Production vs Revenue Efficiency With Limited Tax Capacity

Theory and Evidence From Pakistan

Michael Best, Anne Brockmeyer, Henrik Kleven, Johannes Spinnewijn, Mazhar Waseem

London School of Economics

November 2013

Production Efficiency

- **Production Efficiency Theorem** (Diamond & Mirrlees 1971):

  Any second-best optimal tax system maintains production efficiency

- **Key policy implications**:
  - Permits taxes on consumption, wages and profits
  - Precludes taxes on inputs, turnover and trade

- The theorem has been influential in the policy advice given to developing countries

Production Efficiency vs Revenue Efficiency

- Production Efficiency Theorem **assumes perfect tax enforcement**
  $\rightarrow$ This is violated everywhere, but especially in developing countries
- **Tax evasion** introduces a trade-off between production and revenue efficiency in tax design
- In the context of firm taxation in Pakistan, our contribution is:
  - **Simple model** on the optimal production-revenue efficiency trade-off
  - **Quasi-experimental evidence** on the evasion elasticity w.r.t taxes
  - **Link model & evidence** to quantify optimal policy

Novel Quasi-Experimental Approach

- **Minimum Tax Scheme:** firms are taxed on either profits or turnover (lower tax rate on turnover) depending on which tax liability is larger
  - This production inefficient policy is motivated by tax compliance
- **Non-standard kink** where both the tax rate and the tax base jump
  - Kink changes real and evasion incentives differentially
  - Novel method for estimating tax evasion based on a bunching approach
- **Wide applicability** of our approach since such schemes are ubiquitous

Contributions to Previous Literature

- **Public Finance & Development**: Kleven & Waseem (2013), Pomeranz (2013), Kumler et al. (2013)
- **Optimal taxation with enforcement problems**: Emran and Stiglitz (2005), Gordon & Li (2009), Kleven et al. (2009)
- **Estimating tax evasion**: Andreoni et al. (1998), Slemrod (2007), Kleven et al. (2011)
- **Corporate taxation**: Hassett & Hubbard (2002), Auerbach et al. (2010), Devereux et al. (2013)
- **Bunching methodology:** Saez (2010), Chetty et al. (2011)

Outline

Outline

Firm Behavior: Real vs Evasion Responses

- Real output y, real cost $c(y)$, declared cost $\hat{c}$, penalty $g(\hat{c} - c(y))$
- Tax liability $T = \tau[y - \mu\hat{c}]$
- Maximization of after-tax profits

$$\begin{aligned} c'(y) &= 1 - \omega \\ g'(\hat{c} - c(y)) &= \tau\mu \end{aligned}$$

- **Production wedge** $\omega = \tau\frac{1-\mu}{1-\tau\mu}$:
  - $\omega = 0$ for a profit tax $\mu = 1$ **[production efficiency]**
  - $\omega = \tau$ for a turnover tax $\mu = 0$ **[production inefficiency]**

Proposition [Production Inefficiency]

*With **perfect enforcement**, the optimal tax base is pure profits* $(\mu = 1)$

*With **imperfect enforcement**, the optimal tax base is in between pure profits and turnover* $(0 < \mu < 1)$ *and depends on the evasion-output elasticity ratio*

$$\underbrace{\frac{\tau}{1-\tau} \times \frac{\partial \omega}{\partial \tau}(\mu)}_{\text{effective MTR} \downarrow \text{ in } \mu} = \underbrace{G(\mu)}_{\text{tax gap} \uparrow \text{ in } \mu} \times \underbrace{\frac{\varepsilon_{\hat{c}-c}}{\varepsilon_y}}_{\text{elasticity ratio}}$$

Outline

Minimum Tax Scheme

- Combination of profit tax ($\mu = 1$) and turnover tax ($\mu = 0$):

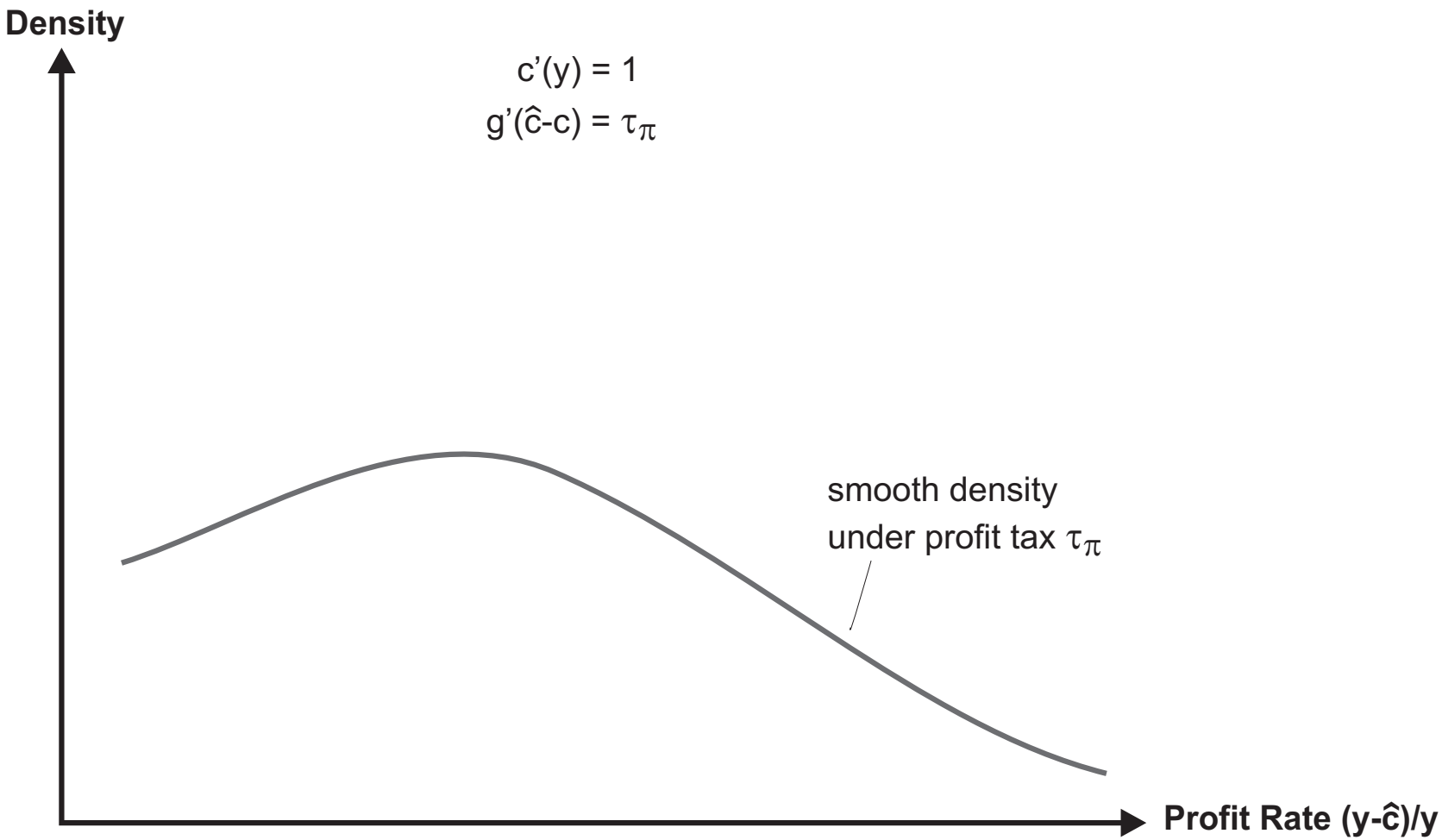
$$T = \max\left\{\tau_\pi\left(y - c\right); \tau_y y\right\}.$$

- Firms switch between the two taxes depending on profit rate p:

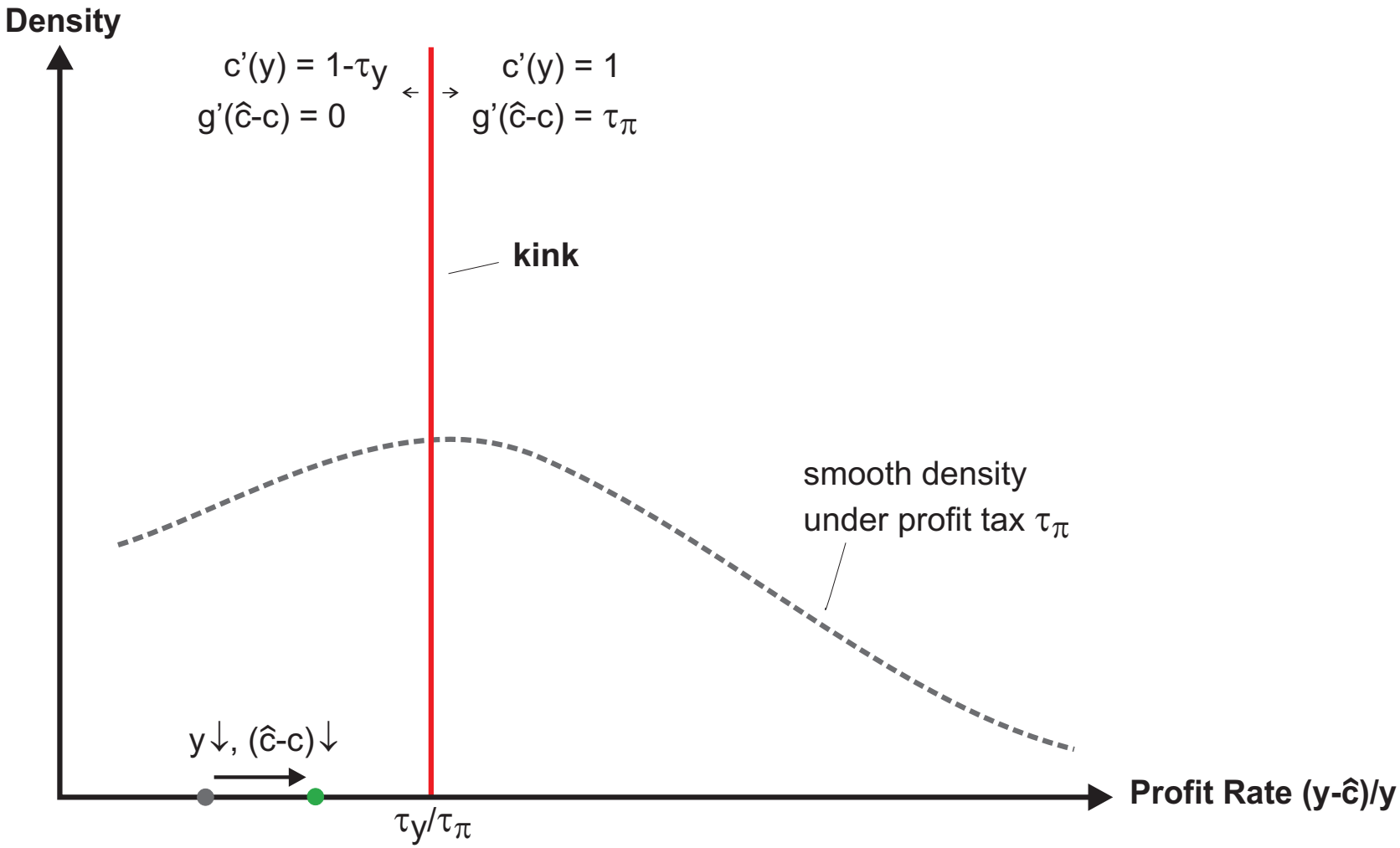
$$\tau_\pi\left(y - c\right) = \tau_y y \quad \Leftrightarrow \quad p \equiv \frac{y - c}{y} = \frac{\tau_y}{\tau_\pi}.$$

- **Kink: tax base and marginal tax rate change discontinuously, but tax liability is continuous**

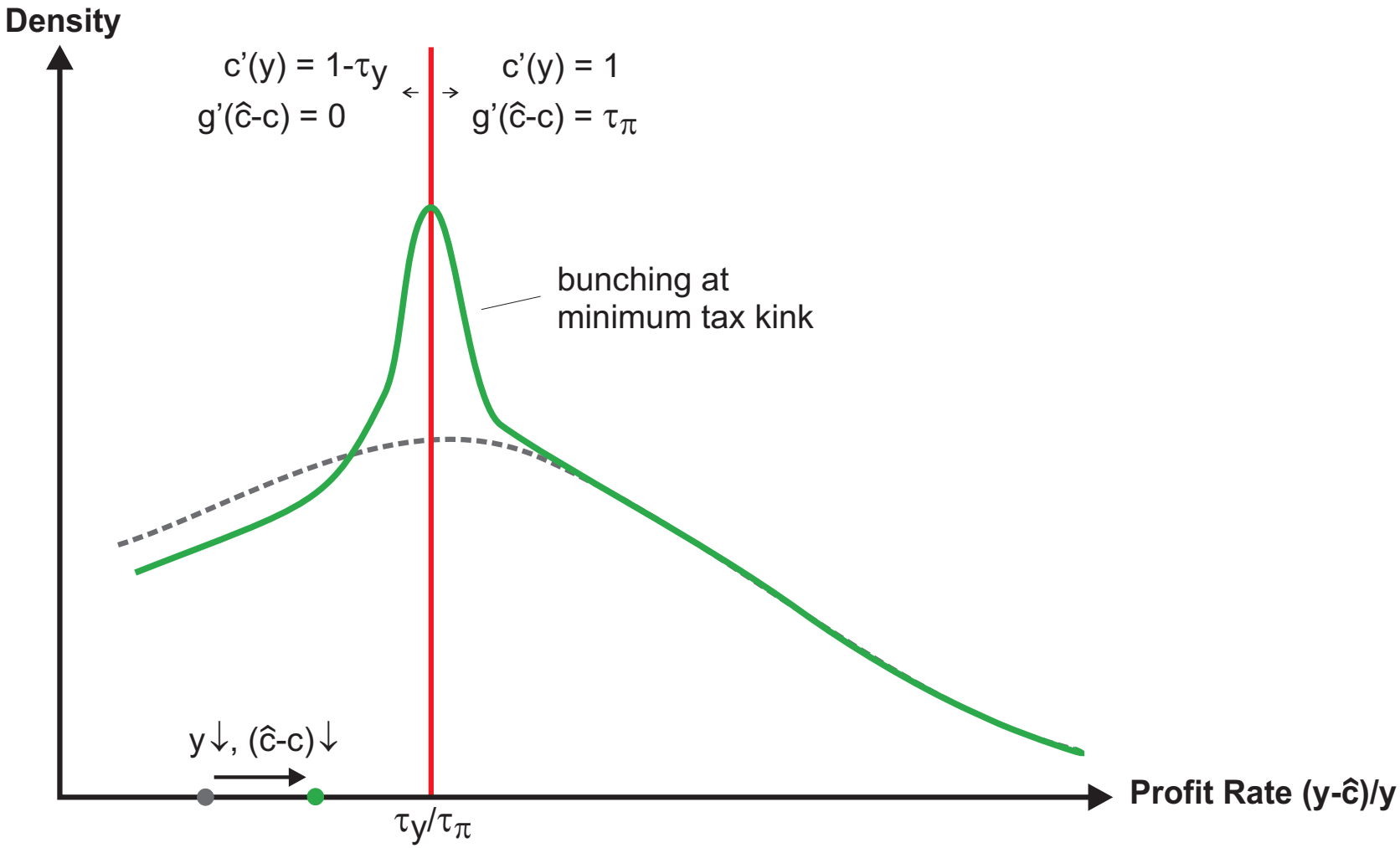
Bunching at the Minimum Tax Kink



Bunching at the Minimum Tax Kink



Bunching at the Minimum Tax Kink



Minimum Tax Kink Ideal for Eliciting Evasion

- **Real output response:**
  - Firms choose real output based on $1-\omega$
  - At the kink, production wedge ω changes from 0 to τ_y (≈ 0)
    $\Rightarrow$ almost no variation and therefore small real response
- **Evasion response:**
  - Firms choose evasion based on $\tau\mu$
  - At the kink, $\tau\mu$ changes from τ_π ($\gg 0$) to 0
    $\Rightarrow$ large variation and therefore large evasion response
- **Bunching B identifies (mostly) evasion:**

$$B \propto \frac{\tau_y^2}{\tau_\pi}\varepsilon_y - \frac{\Delta(\hat{c}-c)}{y}$$

Robustness

- **Distortionary profit tax**
  - If $\omega > 0$ under profit tax, then turnover tax may improve real incentives

    $\Rightarrow$ firms move away from the kink and **create a hole**

- **Distortionary turnover tax**
  - Small τ_y may create big distortions via cascading and extensive margin

    $\Rightarrow$ GE effects and extensive responses **do not affect bunching**

- **Output evasion**
  - If firms can underreport output, the turnover tax reduces output evasion (due to $\tau_y < \tau_\pi$) in addition to cost evasion

    $\Rightarrow$ bunching identifies **differential evasion from output and costs**

Outline

Data

- Administrative data from FBR Pakistan
- All corporate tax returns from 2006-2010 (about 15,000 returns per year)
- New electronic data collection system in place for this time period
- In each year, about half of the firms are turnover taxpayers and half of them are profit tax payers

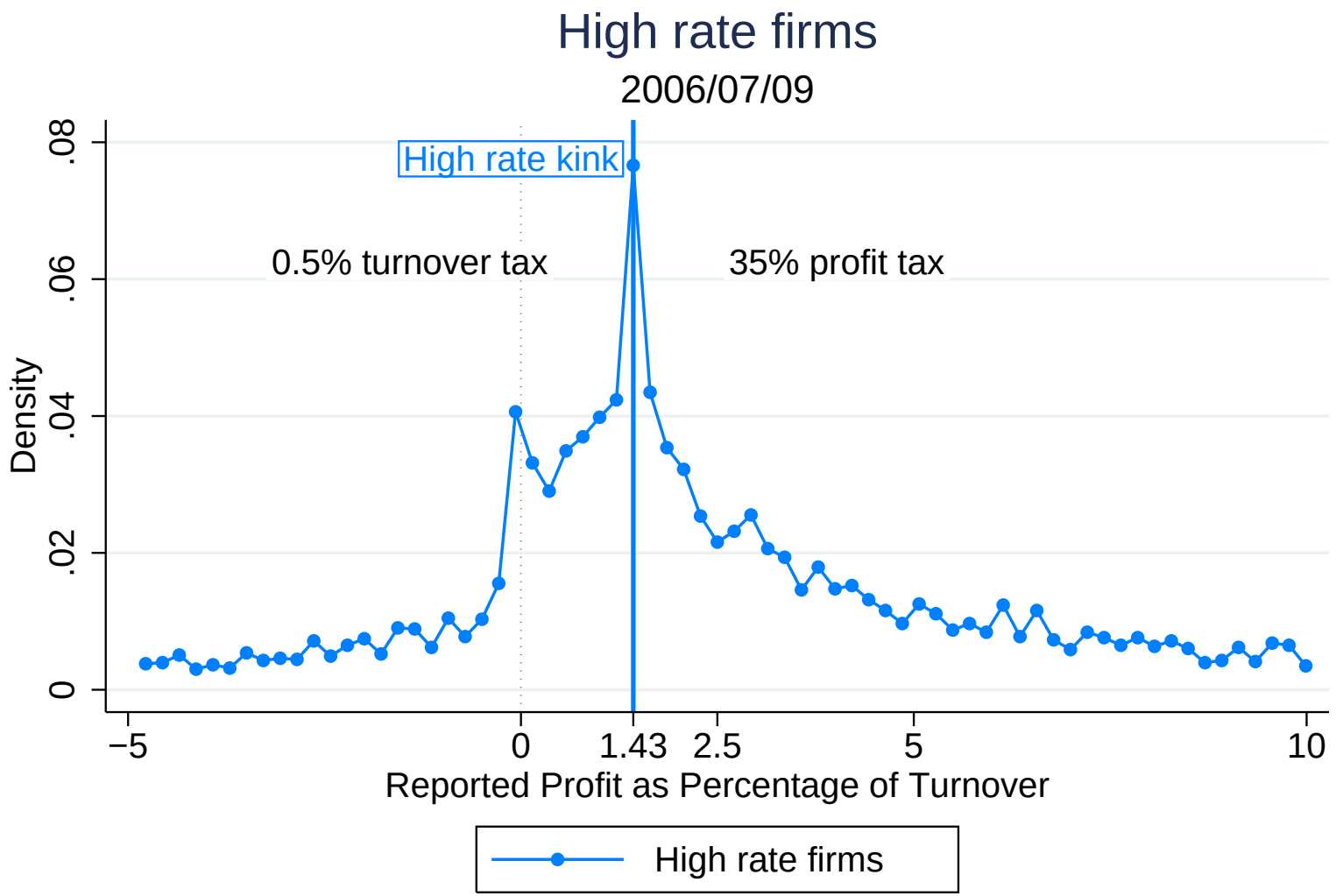
Variation in Minimum Tax Kink

- **Variation in profit tax rate τ_π across firms:**
  - High rate of 35%, low rate of 20%
    [depends on incorporation date, turnover, assets, #employees]

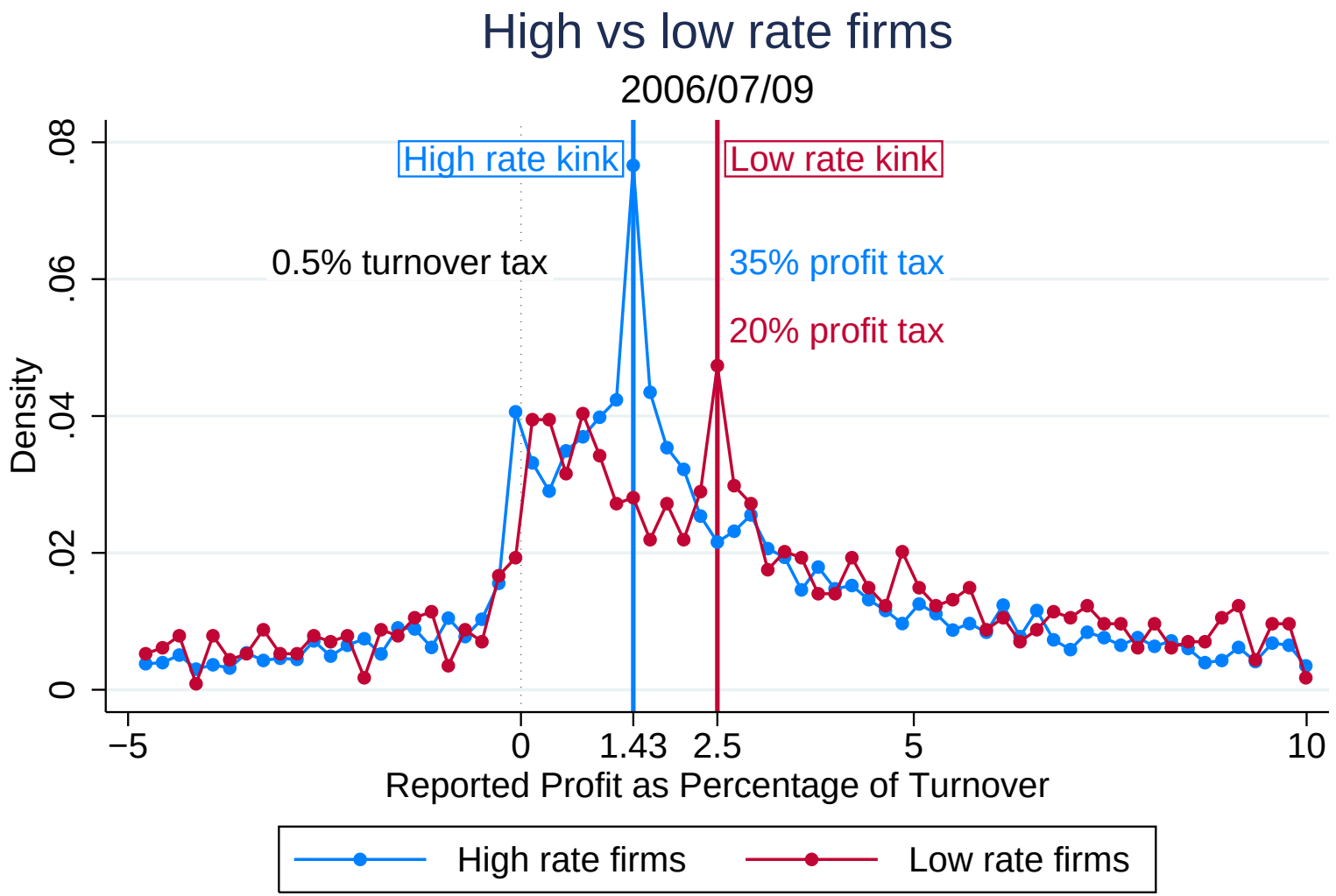
- **Variation in turnover tax rate τ_y over time:**
  - 2006-07: tax rate of 0.5%
  - 2008: turnover tax scheme withdrawn
  - 2009: tax rate of 0.5%
  - 2010: tax rate of 1%

Outline

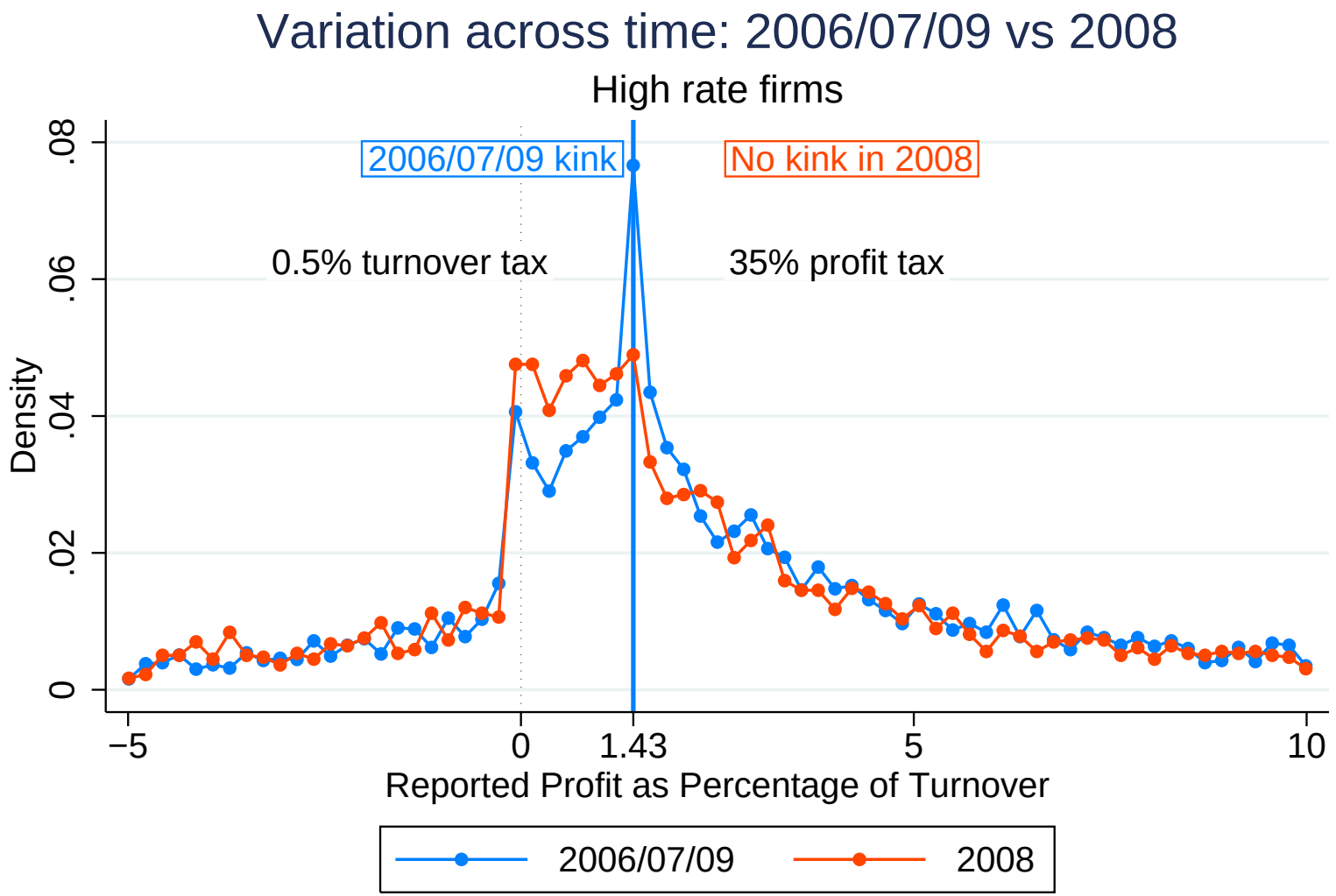
Bunching Evidence



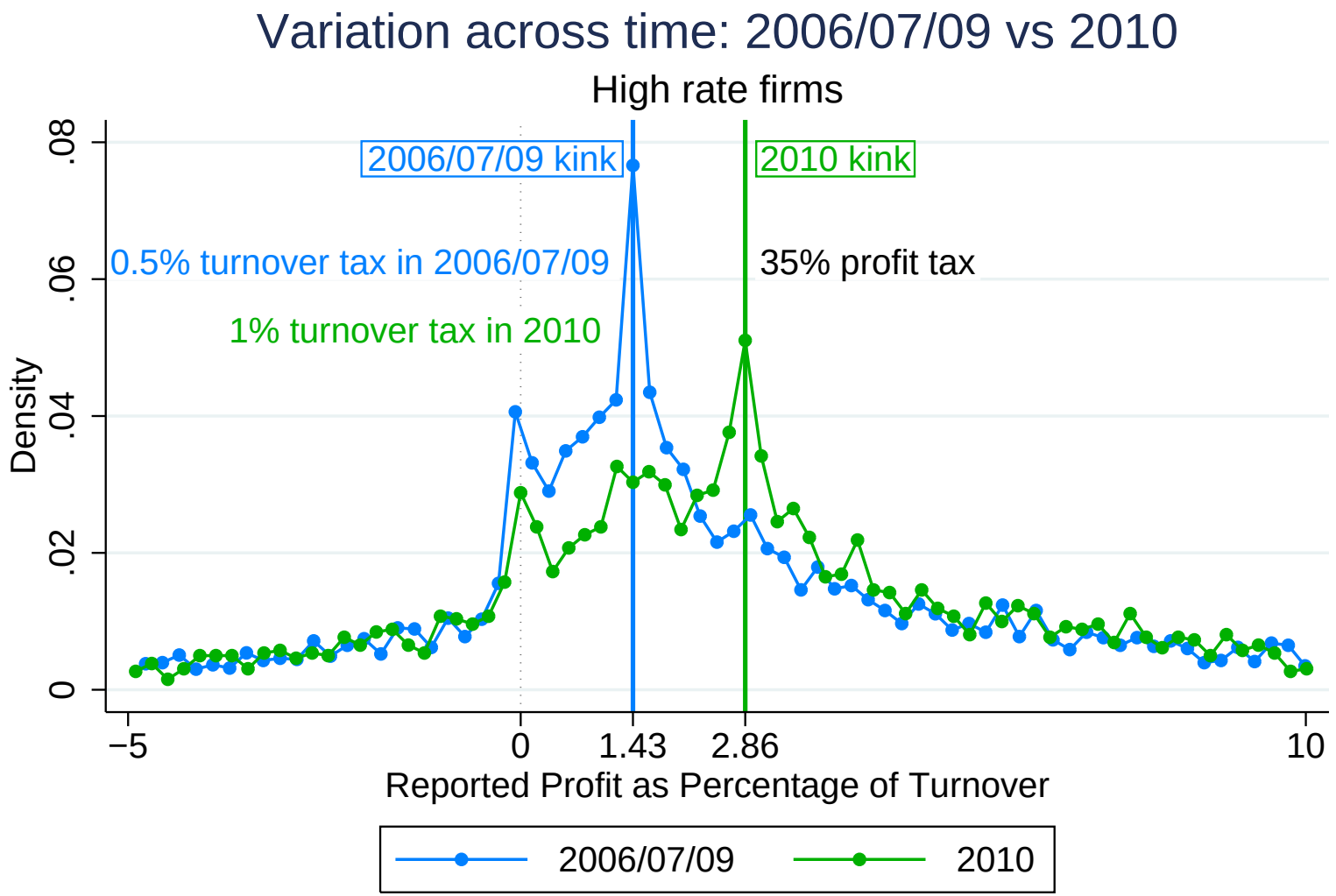
Bunching Evidence



Bunching Evidence



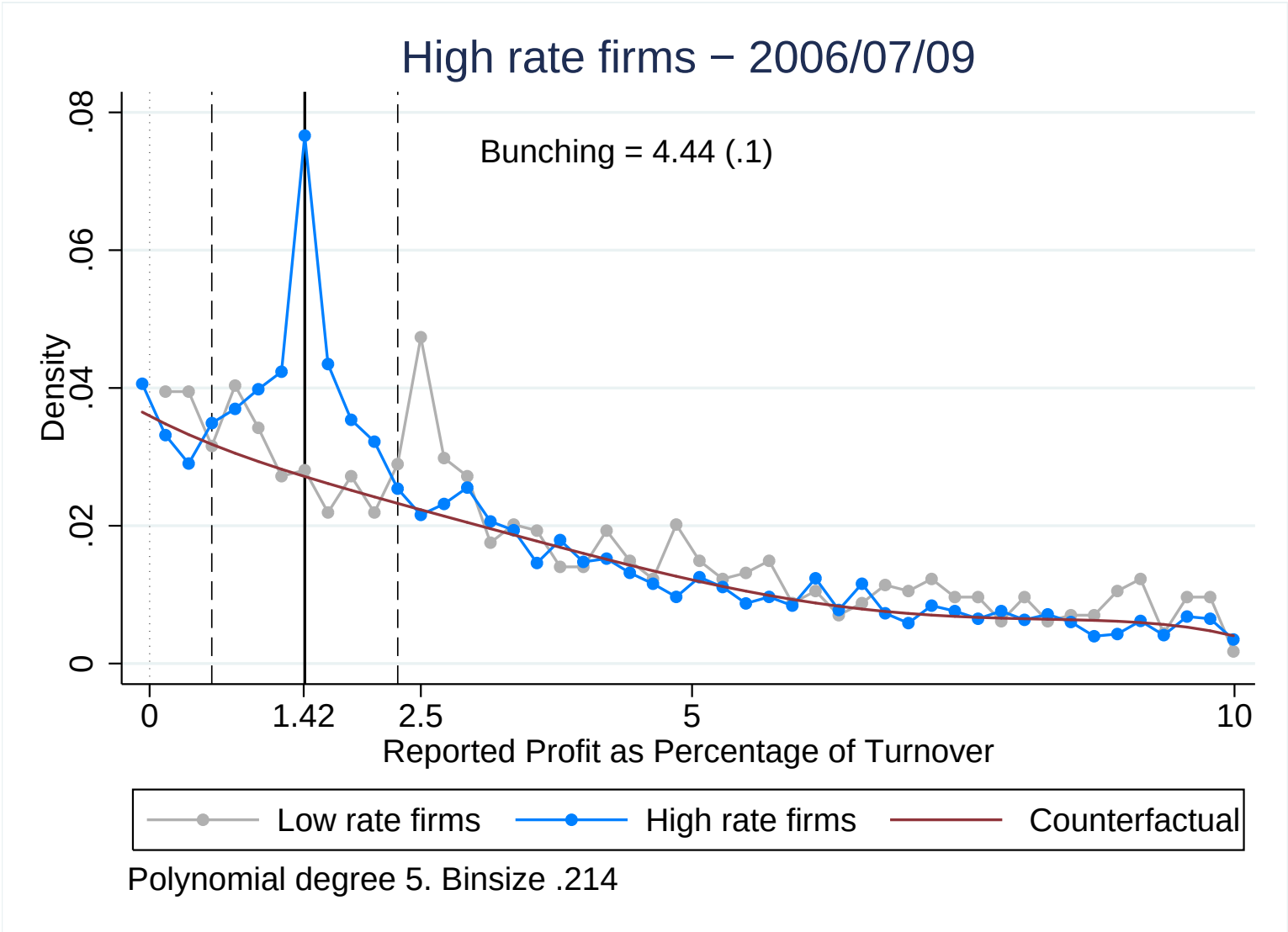
Bunching Evidence



▸ Heterogeneity

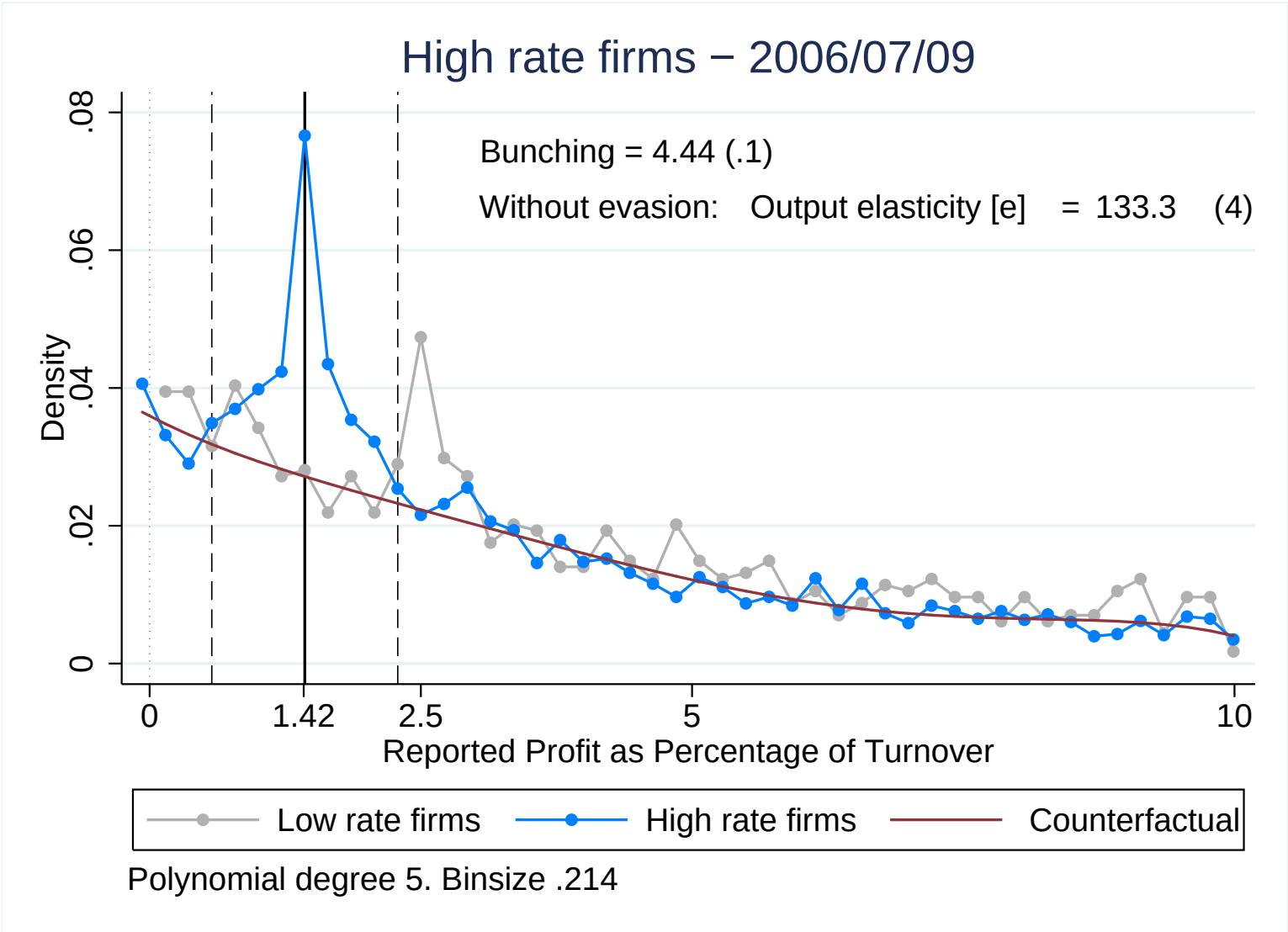
Outline

Estimating Evasion



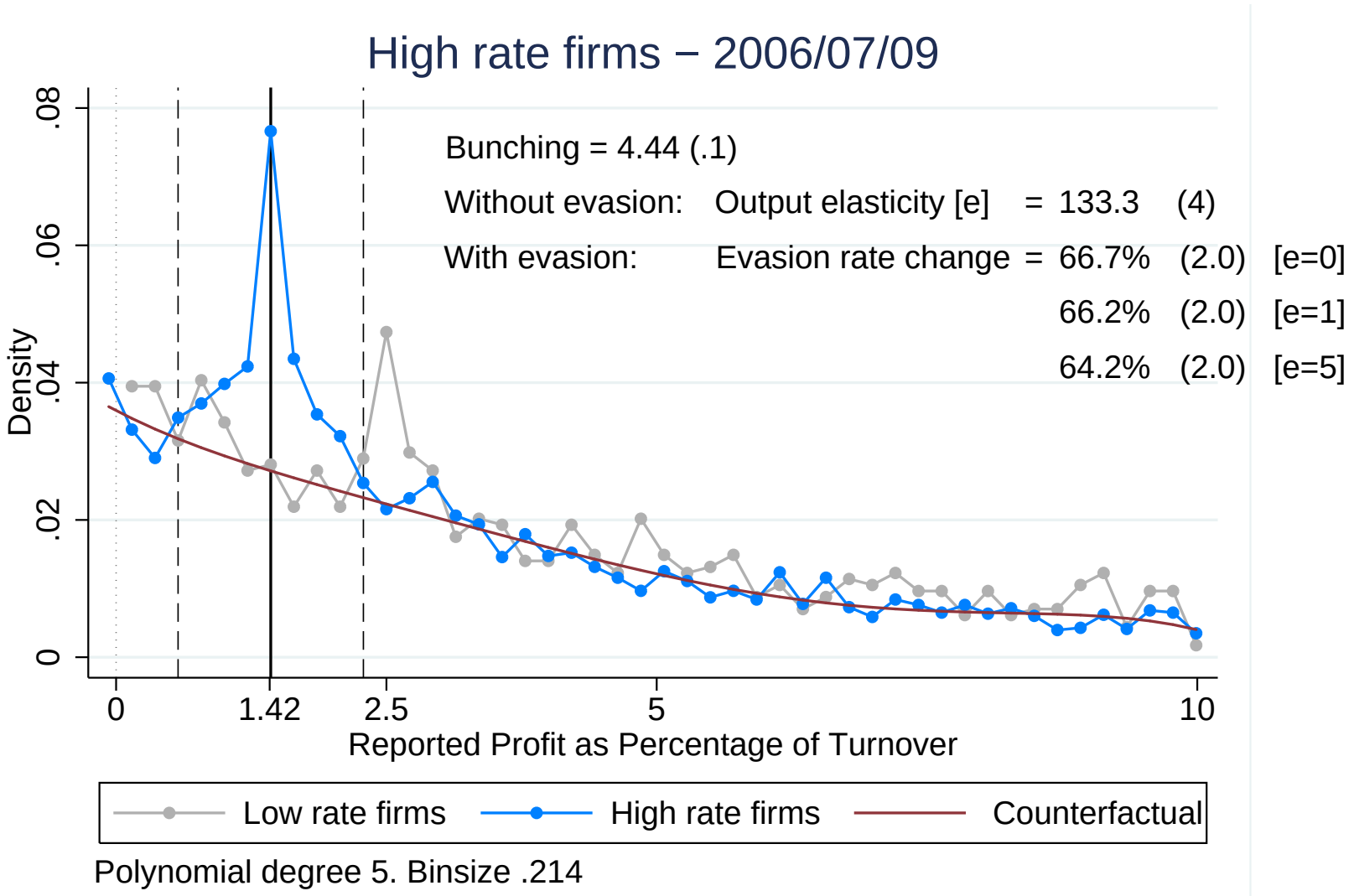
▸ Counterfactual

Estimating Evasion



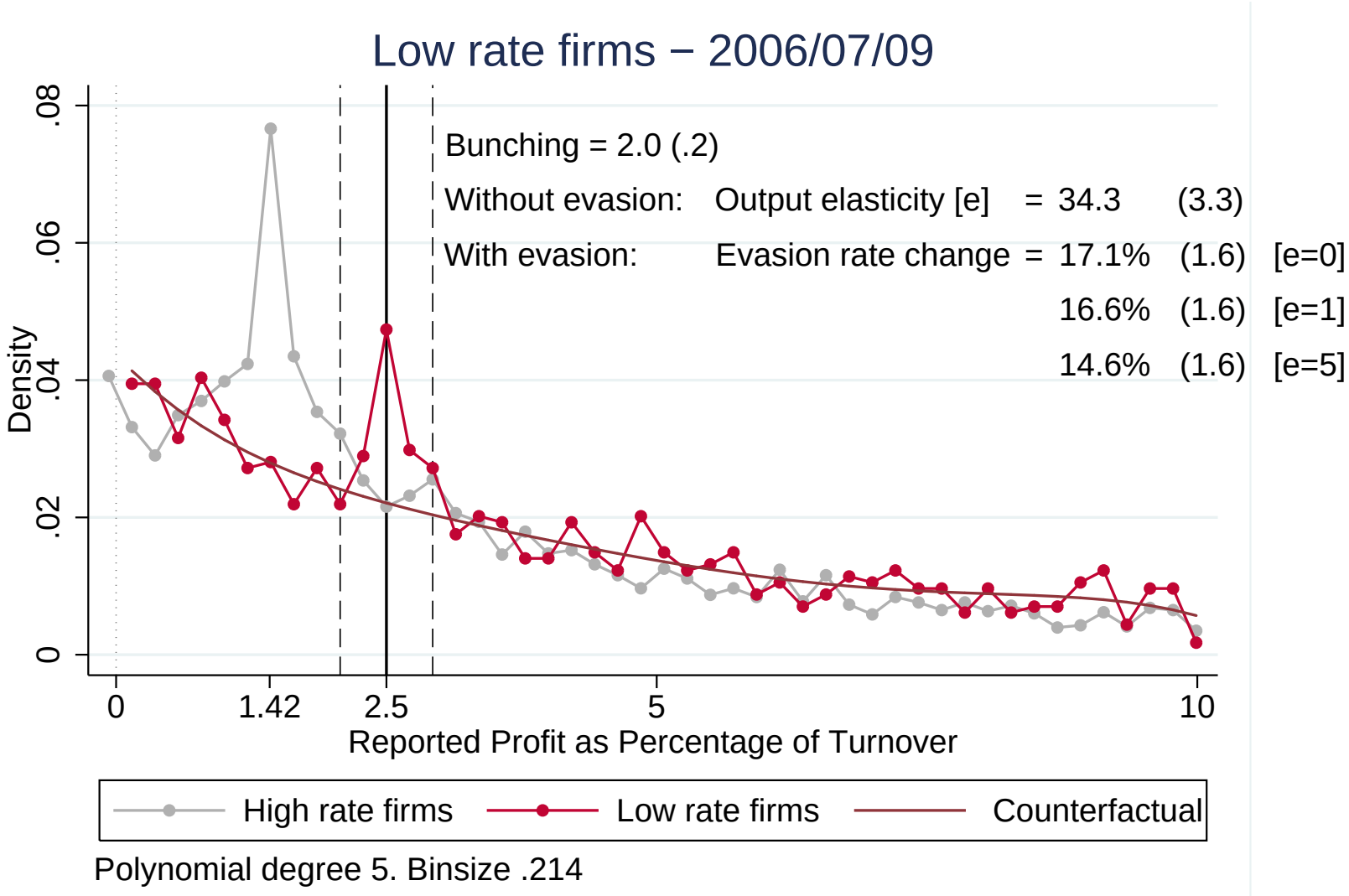
▸ Counterfactual

Estimating Evasion



▸ Counterfactual

Estimating Evasion

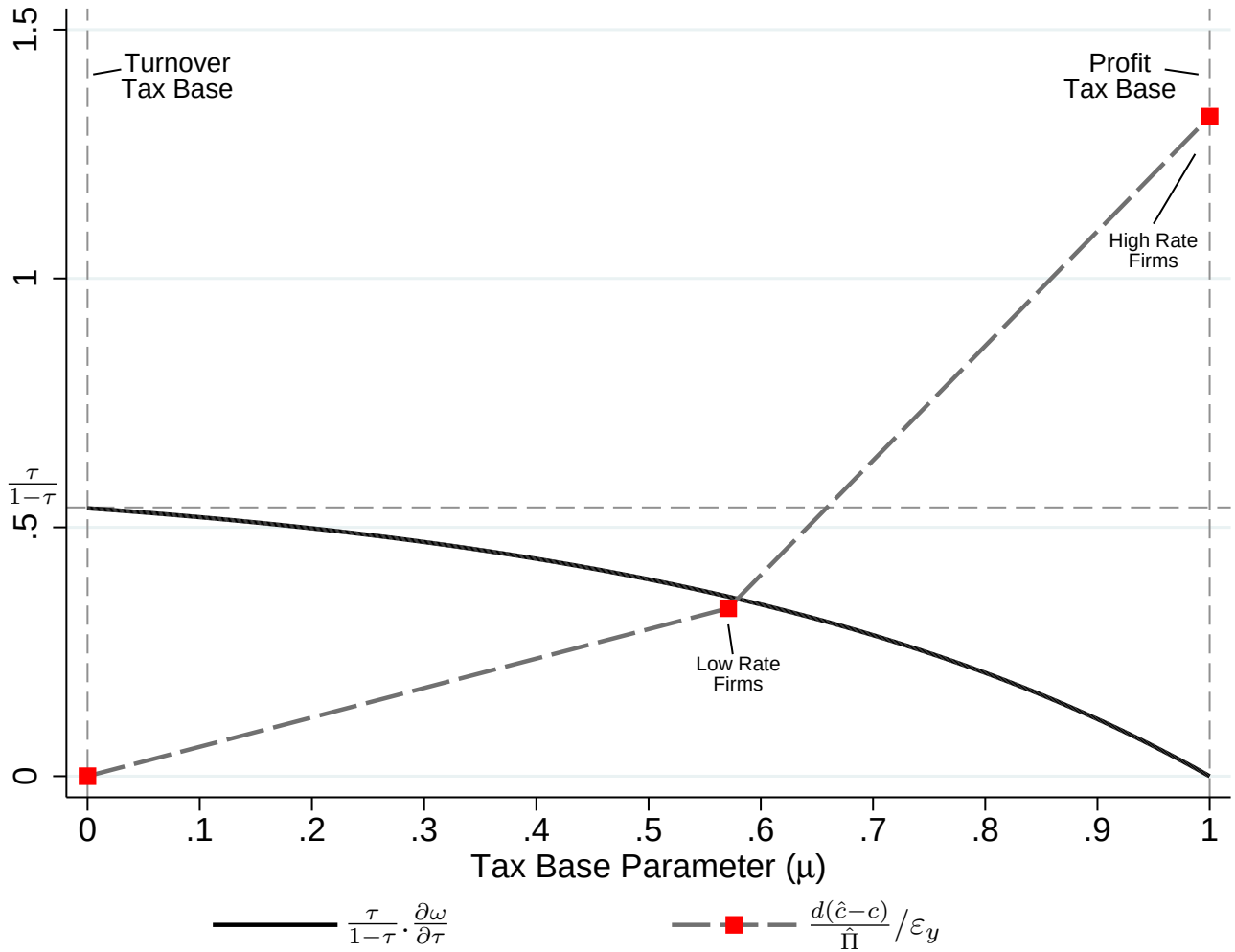


▸ Counterfactual

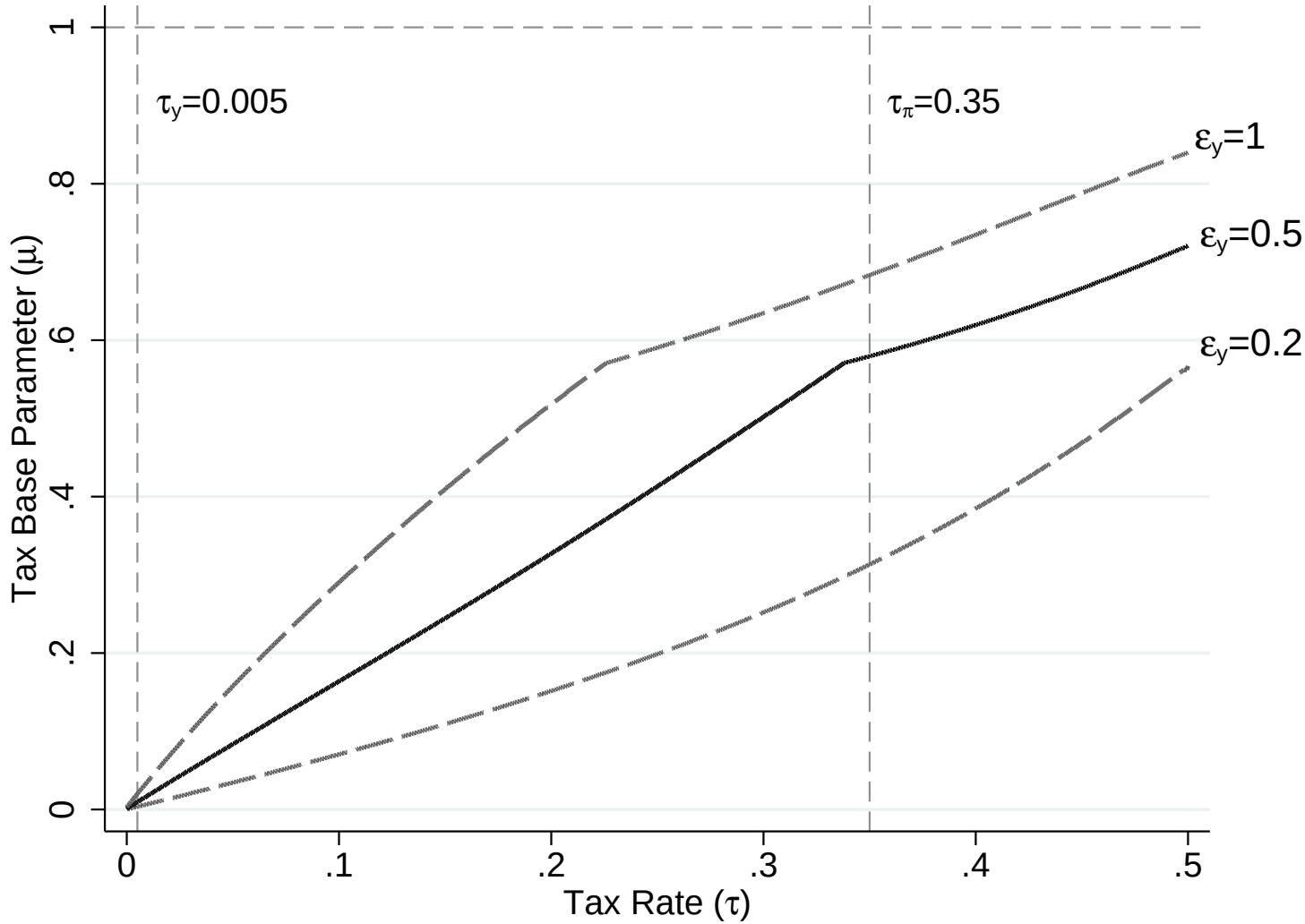
Outline

Optimal Tax Base (Given τ and ε_y)

$$\frac{\tau}{1-\tau} \times \frac{\partial \omega}{\partial \tau}(\mu) \simeq -\frac{\Delta(\hat{c}-c)}{\hat{\Pi}}(\mu) \times \frac{1}{\varepsilon_y}$$



Optimal Tax Base (Varying τ and ε_y)



Conclusion

- Production inefficient policies like turnover taxes may be optimal under imperfect enforcement
- Novel quasi-experimental approach using minimum tax schemes for estimating evasion responses to switches between profit taxes and turnover taxes
- Large evasion responses to such switches in Pakistan, which justify deviations from a production efficient profit tax
- Returns to improved tax enforcement are high: up to 2/3 of profit tax revenues are lost due to underreporting by corporations

Counterfactual Estimation

- Estimate counterfactual density following Chetty et al (2011):

$$d_j = \sum_{l=0}^{q} \beta_l (z_j)^l + \sum_{k=z_L}^{z_U} \gamma_k \cdot \mathbf{1}[z_j = k] + v_j.$$

- Estimate excess mass:

$$b = \frac{\sum_{k=z_L}^{z_U} \hat{\gamma}_k}{\sum_{k=z_L}^{z_U} \hat{d}_k / N_k}$$

- Excess mass indicates the profit rate change Δp for marginal buncher.

Back

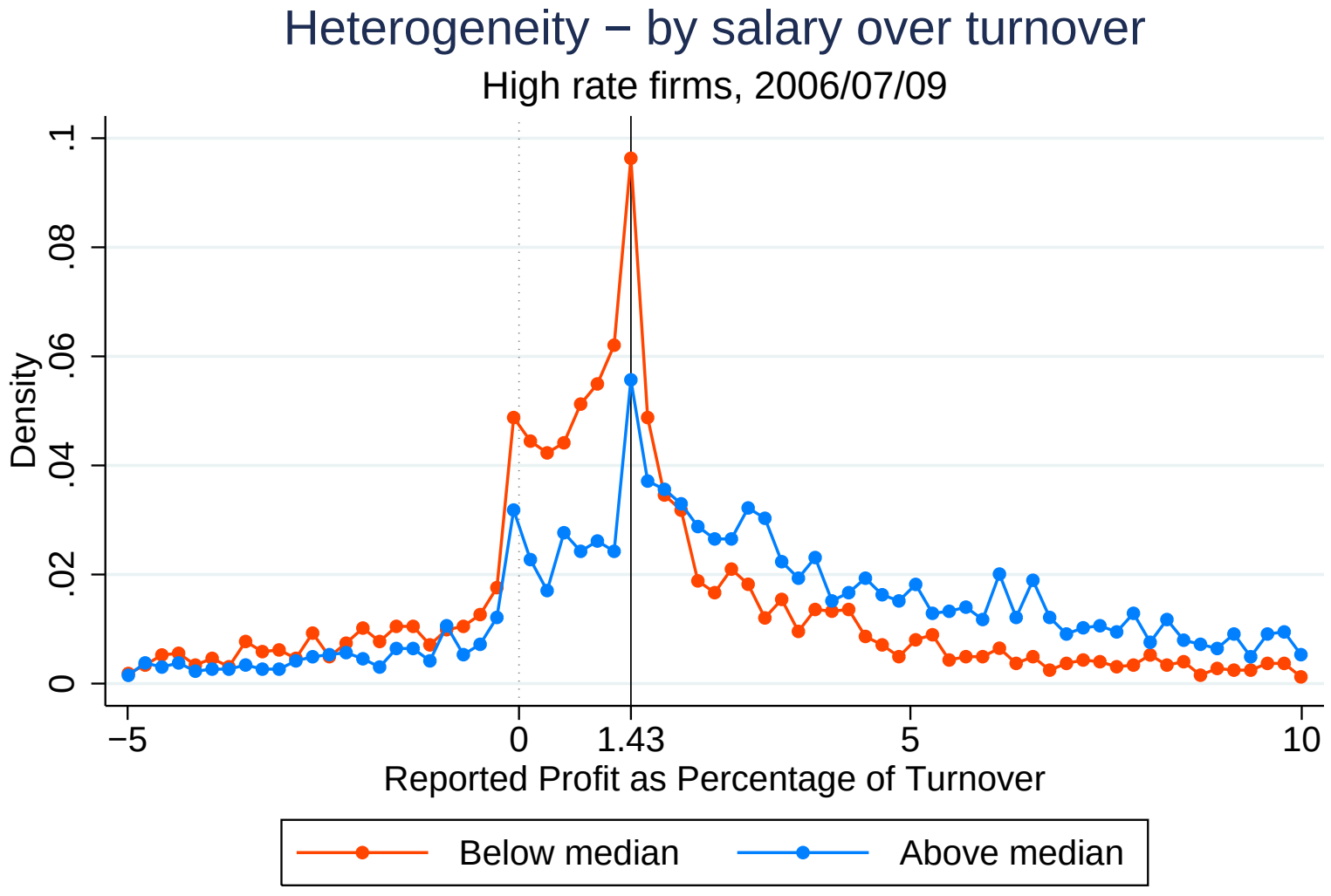
Bunching Heterogeneity by Evasion Proxies

Theory predicts more evasion among firms that are

- **small** in number of employees (Kleven et al. 2009):
  - Collusive evasion is more sustainable in a small group
  - Proxy for firm size: salary payments, turnover
- less dependent on **financial intermediation** (Gordon & Li 2009)
  - Access to formal credit creates a paper trail
  - Proxy for credit needs: interest payments (scaled by turnover)
- selling to **final consumers** (e.g. Pomeranz 2013)
  - Paper trail is lacking for transactions with final consumers
  - Compare "retailers" and "non-retailers"

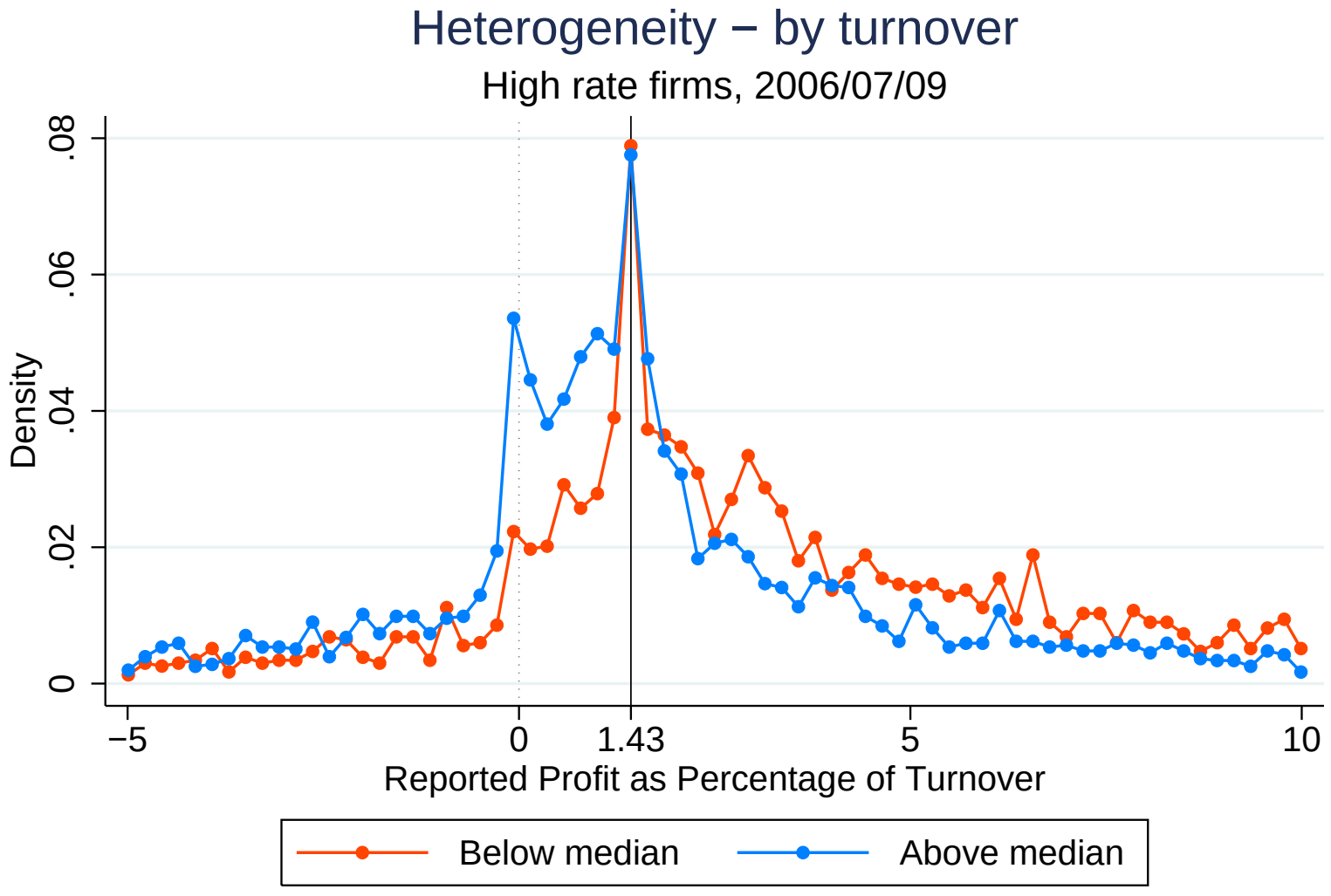
Back

Bunching Heterogeneity



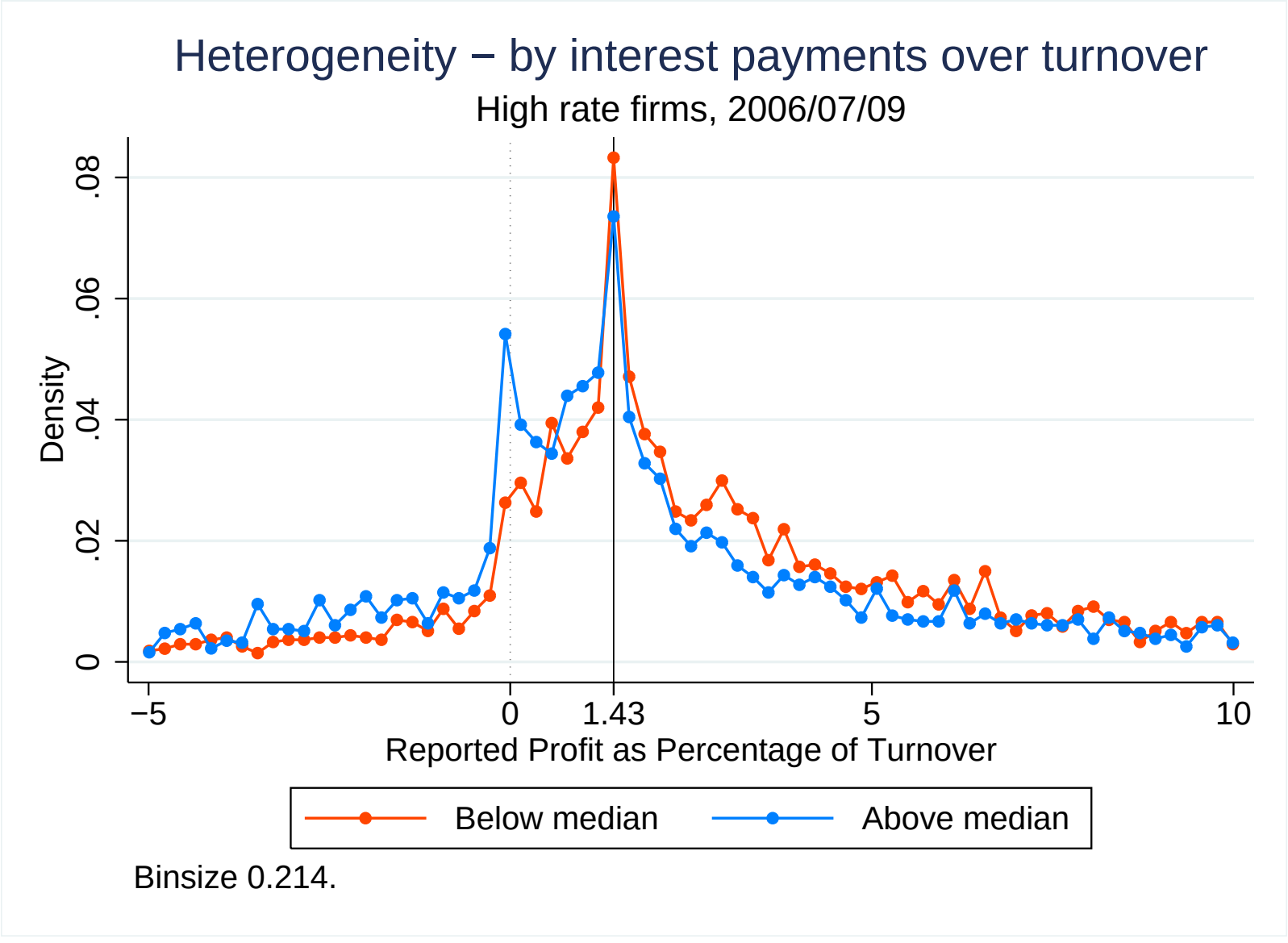
Back

Bunching Heterogeneity



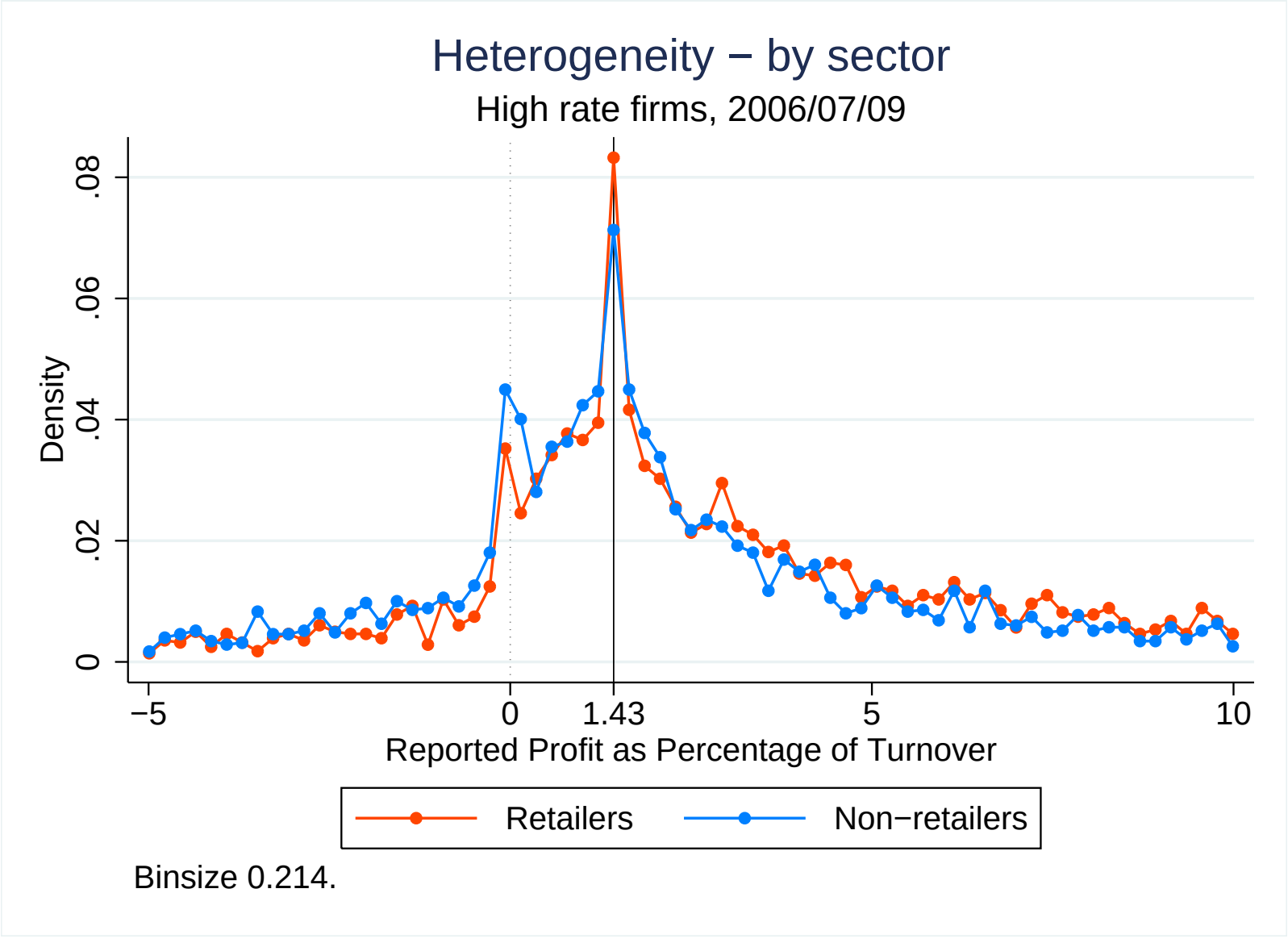
Back

Bunching Heterogeneity



Back

Bunching Heterogeneity



Back