$ this case

Setup

- Local projection (wo./intercept)

$$y_{i,t+h} = \theta_{h,i} x_t + u_{i,h,t+h},$$

- We introduce dimensionality reduction via a linear factor structure

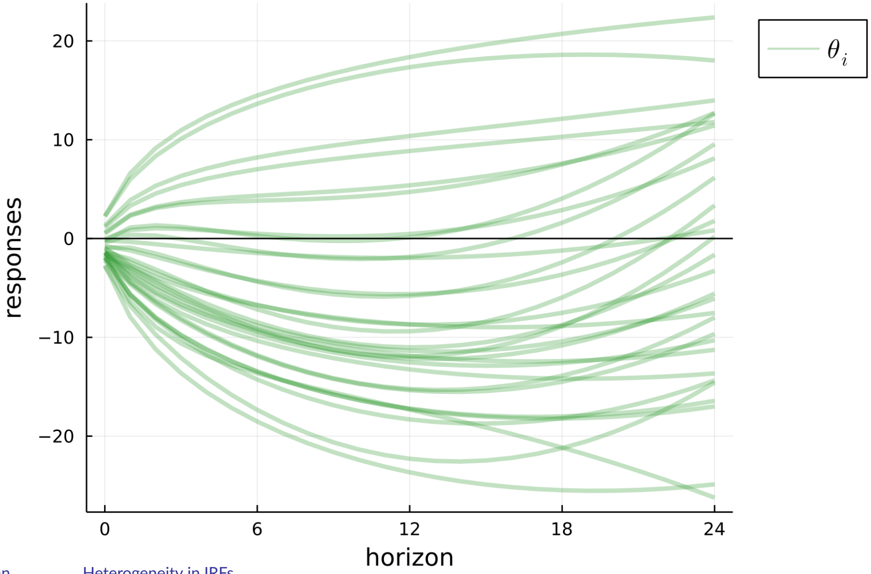
$$\underbrace{\theta_i}_{(H+1)\times 1} = \theta(\eta_i) = \beta_0 + \beta \underbrace{\eta_i}_{D\times 1},$$

w/factors $\eta_{d,i}$ independent, $\eta_{d,i} \sim F_d$, $E\left[\eta_{d,i}\right] = 0$ and $\text{Var}\left(\eta_{d,i}\right) = 1$,

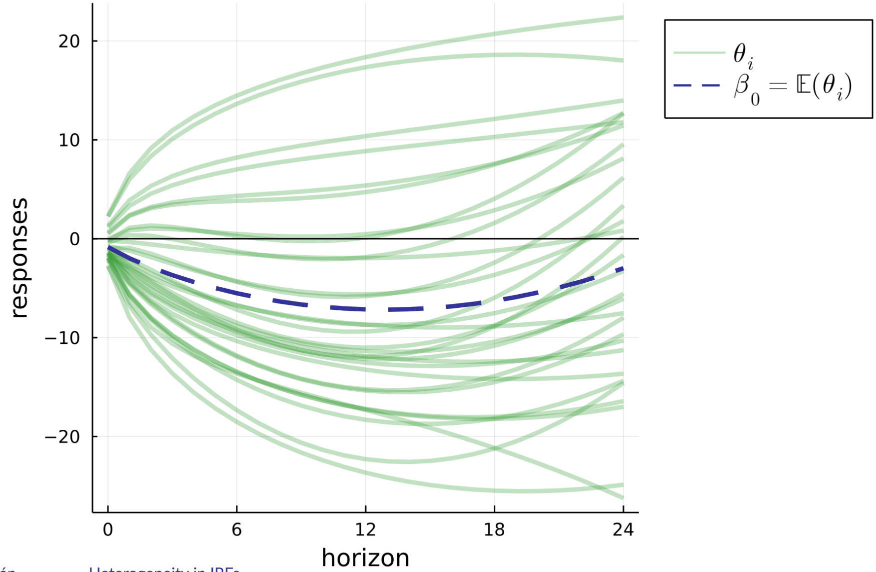
$$E\left[\theta_i\right] = \beta_0,$$
$$E\left[\theta_i \theta_i'\right] = \beta_0 \beta_0' + \beta \beta'.$$

- Cross-section distribution P of θ_i fully determined by $\{\beta_d\}_{d=0}^{D}$., $\{F_d\}_{d=1}^{D}$

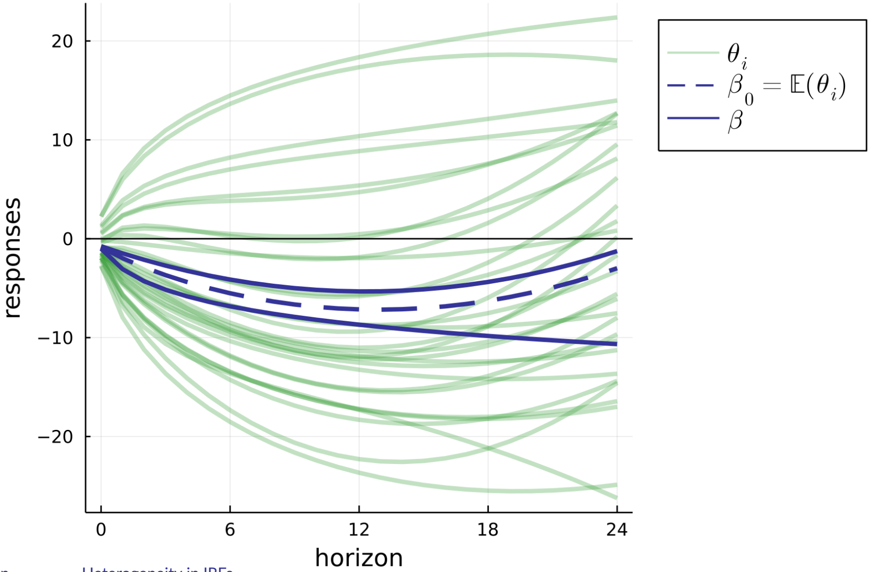
Intuition: $D = 2$



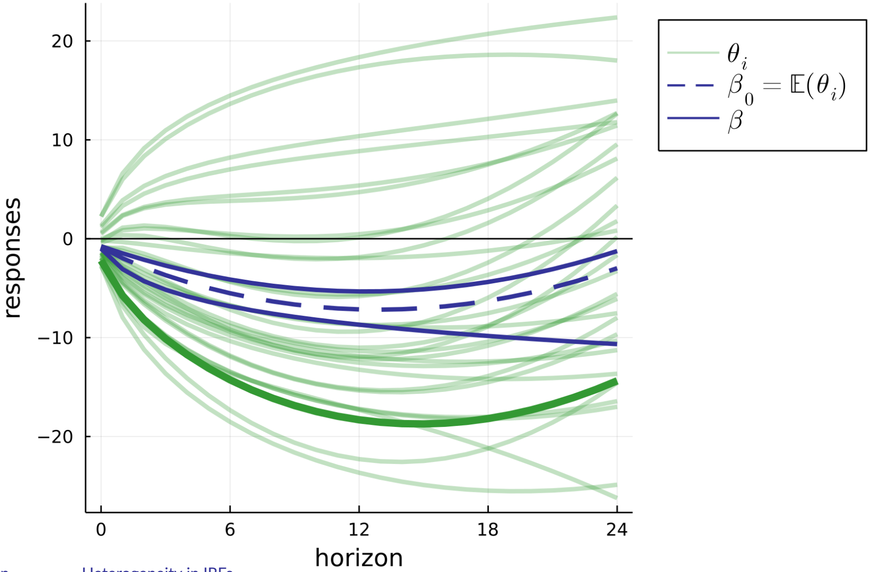
Intuition: $D = 2$



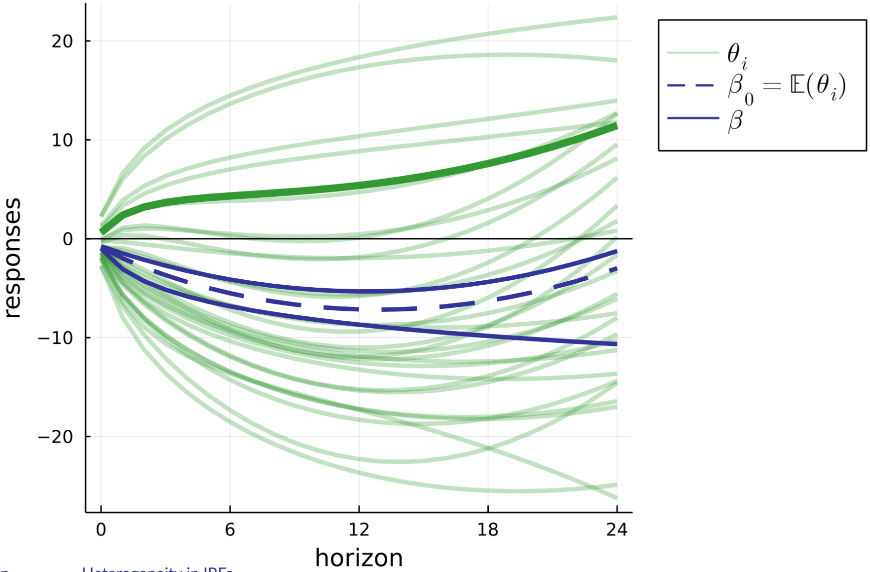
Intuition: $D = 2$



Intuition: $D = 2$



Intuition: $D = 2$



Estimation approach

- Conditional independence assumption:

$$E\left[Y_{i,T} - X_T\theta_i \middle| \theta_i, X_T\right] = 0_{M_T \times 1}$$

$$Y_{i,T} = \begin{pmatrix} y_{i,1} \\ \vdots \\ y_{i,T} \\ y_{i,2} \\ \vdots \\ y_{i,T} \\ \vdots \\ \vdots \\ y_{i,H+1} \\ \vdots \\ y_{i,T} \end{pmatrix}_{M_T \times 1} \qquad X_T = \begin{pmatrix} x_1 & 0 & \dots & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_T & 0 & \dots & \dots & \dots & 0 \\ 0 & x_1 & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & x_{T-1} & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & \dots & 0 & x_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & \dots & 0 & x_{T-H} \end{pmatrix}_{M_T \times (H+1)}$$

Estimation approach

- Model for heterogeneity:

$$\theta_i = \boldsymbol{\beta_0} + \eta_i' \boldsymbol{\beta}, \quad \eta_{d,i} \sim \boldsymbol{F_d}$$

- Estimation method:
  - Steps
    1. Estimate $\{\hat{\theta}_i\}_{i=1}^N$ unit-by-unit
    2. Estimate β_0 and β (+bias reduction)
    3. Estimate F_d by its empirical distribution function (+bias reduction)

    Then form distribution P of θ_i
  - This is a simple algorithm w/separate parametric/nonparametric problems and reduces to a set of minimum distance estimators

Step 2: β_0 and β

- Then
  1. Estimate β_0 via pooled local projections
  2. Estimate β via minimum distance given $\hat{\beta}_0$

$$\hat{\beta} = \arg\min_{\beta} \left\| N^{-1} \sum_{i=1}^{N} (\hat{\theta}_i - \hat{\beta}_0)(\hat{\theta}_i - \hat{\beta}_0)' - \beta\beta' \right\|_{\Omega}$$

  for some weight matrix Ω and additional normalizations to disentangle each $\beta_{h,d}$

Steps 3: F_d

- Given β, for each i, overidentified GMM/MD problem

$$\min_{\eta} \quad \left\| X_T' \left(Y_{i,T} - X_T \left(\hat{\beta}_0 + \hat{\beta} \eta \right) \right) \right\|_W$$

for some weight matrix W

- F_d then estimated by its empirical distributions counterpart,

$$\hat{F}_d(\eta) = \frac{1}{N} \sum_{i=1}^{N} 1 \left[\hat{\eta}_{d,i} \leq \eta \right]$$

after imposing the mean and variance normalizations

Sampling properties

- Mapping from (β, F) to P

$$\begin{aligned} P(c) = P_{\beta,F}(c) &= P_{\beta,F}(\theta_i \leq c) \\ &= \int \cdots \int \prod_{h=0}^{H} 1\left[\beta_{h,0} + \sum_{d=1}^{D} \beta_{h,d}\eta_d \leq c_h\right] dF_1(\eta_1) \ldots dF_D(\eta_D) \end{aligned}$$

- We can get an asymptotic expansion

$$\left(\hat{P} - P\right) \approx \sum_{d=0}^{D} \pi_{d,\beta,F}\left((\hat{\beta}_d - \beta_d)\right) + \sum_{d=1}^{D} \Pi_{d,\beta,F}\left(\hat{F}_d - F_d)\right)$$

Sampling properties

- We can learn a lot from this expansion:
  1. Separate bias reduction for $\hat{\beta}$ and $\hat{F}$ gives bias reduction for $\hat{P}$
     - Without bias reduction, $\hat{P}$ has bias of order T^{-1}
  2. $\sqrt{N}(\hat{P} - P) \implies$ a “linear combination” of Gaussian processes
     - $N, T \to \infty$, H, D fixed
     - Inference via bootstrap can be justified
     - Variance reduction
- Monte Carlo simulation suggests very good small-sample behavior

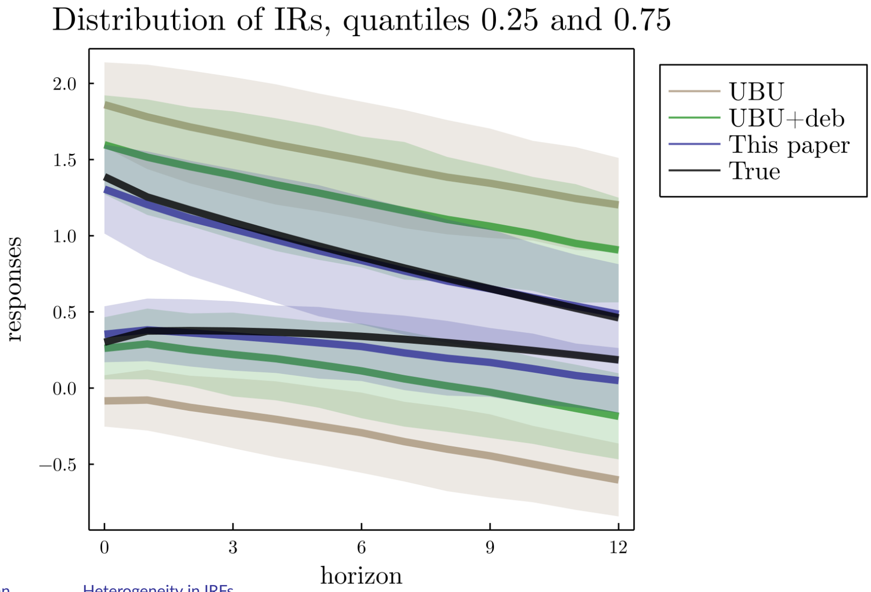
$N = 2000$, $T = 50$, $H = 12$, $D = 2$: $q = 0.25$

	UBU		UBU+JW		This paper	
	Bias	$\sqrt{NSE}$	Bias	$\sqrt{NSE}$	Bias	$\sqrt{NSE}$
$h = 0$	-0.13	0.92	-0.02	1.35	0.02	1.33
$h = 4$	-0.27	0.91	-0.10	1.33	-0.03	1.00
$h = 8$	-0.33	0.97	-0.15	1.45	-0.03	1.03
$h = 12$	-0.37	0.97	-0.17	1.51	-0.01	0.99

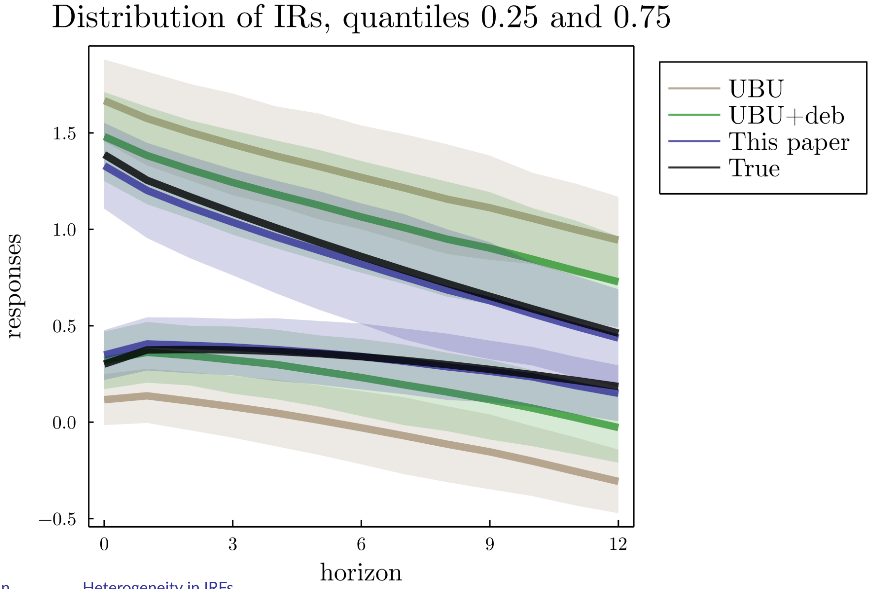
$N = 2000$, $T = 50$, $H = 12$, $D = 2$: $q = 0.75$

	UBU		UBU+JW		This paper	
	Bias	$\sqrt{NSE}$	Bias	$\sqrt{NSE}$	Bias	$\sqrt{NSE}$
$h = 0$	0.06	1.12	-0.02	1.69	-0.14	1.98
$h = 4$	0.12	1.33	-0.01	2.18	-0.05	1.16
$h = 8$	0.22	1.12	0.06	1.62	-0.03	1.13
$h = 12$	0.30	1.06	0.12	1.64	-0.02	1.40

Simulation results



Simulation results

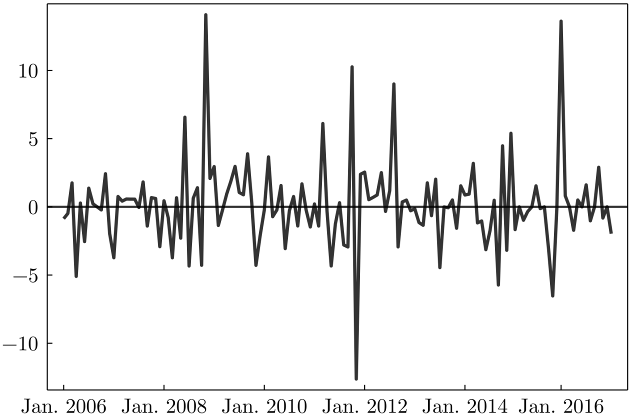


Empirical application

Monetary policy and labor market outcomes in Spain

- Outcomes
  - Worker's side: labor income and job finding probabilities
    - Admin data (MCVL): $N \approx 215,000$
    - 2006-2016 ($T = 132$)
  - Firm's side: turnover (e.g. hiring growth)
    - Admin data (PET)
    - 2013-2016 ($T = 48$)
- Variable of interest $x_{i,t}$
  - We use 'monetary policy surprises', identified as high-frequency movements in a relevant interest rate around ECB meeting dates
  - E.g., Jarociński–Karadi (2020) use Overnight Index Swap rates
- Related literature: Holm–Paul–Tishchbirek. (2021), Singh–Suda–Zervou (2021), Broer–Kramer–Mitman (2021), ...

Monetary policy surprises (Jarociński and Karadi, 2020)



Mean (bps)	0.13
Standard deviation (bps)	3.21

Aggregate responses

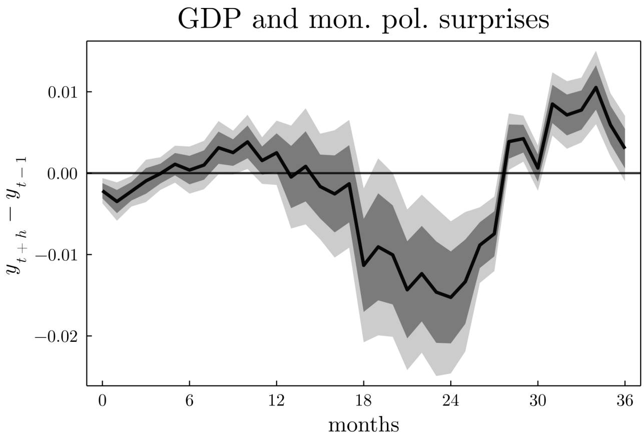


FIGURE. IRFs of the (log) of GDP to a 25 bps shock over the period 2006:1-2016:12; data from Almgren et al. (forthcoming).

Distribution of responses

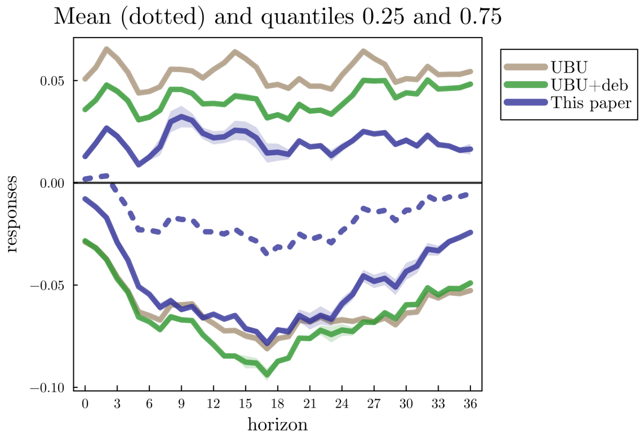


FIGURE. IRFs of the worker's income to a 25 bps shock.

Conclusion

Conclusion

- We propose a method to estimate the distribution of IRs:
- Based on model of heterogeneity that imposes common factor-like structure:
  - IRs are linear combination of a basis of IR shapes
  - Allows us to pool knowledge about IRs at different horizons
- Simple implementation:
  - After unit-by-unit estimation of researcher's choice $\implies$ estimation+bias-reduction of parametric/nonparametric parts
- Sampling properties: bias and variance reduction
- Plenty of potential empirical applications:
  - Monetary policy/fiscal/oil shocks on household/firm-level data

Thank you!