Policy Thresholds as Growth Barriers: Theory and Evidence from a Payroll Tax Notch

Youssef Benzarti[1] **Jarkko Harju**[2] **Sami Jysmä**[2,3]

[1] UC Santa Barbara
[2] Tampere University
[3] Labour Institute for Economic Research (Labore)

NBER Public Economics Program Meeting, Spring 2026

April 16, 2026

Introduction: Motivation

- Discrete policy thresholds are pervasive in tax and regulatory systems.
  - Tax notches, welfare cliffs, size-based regulation, reporting requirements...
- Bunching literature: leverages local responses to identify structural parameters.
- **Separate question:** How do thresholds distort the *overall distribution*?
  - Static theory: intensive/extensive margin responses can shift mass.
  - **Further dynamic channel:** Thresholds that persistently discourage crossing act as **growth barriers** $\rightarrow$ effects compound $\rightarrow$ distortions throughout the distribution.
- Structural models show dynamic effects matter (e.g. Garicano et al. 2016, Gourio & Roys 2014, Aghion et al. 2025), but are setting-specific.

$\rightarrow$ We lack a **general and portable framework** for these distributional impacts.

This Paper

1. **Theory:** Portable framework for how growth barriers reshape the full distribution.
   - Counterfactual characterized by a single sufficient statistic p, identifiable while the notch is active.

2. **Methodology:** *Switcher difference-in-differences* estimator.
   - Standard DiD is biased: units that refrain from crossing are treated but classified as untreated.

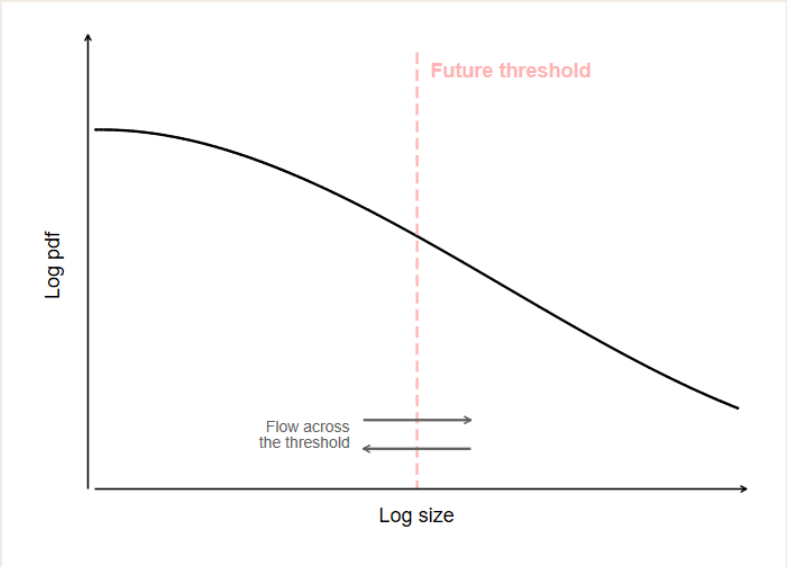
3. **Empirics:** Abolition of a firm-level payroll tax notch in Finland.
   - $\sim$18% fewer firms above threshold; effects extend far into the distribution.
   - Pre-reform counterfactual closely matches post-reform distribution.
   - Repeal increased firm scale by $\sim$10% (employment, capital, value added).

Theory: Overview

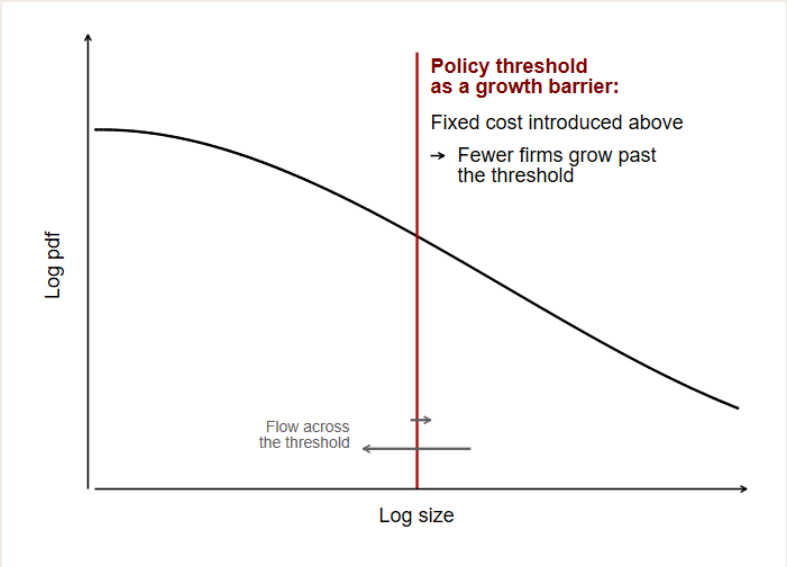
Goal: Portable framework for studying how growth barriers reshape the full distribution.

- Setup: Size (income, wealth, no.employees, capital...) follows a continuous time Itô diffusion with a reflective lower bound.
- Growth barrier: Anything (e.g. a policy notch) that reduces the probability of units crossing it.
- Result: A simple testable sufficient statistics representation of the counterfactual distribution, when this is the only friction.

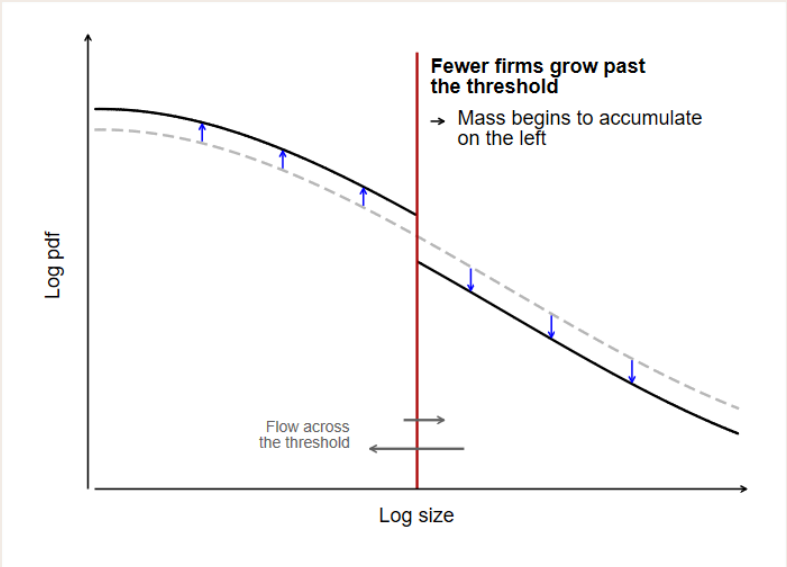
Theory: Undistorted Distribution



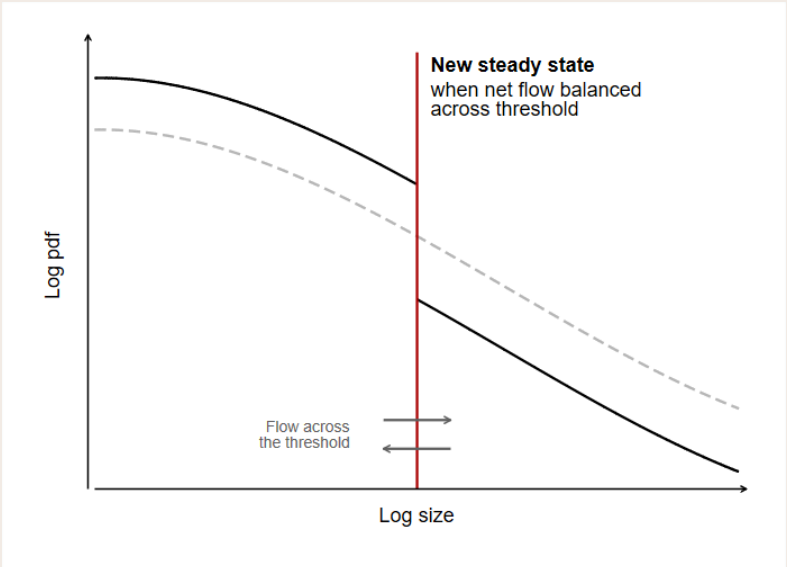
Theory: Notch Introduction



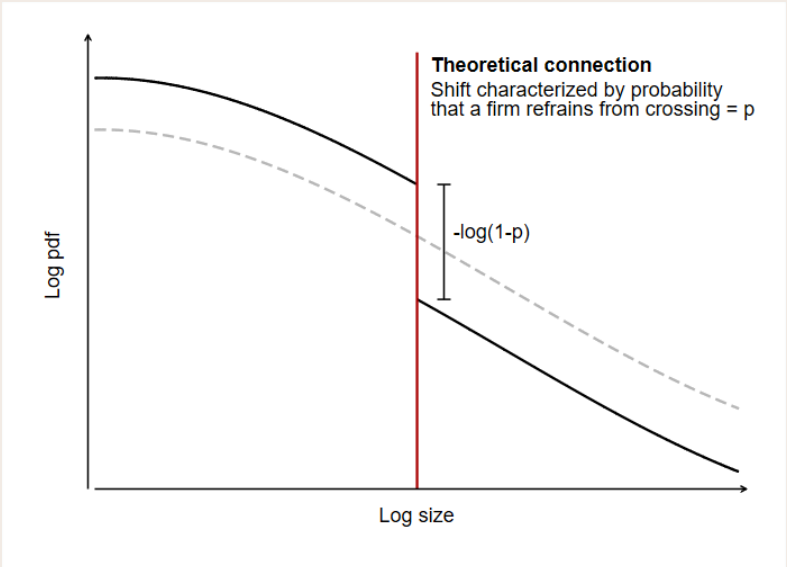
Theory: Transition



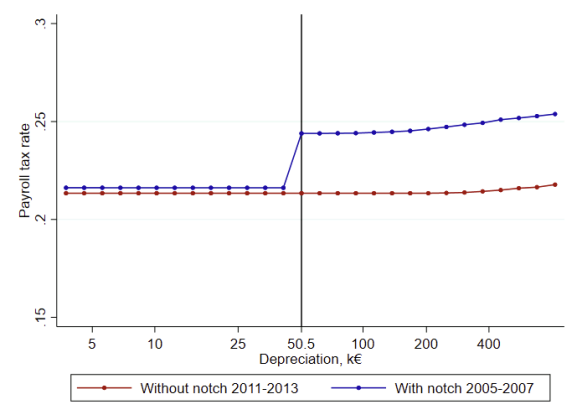
Theory: Steady State with Notch



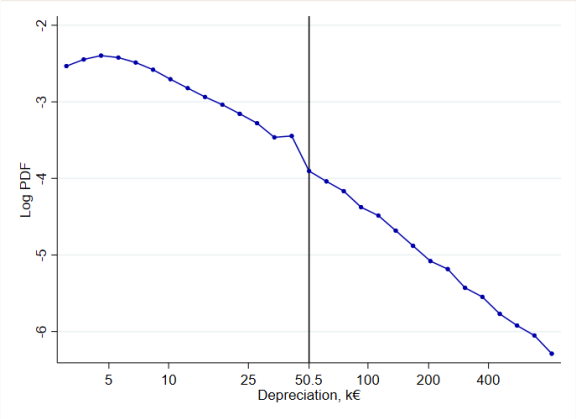
Theory: Sufficient statistic



Testing the Theory: Jump at the Threshold in Payroll Tax Rates



Payroll tax rates before and after repeal: $\sim$2.8 pp. notch



Log firm distribution before repeal

Testing the Theory: Dynamic Bunching

Goal: Estimate a counterfactual distribution, and compare to the observed post-reform distribution.

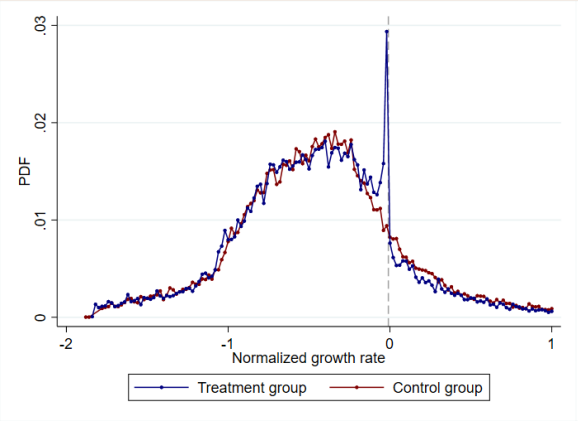
- We use dynamic bunching to estimate $\hat{p}$ (Garbinti et al. (2024))
- Normalized growth rates:

$$\bar{g}^1_{i,t} = \frac{d_{i,t+1} - 50{,}500}{d_{i,t}}$$

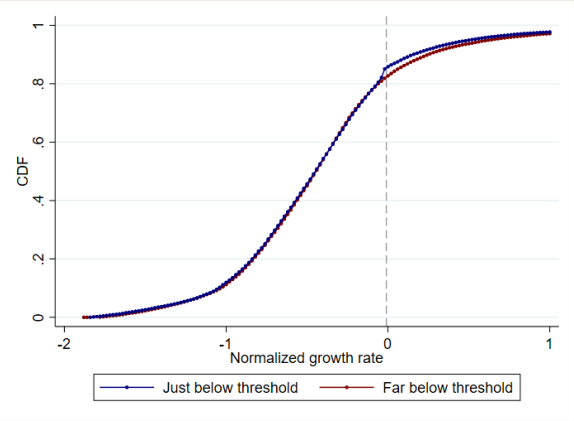
- Treatment group: 3 bins just below threshold (€22,715 – €50,499).
- Control group: 3 bins further below (€15,210 – €22,714), with threshold scaled by $\hat{\alpha} = 0.535$.

$\rightarrow$ Clear bunching in growth rates just below the threshold suggesting that notch impedes growth just below.

Dynamic Bunching: Normalized Growth Rate Distributions



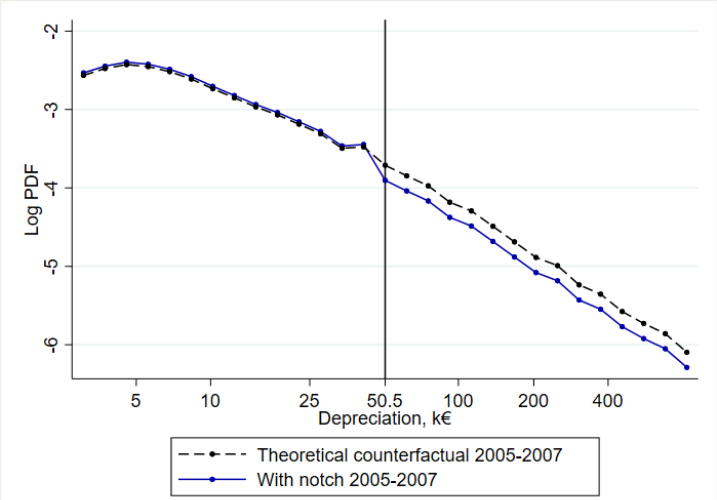
PDF



CDF: $\hat{p} = 0.182\ (0.001)$

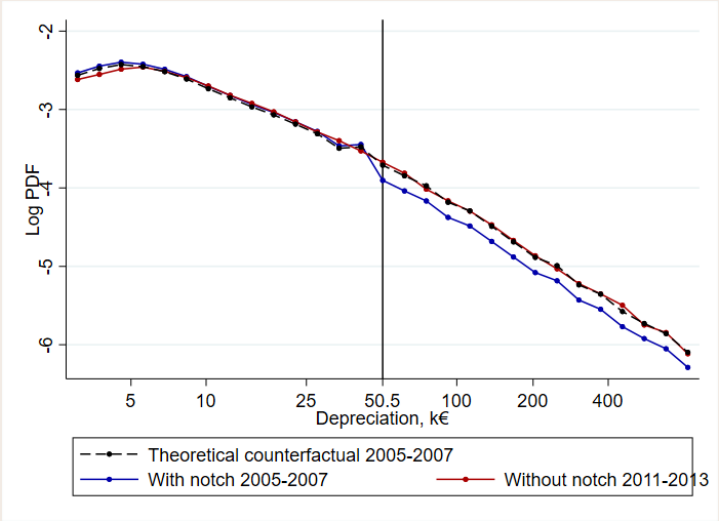
Counterfactual vs Observed Distributions

Using $\hat{p}$ and the pre-reform distribution from **pre-reform data only**, we estimate a theoretical counterfactual.



Counterfactual vs. Observed Distributions

The theoretical counterfactual closely matches the post-reform distribution.



Is This Real?

- Accounting depreciation (3rd-party audited): same pattern. ▸ Figure
- No evidence of firm splitting by owners. ▸ Figure
- Owner-level depreciation distribution. ▸ Figure
- No discontinuity in outsourcing at the threshold. ▸ Figure

→ **Responses seem to reflect genuine changes in firm scale.**

Empirical Methodology: Identification Issue

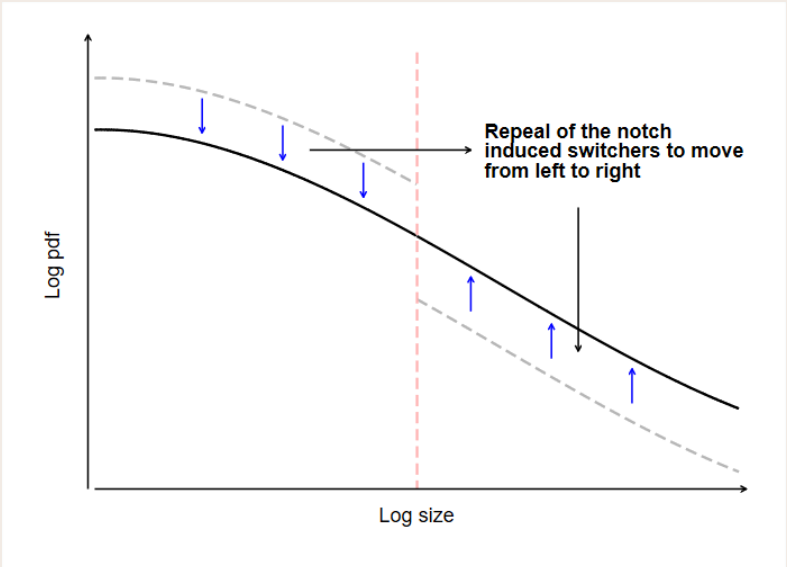
Goal: Estimate causal effects of notch removal when standard DiD fails.

- Standard DiD assumes treatment status is observed and fixed.
- With growth barriers, this fails: some units *refrained from crossing* due to the notch.

$\rightarrow$ These "switcher" units were **treated**, but DiD classifies them as untreated when the notch is in place.

$\rightarrow$ **Control group is contaminated** $\rightarrow$ biased estimates.

Empirical Methodology: Identification Issue



Empirical Methodology

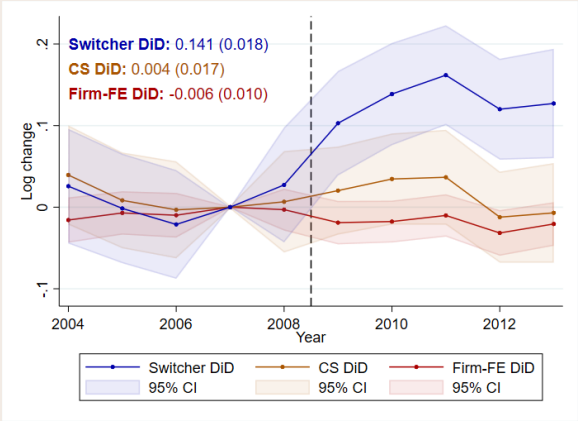
Result: Standard DiD is biased. The bias depends on switchers.

- Under parallel trends (in log counts and outcome means) and SUTVA for below-stayers:

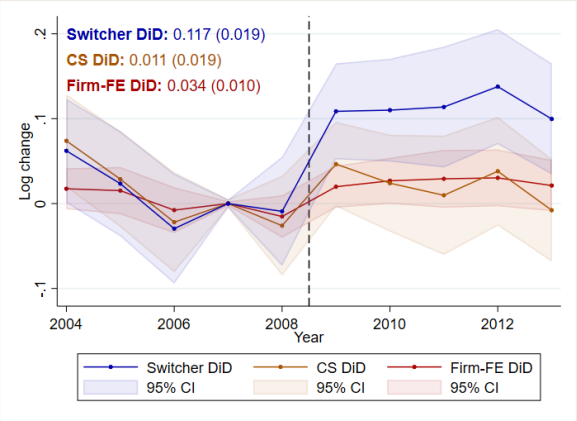
$$ATET = \underbrace{\tau^{DID}}_{\text{Cross-sectional DiD}} + \underbrace{w\left(E[y|\text{ above, pre }] - E[y|\text{ below, pre }]\right)}_{\text{Switcher term (bias)}},$$

  where w is the share of switchers among treated units.
- Switcher term can be estimated in two ways:
  1) Using an repeal/intro of the threshold + repeated cross-sectional data.
  2) Using the theoretical model and panel data with no repeal/intro.

Results: Labor Costs and Employment

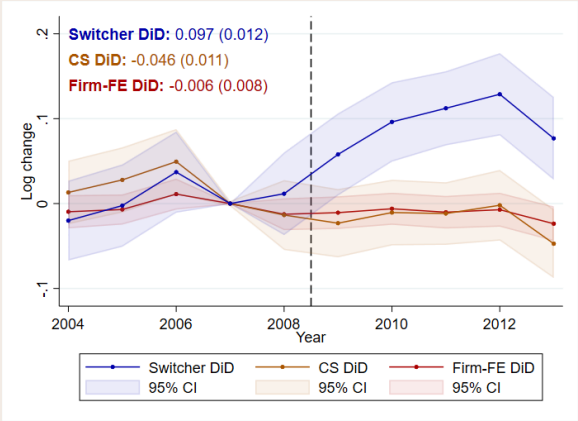


Δ Labor costs

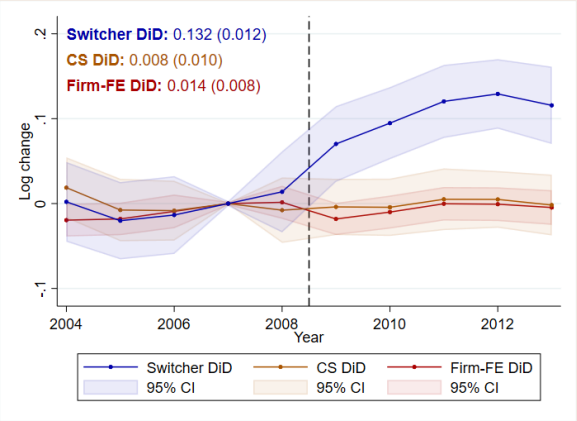


Δ Employees

Results: Capital Stock and Value Added



△ Capital stock



△ Value added

Summary

1. **Theory:** Policy thresholds act as growth barriers, distorting the *entire* distribution – not just locally around the cutoff.
   - Counterfactual distribution characterized by a single sufficient statistic p.
   - Validated out-of-sample: pre-reform $\hat{p}$ predicts the post-reform distribution.
2. **Methodology:** Standard DiD is biased when thresholds induce movement across cutoffs.
   - Two polar-case estimators provide bounds on causal effects.
   - Bias is around 10 log points; standard DiD would suggest a fiscal externality of ≈ 0.
3. **Empirics:** A Finnish payroll tax notch substantially constrained firm growth.
   - $\sim$18% fewer firms above the threshold; effects extend far into the distribution.
   - After repeal: $\sim$10% increases in employment, capital, and value added.
   - Fiscal externality $\approx$ 1/3 of the mechanical effect of payroll taxes (ongoing: other tax bases).

$\rightarrow$ Even modest tax notches can meaningfully distort the size distribution and reduce economic activity.

Thank you!

Sami Jysmä

sami.jysma@labore.fi

Theory: Setup

Setup: Continuous time Itô diffusion with a reflective lower bound.

- Consider unit size S over continuous time t with some minimum size S_{min}
  - Size e.g. income/wealth of individual, no. employees/revenue of firm
- We assume a stochastic growth process for the size dS_t in the:

$$dS_t = g_0(S_t)dt + \sigma(S_t)dW_t, \tag{1}$$

  - $g_0(S_t)$ is the drift parameter, W_t is Brownian motion, $\sigma(S_t)$ determines variance.
- E.g. standard geometric Brownian motion (GBM) is a special case $g_0(S_t) = gS_t$ and $\sigma(S_t) = \sigma S_t$.

Theory: Growth Barrier

Growth barrier: A policy notch is assumed to reduce growth just below the policy threshold.

- A policymaker introduces policy notch at $\bar{S}$ above which costs increase.

$\rightarrow$ Units have an incentive to refrain from crossing the threshold.

- Assumption: growth rate is reduced in a small interval below the threshold $[\bar{S} - \delta, \bar{S}]$.

$$g_1(S_t) = \begin{cases} g_0(S_t), & \text{if } S_t \notin [\bar{S} - \delta, \bar{S}] \\ g_0(S_t) - \gamma, & \text{if } S_t \in [\bar{S} - \delta, \bar{S}] \end{cases} \tag{2}$$

- (An approximation of a general growth reduction, when δ is small.)

Theory: Counterfactual Distribution

Result: A simple testable sufficient statistics representation of the counterfactual distribution.

- Using (1) and (2), we can derive a *counterfactual* pdf distribution $f_0(S)$ in logs:

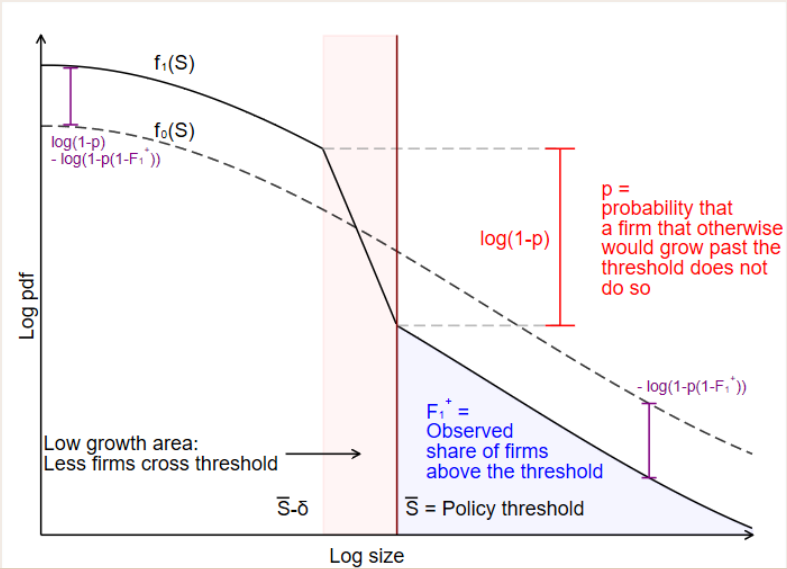
$$\log f_0(S) = \begin{cases} \log(1-p) + \psi + \log f_1(S), & \text{if } S < \bar{S} \\ \psi + \log f_1(S), & \text{if } S \geq \bar{S}. \end{cases}.$$

  where $\psi \equiv -\log\left(1 - p(1 - F_1^+)\right)$, $f_1(s)$ is the observed pdf, and $F_1^+ = \int_{\bar{s}}^{\infty} f_1(s)ds$ is the observed share of units above the threshold.

- p is the probability that a unit refrains from crossing the threshold due to the notch.

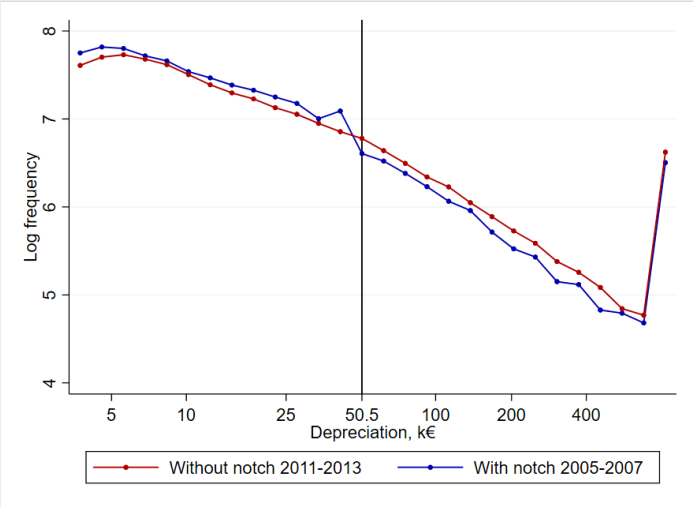
$\rightarrow$ p **can be identified using the dynamic bunching approach of Garbinti et al. (2024)**

Theory: Illustration



Firm distribution: Balanced panel distribution

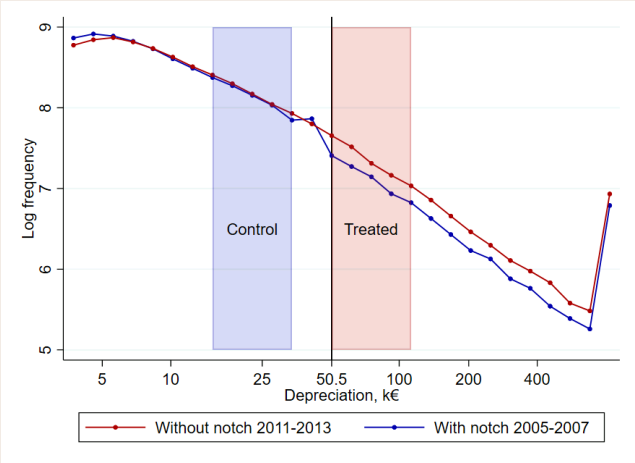
Go back



Firm Distribution: Treatment and control groups

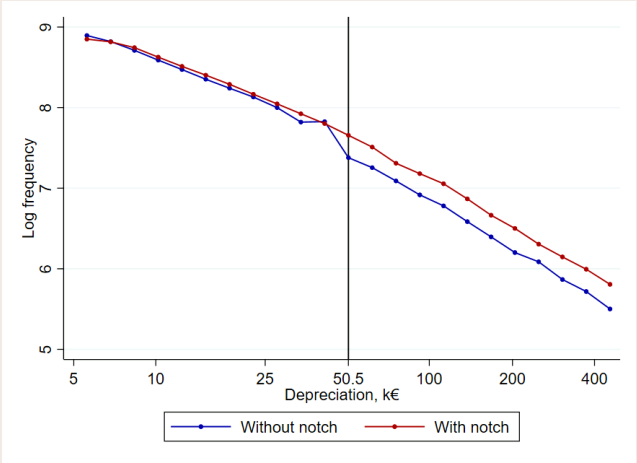
Go back

Treatment: 5 bins above. Control: 5 bins bins below (excluding ”buncher” bin below).



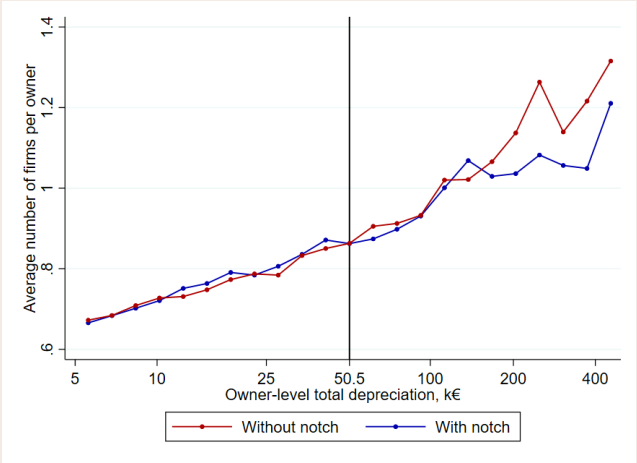
Is this real? Audited accounting capital depreciation

▸ Go back



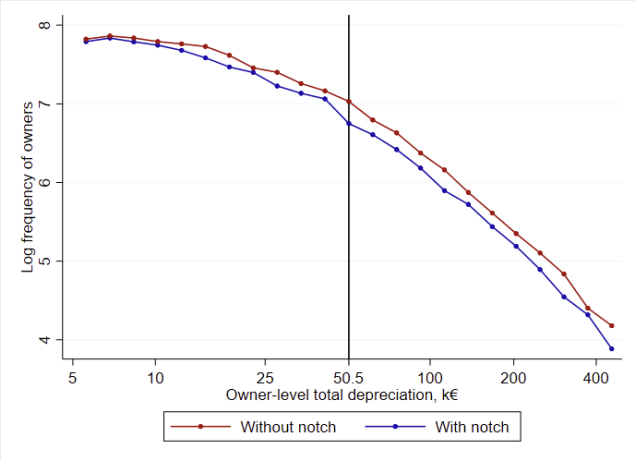
Is this real? Firm splits

Go back



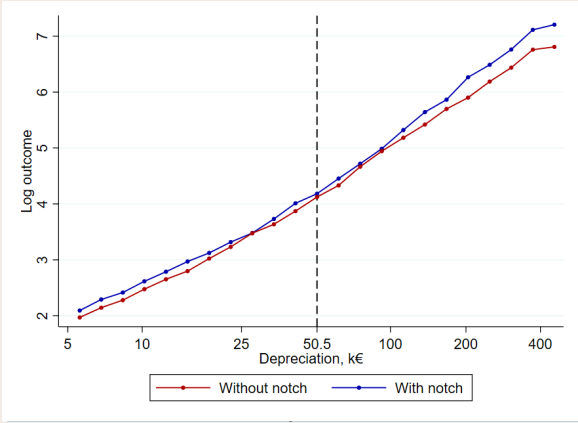
Is this real? Owner-level depreciation

Go back

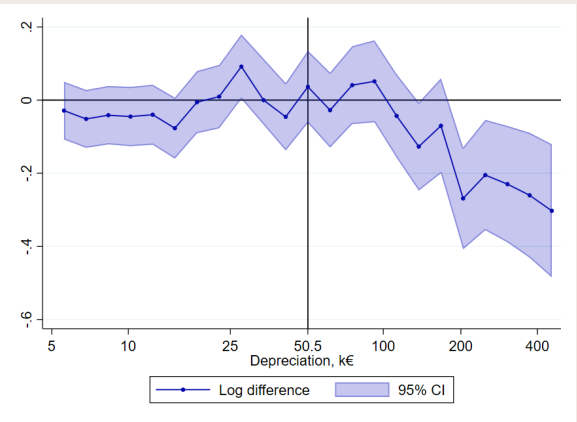


Is this real? Outsourced services

Go back



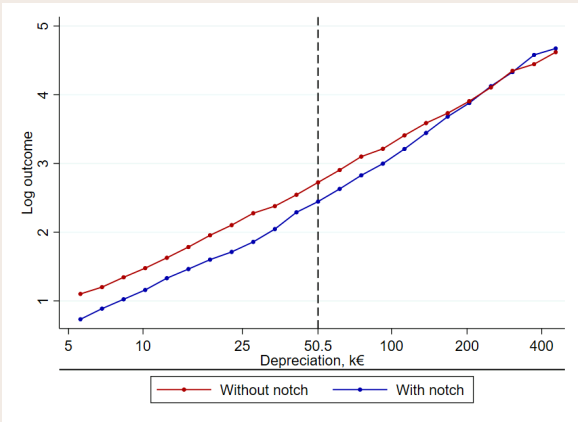
Outsourcing



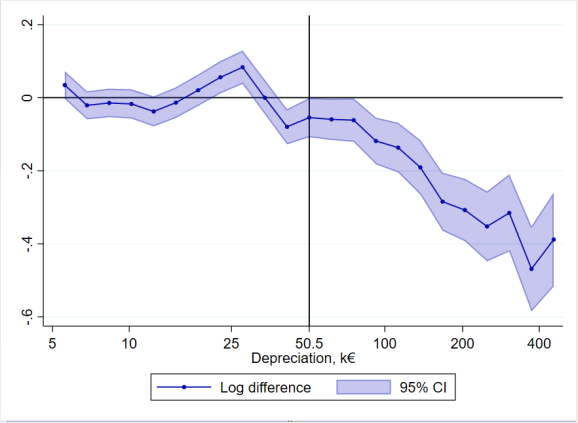
Difference

Is this real? Rental capital

Go back



Rental capital



Difference