Notching R&D Investment with Corporate Income Tax Cuts in China

Replication and Presentation

Shangze Dai, C.H. Lu, Rui Guo

January 27, 2025

Outline

Policy Background

- **China's R&D Tax Incentive Policy (2008): Supporting Innovation through Tax Reductions**:
  - Reduced corporate income tax from 33% to 15% for firms meeting R&D intensity thresholds.
  - Thresholds vary by firm size, for example:
    - Small firms: R&D intensity $\geq 6\%$
    - Medium firms: R&D intensity $\geq 4\%$
    - Large firms: R&D intensity $\geq 3\%$
- Goal: Encourage firms' innovation and social economic growth.
- Potential issue: Firms might cheat (engage in **relabeling**)to meet thresholds.

General Social Concerns

- **For Society and Government**
  - Is this policy being used effectively to drive innovation?
- **For Firms**
  - How do firms react? Are they increasing R&D or cheating?
  - Is the incentive fair? Do different-sized companies benefit?

Specific Research Questions

- **How do firms respond to tax incentives?**
  - Do they adjust behavior to meet R&D intensity thresholds?
- **What is the extent of relabeling?**
  - How much of reported R&D is genuine versus strategic relabeling?
- **Do firm size and resources matter?**
  - Are large firms more likely to increase real R&D?
  - Are small firms more reliant on relabeling?
- **What is the impact of real R&D on productivity?**
  - Does real R&D lead to meaningful economic growth?

Why Is This Study Important?

- **Policy Implications**:
  - Help governments evaluate the effectiveness of tax incentives.
  - Provide insights for improving future policy designs.
- **Academic Contribution**:
  - First to quantify relabeling behavior
  - Highlights firm-size heterogeneity in policy response.
- **Societal Impact**:
  - Ensures fair and efficient use of public resources.
  - Encourages innovation-driven growth while reducing manipulation.
- **......**

Policy Background

Before 2008, firms with an R&D intensity above 5% qualified for high-tech status, enjoying a reduced 15% tax rate instead of 33%. After 2008, thresholds were adjusted to 3%, 4%, and 6% based on firm size, creating strong incentives for firms to report higher R&D through changes in average tax rates.

Table 1—: **Requirements of the InnoCom Program**

Requirement	**Before 2008**	**After 2008**
R&D Intensity	5%	6% if sales $< 50M$ 4% if $50M <$ sales $< 200M$ 3% if sales $> 200M$
Sales of High Tech Products	60% of total sales	
Workers with College Degree	30% of workforce	
R&D Workers	10% of workforce	
Certifying Agency	Local Ministry of Science & Technology	Ministries of Science & Technology, Finance and National Tax Bureau

Note: Size thresholds in millions of RMB, where 50 M RMB ≈ 7.75 M USD and 200 M RMB ≈ 30 M USD.

Practical Phenomenon

The main data come from the Chinese State Administration of Tax (SAT, 2008-2011). They supplement these data with the Chinese Annual Survey of Manufacturing (ASM) (NBS, 2006-2007).

Figure 2. : Bunching at Different Thresholds of R&D Intensity (2011)

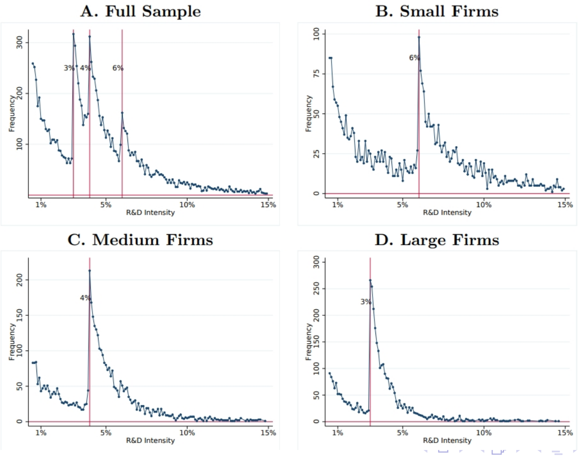
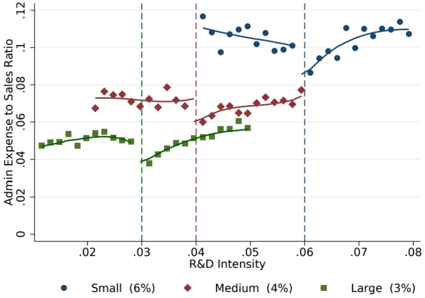
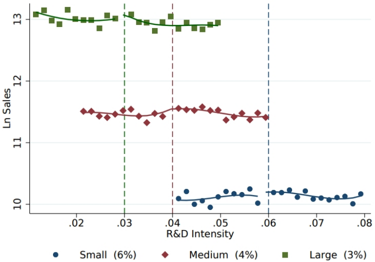


Figure 5. : Empirical Evidence of Relabeling



A. Lack of Sales Manipulation around R&D Intensity Thresholds



Basic Setting

ϕ_{it} is log TFP

$$\text{cost function } c(\phi_{it}, w_t) = w_t \exp\{-\phi_{it}\}$$

$$\phi_{i,t} = \rho\phi_{i,t-1} + \varepsilon \ln(D_{i,t-1}) + u_{it},$$

$D_{i,t-1}$ is R&D investment and $u_{i,t} \sim$ i.i.d. $N(0, \sigma^2)$

We assume that the firm faces a demand function with a constant elasticity: $\theta > 1$. This setup implies that firm sales are given by $\theta\pi_{it}$ and that we can write expected profits as follows:

$$\mathbb{E}[\pi_{it}] = \tilde{\pi}_{it} D_{i,t-1}^{(\theta-1)\varepsilon}$$

where $\tilde{\pi}_{it} \propto \mathbb{E}[\exp\{(\theta-1)\phi_{it}\}|\phi_{i,t-1}]$ measures the non-R&D expected profitability of the firm.

Basic Setting

The firm faces a constant elasticity demand function given by:

$$q_{it} = p_{it}^{-\theta} B_t,$$

where $\theta > 1$ is the demand elasticity and B_t is the aggregate demand shifter. In a given period, the firm chooses p_{it} to:

$$\max_{p_{it}} \quad p_{it}^{1-\theta} B_t - p_{it}^{-\theta} B_t c(\phi_{it}, w_t).$$

The profit-maximizing p_{it} gives the familiar constant markup pricing:

$$p_{it}^{*} = \frac{\theta}{\theta - 1} c(\phi_{it}, w_t),$$

1.25. We also normalize the input cost $w_t \equiv 1$ for the rest of our analysis. We can then write per-period profits as:

$$\pi_{it} = \frac{1}{\theta}\text{Revenue}_{it} = \frac{(\theta-1)^{\theta-1}}{\theta^{\theta}} \left[\exp(\phi_{it})\right]^{\theta-1} B_t.$$

Uncertainty and R&D investment enter per-period profits through the realization of log TFP ϕ_{it}. We can write expected profits as follows:

$$\begin{aligned}
\mathbb{E}[\pi_{it}] &= \frac{(\theta-1)^{\theta-1}}{\theta^{\theta}} B_t \left[\exp(\rho(\theta-1)\phi_{i,t-1} + (\theta-1)^2\sigma^2/2)\right] D_{i,t-1}^{(\theta-1)\varepsilon} \\
&= \tilde{\pi}_{it} D_{i,t-1}^{(\theta-1)\varepsilon},
\end{aligned}$$

Linear Income Tax

firm's inter-temporal problem is given by:

$$\max_{D_1} (1-t_1)(\pi_{i1} - D_{i1} - g(D_{i1}, \theta\pi_{i1})) + \beta(1-t_2)\tilde{\pi}_{i2} D_{i1}^{(\theta-1)\varepsilon},$$

adjustment cost of R&D investment $g(D_{i1}, \theta\pi_{i1}) = b \times \frac{\theta\pi_{1i}}{2}\left[\frac{D_{i1}}{\theta\pi_{1i}}\right]^2$

$$D_{i1}^* = \left[\frac{\beta(1-t_2)(\theta-1)\varepsilon}{1-t_1}\tilde{\pi}_{i2}\right]^{\frac{1}{1-(\theta-1)\varepsilon}}.$$

we express the firm's FOC in terms of the choice of R&D intensity, $d_{i1} = \frac{D_{i1}}{\theta\pi_{i1}}$, such that

$$\underbrace{-(1-t_1)(1+bd_{i1}^*)}_{\text{Increase in Investment Cost}} + \underbrace{\beta(1-t_2)\varepsilon(\theta-1)d_{i1}^{*\,(\theta-1)\varepsilon-1}\frac{\tilde{\pi}_{i2}}{(\theta\pi_{i1})^{1-(\theta-1)\varepsilon}}}_{\text{Productivity Gain from R\&D}} = 0.$$

Heterogeneous Tax

Assume now that the tax in the second period has the following structure, modeled after the incentives in the InnoCom program:

$$t_2 = \begin{cases} t_2^{LT} & \text{if } d_{i1} < \alpha \\ t_2^{HT} & \text{if } d_{i1} \geq \alpha \end{cases},$$

We first calculate the optimal profit of the firm conditioning on bunching at the notch, $\Pi(\alpha\theta\pi_1|t_2^{HT})$:

$$\Pi(\alpha\theta\pi_1|t_2^{HT}) = (1-t_1)\left(\pi_{i1} - \theta\pi_{i1}(\alpha+c) - \frac{b\theta\pi_{i1}}{2}\left[\frac{\alpha\theta\pi_{i1}}{\theta\pi_{i1}}\right]^2\right) + \beta(1-t_2^{HT})(\alpha\theta\pi_{i1})^{(\theta-1)\varepsilon}\tilde{\pi}_{i2}.$$

firms pay a fixed cost of certification: $c \times \theta\pi_{1i}$, where c varies across firms.

A firm that previously chose $d_{i1}^* < \alpha$ will bunch at the notch if $\frac{\Pi(\alpha|t_2^{HT})}{\theta\pi_{i1}} \geq \frac{\Pi(d_{i1}^*|t_2^{LT})}{\theta\pi_{i1}}$.

Heterogeneous Tax

expected cost of misreporting $h(D_{i1}, \tilde{D}_{i1})$ depends on the percentage of misreported $\delta_{i1} = \frac{\tilde{D}_{i1} - D_{i1}}{\tilde{D}_{i1}}$:

$$h(D_{i1}, \tilde{D}_{i1}) = \tilde{D}_{i1}\tilde{h}\left(\delta_{i1}\right),$$

If we want to misreport, of course, we want to misreport at the "notch" $\tilde{D}_1 = \alpha\theta\pi_1$

$$\begin{aligned}\max_{D_{i1}^K} \quad & (1-t_1)\left(\pi_{i1} - D_{i1}^K - \theta\pi_{i1}c - \frac{b\theta\pi_{i1}}{2}\left[\frac{D_{i1}^K}{\theta\pi_{i1}}\right]^2\right) - \alpha\theta\pi_1\tilde{h}\left(\frac{\alpha\theta\pi_1 - D_{i1}^K}{\alpha\theta\pi_1}\right) \\ & +\beta(1-t_2^{HT})\tilde{\pi}_{i2}(D_{i1}^K)^{(\theta-1)\varepsilon}\end{aligned}$$

Structural Estimation: Framework

- Calibrate the model using Survey from Head and Myer (2014)
- Set $\theta = 5$, $\rho = 0.725$, $\sigma = 0.385$, $\beta = 0.925$.
- Assume b, c are i.i.d. distributed where $b \sim \mathcal{LN}(\mu_b, \sigma_b^2)$ and $c \sim \mathcal{EXP}(\mu_c)$.
- Functional form for cost of relabeling: $\frac{\exp(\eta\delta)-1}{\eta}$, where δ is the fraction of reported R&D of relabeling and η is the intensity of regulatory enforcement.
- Target: estimate parameters $\Omega = \{\varepsilon, \eta, \mu_b, \sigma_b, \mu_c\}$

Structural Estimation: Framework

This paper obtains the estimate for Ω by minimizing the difference between moments generated by distribution of simulated firms and data moments, measured by following criterion function:

$$Q(\Omega) = \begin{bmatrix} m^D(\Omega) \\ m^B(\Omega) \end{bmatrix}' W \begin{bmatrix} m^D(\Omega) \\ m^B(\Omega) \end{bmatrix}$$

where W is a bootstrapped covariance weighting matrix. $m^D(\Omega)$ and $m^B(\Omega)$ are moments conditions based on the descriptive statistics and on the bunching estimators.

Structural Estimation: Identification

Table 3—: Structural Estimates

A. Point Estimates

	TFP Elasticity of R&D	Relabeling Cost	Distribution of Adjustment Costs		Distribution of Fixed Costs
	ε	η	μ_b	σ_b	μ_c
Model 1: Excluding Bunching Moments					
Estimate	0.089	5.900	7.989	2.047	0.687
Standard Error	(0.002)	(0.493)	(0.086)	(0.076)	(0.062)
Model 2: All Moments					
Estimate	0.091	6.755	8.011	2.014	0.532
Standard Error	(0.002)	(0.449)	(0.075)	(0.073)	(0.012)

Note: Estimates based on calibrated values of $\theta = 5$, $\rho = 0.725$, and $\sigma = 0.385$. Model 1 estimates the structural parameters using all moments except the bunching estimates. Model 2 uses all the available moments to estimate the structural parameters. See Section IV for estimation details.

B. Simulated vs. Data Moments

	Data	Simulated	
		Model 1: Excluding Bunching	Model 2: All Moments
R&D Dist. Moments: $m^D(\Omega)$			
Below the notch (%)			
$[0.3, 1.2]$	0.373	0.382	0.379
$[1.2, 2.1]$	0.113	0.157	0.146
$[2.1, 3]$	0.067	0.080	0.069
Above manipulated region (%)			
$[5, 6.3]$	0.056	0.055	0.057
$[6.3, 7.6]$	0.026	0.037	0.038
$[7.6, 9]$	0.012	0.026	0.027
Mean R&D intensity $[3\%, 5\%]$	0.037	0.035	0.035
Average TFP below notch	-0.015	-0.017	-0.020
Average TFP above notch	0.027	0.023	0.025
Admin cost ratio break at notch	0.9%	0.8%	0.7%
Bunching Moments: $m^B(\Omega)$			
Bunching Point d^{-*}	0.009	(0.009)	0.010
Increase in Reported R&D: Δd	0.157	(0.124)	0.150
Fraction of firms not bunching	0.641	(0.738)	0.665

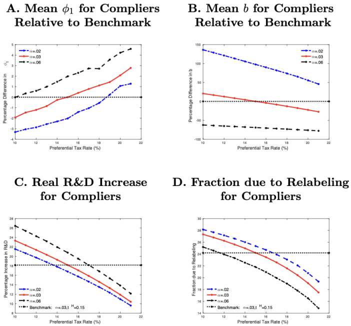
Note: This table compares the moments generated by our simulations with those from the data. The simulation is based on 30,000 firms. The moments that are not targeted by model 1 are in parentheses. The table shows our model does a remarkable job of matching 10 (13) moments from the data using a relatively parsimonious model based on 5 parameters.

Structural Estimation: Robustness and Sensitivity

- Assume that the TFP $\exp(\phi_1) \sim \mathcal{LN}$: OK
- Allow ε to be heterogeneous: cannot match the joint distribution of R&D intensity and TFP
- Allow cost function vary by firm size: OK
- Change the functional form of relabeling cost: Passed
- Allow fixed cost to be related with productivity: OK
- Change the productivity moment measure to Ackerberg, Caves and Frazer (2015)
- Test the OOS prediction: consistent with the reduced form results
- Evaluate the sensitivity of point estimates with each parameters

Counterfactual Analysis: Alternative Notches and Tax Cuts

Figure 8. : Simulated Counterfactual Policies: Selection and Relabeling



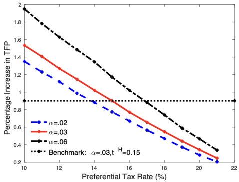
Note: These figures report the effects of different policy parameters on the selection of firms into the InnoCom program and on aggregate outcomes of interest. Panels A and B show that lower preferential tax rates select firms with higher adjustment costs and lower productivity. Panels C and D show how real and relabeled R&D respond to changes in parameters of the policy. See Section V for details on the structural model and the simulation.

Higher notches lead to selection of more productive and less adjustment cost firms. Low tax cut leads to selection of firms with less productive and higher adjustment cost

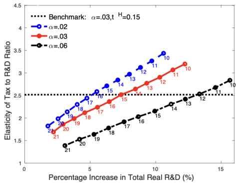
Counterfactual Analysis: Productivity and Fiscal Cost

Figure 9. : Simulated Counterfactual Policies: Productivity and Fiscal Cost of Stimulus

A. Average TFP Increase (Excluded Region)



B. Tax Revenue Cost of Stimulating Real R&D



Note: These figures report the effects of different policy parameters on aggregate outcomes of interest. Panel A shows how different reforms affect TFP. Panel B plots the elasticity of the tax cost to the government to the real R&D increase. This figure represents the fiscal cost curve of incentivizing R&D investment for the government and shows that notches that target larger firms have lower fiscal costs. See Section V for details on the structural model and the simulation.

Average TFP growth and participation rate are induced by tax rate cut.

Higher notches leads to select more productive firms and less tax loss

Counterfactual Analysis: Welfare Implication

In society level, spillover is the source of externality. Assume the TFP follows:

$$\phi_{it} = \rho\phi_{i,t-1} + \varepsilon \ln(D_{i,t-1}) + \zeta S_{t-1} + u_{it}$$

where S_{t-1} is the spillover effect from other firms. The social welfare is then:

$$\Phi_1(L - D_1 - H_1)(1 - \frac{t}{\theta})^{1-\gamma}(\frac{t}{\theta})^{\gamma} + \beta\Phi_2 L(1 - \frac{t}{\theta} + \frac{\tau}{\theta})^{1-\gamma}(\frac{t}{\theta} - \frac{\tau}{\theta})^{\gamma}$$

where Φ_t is the aggregated TFP, D_1 is the aggregated R&D adjustment expenditure and fixed cost, H_1 is the aggregated relabeling cost and τ is the fiscal cost.

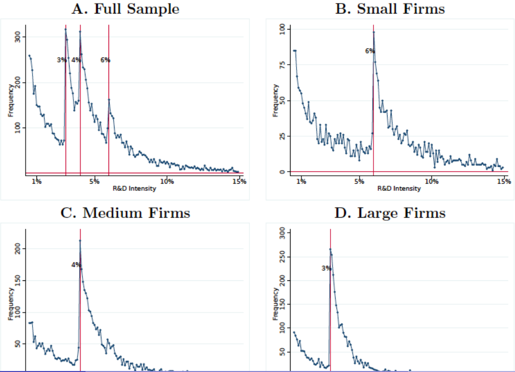
Counterfactual Analysis: Welfare Implication

- When no spillover, $\zeta = 0$, welfare decrease by 0.14% and efficiency Φ_2/Φ_1 increase by 0.12%.
- To break even the welfare, the spillover $\hat{\zeta} \geq 0.069$.
- Literature show that the InnoCom project is efficient $\zeta \approx 0.2$ and welfare increased by 0.27% and aggregated TFP increased by 0.53%

Policy Responses at Thresholds

- **Finding 1: Firms Adjust Behavior to Meet Thresholds**
  - Significant **bunching effect** near R&D intensity thresholds.
  - Firms increase R&D reporting just enough to qualify for tax incentives.
  - Different responses by firm size:
    - ★ Large firms: smoother distribution; rely more on real R&D.
    - ★ Small firms: sharp bunching; rely more on relabeling.

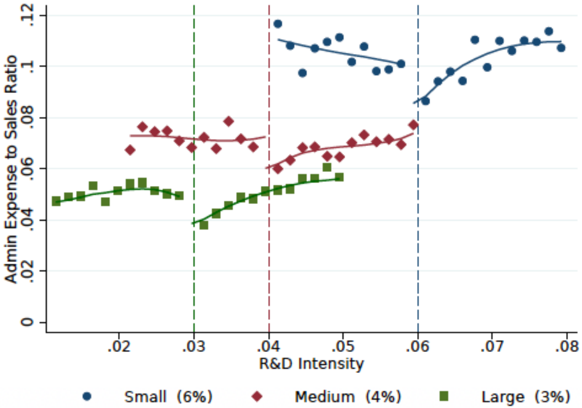
Figure 2. : Bunching at Different Thresholds of R&D Intensity (2011)



Extent of Relabeling

- **Finding 2: 24.2% of Reported R&D is Relabeled**
  - Administrative expenses decline significantly at thresholds.
  - Relabeled expenses make up a significant share of reported R&D.
  - Small firms rely more heavily on relabeling compared to large firms.

Figure 5. : Empirical Evidence of Relabeling



Firm Size Heterogeneity in Policy Response

- **Finding 3: Large vs. Small Firms**
  - Large firms:
    - More resources for real R&D investment.
    - Relatively lower reliance on relabeling.
  - Small firms:
    - Higher adjustment costs for real R&D.
    - Relatively higher reliance on relabeling.

Productivity Impact of Real R&D

- **Finding 4: Real R&D Significantly Increases Productivity**
  - Real R&D elasticity: Doubling R&D increases productivity by **9%**.
  - Implication: Encouraging real R&D is critical for long-term economic growth.

The author's first extension: Comparing Incentive Designs

- **Research Goal**:
  - Compare the effectiveness of **Notch-Based** and **Linear** tax incentives.
- **Results**:
  - **Notch-Based Incentives**: Reduce relabeling by **10-15%** and increase real R&D efficiency by **20-30%**.
  - **Linear Incentives**: Broader coverage but higher relabeling behavior, leading to lower fiscal efficiency.
- **Policy Insight**:
  - Targeted monitoring for firms near thresholds optimizes fiscal resources. So Notch-Based incentives are more effective at reducing manipulation.

The author's second extension: Welfare Analysis, Quantifying Social Gains

- **Framework**:
  - Uses a **quasi-linear utility model** to measure welfare impacts on consumers and firms.
- **Key Findings**:
  - **Consumer Welfare**: Benefits from innovation-driven price reductions.
  - **Firm Welfare**: Gains from reduced tax burdens and increased R&D.
  - **Total Welfare**: Net positive effects if spillovers exceed **6.9%**.

Extension 1: New Data Collection and Long-Term Analysis

- **Objective**:
  - Apply new dataset
  - Analyze short-term and long-term effects of tax incentives using new datasets.
- **Motivation**:
  - Short-term effects (e.g., immediate changes in R&D reporting) may differ from long-term impacts (e.g., sustained innovation and productivity growth).

Extension 2: Tax Revenue Changes and Spillovers

- **Objective**:
  - Quantify the short-term fiscal burden and long-term societal benefits/loss of tax incentives.
- **Motivation**:
  - Clarify tax revenue changes into:
    - Immediate revenue loss from tax reductions.
    - Long-term revenue gains from productivity improvements and innovation.

Extension 3: Mediation Analysis: Why Notch-Based Incentives Are More Effective?

- **Objective**:
  - Conduct mediation analysis to identify mechanisms driving the effectiveness of notch-based incentives.
- **Motivation**:
  - Verify if higher relabeling costs (related to oversight) are the primary reason for better effectiveness.
  - Explore alternative explanations.

Extension 4: Industry Heterogeneity

- **Objective**:
  - Extend the analysis to explore cross-industry differences in policy response.
- **Motivation**:
  - Different industries (e.g., high-tech electronics, construction, automotive) may have different levels of relabeling and real R&D investment. Also may get different effect form this policy.

Extension 5: Robustness Test TBD

- **Objective**:
  - Validate the key findings through additional robustness checks.
- **Motivation**:
  - Strengthen confidence in results by addressing potential concerns with data limitations or model assumptions.