

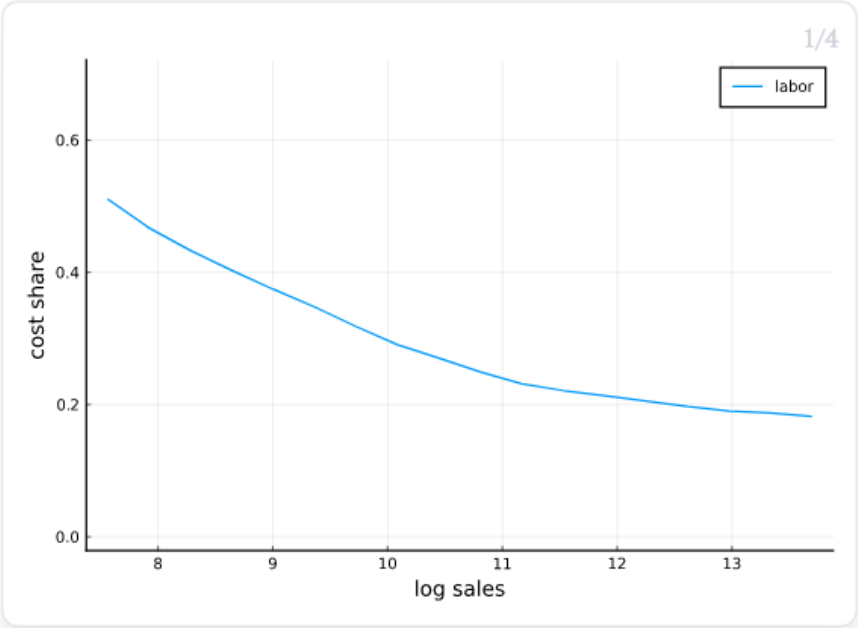
Why Do Larger Firms Have Lower Labor Shares?

L. Henry de Frahan, T. Lamadon, T. Meling, and M. Mogstad

NBER Spring meeting 2022

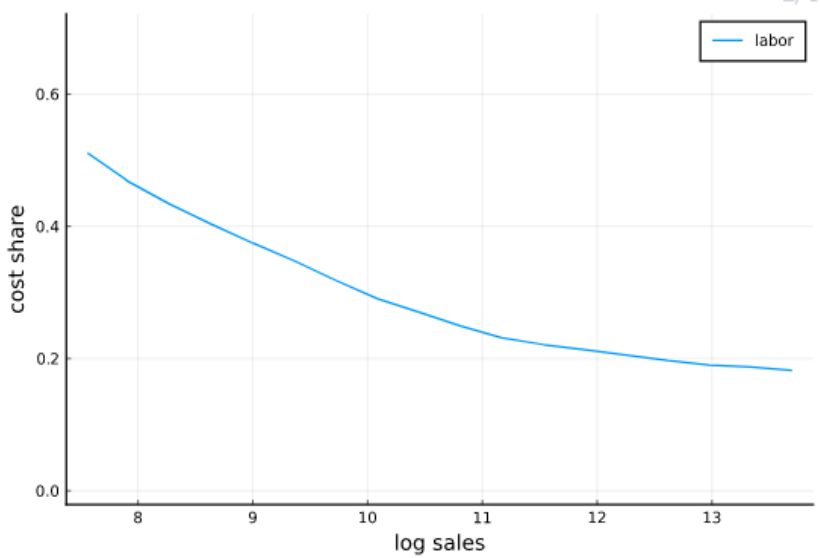
Motivation

- In the cross-section of firms in many countries (including the U.S.)
  - wage bill as a share of total cost is decreasing in sales
  - wage bill as a share of sales (or value added) is decreasing in sales
- In Norwegian data more



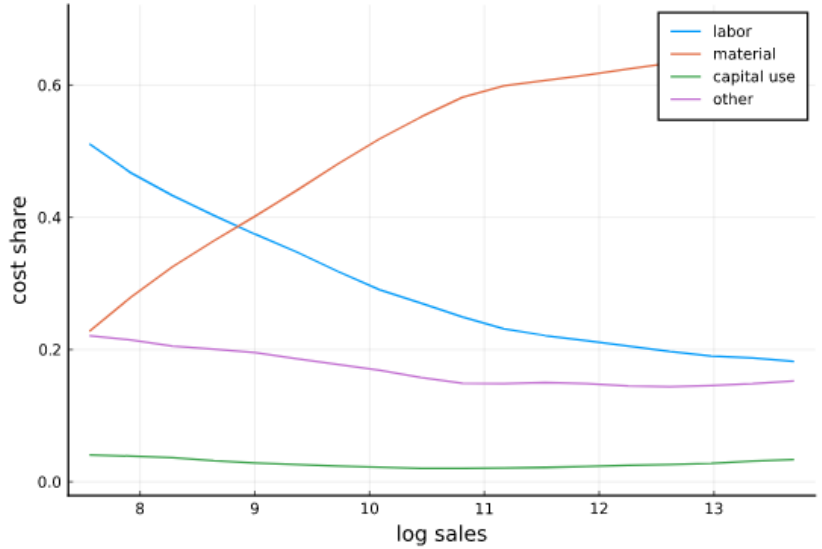
Motivation

- In the cross-section of firms in many countries (including the U.S.)
  - wage bill as a share of total cost is decreasing in sales
  - wage bill as a share of sales (or value added) is decreasing in sales
- In Norwegian data more



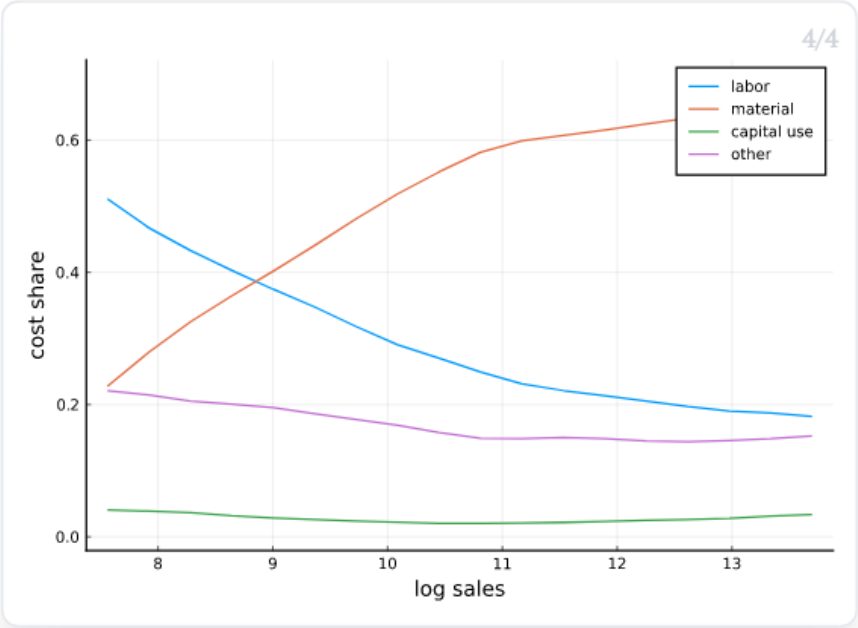
Motivation

- In the cross-section of firms in many countries (including the U.S.)
  - wage bill as a share of total cost is decreasing in sales
  - wage bill as a share of sales (or value added) is decreasing in sales
- In Norwegian data more



Motivation

- In the cross-section of firms in many countries (including the U.S.)
  - wage bill as a share of total cost is decreasing in sales
  - wage bill as a share of sales (or value added) is decreasing in sales
- In Norwegian data more



Channels we study

1. Upward-sloping labor supply curves
   - labor units become more expensive with employment size
   - labor supply becomes more inelastic, wage markdowns get larger
2. Within-firm scale effects in production
   - input elasticity of revenue increases (e.g. increasing product market markups)
   - substitution effect: relative marginal revenue of labor decreases
3. Fixed firm heterogeneity that correlates with TFP
   - production heterogeneity (e.g. automation)
   - labor costs (e.g. amenities)

Channels we study

1. Upward-sloping labor supply curves
   - labor units become more expensive with employment size
   - labor supply becomes more inelastic, wage markdowns get larger
2. Within-firm scale effects in production
   - input elasticity of revenue increases (e.g. increasing product market markups)
   - substitution effect: relative marginal revenue of labor decreases
3. Fixed firm heterogeneity that correlates with TFP
   - production heterogeneity (e.g. automation)
   - labor costs (e.g. amenities)

Importantly, different channels have different efficiency and policy implications.

Our approach

A model of heterogeneous monopsonic firms with

- a non-homothetic revenue function with multiple inputs
- upward-sloping labor supply curves with varying elasticity
- unobserved heterogeneity in TFP, production, and labor cost

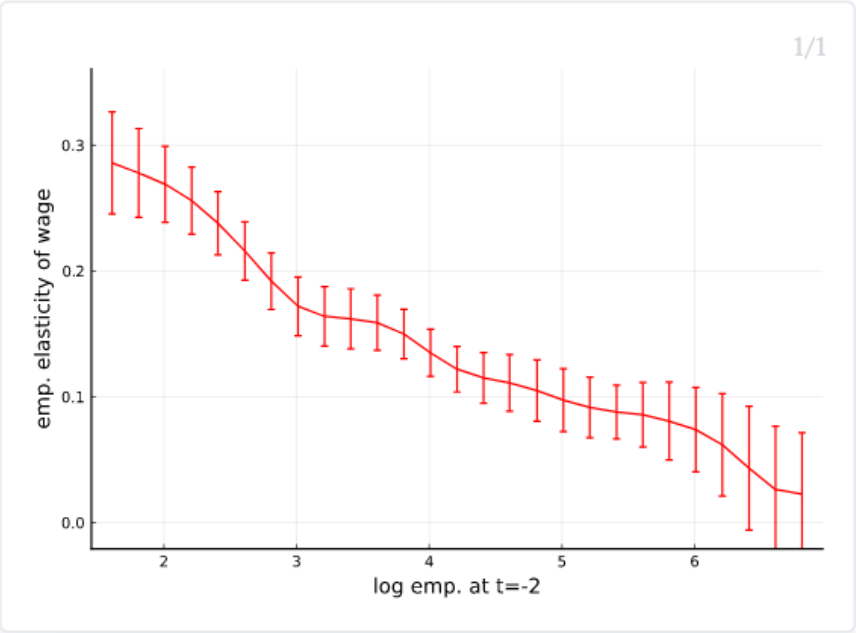
Properties and identification of the model

- express primitives as functions of elasticities in response to TFP (or product demand) shocks
- elasticities conditional on size allow for semi-parametric identification and estimation

Panel data on the universe of Norwegian firms

- rely on firm earnings statements (revenue, intermediate inputs, ...) and individual worker earnings
- use panel to run analysis within firm and worker
- use IV to isolate productivity shocks
  - internal IV: panel data restrictions
  - external IV: procurement auction wins

Preview of main findings



1. Employment elasticity of wages decreases as firms get larger
2. Labor faced by large firm looks more or less like a competitive market

Literature

- Karabarbounis and Neiman (2019); Kehrig and Vincent (2021); Autor et al. (2020); Grossman and Oberfield (2021)
- Lamadon, Mogstad and Setzler (2021); Kroft, Luo, Mogstad and Setzler (2021); Berger, Herkenhoff and Mongey (2021); Jarosch, Nimczik and Sorkin (2021)
- Acemoglu and Restrepo (2020); Hubmer and Restrepo (2021); Edmond, Midrigan and Xu (2021); Bilal and Lhuillier (2021); Baqaee and Farhi (2018)
- Hall (1988); De Loecker and Warzynski (2012); De Loecker, Eeckhout and Unger (2020); Chen, Macaluso and Hershbein (2021)

This talk

1. Model of the firm

2. Model implied elasticities and identification

3. Data and empirical approach

4. Results
   - labor supply curves
   - use model for broader lessons

5. Take-away and conclusion

Input factors

Firms are indexed by j and use 2 input factors
Consider firm j in isolation

1. Labor input L with cost

$$C_{jt} = \delta_j C(L_{jt})$$

   - $C(\cdot)$ allows for flexible upward-sloping curves
   - δ_j allows for firm-specific intercept (e.g. amenities)
   - $\delta_j C(\cdot)$ can be supported by a Kimball preference aggregator for workers

2. Other inputs V treated as expenditure (constant marginal cost)

Revenue function

$$Y_j(A_{jt}, L_{jt}, V_{jt}) = A_{jt} \cdot F\Big(\omega_j H(L_{jt}) + (1-\omega_j) G(V_{jt})\Big)$$

- A_{jt} is firm TFP
- L_{jt} and V_{jt} are labor and other inputs
- ω_j allows for firm-specific heterogeneity (e.g. better average labor quality)

Revenue function

$$Y_j(A_{jt}, L_{jt}, V_{jt}) = A_{jt} \cdot F\Big(\omega_j H(L_{jt}) + (1 - \omega_j) G(V_{jt})\Big)$$

- A_{jt} is firm TFP
- L_{jt} and V_{jt} are labor and other inputs
- ω_j allows for firm-specific heterogeneity (e.g. better average labor quality)

Properties

- nests CES and Cobb-Douglas production functions
- $H(\cdot)$ and $G(\cdot)$ allow for different returns to L and V
- $F(\cdot)$ allows for flexible input elasticity of revenue
- elasticity of substitution can vary with revenue
- additional inputs can be included (we currently focus on materials)

Firm problem

- Firms maximize profits

$$\Pi_j(A_{jt}) = \max_{L,V} Y_j(A_{jt}, V, L) - \delta_j C(L) - V$$

- Curvature in $C(L)$
  - taken into account when choosing L
  - allows for monopsony power
  - can vary with size
- Revenue function reflects
  - product market power
  - returns to scale in production

Identification overview

- Objects of interest: $C(\cdot), G(\cdot), H(\cdot), F(\cdot)$ and δ_j, ω_j for each j
- Ideal experiment: access to firm-specific shock to TFP A
- Using model structure, identify primitives from elasticities of observed quantities in three steps

1. ΔL and ΔW at each L
   - $\rightarrow$ trace out labor supply curve $C(\cdot)$
   - $\rightarrow$ levels pin down δ_j
2. ΔL and ΔV at each (L, V)
   - $\rightarrow$ give $H(\cdot)$ and $G(\cdot)$
   - $\rightarrow$ levels pin down ω_j
3. ΔY and $\Delta \Pi$ at each composite input level
   - $\rightarrow$ give $F(\cdot)$

Step 1: Labor supply curve

- Firm j hires labor on its labor supply curve

$$\delta_j C(L_{jt}) = W_{jt} \cdot L_{jt}$$

- Given a shock to A keeping $\delta_j C(\cdot)$ fixed, we get

$$\frac{\partial \log C(L_{jt})}{\partial \log L} = 1 + \frac{d_A \log W_{jt}}{d_A \log L_{jt}} = 1 + \varepsilon_{W,L}(L_{jt})$$

Step 1: Labor supply curve

- Firm j hires labor on its labor supply curve

$$\delta_j C(L_{jt}) = W_{jt} \cdot L_{jt}$$

- Given a shock to A keeping $\delta_j C(\cdot)$ fixed, we get

$$\frac{\partial \log C(L_{jt})}{\partial \log L} = 1 + \frac{d_A \log W_{jt}}{d_A \log L_{jt}} = 1 + \varepsilon_{W,L}(L_{jt})$$

- Conditional on L_j the elasticity does not depend on j
- Recovering such elasticity for each L_j identifies $C(\cdot)$
- $C(\cdot)$, wage bill and L_j pin down δ_j for each firm

Step 2: Input functions

- The ratio of FOCs gives

$$\frac{\omega_j}{1-\omega_j}\frac{\cancel{A_{jt}F'(\cdot)}H'(L_{jt})}{\cancel{A_{jt}F'(\cdot)}G'(V_{jt})} = \delta_j C'(L_{jt})$$

- The log change due to a shock to A implies

$$\frac{\partial \log H'(L)}{\partial \log L} - \varepsilon_{V,L}(V,L)\frac{\partial \log G'(V)}{\partial \log V} = \frac{\partial \log C'(L)}{\partial \log L}$$

Step 2: Input functions

- The ratio of FOCs gives

$$\frac{\omega_j}{1-\omega_j}\frac{\cancel{A_{jt}F'(\cdot)}H'(L_{jt})}{\cancel{A_{jt}F'(\cdot)}G'(V_{jt})} = \delta_j C'(L_{jt})$$

- The log change due to a shock to A implies

$$\frac{\partial \log H'(L)}{\partial \log L} - \varepsilon_{V,L}(V,L)\frac{\partial \log G'(V)}{\partial \log V} = \frac{\partial \log C'(L)}{\partial \log L}$$

  - Conditional on L_{jt} and V_{jt} the elasticity does not depend on j
  - Under some additional restrictions, this identifies $H(\cdot)$ and $G(\cdot)$
  - Given $H(\cdot), V(\cdot)$ and $C(\cdot)$ we identify ω_j from L_{jt}, V_{jt}

Norwegian data

Firm accounting data (1999-2018)

- all incorporated firms (main sample: 43k firms)
- full earnings statement (revenue, wage bill, material cost, depreciation and amortization, and other operating costs)

Worker data (1999-2018)

- all employment relationships (main sample: 4.5m workers)
- wage payments, hours worked, occupation, industry, and other demographics

Procurement data (2003-2019)

- scraped from Doffin.no, a website announcing government procurement contract wins in Norway
- winning firm, contract size, product type, and contract location
- standard name matching algorithm to link winner names to firm tax IDs
- 5k unique first-time winners in our sample

Empirical strategy

- We use the panel dimension to control for firm heterogeneity
- To get $\varepsilon_{W,L}(L)$, we rely on the first order expansion around a given time τ_0

$$\begin{aligned}\log W_{jt} &= \log W_{j\tau_0} + \log \delta_{jt} + \frac{\partial \log \tilde{W}(L_{j\tau_0})}{\partial \log L}(\log L_{it} - \log L_{j\tau_0}) + \epsilon_{jt} \\ &= \alpha_j + \lambda_t + \varepsilon_{W,L}(L_{j\tau_0}) \cdot \log L_{jt} + \epsilon_{jt}\end{aligned}$$

  - $W_{jt} = \delta_{jt}\tilde{W}(L_{jt}) \equiv \delta_{jt}C(L_{jt})/L_{jt}$
  - amenity term has a time-specific market component $\delta_{jt} = \delta_j \cdot \delta_{m(j),t}$
- A valid instrument Λ_{jt} conditional on size $L_{i\tau_0}$ requires

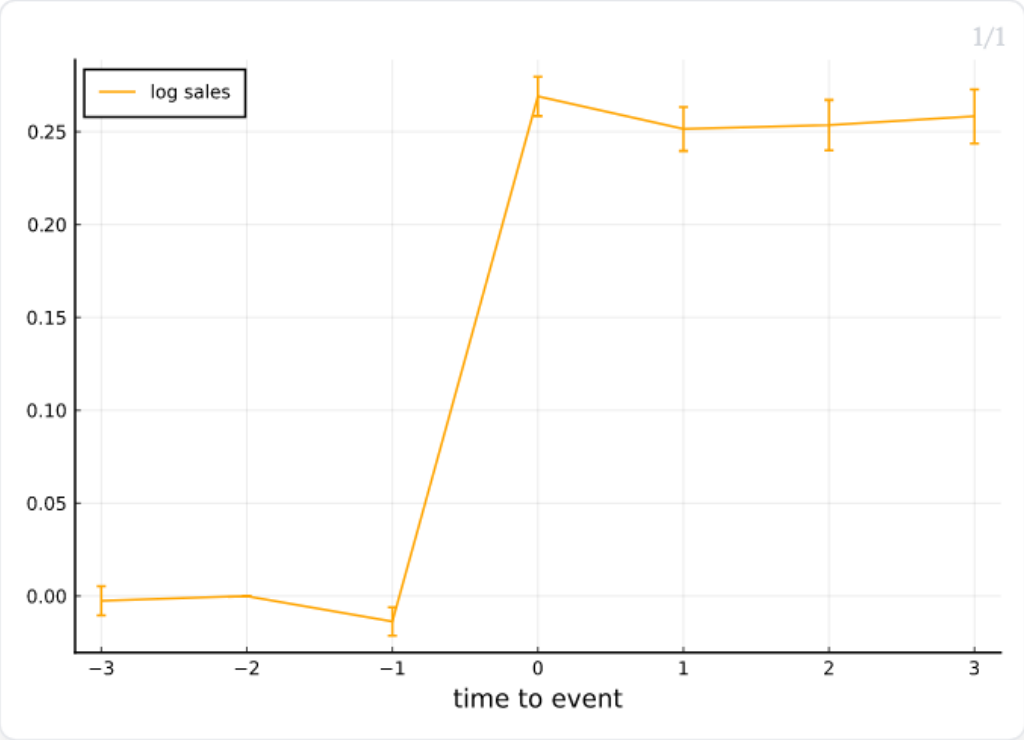
$$\mathbb{E}[\Lambda_{jt}\epsilon_{it}|L_{j\tau_0}, \alpha_j, \lambda_t] = 0, \ \mathbb{E}[\Lambda_{jt}\log L_{it}|L_{j\tau_0}, \alpha_j, \lambda_t] \neq 0$$

Internal instrument

- We instrument long-run changes in L_{jt} with short-run changes in Y_{jt} under the assumptions that
  - TFP is a unit-root process ▸ details
  - $\omega_j, \delta_j, C, H, G, G$ remain fixed
  - transitory errors in different variables are independent
- We run the analysis within cohorts of firms around a year t_0
  - $\Lambda_j = \mathbf{1}\{\Delta Y_{jt_0} > m\}$ where m is median sales growth in the market at t_0
  - L_{jt} is total employment
  - W_{jt} is the wage of stayers with full-year employment
  - include market-time-cohort FE and firm-cohort FE
  - we condition on L_{j,t_0-2} (i.e. $\tau_0 = t_0 - 2$) using a Gaussian kernel

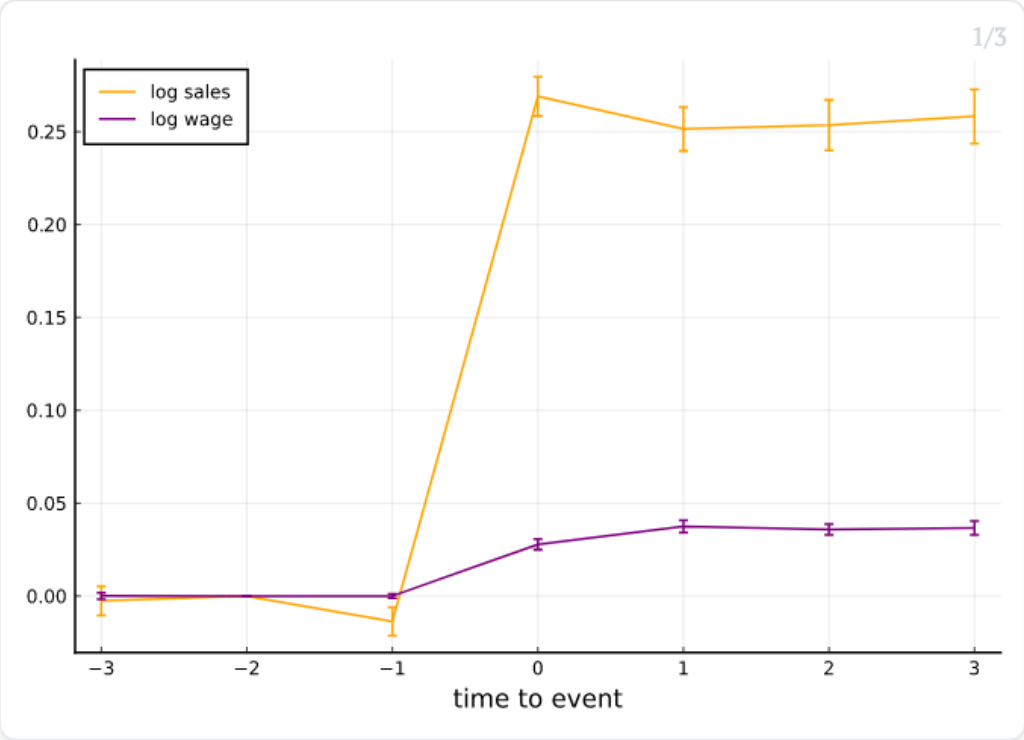
- The rank condition can be tested with the first stage
- Overidentification can be tested with pre-trends

Effect on sales - unconditional



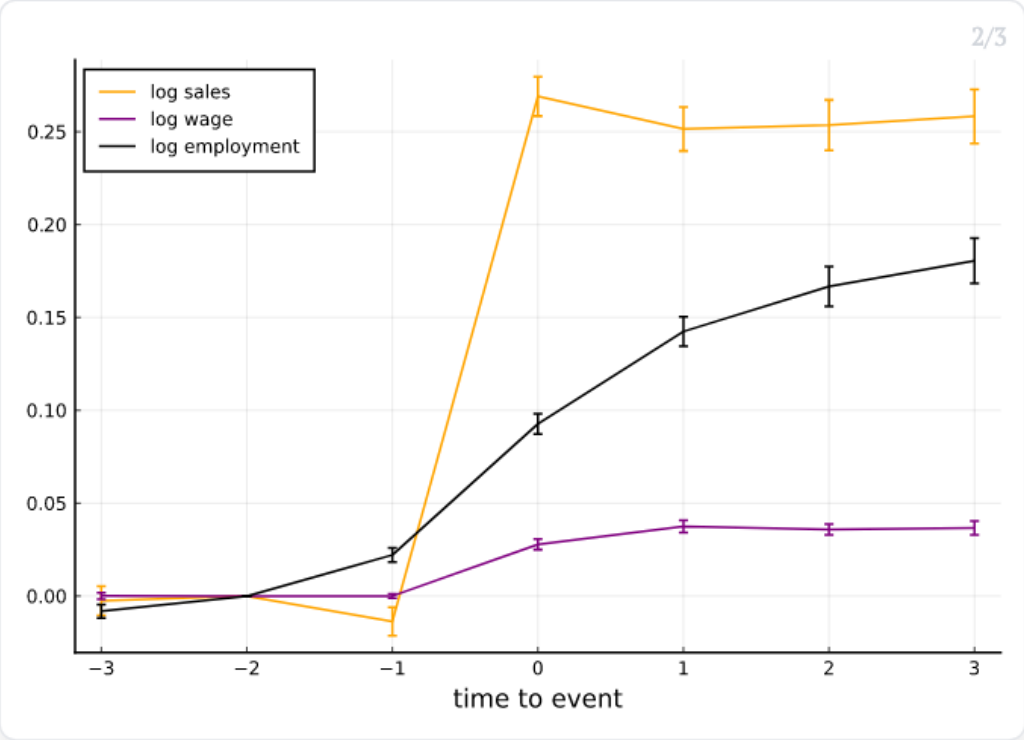
- Treatment assignment uses change between -1 and 0
- no apparent anticipation
- shock appears permanent

First stage and reduced form - unconditional



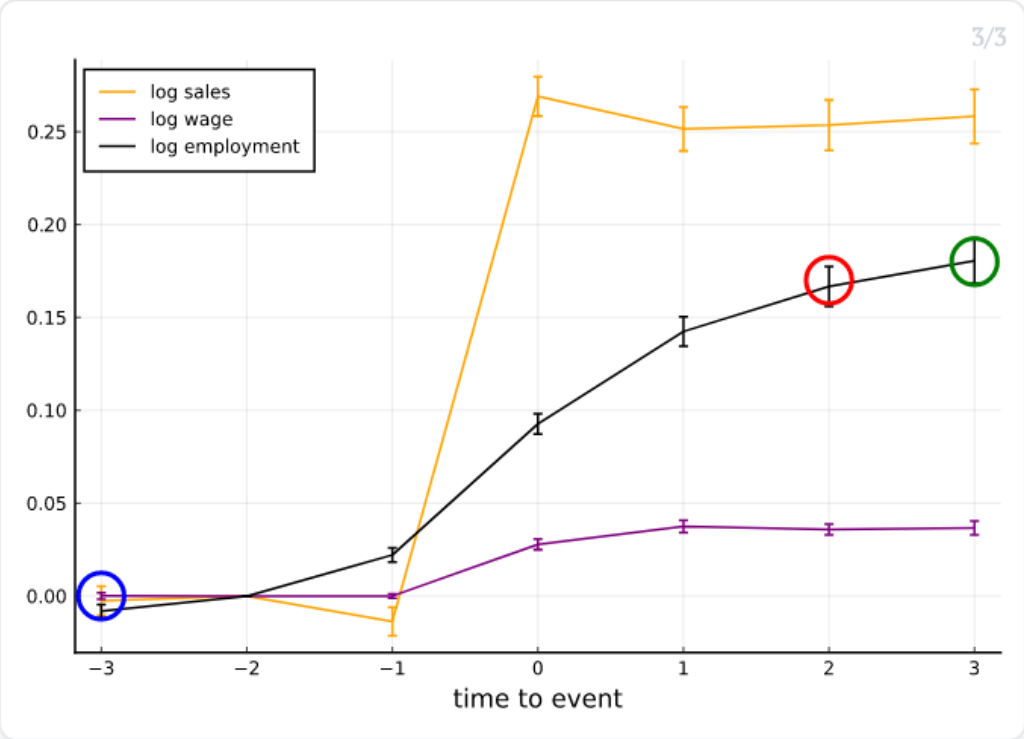
- We compare $t{=}{-}2$ to $t{=}2, 3$ for the effect of the shock ✅
- We compare $t{=}{-}2$ to $t{=}{-}3$ for the pre-trend ✅

First stage and reduced form - unconditional



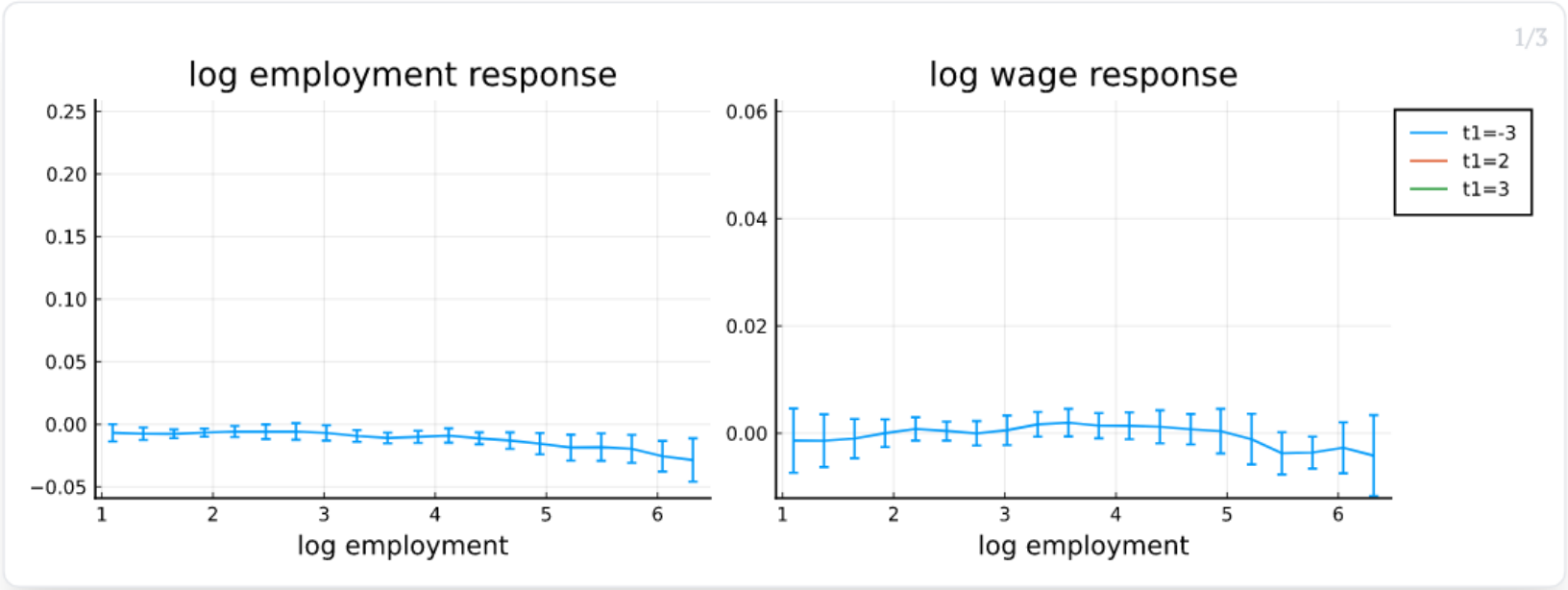
- We compare $t{=}{-}2$ to $t{=}2, 3$ for the effect of the shock ✅
- We compare $t{=}{-}2$ to $t{=}{-}3$ for the pre-trend ✅

First stage and reduced form - unconditional



- We compare $t{=}{-}2$ to $t{=}2, 3$ for the effect of the shock ✅
- We compare $t{=}{-}2$ to $t{=}{-}3$ for the pre-trend ✅

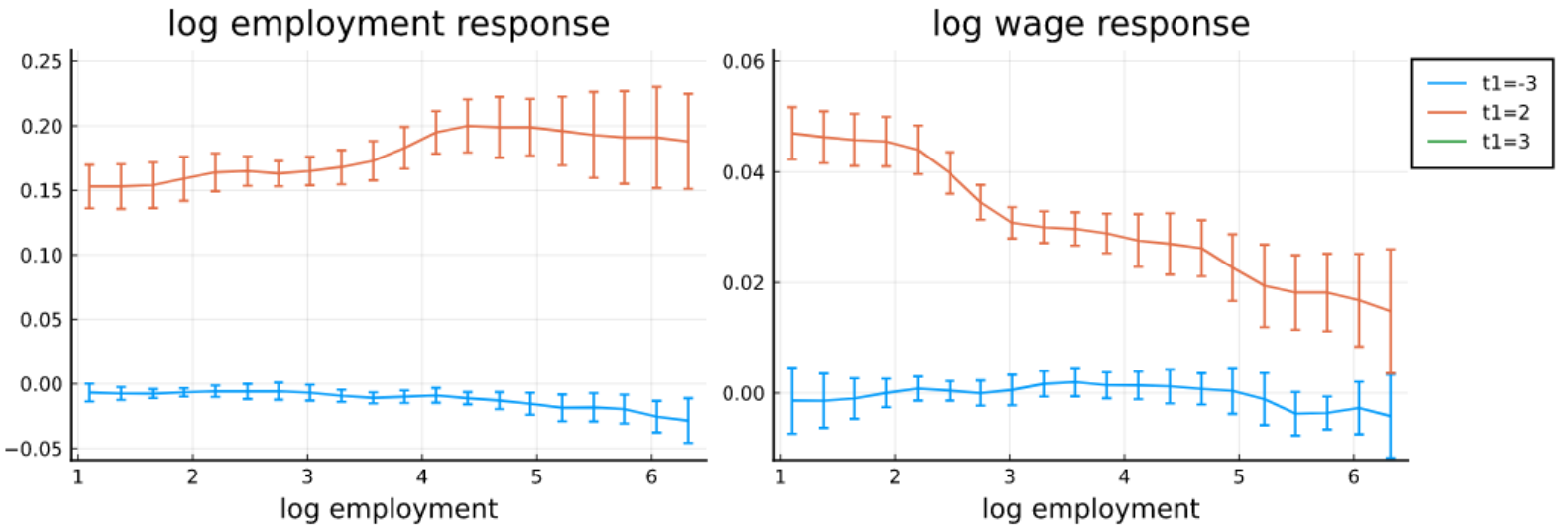
First stage and reduced form - conditional on size



- Employment response doesn't change much with firm size
- Wage response decreases as firms get larger
- All first stage regressions are strong

Sales Distr.

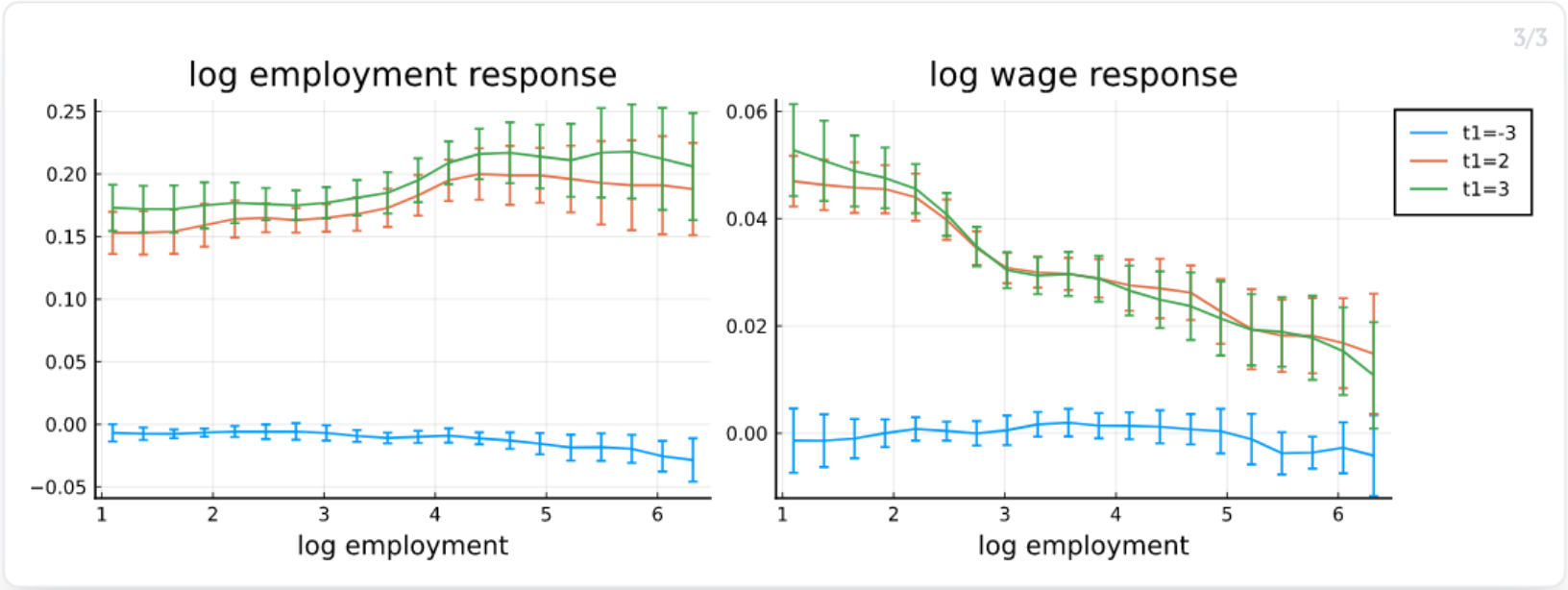
First stage and reduced form - conditional on size



- Employment response doesn't change much with firm size
- Wage response decreases as firms get larger
- All first stage regressions are strong

Sales Distr.

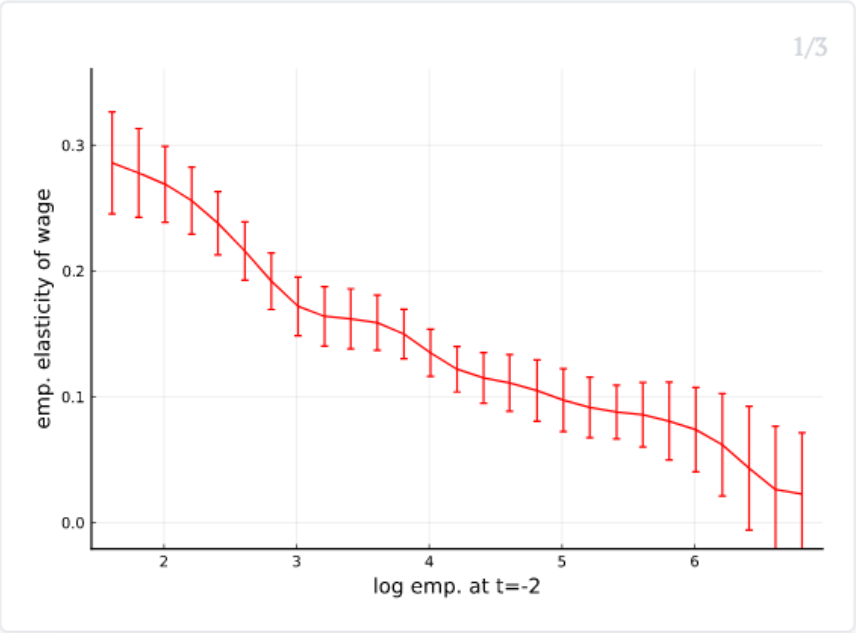
First stage and reduced form - conditional on size



- Employment response doesn't change much with firm size
- Wage response decreases as firms get larger
- All first stage regressions are strong

Sales Distr.

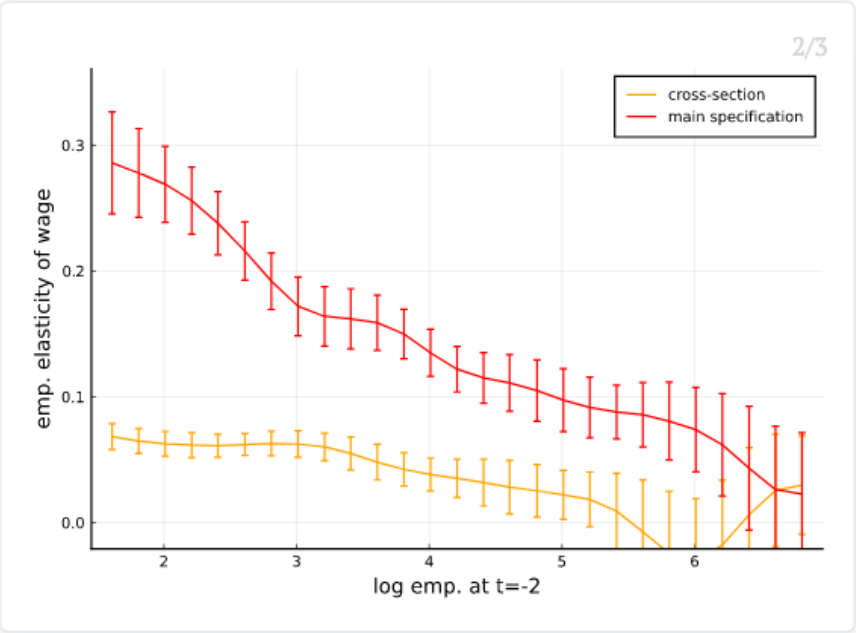
Size elasticity of wages - conditional on size



Main findings

1. Employment elasticity of wages decreases as firms get larger
2. Level is much larger than estimates from the cross-section

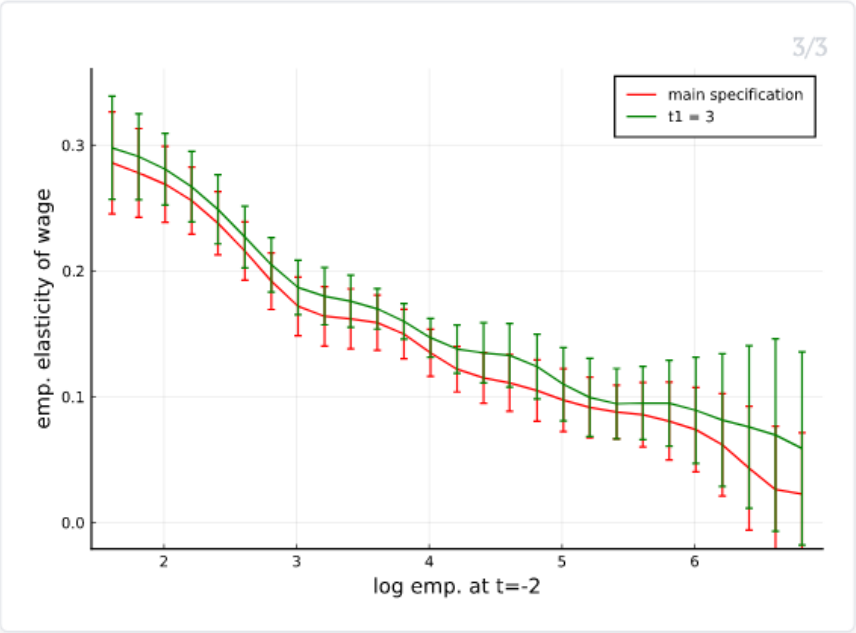
Size elasticity of wages - conditional on size



Main findings

1. Employment elasticity of wages decreases as firms get larger
2. Level is much larger than estimates from the cross-section

Size elasticity of wages - conditional on size



Main findings

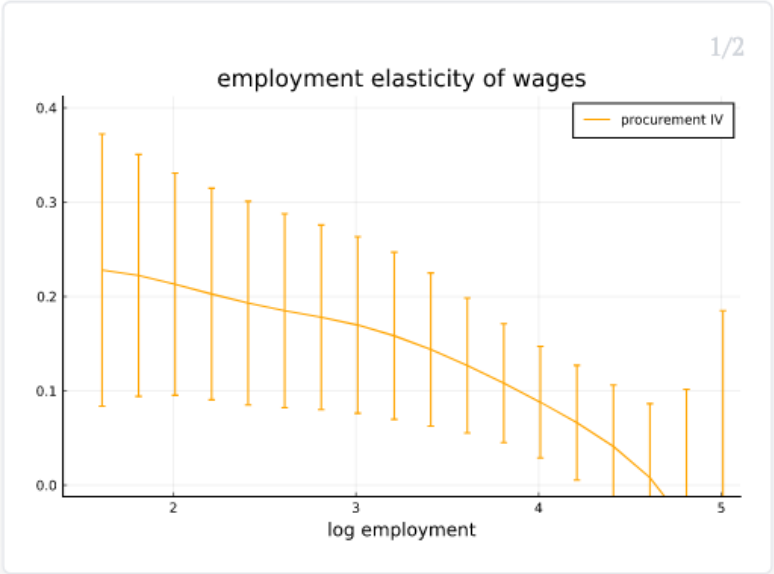
1. Employment elasticity of wages decreases as firms get larger
2. Level is much larger than estimates from the cross-section

Size elasticity of wages - robustness

1. Control for market size
   - we estimate elasticities by firm size and overall market size
2. Control for worker composition in L
   - we compute AKM worker effects (including firm-time FE) and construct effective firm labor size that takes into account worker heterogeneity
3. Run separately by education groups
   - estimate on college and non-college workers separately
4. Run for low and high labor-to-material ratio firms
5. Focus on small shocks
6. Use hourly wage rather than yearly earnings

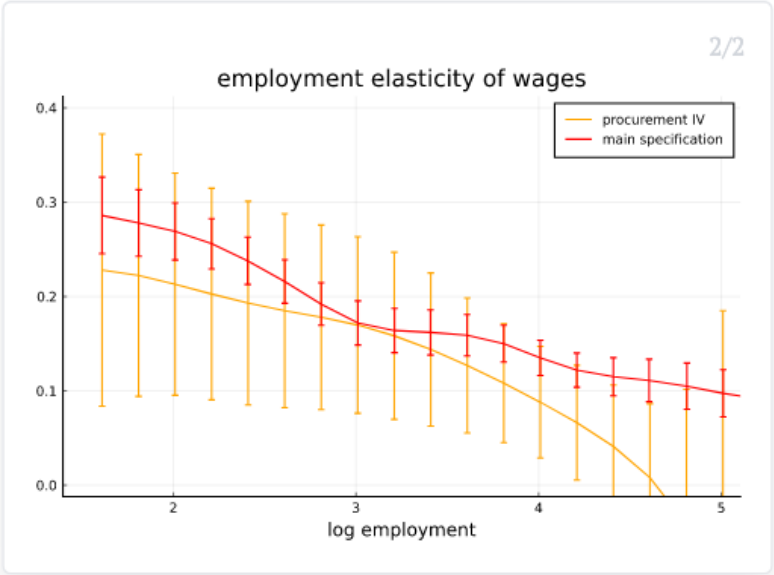
External IV

- Within each cohort t_0
  - treatment group includes firms which win their first procurement auction in t_0
  - control group includes firms which win their first procurement auction later than $t_0 + 3$ or never
- Other elements of the analysis remain the same



External IV

- Within each cohort t_0
  - treatment group includes firms which win their first procurement auction in t_0
  - control group includes firms which win their first procurement auction later than $t_0 + 3$ or never
- Other elements of the analysis remain the same



Model and counterfactuals

Parametrizing the model

We choose simple forms for $C(\cdot), H(\cdot), G(\cdot)$ and $F(\cdot)$:

- Labor costs: the Kimball aggregator from Kleenow and Willis (2016) fit

$$C(L) = L \exp\left(\frac{a}{b} L^b\right)$$

- We allow for different returns and a common scale effect in H() and G() fit

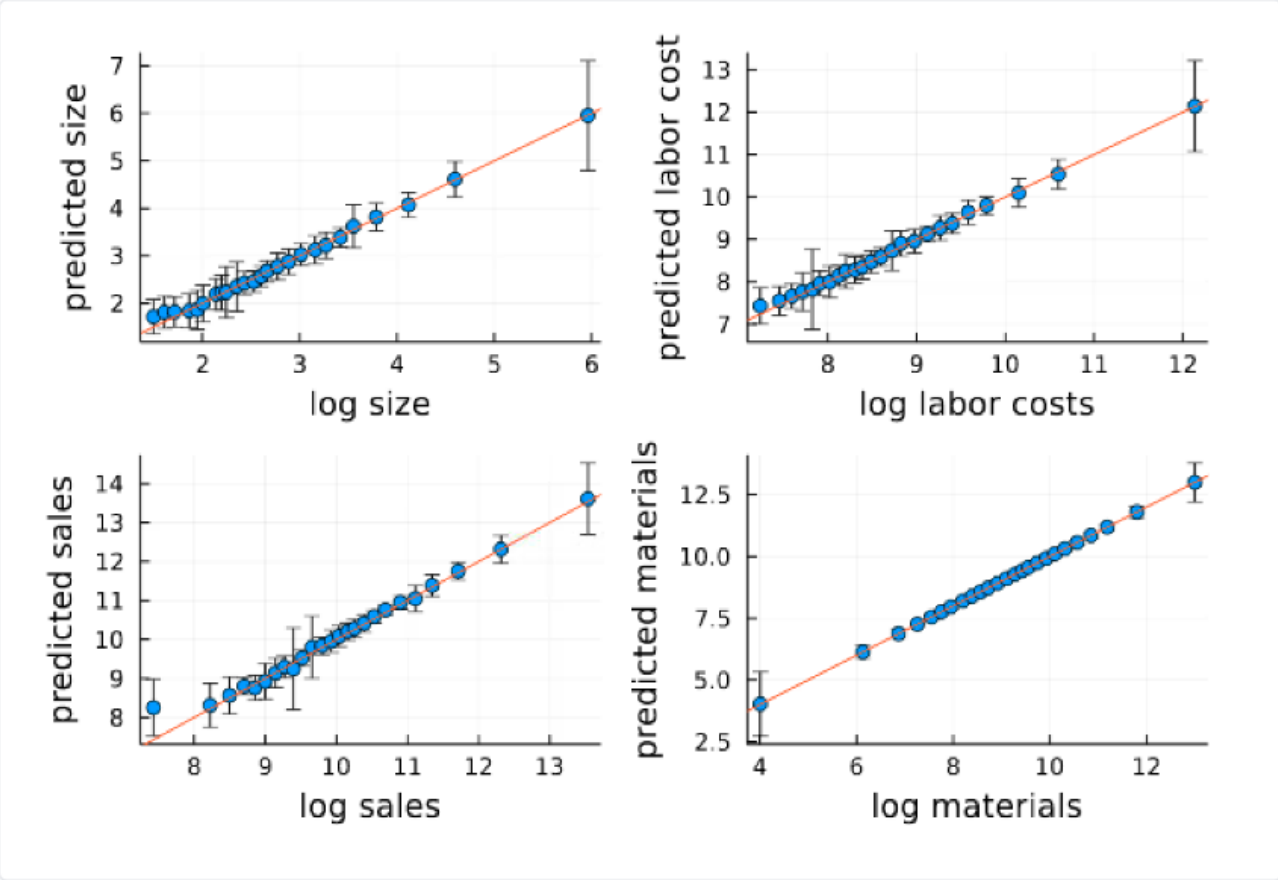
$$G(V) = V^{\beta_V} \qquad H(L) = e^{\frac{\beta_L}{\gamma} L^\gamma}$$

  - as $\gamma \to 0$ we get $(1-\omega_j)V^{\beta_V} + \omega_j L^{\beta_L}$ and nest the CES, and hence also Cobb-Douglas
  - γ allows for the complementarity between L and V to vary with scale

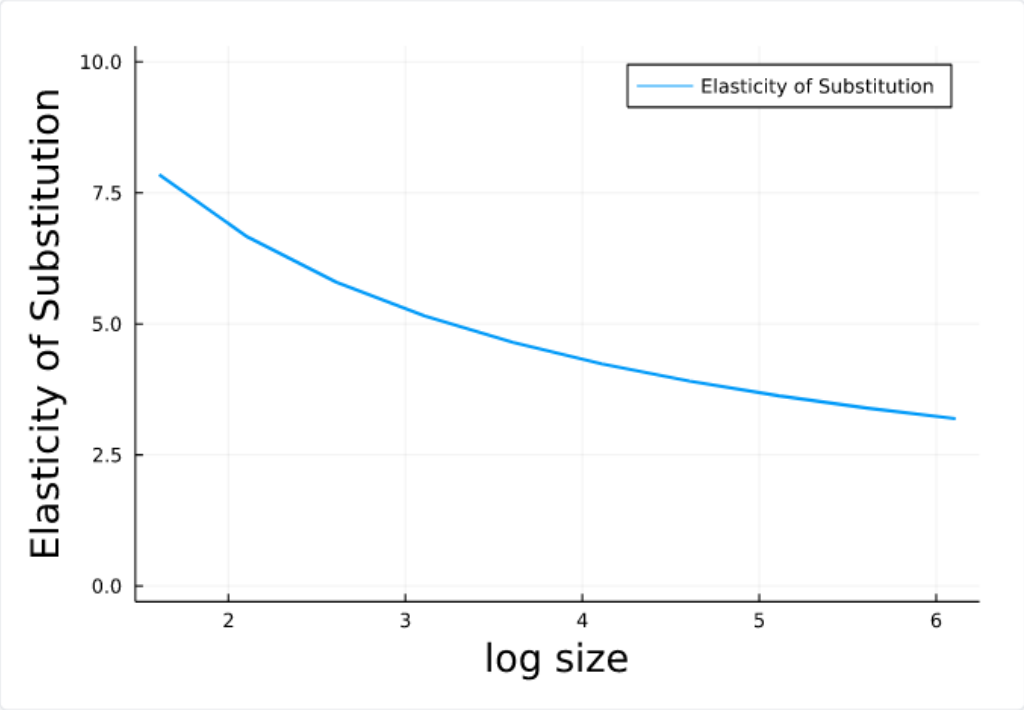
- $F(\cdot)$ use a constant elasticity fit

$$F(S) = S^{1-\eta}$$

Model fit in the cross-section



Elasticity of substitution between L and V



Main decomposition

Table 1: Actual and counterfactual responses to TFP shocks.

	dS_L/dVA	dV/dL	dW/dL	Avg. μ	Stdev μ
Cross-section					
	-0.10	1.02	0.03	0.03	0.01
	(0.00)	(0.01)	(0.00)	.	.
Pooled IV					
	-0.06	1.59	0.20	0.19	0.05
	(0.01)	(0.03)	(0.01)	.	.
Model Implied					
Total Effect	-0.08	1.60	0.21	0.17	0.04
Iso-elastic LS	-0.15	1.85	0.21	0.17	0.04
Homogenous & iso-elastic LS	-0.14	1.84	0.21	0.17	0.00
Perfect Competition	0.03	0.85	0.00	0.00	0.00

Note: $S_L \equiv \frac{C}{VA}$ is the log labor share, C is log total wagebill, VA is log value-added, W is log wages, L is log labor, $\mu \equiv 1 - W/MPL = \frac{\varepsilon_{w,L}}{1+\varepsilon_{w,L}}$ is the mark-down on wages. The table displays ratios of log responses to a small TFP shock. The cross-sectional variance in mark-downs is based on regressing $log(W)$ on a quadratic in $log(L)$. All standard errors are clustered at two-levels: firm ID and year. All regressions are within market $\times$ year.

Preliminary take-away

- Labor costs
  - costs are upward-sloping $\longrightarrow$ substitute away from L ✅
  - wage elasticities decrease $\longrightarrow$ markdowns don't increase with size ❌
- Production (preliminary)
  - relative returns to labor and material: appear to favor labor ❌
  - ω_j decreases and larger elasticity of substitution suggests relatively less labor ✅
  - revenue elasticity increases with size: not sure yet ❓

Conclusion and roadmap

- What we have done
  - propose parsimonious model of firm production
  - separately identify and estimate properties of labor cost and production
  - use external and internal instruments
  - main finding: elasticity of wages decreases with employment
- What to do next
  - put the model to work: quantitative counterfactuals
  - study other inputs
  - ...

Appendix Slides

Internal instrument - details

- Consider the following unit root passthrough model

$$\begin{aligned}
\tilde{y}_{jt} &= \alpha_i + y^p_{i\tau_0} + \frac{dy}{da} \cdot da_{jt} + v_{jt} \\
\tilde{w}_{it} &= \delta_i + w^p_{i\tau_0} + \frac{dw}{da} \cdot da_{jt} + \xi_{it} \\
da_{jt} &= da_{jt-1} + \nu_{it}
\end{aligned}$$

- When $v_{it}, \nu_{it}, \xi_{it}, u_{it}$ are independent innovations, we have

$$E\left[(y_{it} - y_{it-1})\left(w_{it+p} - w_{it-q} - \frac{dw}{dy}(y_{it+p} - y_{it-q})\right)\right] = 0$$

- Typical internal instrument approach as in Guiso, Pistaferi and Schivardi (2003), but also in income and consumption literature

Internal instrument - details 2

- It also implies

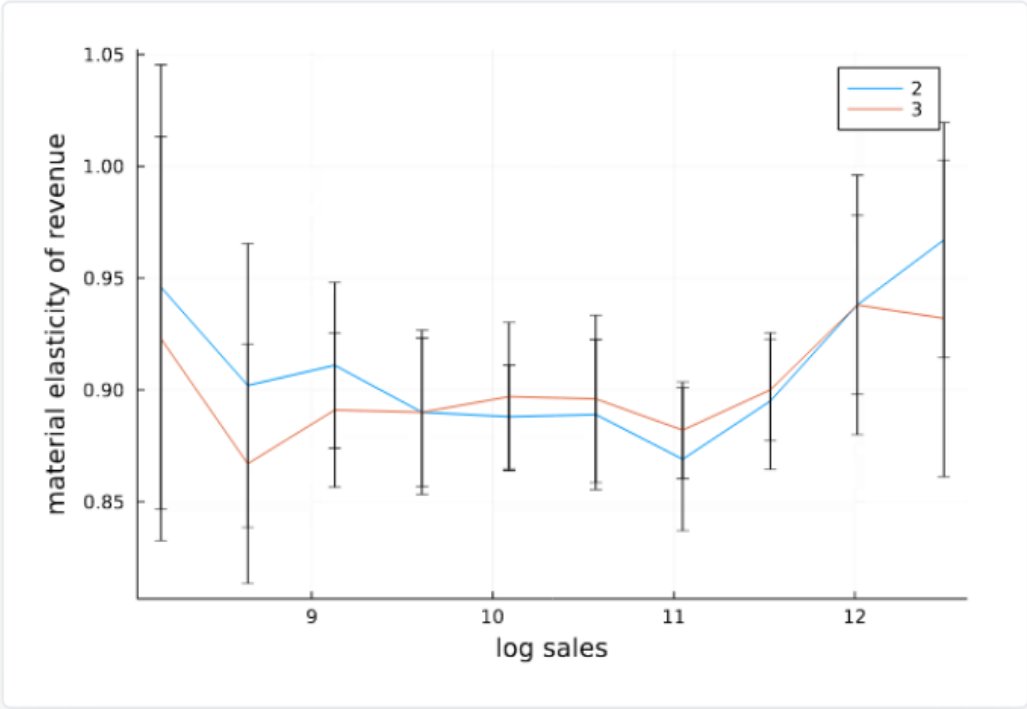
$$
\begin{aligned}
E\left[y_{it+p} - y_{it-q} | (y_{it} - y_{it-1}) > m\right] = \\
E\left[\nu_{it} | \nu_{it} + \epsilon_{it} - \epsilon_{it-1} > m\right] \\
E\left[w_{it+p} - w_{it-q} | (w_{it} - w_{it-1}) > m\right] = \\
\frac{dw}{dy} E\left[\nu_{it} | \nu_{it} + \epsilon_{it} - \epsilon_{it-1} > m\right]
\end{aligned}
$$

- Hence we can express our parameter as the ratio of two DiD

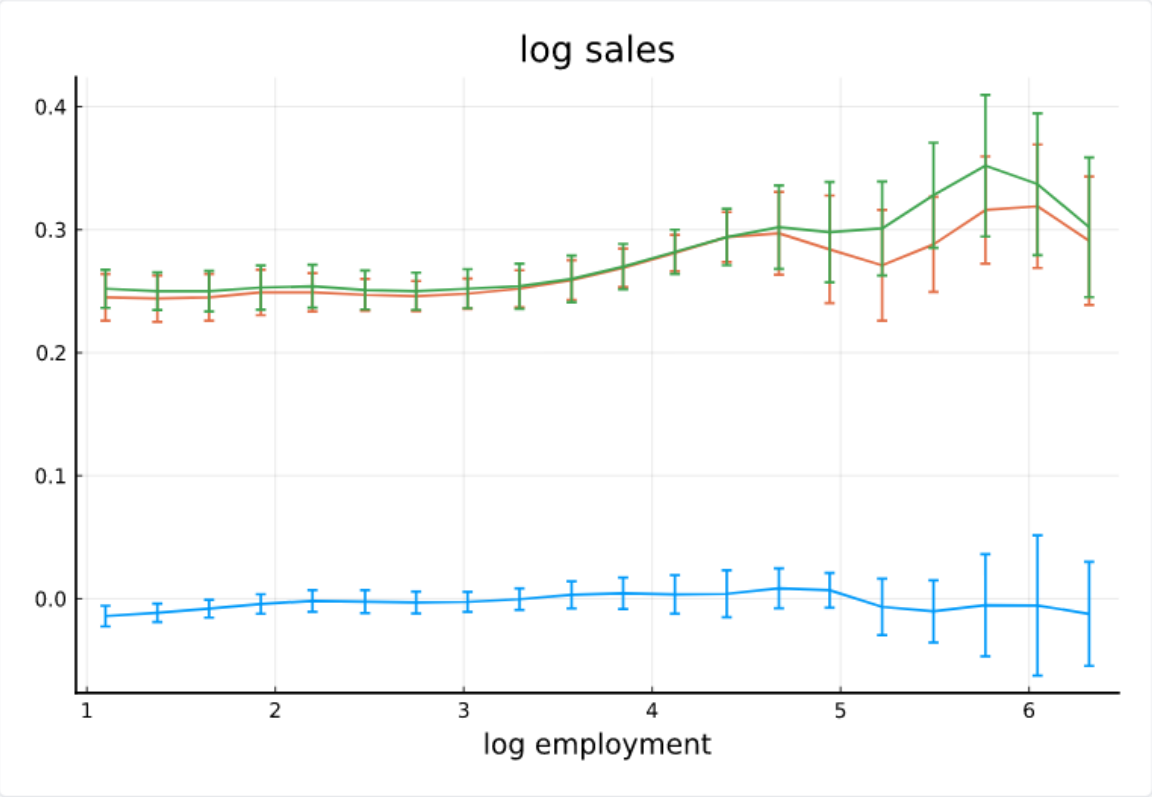
$$
\frac{dw}{dy} = \frac{E\left[w_{it+p} - w_{it-q} | (y_{it} - y_{it-1}) > m\right] - E\left[w_{it+p} - w_{it-q} | (y_{it} - y_{it-1}) < m\right]}{E\left[y_{it+p} - y_{it-q} | (y_{it} - y_{it-1}) > m\right] - E\left[y_{it+p} - y_{it-q} | (y_{it} - y_{it-1}) < m\right]}
$$

- Allows us to provide a graphical representation of the design

Revenue elasticity of materials

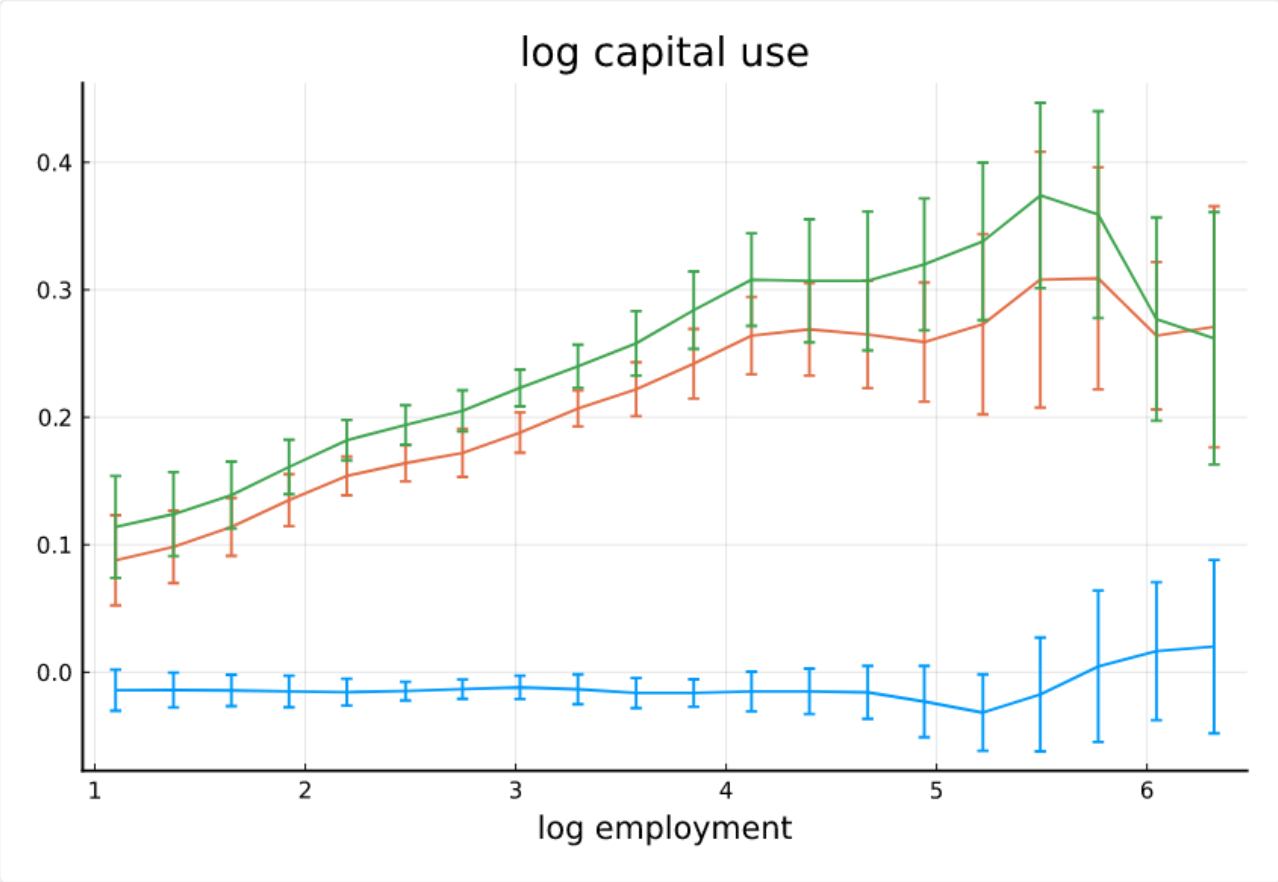


Sales response - by size



back

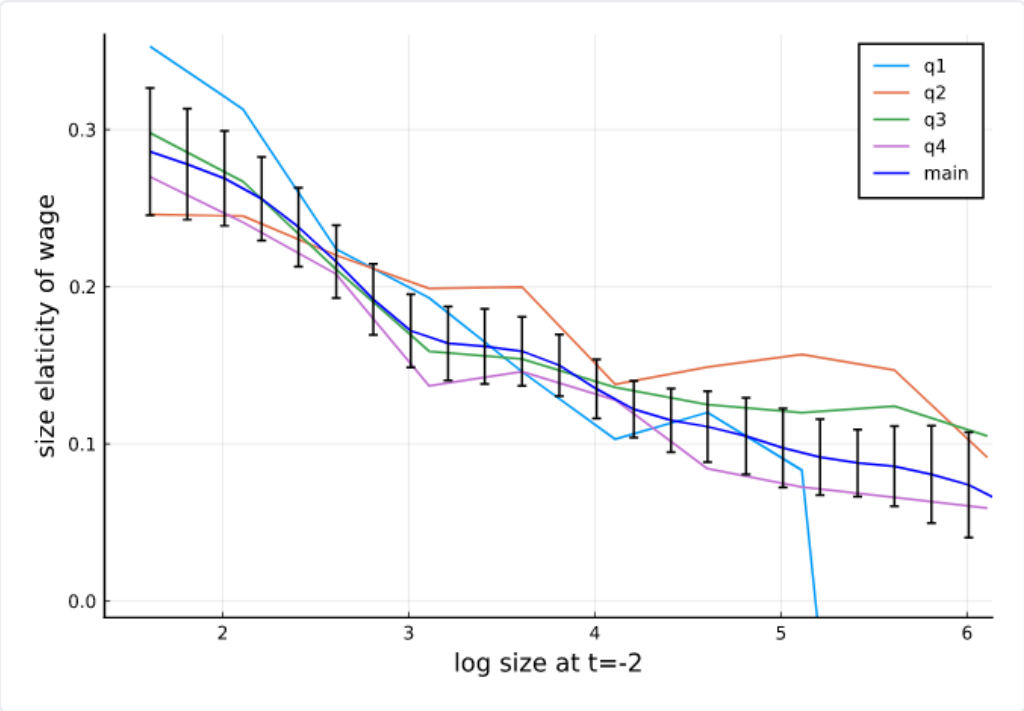
Distribution and treatment probability



back

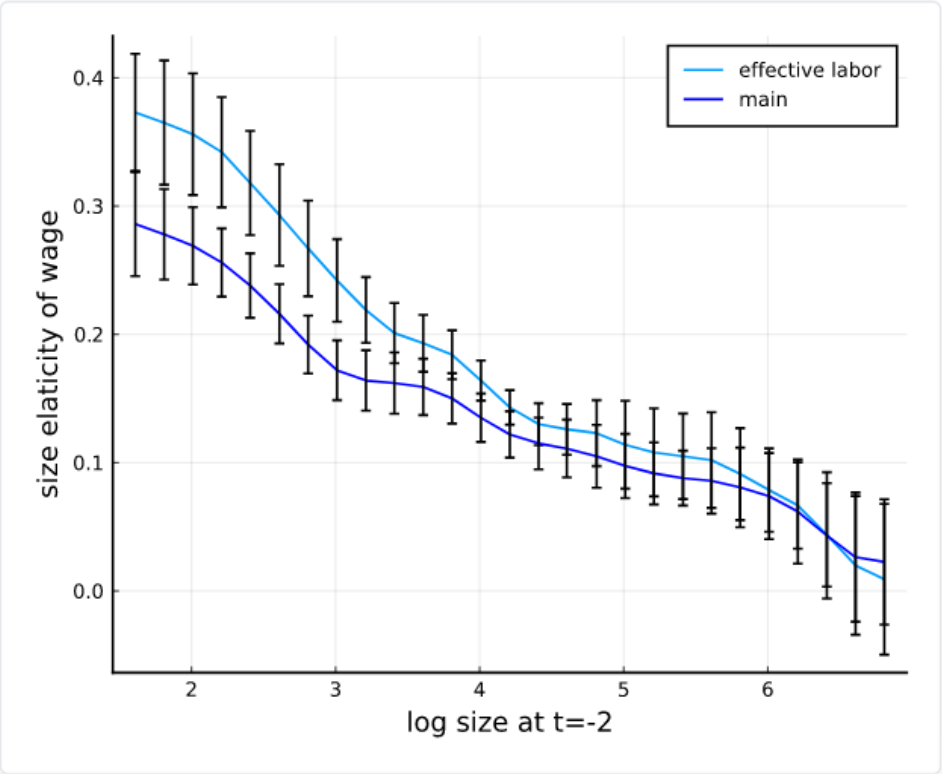
Robustness - by market size

- We run the estimator in each quartile of initial market size



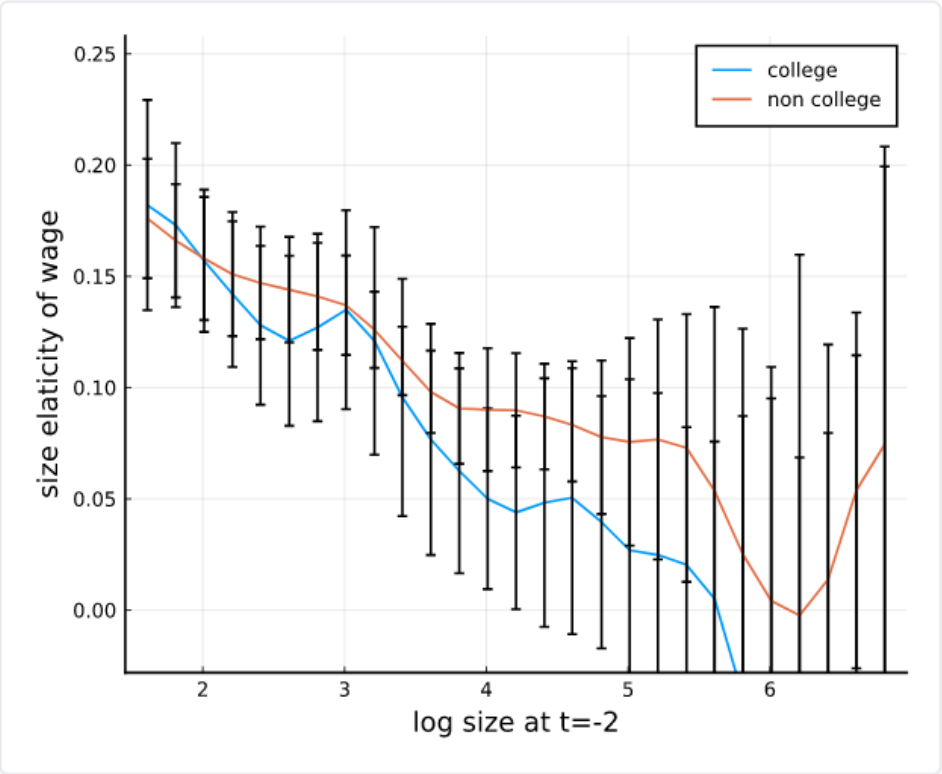
back

Robustness - using effective labor



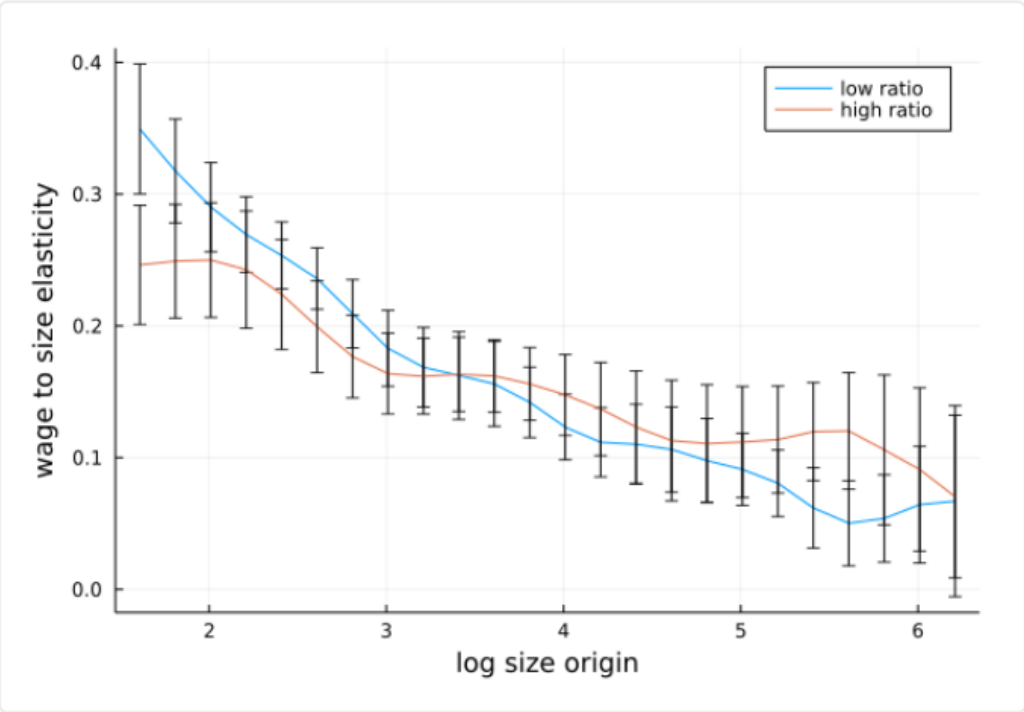
back

Robustness - by education



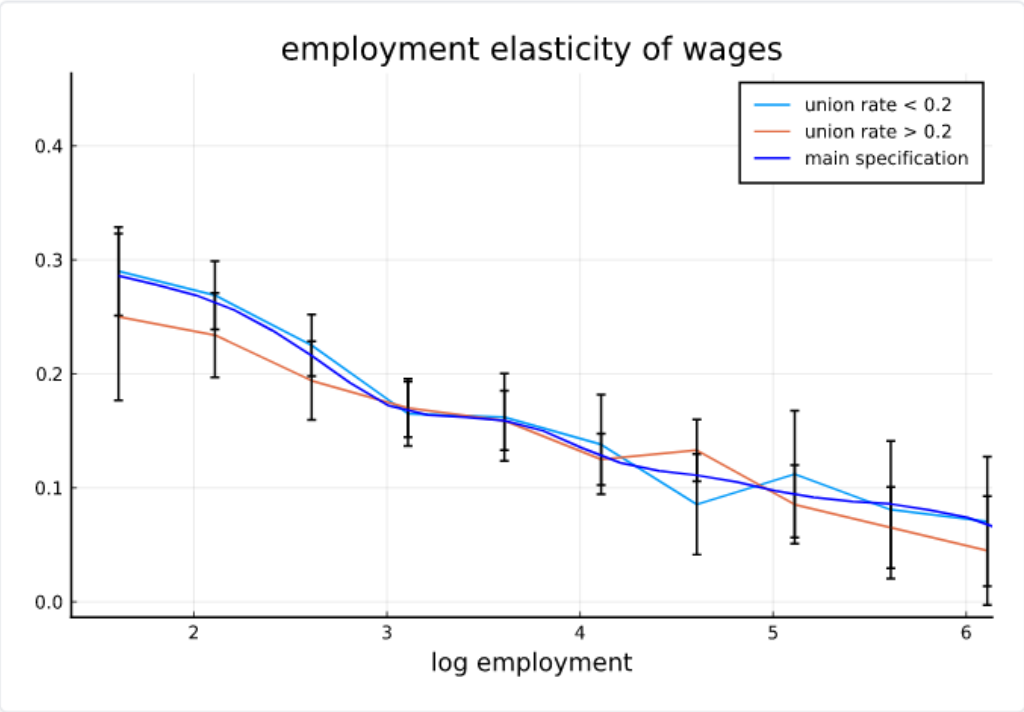
back

Robustness - by labor-to-material ratio



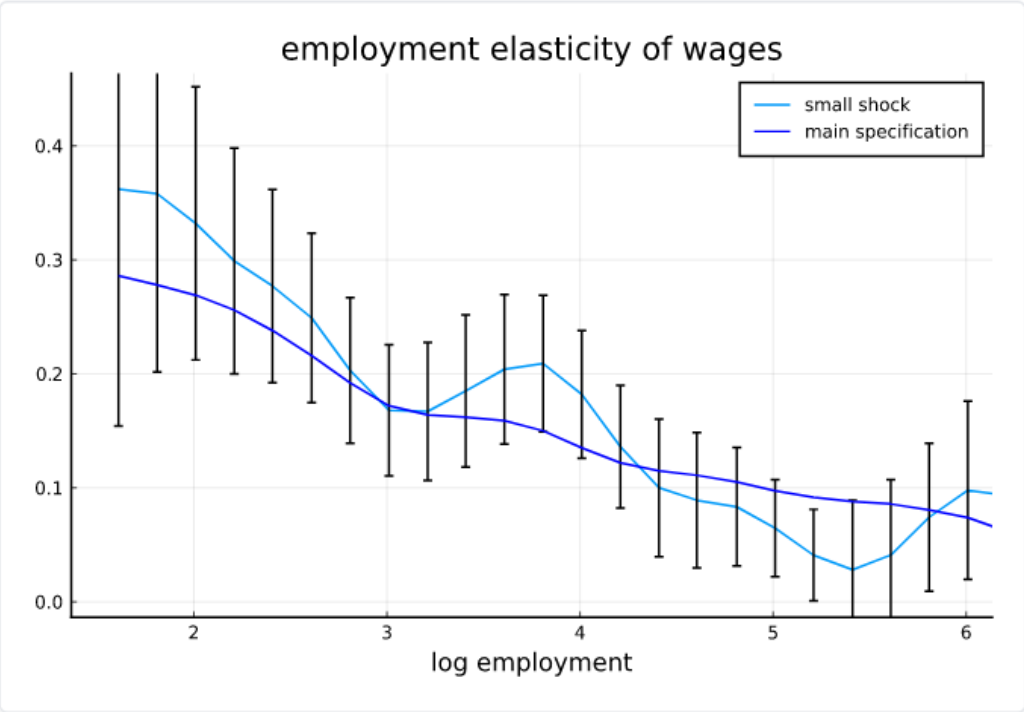
back

Robustness - by unionization ratio



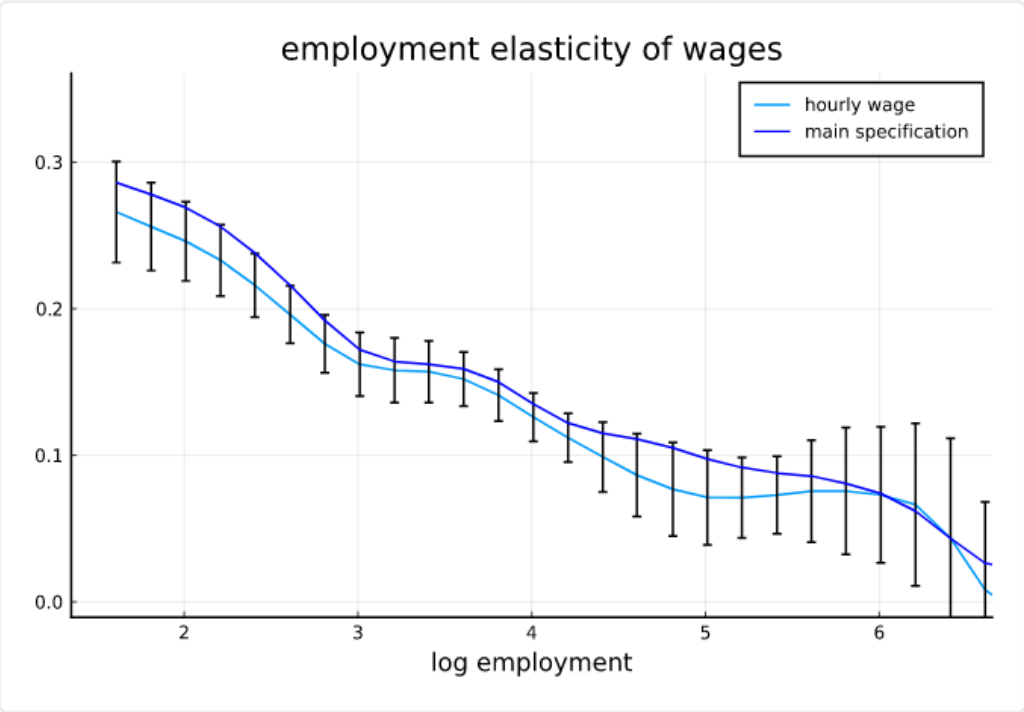
back

Robustness - small shocks



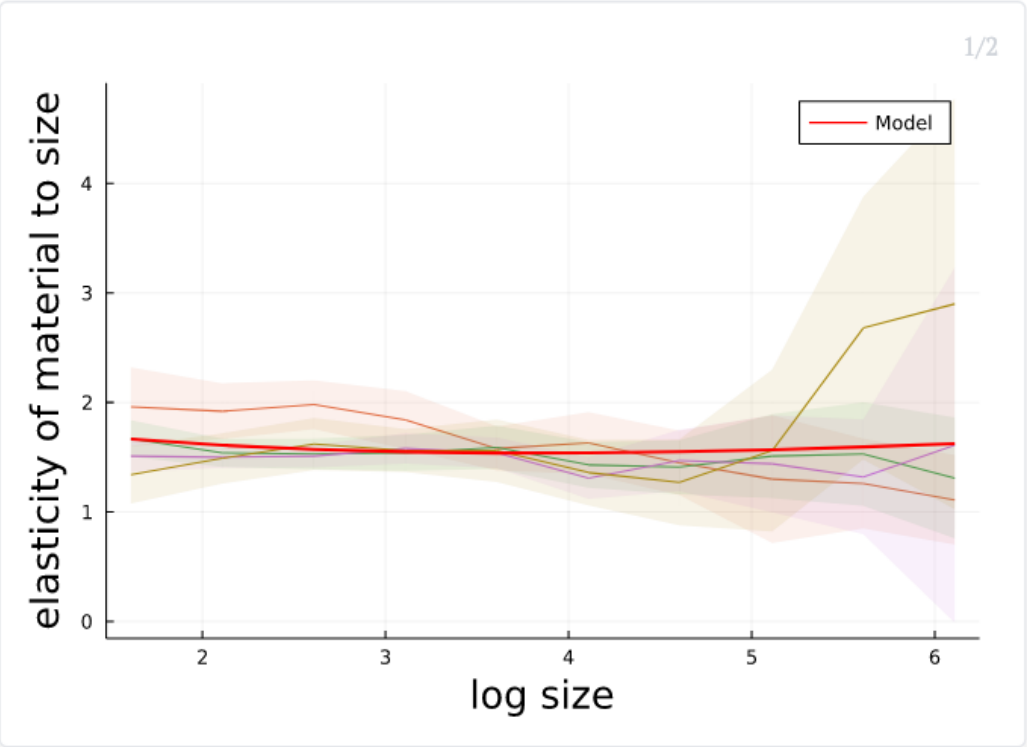
back

Robustness - hourly wage



back

Production function fit



back