Survival and Growth of Small and Midsized Enterprises: Evidence from Corporate Income Tax Cuts in China *

Junxue Jia † Wenzhuo Lu ‡ Guangrong Ma § Xiaochen Xie ¶

July 31, 2024

Abstract

Our study investigates China's notch-based corporate income tax cuts program implemented from 2014 to 2018 to support small and midsized enterprises (SMEs). Utilizing confidential administrative data from corporate tax returns, we show that these tax cuts encourage small businesses to expand production while causing some above-notch firms to reduce production for tax arbitrage. We then develop a structural model to evaluate the impact of the notch-based tax cuts on the overall economic environment and the growth of SMEs. Finally, we compare the effectiveness of various size-dependent policies – including notch-based tax cuts, kink-based tax cuts, and output-based subsidies – in boosting aggregate economic growth and maintaining tax revenue.

Keywords: corporate income tax, size-dependent policy, SME

*We express our gratitude to Costas Arkolakis, Shengmao Cao, Jonathan Eaton, Jingting Fan, Kala Krishna, Yue Liu, Chang Sun, Jim Tybout, Felix Tintelnot, Shaoda Wang, and Daniel Yi Xu for their invaluable insights and comments. We are grateful to seminar participants at the Central University of Finance and Economics, Liaoning University, Peking University, Renmin University of China, Shanghai University of Finance and Economics, Tsinghua University, RUC-Xiamen U Public Finance Forum, the Penn State Alumni Academic Conference, and the Asian Meeting of Econometric Society 2023 for constructive feedback. Xuan Liang and Zijie Tang provided outstanding research assistance. Any errors that appear are solely our responsibility.

†Renmin University of China. Email: jiajunx@ruc.edu.cn.

‡University of International Business and Economics. Email: luwenzhuo@uibe.edu.cn.

§Department of Public Finance, School of Finance, Renmin University of China. 59 Zhong Guan Cun Ave, 835 Mingde Main Building, Haidian District, Beijing, China, 100872. Email: grma@ruc.edu.cn.

¶Department of Public Finance, School of Finance, Renmin University of China. 59 Zhong Guan Cun Ave, 831 Mingde Main Building, Haidian District, Beijing, China, 100872. Email: xiexiaochen@ruc.edu.cn.

1 Introduction

Governments worldwide prioritize the stability and growth of small firms and have introduced various support programs to assist them. Common measures include tax cuts, subsidies, and simplified operational procedures for firms below a certain size threshold. While these policies are popular, they remain controversial. On one hand, small firms are vital for job creation and often face growth challenges due to market failures and capital market imperfections (Decker et al., 2014). On the other hand, size-dependent policies may incentivize firms to stay below the threshold and lead to misallocation by crowding out larger firms (Guner et al., 2008; Garicano et al., 2016; Martin et al., 2017; David and Venkateswaran, 2019). Therefore, understanding the overall economic impact and efficiency of various size-dependent policies is crucial.

This study investigates the impact of China's Halve-Levy program, one of the most significant and impactful corporate income tax cuts in China over the past 20 years, on economic output and tax revenue. The program offers a much lower corporate tax rate for firms with annual taxable income below a specified threshold, resulting in sudden increases in the average tax rate when firms surpass this threshold. Meanwhile, China consistently increased this taxable income threshold annually between 2014 and 2018, providing a natural experiment for studying the effectiveness and efficiency of size-dependent tax policies.

Analyzing the effects of this program offers valuable insights for other countries. As shown in Table 1, many countries, except those with a linear corporate tax structure, provide preferential tax rates for small and midsized enterprises (SMEs). These countries can be categorized into two groups based on their tax reduction methods: notch-based and kink-based. In the notch-based tax system (China's case), the average tax rate a firm pays can suddenly increase when the firm's income surpasses certain thresholds, creating discontinuous jumps in the average tax rate. In contrast, the kink-based tax system allows a firm earning just below a certain threshold to pay a lower tax rate on its entire income, but once it crosses that threshold, the rate on any additional income jumps to a higher level. This leads to kink points when graphing the marginal tax rate against taxable income. Constructing a quantitative framework to capture how tax policy design affects firms' output will help understand the economic impacts of the notch-based tax system in China and can be extended to analyze the effects of similar size-dependent tax policies in other countries.

Using confidential administrative data from corporate tax returns in Shanxi Province, which includes changes in eligibility thresholds for the Halve-Levy program, we provide evidence of how Chinese small-scale firms react to the program. We then develop a structural model to capture firms' production behaviors, emphasizing the influence of tax cuts on their output. Based on this model, we conduct counterfactual analyses to answer the following

Table 1: Corporate Income Tax Cuts for Firms of Different Sizes by Country

Linear tax (no cuts)		Kink-based cuts	Notch-based cuts
Austria	Italy	Argentina	Australia
Czech Republic	Mexico	Brazil	Chile
Denmark	New Zealand	Belgium	China
Finland	Norway	Canada	India
Germany	Slovenia	France	Poland
Greece	Sweden	Japan	Slovakia
Hungary	Switzerland	South Korea	Spain
Ireland	Turkey	Luxembourg	
Israel	USA	Holland	
		Portugal	
		South Africa	
		UK	

Data Source: Wolters Kluwer Taxation Database.

key questions:1. How do firms make production-scale decisions, and how do they respond to tax cuts? 2. What is the impact of China's notch-based tax cuts on overall economic output, and how do different types of firms, affected by the policy shift, collectively contribute to the net aggregate change? 3. Can alternative policies, such as kink-based tax cuts and output-based subsidies, achieve the same economic stimulus effect while avoiding larger reductions in tax revenue? To what extent does the more effective policy reduce the tax expenditure associated with the preferential program?

In Section 3, we begin with three stylized facts on the impact of corporate tax cuts under the Halve-Levy program on firms' taxable income. First, using bunching estimates proposed by Kleven and Waseem (2013), we show that the notch-based Halve-Levy program causes some firms to reduce their output just below the notch thresholds to benefit from tax reductions, leading to pronounced bunching patterns in the distribution of taxable income. Specifically, firms near the thresholds reduced their reported taxable income by 4.6%, 4.5%, 7.0%, 7.9%, and 4.3% for each year between 2014 and 2018, respectively. However, many firms that could benefit from reducing output to reach the notch threshold do not do so due to optimization frictions, such as unawareness of the tax cuts program or the existence of adjustment costs. The estimates of the proportion of non-responders fluctuate between 58% and 82% for each year during the sample period.

Next, we examine the positive impact of the Halve-Levy program on production expansion. As the program raises the thresholds for tax cuts annually, we identify the taxable income ranges newly covered by tax cuts each year. We divide SMEs into five groups based on these newly covered ranges and plot the share of taxable income for each group relative to the total taxable income among the five groups. The results show that each group ex-

periences an increase in its taxable income share in the first year they receive tax benefits, which then decrease in the following years as other groups are successively included in the Halve-Levy program. These observations indicate that the tax reductions encourage eligible firms to expand production.

Third, we demonstrate that the positive impact of tax reductions is more significant for firms with lower taxable income. For each group representing the newly covered taxable income range, we further divide firms into low-productive and high-productive subgroups based on whether their taxable income was below or above the median taxable income within the subgroup in the previous year. We then calculate the average growth rate of taxable income for each subgroup, finding that the average growth rate of low-productive firms is consistently higher than that of high-productive firms.

Motived by the stylized facts, we develop a structural model that examines a firm's adjustments to production decisions in response to changes in costs and tax incentives in Section 4. Our model builds upon Arkolakis' entry model (2010) by incorporating the impact of corporate income tax relief policy on firms' production decisions. Specifically, our model considers factors such as after-tax marginal revenue at a given production scale, the marginal opportunity cost faced by management in investing or forfeiting at that scale, and the influence of tax incentives. By incorporating China's corporate tax cuts into our model, we propose that a reduced corporate income tax rate can amplify after-tax returns, prompting policy-covered small business to expand production. Simultaneously, the mechanism of notch-based tax cuts encourages firms near the notch threshold to optimize their profits strategically by reducing production to qualify for a lower tax rate.

Furthermore, our model introduces the concept of heterogeneous firms characterized by varying productivity levels, resulting in diverse production decisions. Under a linear tax regime, different firms have distinct production sizes. When considering the Halve-Levy program, we observe that the smallest firms may exit the market under a 25% linear tax system; however, they may survive under the current policy. Similarly, small firms that would have endured under a 25% tax system can now expand their output owing to the lower corporate income tax rate. Some firms positioned near the higher notch point experience tax-incentivized motivations to reduce their production, while specific larger firms remain relatively unaffected by the tax cuts policy designed for small and micro firms.

In Section 5, we describe our two-step parameter estimation process. First, we estimate supply- and demand-related parameters to reconstruct firms' production curves using a non-linear least squares approach within a standard entry framework. In the second step, employing a non-parametric estimation method and based on the retrieved production curve, we estimate the parameters associated with firms' bunching behavior. The structural estimates of firms' bunching behavior reveal that during the first five years of the Halve-Levy

program, over 50% of firms that would benefit from bunching just below the notch threshold do not do so. These results, consistent with the non-responding estimates in Section 3.1, underscore the presence of significant frictions in firms' awareness of tax cuts program and adjustment costs.

Next, we explore the role of different groups of firms affected by the Halve-Levy program in determining the overall economic outcome. We categorize firms based on their responses to the tax relief policy into four groups: new entrants influenced by the policy, firms covered by the policy that expand production, firms strategically reducing production to benefit from the notch, and firms not covered by the policy that do not engage in such arbitrage. Our counterfactual analysis reveals that the primary driver of increased economic output is the production expansion by existing firms. In contrast, the contributions from new entrants and the production reductions due to bunching behavior of the third group have minimal impact on the overall output shift in our sample year.

We also examine the time trend of the policy's effects. As the taxable income threshold of the Halve-Levy program increases, the newly covered firms are larger and experience a smaller boost in output from the tax cuts policy compared to smaller firms. At the same time, as firms become more adept at reducing output to benefit from the notch, the negative impact of this bunching behavior intensifies. Consequently, the growth rate of aggregate production output, due to the preferential policy, slows down each year. These simulation results provide valuable insights for determining optimal taxable income thresholds to effectively stimulate the overall economy.

Furthermore, we evaluate the merits and effectiveness of various government size-dependent policies for small businesses compared to the benchmark notch-based tax cuts. Specifically, we compare kink-based tax cuts and output-based subsidies, where the opportunity cost of operating is subsidized. Our assessment focuses on their ability to stimulate economic output while preserving tax revenues, all within a consistent framework of tax rates and taxable income thresholds. Simulations demonstrate that if firms are well-informed about policy changes and adapt accordingly, the notch policy increases total economic output by approximately 0.04–0.05% annually. Conversely, owing to the absence of policy distortions caused by bunching behaviors, the kink and subsidy policy result in a more substantial increase, ranging from 0.1 to 0.13%.

In terms of minimizing tax revenue reduction, all three types of policies promote economic growth but reduce tax revenues. Among these, the kink policy generally reduces annual tax revenues by approximately twice as much as the notch policy, as the tax breaks apply to the portion of all firms' incomes below the threshold. Contrarily, an output-based subsidy policy stimulates economic output and minimizes the reduction in tax revenues. This is because a subsidy only needs to compensate for the opportunity cost of engaging in business

activities to promote production expansion, whereas a tax incentive involving a lower tax rate would require forfeiting a significant proportion of tax revenues. However, implementing such an efficient subsidy policy requires establishing a precise mapping between the subsidy amount and the operational profit of each specific firm, necessitating advanced forecasting and modeling capabilities.

Finally, we compare notch-based tax cuts, kink-based tax cuts, and output-based subsidy policies that achieve the same effect of stimulating economic output. Moreover, we assess the extent to which each policy type minimizes the reduction in tax revenues. Our counterfactual analysis demonstrates that if the government optimizes aggregate output, kink-based tax cuts achieve the same policy effect while causing a reduction in tax revenues only 1/8–1/3 of that caused by notch-based tax cuts. Furthermore, the reduction in tax revenue from the subsidy policy for small firms is merely one-tenth of that resulting from notch-based tax cuts.

We build a counterfactual framework to compare the merits and drawbacks of various SME preferential policies. Notably, while notch-based corporate tax cuts are widely adopted by various governments, including China, owing to their targeted support for SMEs and straightforward administration, they exhibit inefficiencies related to firms' production cuts near the notch point. Consequently, the policy effects are notably less efficient compared to kink-based tax cuts and output-based subsidies to firms. Our counterfactual results are based on China's policy experience, while the structural framework we develop in this study holds potential for broad application and extension to study various size-dependent policies in other countries.

Our study contributes significantly to the existing literature in several ways. First, we contribute to the ongoing debate regarding the effectiveness of preferential tax policies in supporting small businesses. While some research has found positive effects of tax cuts in stimulating small firms (Harju et al., 2020), others express concerns about policy distortions. Many studies have observed bunching around the cutoff of size-dependent policy thresholds (Boonzaaier et al., 2017; Bachas and Soto, 2020; Devereux et al., 2014). Conversely, our comprehensive evaluation of size-dependent policies considers the benefits from incentive effects on incumbent firms and new entrants, as well as the losses from bunching around bracket cutoffs and the potential crowding-out effects on large firms' businesses. By building a quantitative framework that incorporates both aspects of tax policies, we conduct counterfactual analysis to compare the economic efficiencies of alternative size-dependent policies for SMEs.

To better understand how tax cuts influence the decision-making of SMEs, we draw on methods from Arkolakis (2010) in the trade literature, adapting them to model simplified trade elements and incorporating tax incentives into firms' decision-making processes. This family of trade entry models originated from Melitz (2003) and Chaney (2008), who

introduced the concept of firms' heterogeneous productivity following Pareto distribution. Building upon this framework, Arkolakis (2010) and Eaton, Kortum, and Kramarz (2011) expanded the Melitz-Chaney model by introducing market penetration costs to capture the performance of small and medium-sized exporters within the trade market. In our model setup, we integrate the incentive of corporate income tax for SMEs into the market expansion cost function of Arkolakis (2010) to analyze the impact of current policies on SMEs' production decisions. Our model extends these classic models without additional complexity and effectively aligns with SMEs' production decisions, seamlessly facilitating counterfactual analyses aimed at assessing the effects of different types of policies on SMEs and aggregate output.

This study contributes to the broader literature concerning the impact of corporate taxes on business behavior, particularly their impact on corporate investment and output expansion (Hassett and Hubbard, 2002). Extensive research has explored tax reforms that modify tax rates or introduce changes in tax burdens, often involving adjustments in depreciation incentives. Additionally, various studies have analyzed the elasticity of taxable income (ETI) in response to changes in corporate income tax rates (Devereux et al., 2014; Gruber and Rauh, 2007; Dwenger and Steiner, 2012; Patel et al., 2015; Baches and Soto, 2020). In line with this context, our analysis also reveals a positive effect of corporate income tax rate reductions on firm production.

The subsequent sections of this paper are organized as follows. Section 2 provides a review of the policy background and describes the dataset, while Section 3 presents the descriptive evidence. Subsequently, Sections 4 and 5 describe the model and its empirical implementation, and Section 6 introduces the counterfactual analysis. Finally, Section 7 outlines the conclusions.

2 Policy Background and Data

2.1 Tax reform

Since 2008, regular firms in China have been subject to a flat 25% tax rate under the corporate income tax law. In response to the Global Financial Crisis, the Chinese government introduced the Halve-Levy program, one of the most significant and impactful corporate income tax cuts in China over the past 20 years, in December 2009. This program allowed firms with annual taxable income of less than RMB 30,000 to pay a reduced corporate income tax rate of 10%. Over the years, the upper limit for taxable income under this program has gradually increased: RMB 60,000 for 2012-2013, RMB 100,000 for 2014, RMB 200,000 for the first three quarters of 2015, RMB 300,000 for the fourth quarter of 2015 and 2016, RMB 500,000 for 2017, and ultimately RMB 1,000,000 for 2018. According to the Chinese

government, this corporate tax cut program saved firms RMB 89.5 billion (approximately USD 12.3 billion) in 2018, accounting for 2.5% of total corporate income tax revenue.[1] Hence, due to the Halve-Levy program, China's notch-based corporate income tax system exhibits discontinuous jumps in the average tax rate at specific thresholds, similar to the tax systems in Australia, Spain, India, and other countries.

2.2 Data

In this paper, we use confidential administrative data from corporate tax returns in Shanxi Province, covering the years 2014 to 2018. Our dataset includes any firm that pays a positive amount of corporate income tax, even as little as 1 cent, to the Shanxi tax administration, making it ideal for studying small and micro firms. In contrast, one of the commonly used datasets for studying Chinese firms – the Annual Survey of Manufacturing Firms – only includes manufacturing firms with annual sales exceeding 5 million yuan (increased to 20 million yuan since 2011).

Our dataset is provided by the Shanxi Provincial Tax Bureau of the Chinese State Administration of Tax, an entity that handles tax collection and auditing for all registered firms within the province.[2] In 2018, Shanxi Province, located in northern China, had a population of 37.2 million, ranking 18th among Mainland China's 31 provinces—approximately equivalent to Canada's population. Additionally, Shanxi's GDP in 2018 reached 1682 billion yuan (USD 262.7 billion), similar to Pakistan's GDP for the same year.

The dataset includes detailed records of tax payments and other financial statement information, such as taxable income, profit, sales, and employment, used for tax-related calculations. Table 2 shows that the dataset comprises approximately 470,000 observations, with the count progressively increasing over time from 80,341 in 2014 to 116,572 in 2018. As the Halve-Levy threshold is gradually raised, the share of firms eligible for corporate income tax reductions increases from 84.6% in 2014 to 95.6% in 2018. Overall, around 91.6% of firms halve their corporate income tax during the sample period.

In addition to the restriction on taxable income, the Halve-Levy program requires eligible

[1]The Halve-Levy program is not the only tax cut policy during the sample period. Since 2008, firms with annual taxable income below 300,000 RMB benefit from a reduced corporate tax rate of 20%. Thus, during the period from 2008 to the fourth quarter of 2015, there are two thresholds: firms with taxable income below the Halve-Levy threshold enjoy a flat 10% tax rate, firms with income above the Halve-Levy threshold but below 300,000 RMB are eligible for a flat 20% tax rate, and firms with taxable income above 300,000 RMB pay the standard 25% tax rate. After the fourth quarter of 2015, these two thresholds merged, so firms with income below the Halve-Levy threshold benefit from a 10% tax rate, while all others pay a 25% tax rate. While it is not immediately obvious when plotting the distribution of firms' taxable income over time, we do consider the existence of these two thresholds when conducting the counterfactual analysis.

[2]Prior to 2018, China's tax administration comprised two entities: the National Taxation Bureau and the Local Taxation Bureau. The former handled corporate income tax revenue from firms established before 2002, while the latter managed revenue from those established afterward. In 2018, China merged these two, creating a unified tax system. Our dataset, acquired in 2019 post-merger, includes firms from both periods.

Table 2: Overview

Year	Halve-levy Threshold (yuan)	Observations	Eligible firms	Share
2014	100,000	80,341	67,998	84.6%
2015	200,000	80,205	71,891	89.6%
2016	300,000	87,459	80,382	91.9%
2017	500,000	100,355	94,333	94.0%
2018	1000,000	116,572	111,481	95.6%
All	-	464,932	426,085	91.6%

Notes: Column 4 reports the number of firms that are eligible for corporate income tax cut, and Column 5 is the ratio of the number in Column 4 to that in Column 3.

firms to have no more than 100 employees for manufacturing firms or 80 employees for other firms to qualify for tax cuts. We find that, conditional on taxable income falling below the Halve-Levy thresholds, more than 94.1% of firms also meet this employee requirement. Additionally, the program requires eligible manufacturing firms to have no more than 30 million RMB in total assets. Though our data lacks information on firms' total asset values, we use the Tax Survey Data and find that over 90% of firms below the Halve-Levy thresholds also meet the total asset requirement. Hence, during estimation, we treat the total amount of taxable income as the sole criterion to distinguish eligible and non-eligible firms for tax reductions.

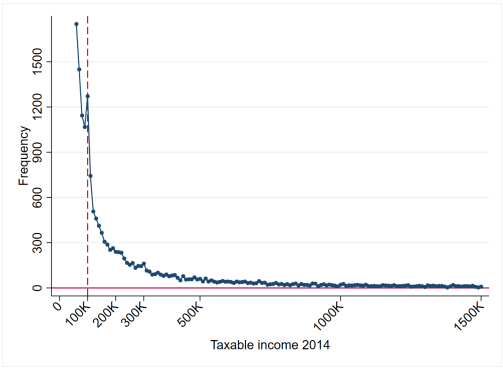
3 Descriptive Evidence

In this section, we present three key findings on the impact of tax cuts under the Halve-levy program on firms' taxable income. We show that the program causes some firms near the threshold to reduce their output to benefit from lower taxes, while encouraging some others to increase production. Besides, the positive effects of tax reductions are more significant for firms with lower taxable income. This descriptive evidence provides a preliminary understanding of the tax policy's effects on firms and lays the groundwork for developing a structural model for a more detailed counterfactual analysis.

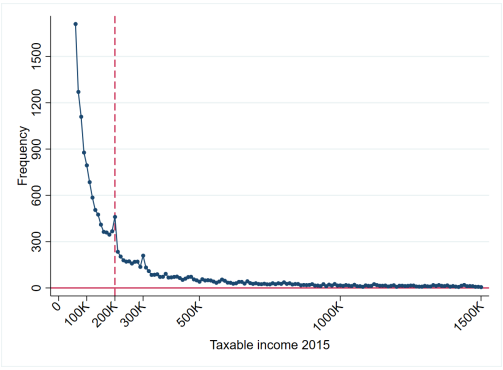
3.1 Fact 1: Bunching

The Halve-levy program creates significant discontinuous jumps in the average tax rate at program thresholds, inducing firms to bunch just below these thresholds to benefit from tax reductions. Figure 1 illustrates the distribution of taxable income across different years. Panel A shows that in 2014, there was a pronounced and sharp excess bunching below 100 thousand RMB, the program threshold for that year, accompanied by a missing mass above the threshold. No evidence of bunching was found at other points. Similarly, Panels B to E show evident bunching at 200 thousand, 300 thousand, 500 thousand, and one million RMB

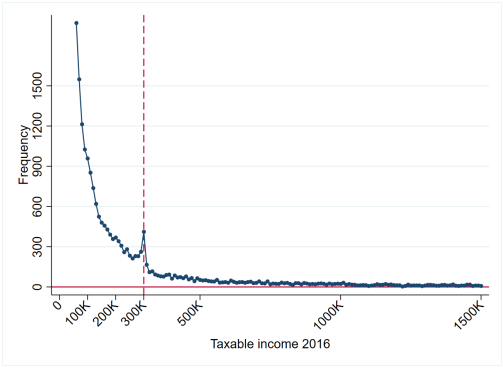
Figure 1: Bunching at Different Thresholds of Taxable Income



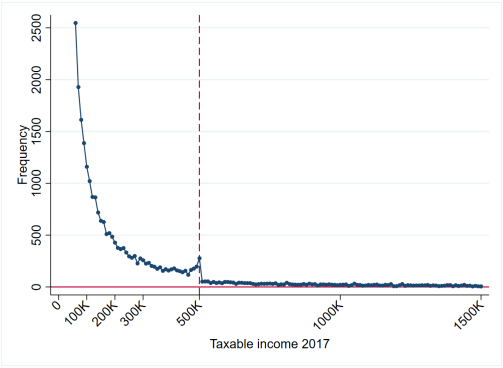
(a) Panel A. 2014



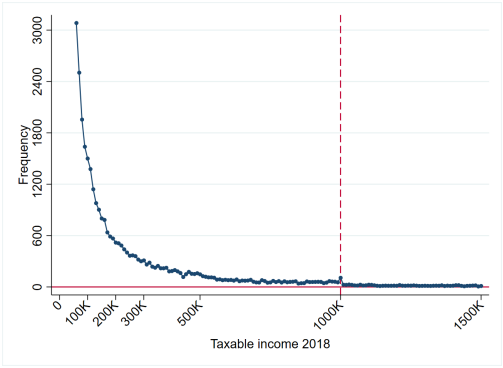
(b) Panel B. 2015



(c) Panel C. 2016



(d) Panel D. 2017

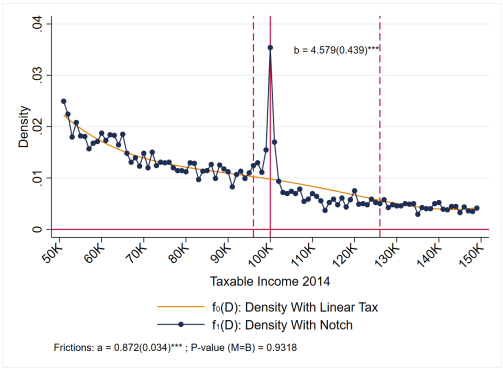


(e) Panel E. 2018

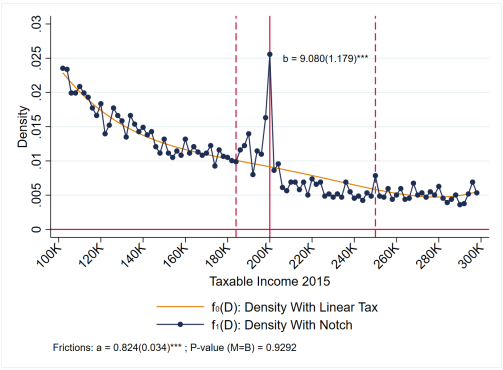
Note: This figure plots the distribution of the taxable income for firms with taxable income between 50,000 RMB and 1.5 million RMB. Panels A, B, C, D, and E report the distribution for each year from 2014 to 2018, respectively. Notably, large fractions of firms bunch immediately below the thresholds (100,000 RMB in 2014, 200,000 RMB in 2015, 300,000 RMB in 2016, 500,000 RMB in 2017, and 1 million RMB in 2018) at which they can enjoy the reduced tax rate.

Source: Shanxi Administrative Tax Return Database.

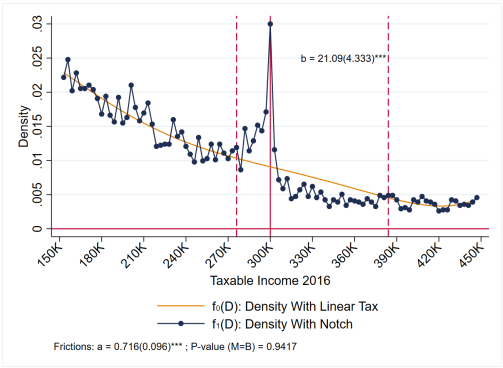
Figure 2: Estimated Counterfactual Densities of Taxable Income



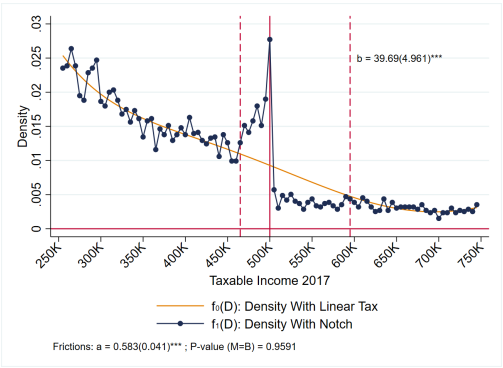
(a) Panel A. 2014



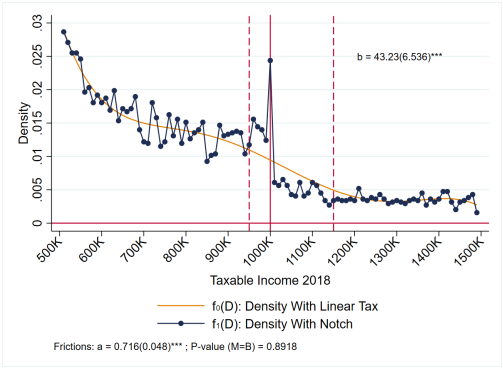
(b) Panel B. 2015



(c) Panel C. 2016



(d) Panel D. 2017



(e) Panel E. 2018

Note: This figure reports the results of bunching estimation for each year from 2014 to 2018. Within each panel, we plot the empirical density of taxable income in blue and the estimated counterfactual taxable income in orange. The lower bound TI_l and upper bound TI_u for the excluded region is indicated by vertical short-dashed lines. α^* represents the fraction of firms constrained from participating in the program. We provide the p-value of the test assessing the equality of the missing mass and the excess mass. b represents the average decrease in taxable income in the exclusion regions. The horizontal coordinate is in thousands of RMB.
Source: Shanxi Administrative Tax Return Database.

for the years 2015, 2016, 2017, and 2018, respectively.[3] Overall, Figure 1 indicates that the Halve-levy program led to pronounced bunching patterns in the distribution of taxable income.

Next, to quantify the overall response in taxable income, we follow the approach of Kleven and Waseem(2013) and Chen et al.(2021), which uses the observed distribution of taxable income, $f_1(\cdot)$, to infer the counterfactual distribution without the Halve-levy program, $f_0(\cdot)$. This approach relies on the assumption that only firms with taxable income in the *excluded region* $[TI_l, TI_u]$ respond to the program. This assumption allows us to use firms drawn from $f_1(\cdot)$ that are outside this region to recover $f_0(\cdot)$. Specifically, we first divide the data into bins of taxable income, TI, and then estimate the following polynomial:

$$c_j = \sum_{i=0}^{p} \beta_i \cdot (TI_j)^i + \sum_{s=TI_l}^{TI_u} \gamma_s \cdot \mathbb{I}(TI_j = s) + \varepsilon_j,$$

where c_j is the number of firms in the bin corresponding to taxable income TI_j, p is the order of the polynomials, TI_l and TI_u are the lower and upper bound of the excluded region, and the γ_s's are intercept shifters for each of the bins in the excluded interval. Then $\hat{c}_j = \sum_{i=0}^{p} \hat{\beta}_i \cdot (TI_j)^i$ is an estimate for $f_0(TI)$. Intuitively, when only firms in the exclusion region $[TI_l, TI_u]$ respond to the program, $\hat{f}_0(TI)$ will be equal to $f_1(\cdot)$ outside this region. To ensure that the estimation is not contaminated by firm response to the program, TI_l and TI_u are determined by a data-driven procedure that $\hat{f}_0(\cdot)$ has the same mass over the excluded region as $f_1(\cdot)$.[4]

Comparing this counterfactual distribution to the observed distribution we can estimate the excess bunching mass to the left of the threshold: $\hat{B} = \sum_{j=TI_l}^{TI^*} (\hat{c}_j - c_j) \geq 0$ with the Halve-levy threshold TI^*. Then, we derive the standard bunching estimator $\hat{b} = \frac{\hat{B}}{\frac{1}{2}[\hat{f}_0(TI^*) + \hat{f}_0(TI_u)]}$, which is defined as the ratio of excess bunching mass to the average height of the counterfactual density at the threshold. Intuitively, the estimator $\hat{b}$ characterizes the percentage decrease in the average taxable income in the exclusion region.

Figure 2 displays the results of this estimation. In each panel, the blue line with circle markers represents the observed distribution of taxable income $f_1(\cdot)$, the orange line represents the estimated counterfactual distribution $\hat{f}_0(\cdot)$, and the vertical short-dashed lines displace the excluded region. Panel A shows that the bunching estimator is 4.579 (s.e. 0.439)

[3] There was a small and insignificant bunching at 300 thousand in 2015, aligning with the fact that the threshold was raised from 200 thousand to 300 thousand in the fourth quarter of 2015.

[4] We follow Diamond and Persson (2016) to select the width of the excluded region (i.e., $[TI_l, TI_u]$) and the order of polynomial. In particular, we use K-fold (K = 5) cross-validation to search over a triple (p, TI_l, TI_u) that determines the excluded region. We first test the hypothesis that $f_0(\cdot)$ and $f_1(\cdot)$ have equal mass over the exclusion restriction and keep the combination of parameters that do not reject the test of equality at the 10% level. Then, we compute the sum of squared errors across the test subsamples and select a triple (p, TI_l, TI_u) that minimizes the sum of squared errors. Finally, we bootstrap the entire procedure to calculate standard errors.

for the year 2014 and is statistically different from zero. This point estimate indicates that firms originally within the bunching interval *reduce* their reported taxable income by an average of 4.58 thousand RMB, which is approximately 4.6% of the total taxable income, in response to the tax notch. Note that the magnitude of these bunching responses is constrained by the fact that many firms within the bunching region that should logically respond to the notch but do not, owing to optimization frictions, such as neglect of the tax cuts program or adjustment costs. The fraction of firms situated in the exclusive region that fail to response to the program in 2014 is $\alpha^* = 87\%$.[5] Moreover, Panels B to E further show that firms in the exclusion region reduce their reported taxable income by 4.5%, 7.0%, 7.9%, and 4.3% for each year between 2015 and 2018, respectively, and the estimates of the proportion of non-responders fluctuate between 58% and 82% during the same period.[6]

3.2 Fact 2: Production Expansion

Based on the tax incentives of the Halve-Levy program, we focus on firms with annual taxable income below 1 million RMB and divide them into five groups: Group 1 (0 to 100 thousand), Group 2 (100 thousand to 200 thousand), Group 3 (200 thousand to 300 thousand), Group 4 (300 thousand to 500 thousand), and Group 5 (500 thousand to 1 million).[7] We then plot the share of taxable income for each group relative to the total taxable income in Panel A of Figure 3.

In 2014, only Group 1 benefited from the lower tax rates of the Halve-Levy Program. As other groups were gradually covered by the program, Group 1's share of total taxable income declined. Group 2 experienced a significant increase in its taxable income share in 2015 due to the program's coverage, but this share decreased in the following years as Groups 3 to 5 were successively included in the Halve-Levy Program. Similar patterns were observed for Groups 3, 4, and 5. These observations suggest that the reduced tax rates under the Halve-Levy Program encourage eligible firms to expand production.

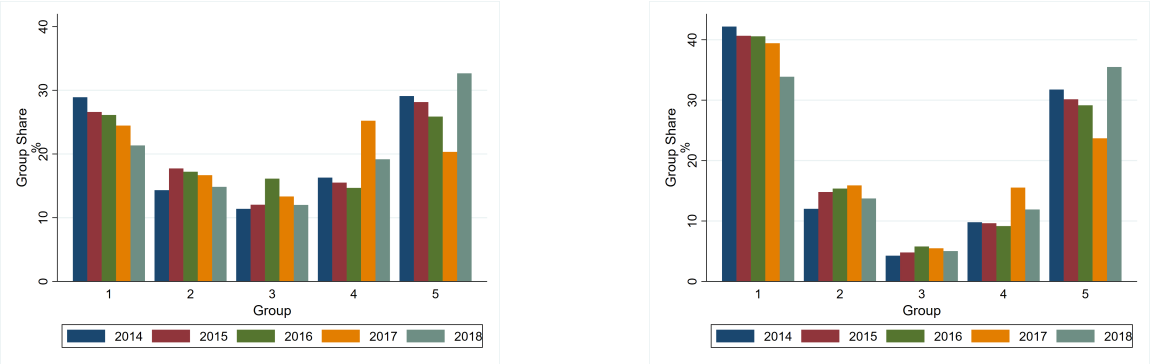
We note that the pattern shown in Panel A of Figure 3 might be influenced by firms' bunching behavior: the increase in each group's share might not result from production expansion within the group but rather from firms in other groups lowering their taxable income to qualify for a lower tax rate, thereby increasing the share of the newly covered group. Hence, we drop firms from the bunching excluded regions (from Section 3.1) and plot

[5] Assume that the proportion of firms that should have incentives to bunch but do not react because of prohibitive adjustment costs is constant in the bunch range. Then we follow Kleven and Waseem (2013) to define the estimator $\alpha^* \equiv \frac{\int_{TI^*}^{TI_d} f_1(TI)d(TI)}{\int_{TI^*}^{TI_d} \widehat{f_0}(TI)d(TI)}$, where TI_d is the indifference point of firms under the two tax systems—the tax liability at TI_d is the same as that at the threshold TI^*. In a frictionless world, the range (TI^*, TI_d) will be dominated by the notch, and there would be zero mass in this range. Thus, we refer (TI^*, TI_d) as the strictly dominated region.

[6] Appendix A.1.2 shows that these bunching estimates are robust to a battery of specification tests.

[7] Hence, a firm might be classified in Group 1 today and in Group 2 next year.

Figure 3: Production Expansion of Newly Eligible Firms



(a) Panel A: Bunching regions included

(b) Panel B: Bunching regions excluded

Note: Panel A plots the share of taxable income for each group among firms with taxable income below 1 million RMB without any adjustments. Panel B shows the share for each group after excluding firms from the excluded regions.

Source: Shanxi Administrative Tax Return Database.

the annual share for each group in Panel B.[8] We find the share changes in Panel B of Figure 3 are largely consistent with those in Panel A.

In Appendix A.2, we consider alternative excluded regions to mitigate bias from firms' bunching behavior. The trends for each group remain consistent with those in Figure 3. Additionally, in Appendix A.3, we use the DID method to robustly examine the effects of tax cuts on the growth of eligible firms. The regression results in Table A1 support the empirical findings that the lower tax rate of the Halve-Levy program leads to production expansion for eligible firms.

3.3 Fact 3: Heterogeneity of Production Expansion

In Fact 2, we demonstrate that a lower corporate tax rate leads to higher taxable income. Here, we want to show that this positive impact of tax reductions is more significant for firms with lower taxable income. Since the Halve-levy threshold increased from 100 thousand to 200 thousand yuan in 2015, we begin with how firms with taxable income between 100 and 200 thousand yuan in 2014 respond to this policy change in 2015.

We recognize that some firms may have strategically reduced their taxable income to slightly below 200 thousand yuan in 2015 to minimize their total tax payment, while others may have attempted but failed to reduce their taxable income below 100 thousand yuan in

[8] In 2014, the notch point was 100 thousand RMB, with the excluded region from 96 thousand to 126 thousand. In 2015, the notch point was 200 thousand, with the excluded region from 184 thousand to 250 thousand. In 2016, the notch point was 300 thousand, with the excluded region from 276 thousand to 384 thousand. In 2017, the notch point was 500 thousand, with the excluded region from 465 thousand to 595 thousand. In 2018, the notch point was 1 million, with the excluded region from 950 thousand to 1.15 million.

Table 3: Firms by Categories of Exposure to Policy Shocks

	2014	2015	2016	2017
Panel A: one-year growth rate				
Low-productive firms	35.91%	49.68%	69.75%	48.96%
High-productive firms	35.07%	36.76%	40.53%	26.30%
Num of observations	1466	341	469	900
Panel B: two-year growth rate				
Low-productive firms	60.69%	97.71%	186.47%	-
High-productive firms	54.83%	93.88%	106.92%	-
Num of observations	1224	288	393	-
Threshold in thousand yuan	200	300	500	1,000
Income range in thousand yuan	126-184	250-276	384-465	595-950

Note: Panel A documents the average one-year growth rate of firms in each subgroup, and Panel B shows the average two-year growth rate of selected firms.

Source: Shanxi Administrative Tax Return Database.

2014 to benefit from the lower tax rate. To eliminate bias from such "bunching" behavior, we exclude firms from bunching-excluded regions and focus on those with taxable income between 126 and 184 thousand yuan in 2014. As shown in Table 3, there are 1,466 firms that remained active in both 2014 and 2015 and had taxable income within this range.

We then divide these firms into low-productive and high-productive subgroups based on whether their 2014 taxable income was below or above the median taxable income among these firms. We calculate the growth rate of taxable income for each firm, truncate the largest 1 percent of growth rates, and report the simple average one-year growth rates of firms in each subgroup in Column 2 of Panel A. The data suggests that the average growth rate of low-productive firms is higher than that of high-productive firms.

We conduct a similar calculation for firms with selected taxable incomes, as indicated in the last row of Table 3, and compare the differences in the average growth rates between low-productive and high-productive firms. We find that in response to a tax cut, low-productive firms consistently grow faster than higher-productive firms. This finding aligns with empirical evidence in the public finance literature, which shows that smaller businesses tend to respond more strongly to tax incentives, such as in investment behaviors (e.g., Zwick and Mahon, 2017).

Finally, recognizing that it may take time for firms to adjust their total output, we calculate and report the average two-year growth rate of firms in each subgroup in Panel B of Table 3. For example, Column 2 in Panel B shows that there are 1,224 firms that remained active in both 2014 and 2016 and had taxable income ranging between 126 and 184 thousand yuan in 2014. After dividing these firms into two subgroups based on their

2014 taxable income, we show that the simple average two-year growth rate of low-productive firms is consistently higher than that of high-productive firms. Columns 3-5 exhibit similar data patterns using different samples.

4 Model

In this section, we develop a quantitative framework to model firms' production and entry decisions, with a focus on understanding the influence of changes in tax incentives on SMEs' output. The construction of this model is motivated by empirical findings and provides model propositions and structural parameters that align with observed patterns in firm behaviors.

The fundamental structure of our model is based on a standard Melitz (2003) model. In this framework, each firm produces a differentiated product, and firms exhibit diversity in productivity. Under monopolistic competition, firms make decisions regarding market entry by assessing after-tax profits derived from their business activities, along with unamortizable pre-tax entry opportunity costs independent of their direct business operations.

To address the impact of corporate income tax incentives on changes in a firm's production scale, we introduce the Arkolakis (2010) fixed cost structure for market entry. Within our framework, the opportunity cost incurred in establishing a firm's business is a convex function of the firm's market penetration. This modeling approach ensures the alignment of our theoretical propositions with Fact 2 and Fact 3 of descriptive evidence: (1) reduced corporate income tax rates prompt incentivized firms to expand their production levels; (2) smaller treated firms grow more than larger treated firms. Additionally, we incorporate incentives for firms' bunching behaviors to scale down production to the notch point, achieving a decreased tax rate in our theoretical model, consistent with Fact 1 confirmed from the bunching estimation.

Consequently, our model captures the positive effect of tax incentives on economic output, reducing costs for beneficiary SMEs and promoting their growth, and the negative effect on economic output, prompting specific firms to scale down production to exploit the tax incentives. Equipped with this theoretical foundation, we can analyze the overall impact of existing notch tax policies on economic output and assess the effectiveness of different policies in supporting the development of SMEs.

We proceed by first addressing the demand side of the model and subsequently introduce the model setup pertaining to the supply side.

4.1 Demand

We utilize a standard Constant Elasticity of Substitution preference for the aspect of the demand. In China, each firm produces a single and differentiated product, denoted as j. A representative Chinese consumer optimizes their utility by purchasing individual goods j in quantities q_j according to the utility function:

$$U = [\int_{j\in\Omega} f_j^{\frac{1}{\sigma}} q_j^{\frac{\sigma-1}{\sigma}} dj]^{\frac{\sigma}{\sigma-1}} \tag{1}$$

where $\sigma > 1$ represents the elasticity of substitution between goods, and Ω represents the set of all available products in the Chinese market. The parameter f_j denotes the market penetration ratio of product j in China, serving as a preference shifter within consumers' preferences. Higher market exposure of firm j's products implies a higher likelihood that a representative consumer gains access to its product, resulting in increased utility from purchasing more of good j for the representative consumer. Consumers optimize their consumption q_j to maximize their utilities. We define the aggregate expenditure in China as X_n. Hence, the quantity demanded of good j at price p_j can be expressed as follows:

$$q_j = f_j \frac{p_j^{-\sigma}}{P_n^{1-\sigma}} X_n \tag{2}$$

where we define the price index in China as $P_n = [\int_{j\in\Omega} f_j p_j^{1-\sigma} dj]^{\frac{1}{1-\sigma}}$. The expenditure on good j is then given by

$$r_j = f_j (\frac{p_j}{P_n})^{1-\sigma} X_n \tag{3}$$

4.2 Producer heterogeneity and unit cost

The firm producing goods labeled as j in Shanxi utilizes resources and manufacturing inputs for its production activities. Each firm denoted as j possesses a distinct productivity denoted as φ_j, and the cost of its bundle of resources and manufacturing inputs is denoted as w_s in Shanxi. We assume a constant return to scale production function, and the marginal cost of good j is represented as follows:

$$c_j = \frac{w_s}{\varphi_j} \tag{4}$$

Each firm j's productivity φ_j is drawn from a Pareto distribution [9] as

$$\mu(\varphi_j \geq \varphi) = T_s(\varphi)^{-\theta} \tag{5}$$

where T_s represents the technology level of Shanxi firms, and θ controls the dispersion of firm productivity.

4.3 Opportunity cost

The firm manager initially observes the firm's productivity and subsequently makes a decision on whether to enter the market and compete to generate profits. This decision considers the firm's productivity level, market competition, market demand, and the opportunity cost of establishing a firm.

We consider the opportunity cost of establishing a business as an independent cost not directly related to the firm's primary business activities and not deductible in the calculation of corporate income tax. To account for the fact that larger product market penetration ratios and production scales require a larger management board and more time investment from managers, we define the opportunity cost F_j incurred by firm j for selling and penetrating a fraction f_j of potential consumers in the Chinese market[10]. The specification is given by

$$F_j = F_s \frac{1-(1-f_j)^{1-\frac{1}{\lambda}}}{1-\frac{1}{\lambda}} \tag{6}$$

where F_s represents the opportunity cost shock experienced by all firms in Shanxi, while the parameter λ reflects the increasing cost of reaching a larger fraction of potential buyers. The variable f_j assumes values between 0 and 1, measuring the exposure of firm j's products to the total buyers in China. We assume that each consumer in China encounters each firm's product with an equal probability, and the impact of the penetration degree f_j is already incorporated into the utility function of the representative consumer.

This functional form for the opportunity cost possesses desirable characteristics to ap-

[9] The distribution choice is often employed in influential papers in macro and trade literature, such as those by Gabaix (1999), Luttmer (2007), Helpman, Melitz, and Yeaple (2004), Chaney (2008), Arkolakis (2010), and Eaton, Kortum, and Kramarz (2011), among others. The selection of the Pareto distribution of firm productivity enables us to obtain a solid alignment with the distribution of firms' sales, considerably simplifying our estimation procedure.

[10] A similar form of fixed costs is also present in studies by Arkolakis (2010) and Eaton, Kortum, and Kramarz (2011) to model small exporters' behaviors. Arkolakis (2010) interprets the fixed cost as marketing cost, while Eaton, Kortum, and Kramarz (2011) introduced a firm-specific fixed cost shock into the framework of Arkolakis (2010).

[11] In contrast to conventional trade entry models, our analysis specifically concentrates on a one-sector closed economy. In essence, our fundamental model setup aligns with Hopenhayn's (1992) model, incorporating Arkolakis's (2010) formulation of opportunity cost. The incorporation of Arkolakis's (2010) market penetration cost structure in our opportunity cost framework enables us to model how treated firms respond to tax incentives. Unlike the conventional Hopenhayn (1992) or Melitz (2003) fixed cost settings, our approach accounts for the nuanced adjustments in intensive margin production by firms when subjected to corporate income tax shocks, contingent upon their market presence.

proximate firms' opportunity costs and sales decisions. First, the fixed cost is 0 if the firm reaches $f = 0$ consumers, and it monotonically increases with the penetration ratio f_j. Second, as shown in the subsequent subsection, the convex function of penetration degree f_j implies a negative relationship between the corporate income tax rate faced by firm j and the production scale it decides upon. This property is crucial for our model to align with the evidence that reductions in corporate income tax result in production expansion for firms affected by this policy. Furthermore, this function form effectively captures the behaviors of small firms: small businesses correspond to low opportunity costs. As we will demonstrate in the model propositions, small firms exhibit stronger reactions to production expansion following the impact of corporate tax cuts than large firms, aligning with the patterns illustrated in Fact 3.

4.4 Entry and competition

Conditional on entry, firm j characterized by a unit cost of c_j, pricing at p_j, and penetrating a fraction f_j of consumers generates a net profit under its applicable corporate income tax rate τ_j as follows:

$$\pi_j(p_j, f_j) = (1 - \tau_j) f_j (p_j - c_j) \frac{p_j^{-\sigma}}{P_n^{1-\sigma}} X_n - F_s \frac{1 - (1 - f_j)^{1 - \frac{1}{\lambda}}}{1 - \frac{1}{\lambda}} \tag{7}$$

For optimization, firm j prices its product at the constant markup over unit cost, following the framework of Dixit and Stiglitz's (1977):

$$p_j = \frac{\sigma}{\sigma - 1} c_j. \tag{8}$$

To simplify the notations, we define $\overline{m} = \frac{\sigma}{\sigma - 1}$. By taking the F.O.C. of Equation (7) regarding the penetration ratio f_j, incorporating Equation (8), we obtain the optimal market penetration ratio for firm j:

$$f_j = max\{1 - [(1 - \tau_j) \frac{X_n}{\sigma F_s} (\frac{\overline{m} c_j}{P_n})^{1-\sigma}]^{-\lambda}, 0\} \tag{9}$$

The maximum unit cost $\overline{c}$ incurred by a firm to earn a positive profit and exist on the market, making $f > 0$, is given by

$$\overline{c} = [(1 - \bar{\tau}) \frac{X_n}{\sigma F_s}]^{\frac{1}{\sigma - 1}} \frac{P_n}{\overline{m}} \tag{10}$$

where $\bar{\tau}$ is the lowest level of corporate income tax rates for the smallest firms. Correspondingly, the productivity threshold for the least productive firm to survive in the market

is

$$\bar{\varphi} = w_s[(1-\bar{\tau})\frac{X_n}{\sigma F_s}]^{\frac{1}{1-\sigma}}\frac{\overline{m}}{P_n} \tag{11}$$

Firm j enters the market if and only if its productivity satisfies $\varphi_j \geq \bar{\varphi}$. With goods markets clearing, by substituting Equations (8), (9), and (11) into (3), we arrive at the functional representation of the taxable income of firm j as follows:

$$\pi_j^{tax} = \frac{r_j}{\sigma} = \frac{X_n}{\sigma}(\frac{\overline{m}c_j}{P_n})^{1-\sigma}\{1-[\frac{1-\bar{\tau}}{1-\tau_j}(\frac{c_j}{\bar{c}})^{\sigma-1}]^{\lambda}\} = \frac{F_s}{1-\bar{\tau}}(\frac{\varphi_j}{\bar{\varphi}})^{\sigma-1}\{1-[\frac{1-\bar{\tau}}{1-\tau_j}(\frac{\varphi_j}{\overline{\varphi}})^{1-\sigma}]^{\lambda}\} \tag{12}$$

Taking the logarithm of both sides of the equation and incorporating time subscript t, we have

$$log(\pi_{jt}^{tax}) = log(F_{st}) - log(1-\bar{\tau}_t) + (\sigma-1)(log(\varphi_{jt}) - log(\bar{\varphi}_t)) + log\{1-[\frac{1-\bar{\tau}_t}{1-\tau_{jt}}(\frac{\varphi_{jt}}{\overline{\varphi}_t})^{1-\sigma}]^{\lambda}\} \tag{13}$$

Equations (10) and (11) govern entry, while Equations (9) and (12) govern taxable income conditional on entry. These equations link our theory on how exogenous economic shocks, including tax cuts, affect firms' production to the data.

These equations are crucial for our counterfactual analysis. They enable us to draw inferences consistent with our empirical findings. By merging Equations (9) and (3) and subsequently taking the first-order derivative of the corporate income tax rate τ_j, we obtain the following result.

Proposition 1: If $\lambda > 0$, firms covered by a lower corporate income tax rate expand their production.

Proof:

$$\frac{\partial r_j}{\partial \tau_j} = \frac{\partial f_j}{\partial \tau_j} X_n P_n^{\sigma-1}(\bar{m}c_j)^{1-\sigma} = -\lambda(1-\bar{\tau})^{-\lambda-1}(\sigma F_s)^{\lambda}[X_n P_n^{\sigma-1}(\bar{m}c_j)^{1-\sigma}]^{1-\lambda} < 0.$$

In our model, a firm's taxable income is determined by a constant return to scale of the firm's market penetration. Simultaneously, the opportunity cost of establishment faced by the firm follows a convex function of the market penetration ratio. In this framework, the optimal market penetration ratio chosen by the firm depends on the marginal taxable income gain and the marginal opportunity cost from an additional unit of market penetration. When the corporate income tax is lower, the marginal benefit of penetrating the market is higher; hence, firms increase their market penetration and the final production output. This model proposition aligns with our empirical evidence discussed in Section 3.2. This proposition within our model ensures that it provides a mechanism for tax preferential policies to facilitate the expansion of covered firms' production scale. This aspect becomes crucial when testing the economic effects of existing policies. Furthermore, the following proposition

demonstrates how our model is consistent with Fact 3 as we show in Section 3.3.

Proposition 2: If $\lambda > 0$, smaller firms will expand their production size more substantially as a result of being exposed to lower tax rates than larger firms.

Proof:

$$\frac{\partial log(r_j)}{\partial \tau_j \partial \varphi_j} = \frac{1}{f_j}\frac{\partial f_j}{\partial \tau_j \partial \varphi_j} = \frac{1}{f_j}\lambda^2(\sigma-1)(1-\tau_j)^{-\lambda-1}\varphi_j^{-(\sigma-1)\lambda-1}[\frac{X_n}{\sigma F_s}(\frac{\bar{m}w_s}{P_n})^{1-\sigma}]^{-\lambda} > 0.$$

In our model, a firm's marginal benefits derived from its market penetration are also related to the firm's productivity and manufacturing costs. Larger and more productive firms demonstrate greater market expansion size in our model. In a setting where a firm's opportunity cost follows a convex function relative to the market penetration ratio, larger firms experience a steeper increase in the opportunity cost associated with further expansion than smaller firms. Thus, lower productivity and smaller firms expand more when being exposed to a lower tax rate. As a result, Proposition 2 predicts that the increasing size requirement threshold corresponds to the fact that smaller treated firms grow more than the larger treated firms as we show in Section 3.3.

The aforementioned model propositions explore the impact of changes in the corporate income tax rate on a firm's production scale regarding the intensive margin. The following proposition analyzes the impact of lower corporate income tax rates on economic output regarding the extensive margin.

Proposition 3: A lower corporate income tax rate leads to more micro firms surviving in the market.

Proof:

$$\frac{\partial \bar{\varphi}}{\partial \bar{\tau}} = \frac{1}{\sigma-1}w_s\frac{\bar{m}}{P_n}(\frac{X_n}{\sigma F_s})^{\frac{1}{1-\sigma}}(1-\bar{\tau})^{\frac{\sigma}{1-\sigma}} > 0.^{12}$$

A lower tax rate provides micro and small firms with a larger marginal gain in production and lowers the productivity threshold requirement for firms to survive in the market. This proposition aligns with the corollary of the extensive effect of taxes, tariffs and other economic shocks in the margin on economic output and trade volumes in many macro, public finance and trade models.

Our model setup for firm output is derived from existing literature, and the model's predictions of how corporate income tax rate policy affects firm production size are consistent

[12]In traditional trade entry models, lower trade costs resulting from trade liberalization lead to more entrants and market competition, and thus increase the productivity threshold for survival in the market. In our model, lower corporate taxes would also encourage firms enjoying lower tax rates to penetrate a larger share of the market, thus increasing market competition. However, it is crucial to note that the economic output of SMEs covered by the Halve-Levy program accounts for less than 3% of total output during our sample year. Therefore, it is unlikely that the increase in market competition triggered by the expansion of production by these firms poses a greater challenge to the survival of microenterprises than the incentive for expansion that a lower tax rate directly creates for microenterprises. Thus, unlike the conclusions drawn from the classic Melitz model, our model underscores that reduced corporate taxes lead to the survival of a greater number of firms.

with empirical findings from real data. To complete our analysis of tax incentives' impact on firms' production, we model firms' bunching responses to notch-based tax policies in the next subsection.

4.5 Notch and bunching behaviors

The corporate income tax in China's Halve-Levy program has the following structure:

$$\tau = \begin{cases} \tau_0, & \text{if } \pi_j^{tax} > \alpha \\ \bar{\tau}, & \text{if } \pi_j^{tax} \leq \alpha \end{cases}$$

Here, α represents the policy threshold, τ_0 denotes the regular corporate income tax rate (25%) applied to all firms before the Halve-Levy program started, and $\bar{\tau}$ denotes the preferential tax rate for the small treated firms (10%) in the Halve-Levy program.

A firm decides whether to engage in bunching behavior by comparing its net profit from bunching, achieved by setting f^α to make $\pi^{tax} = \alpha$, to the net profit of the firm at its optimal production scale above the notch—f_j from Equation (9). Specifically, firm j decides to reduce its production to α if $\pi_j^{tax}(\tau_j = \bar{\tau}) > \alpha$ and $(1-\tau_0)\pi_j^{tax}(\tau_j = \tau_0) - F(f_j(\pi_j^{tax}(\tau_j = \tau_0))) < (1-\bar{\tau})\alpha - F(f_j(\pi_j^{tax} = \alpha))$.

Following our aforementioned theoretical framework, there is a point at which all firms with taxable income under the original tax structure before the incentive is issued decide to reduce production. However, in reality, many firms in this interval did not scale back. In our model, we interpret these firms' failure to bunch as a result of missing information about favorable tax rates or the inability to adjust production or financial management promptly. We treat these missed bunches as exogenous and stochastic, and we assume that the probability of exogenously missing a bunch is δ_t^m for each year t for a firm with a potential bunch.

Additionally, not all firms that can enjoy a lower tax rate by reducing production can precisely reduce their taxable income to the notch point α. We assume that if a firm attempts to reduce production to the notch point, the scale of production it eventually reduces to, π^α, follows an exogenous probability distribution, $G^b(\alpha)$.

The introduction of firms' bunching behavior in our model is based on the bunching estimates in Section 3.1. This enables our theory to capture the distortions and inefficiencies brought to the economy by the existing notch-based tax cuts. By incorporating these elements, our model provides a comprehensive analysis of the impact of tax incentives on firms' production, accounting for firms' bunching responses to the policy.

4.6 Misreporting

Thus far, our discussion of the Halve-Levy program's impact has focused on firms' reported behaviors, particularly the amount of taxable income they report. Economists studying the economic effects of taxation place great interests on distinguishing between the actual decisions made by firms and the decisions they report to tax authorities. In this section, we theoretically and empirically explain why the Halve-Levy program is unlikely to induce additional misreporting behaviors among covered firms.

4.6.1 Theory

Following Chen et al. (2021) and the penalty mechanism of Chinese tax authorities to deter tax evasion, we assume that the potential profit a firm gains from tax evasion is less than the expected penalty cost if caught by the government. More specifically, if a firm underreports income to reduce taxes by $evasion_i$, the expected penalty cost of this underreporting is $h(evasion_i)$, where $h(evasion_i)$ is a function related to tax penalties and $h(evasion_i) > evasion_i$. Additionally, we assume that firms are rational and have sufficient time to determine their reported and actual income in response to taxation policy changes.[13]

In this setup, if a firm analyzed in our study does not engage in bunching—it does not report its taxable income as α—it naturally lacks an incentive to misrepresent its taxable income because the potential benefits from such actions are outweighed by the expected penalty cost of getting caught. Similarly, if a firm i decides to bunch and report at the notch point α, it has no incentive to misreport. Suppose firm i has already produced up to its reported taxable income of α. In that case, if it chooses to produce at a higher level of π_i, surpassing α, and then misreports its income to α, the expected penalty cost $h(\pi_i - \alpha)$ incurred would exceed the profit $\pi_i - \alpha$ made from the additional production and misreporting. [14]

Even if we further relax the functional form to allow $h(evasion_i) < evasion_i$ for $evasion_i$ within a small interval $(0, \bar{evasion})$ where $\bar{evasion}$ represents the threshold at which the expected penalty equals the evasion value, indicating that firms may engage in a small level of misreporting under the Chinese tax administration system.[15] Importantly, the Halve-Levy program still does not introduce incentives for firms to engage in additional misreporting and

[13] During the study period, tax-cut policies were announced several months prior to implementation. This provided firms with ample time to adjust their production and tax reporting in line with rational strategies.

[14] Our misreporting outcome differs from Chen et al. (2021) due to differing policies. The R&D tax cuts policy in their study required firms to exceed a specific R&D threshold, potentially incentivizing misreporting to avoid substantial R&D costs. In contrast, the SME policies we study require firms to have taxable income up to a specific threshold. Firms have no additional costs to bunch but face stricter penalties for producing more than the threshold and then misreporting; hence, they have little motivation to misreport.

[15] Under these conditions, optimizing firms would equate the marginal cost of sheltering \$1 of income from taxation with the net marginal cost of reducing real production by \$1, as suggested by Chetty (2009).

even empowers covered firms to better regulate their evasion behaviors. Specifically, once firms earn taxable income up to the policy threshold α, the cost of continuing production and then misreporting their income includes two sources of penalty costs: the expected penalties that every misreporting firm faces, and the additional penalty cost of being removed from the SME preferential tax rate. The additional penalty cost would make firms more cautious in regulating their misreporting behavior.

Even under these relaxed assumptions, while all firms might engage in some degree of misreporting, as long as the Halve-Levy program does not provide additional incentives for misreporting, this misreport behavior will not interfere with the results of our counterfactual analysis. Thus, our assessment of the economic impact of the Halve-Levy program and our comparison of alternative SME support policies in the following sections remain valid.

4.6.2 Empirical Evidence

To consolidate our theoretical discussion, we further conduct an empirical check of firms' misreporting behaviors by analyzing the number of employees of firms near the notch point.

Figure 4 illustrates the ratio of taxable income to the number of employees for different taxable income size groups. The figure groups firms into bins of taxable income and plots the mean of this ratio for each bin. Additionally, we provide an estimated cubic regression on the ratio of taxable income with heterogeneous coefficients above and below the notches.

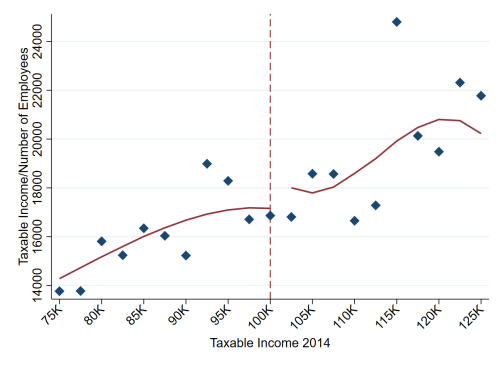
If a substantial number of firms underreport their taxable income by manipulating their tax returns, we would expect to observe a clear and discontinuous jump upward at the notch. This is because while bunchers may underreport their taxable income, they would not have incentives to underreport the number of employees. Consequently, bunchers (immediately below the notch) would demonstrate a lower ratio of taxable income to the number of employees than non-bunchers (immediately above the notch).

However, Figure 4 indicates that no evident discontinuous jump upward at the notch is observed in most years. This suggests that most bunchers may not manipulate their tax returns but opt to reduce their real output. This evidence indicates that the focus of manipulation, if any, is likely on reducing production rather than underreporting taxable income.

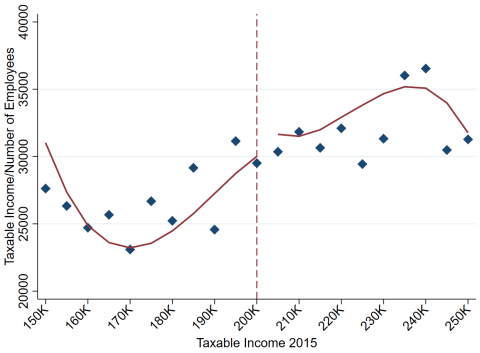
In Appendix A.4, we use employment numbers to replace the ratio of taxable income to the number of employees and re-plot Figure 4. The results provide additional empirical evidence that the SME policies we study do not incentivize firms to misreport their income.

Overall, in this section, we theoretically and empirically demonstrate that firms' bunching behaviors induced by the Halve-Levy program are more inclined to reduce production rather than engage in misreporting.

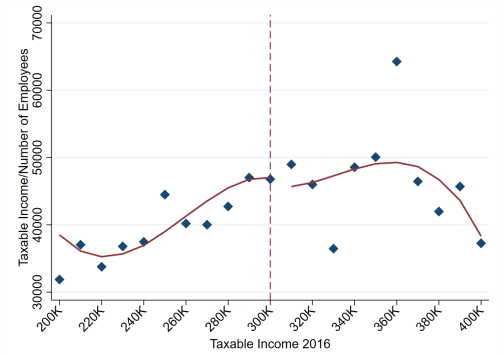
Figure 4: Alternative Measure of Firm Production at the Notch



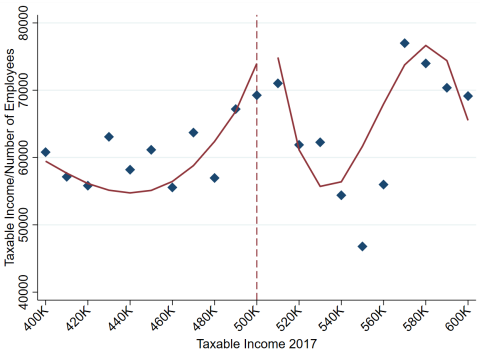
(a) Panel A. 2014



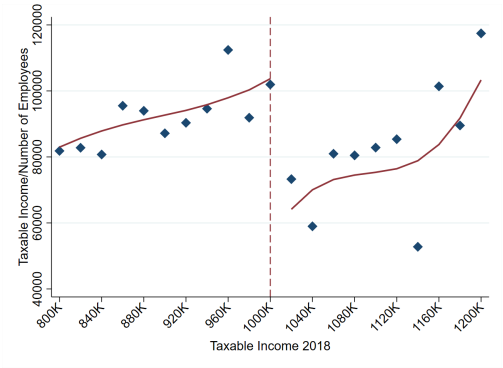
(b) Panel B. 2015



(c) Panel C. 2016



(d) Panel D. 2017



(e) Panel E. 2018

Note: This figure plots the ratio of taxable income to the number of employees at each size category, along with an estimated cubic regression of the ratio of taxable income with heterogeneous coefficients above and below the notches.

Source: Shanxi Administrative Tax Return Database.

5 Estimation

In the preceding section, we construct a model to capture the empirical pattern of how firms respond to tax incentives. In this section, we present our estimation method for recovering the parameters of the structural model, which will enable us to perform our counterfactual analysis.

5.1 Estimation Procedure

The set of parameters to be calibrated and estimated includes the following:

$$\{\sigma, \theta, \lambda, F_{st}, \delta_t^m, G_t^b(\alpha_t)\}.$$

We begin by calibrating the elasticity of substitution σ. Following previous literature (Melitz and Redding, 2015; Chen, Chen, Liu, Serrato, and Xu, 2021), we set σ to 4.

Next, we employ the non-linear least square method to estimate θ, λ, and F_{st}. Our estimation process follows the algorithm outlined below:

1. Given the value of θ, we simulate N_t firm heterogeneity draws $\frac{\hat{\varphi_{nt}}}{\bar{\varphi}_t}$, where N_t denotes the number of firm observations for each year t in the data. Under the assumption that firm productivity follows a Pareto distribution of dispersion θ, conditional on a firm j surviving on the market and being observed by the economist, the firm's productivity relative to the threshold productivity $\frac{\varphi_{jt}}{\bar{\varphi}_t}$ follows a Pareto distribution of scale 1 and dispersion θ. [16]

2. With the heterogenous firm draws, we then simulate the taxable income of the N_t artificial firms using Equation (13), along with the values of θ, λ, and F_{st}.

3. Finally, we rank the artificial and actual taxable incomes from smallest to largest for each year and estimate the parameters θ, λ, and F_{st} by minimizing the objective function:

$$\min \sum_t \sum_r \| log(\pi_{rt}^{tax,\ arti}) - log(\pi_{rt}^{tax,\ real}) \|^2$$

   where $\pi_{rt}^{tax,\ arti}$ represents the rth value of the taxable income of the artificial firms, and $\pi_{rt}^{tax,\ real}$ represents the rth value of the taxable income in the actual data. Additionally, to mitigate estimation bias arising from firms' bunching behaviors, we exclude a significant range of observations close to the notch and their corresponding artificial firms.

[16] For a detailed proof, see Appendix A.5.

Our estimation strategy is similar to the method of simulated moments employed in recent structural studies of Chinese policies (e.g., Chen et al., 2021; Chen, Chen, Liu, Serrato, and Xu, 2021). However, our estimation is akin to an SMM estimation that maps each firm's sales to its counterpart as a moment to fit the overall distribution shape rather than targeting several sets of moments to capture the distribution shape.

Subsequently, we identify the probability parameter δ_{mt} and the probability distribution $G_t^b(\alpha_t)$ associated with firms' bunching behavior. Using the demand and production-related parameters θ, λ, and F_{st} estimated in the previous steps, we can recover the distribution of firm production under the linear tax. Thus, we can determine the interval $[\alpha, bunchbound_t]$, where $bunchbound_t$ is the upper bound of the taxable income of firms that have an incentive to reduce production for higher returns. We then introduce into the model the number of firms falling within this interval. By comparing the model simulation with data observations, we can infer the exogenous probability δ_{mt} that a firm misses the opportunity to reduce production to the notch.

Similarly, we divide the interval $[0.9\alpha_t, \alpha_t]$ into ten small segments. Based on the distribution of corporate taxable income under linear tax rates that we recover, we examine the extent to which more of the data are observed within these ten small bands compared to what is simulated by the model. Employing this non-parametric approach, we estimate the distribution of the final reduction outcome $G_t^b(\alpha_t)$ for the firm.

5.2 Estimation Results

Table 4 presents the calibration and estimation results for our structural parameters related to production and demand. Tables 5–7 display the estimation of parameters related to bunching behavior.

We first examine the parameter λ in the model, which is particularly relevant to the expansion of SMEs. Previous studies, such as those by Arkolakis (2010) and Eaton, Kortum, and Kramarz (2011), estimated coefficients around 1.093 and 0.91, respectively, using data for French exporters and fixed cost functions for market penetration. In our study, using data from manufacturing firms located in Shanxi Province, China, we estimate λ to be 0.623, which is smaller than the coefficients reported in the previous literature. This suggests a significant increase in fixed costs for market expansion, indicating that Chinese SMEs may encounter more resistance when expanding.

We also analyze the estimated dispersion parameter θ of the firm productivity distribution. Our estimated θ is smaller compared to the estimates in the studies by Arkolakis (2010) and Eaton, Kortum, and Kramarz (2011) and even smaller than $\sigma - 1$. One possible explanation is that the firms in Shanxi Province exhibit substantial variation differences. Shanxi Province is a relatively backward province in China. However, the province is known for its

unique natural resources, which have contributed to the emergence of many affluent firms and billionaires in China. The dataset used in our analysis encompasses all tax-administrative firms in Shanxi Province. If we were to incorporate data from all firms in China, we would derive a higher estimate for the parameter θ.

Table 4: Estimated Parameters - Demand and Supply

Parameters	Calibrated / Estimated Values	Target
σ	4	Melitz and Redding (2015) Chen, Chen, Liu, Serrato, and Xu (2021)
θ	1.526	dispersion shape of taxable income
λ	0.623	dispersion shape of taxable income
$F_s(2014)$	3031.2	production level in 2014
$F_s(2015)$	2990.77	production level in 2015
$F_s(2016)$	3979.37	production level in 2016
$F_s(2017)$	5186.81	production level in 2017
$F_s(2018)$	6081.16	production level in 2018
Observation	453,651	
R^2	0.999	

Note: The F_{st} is measured in units of Chinese RMB. We delete the top and last 0.5% of the sample in the estimation to alleviate the bias of the extreme observations.

Next, we focus on the estimated F_{st}. Our results show that the market penetration cost base increases annually, except for 2015. This trend is consistent with the growth rate of Shanxi's regional GDP, which experienced a small decrease in 2015 and then increased annually, especially in 2016–2017.[17] By integrating our estimated λ, we can infer the market penetration costs for firms within Shanxi Province under the context of the specific Halve-Levy program policy. For instance, in 2016, a firm penetrating 50% of the national market corresponds to a market penetration cost of 3,427 RMB, while penetrating 99% of the national market incurs a cost of 100,141 RMB. This range of opportunity costs implies that large firms undergo smaller changes in production scale before and after the tax policy change. However, small and medium-sized firms experience a significant tax impact, which aligns with our empirical observations and model propositions.

Table 5: Bound of Potential Bunching Firms' Taxable Income

Year	2014	2015	2016	2017	2018
Upper bound	132.84	258.43	386.09	638.58	1259.13
Lower bound (threshold, α_t)	100	200	300	500	1,000

Note: The values above are measured in units of thousands of Chinese RMB.

[17] According to statistics from the Shanxi Bureau of Statistics, the GDP growth rate of Shanxi Province was 4.9% in 2014, 3.0% in 2015, 4.1% in 2016, 6.8% in 2017, and 6.7% in 2018.

With the coefficients estimated above, in conjunction with the specific Halve-Levy program, we can deduce the taxable income interval $[\alpha_t, \text{bunchbound}_t]$ applicable to firms incentivized to profit by cutting production. Table 5 showcases our estimated values for bunchbound_t for each year. Our finding reveals that the bunch bound for each year surpasses the intuitive bound $1.2\alpha_t$.[18] Under the current tax system, the accounting tax benefits of producing taxable income at both α_t and $1.2\alpha_t$ are the same.[19]. Nevertheless, achieving production at $1.2\alpha_t$ requires greater management efforts, management time, and additional opportunity costs that cannot be accounted for in the accounting profit, resulting in a net benefit lower than that of reducing production to the bunch point α_t. Consequently, we establish an upper bound on the taxable income of firms with an incentive to reduce production, which also exceeds $1.2\alpha_t$ per year.

Table 6: Bunching Parameters - Miss Bunching Probability

	Miss Bunching Probability δ_t^m (%)				
Year	2014	2015	2016	2017	2018
Rate	86.68	84.92	73.71	59.99	67.17

Table 6 presents our estimates of the probability δ_t^m that a firm misses the opportunity to profit from a production reduction each year. Our structural parameter estimates indicate that in the first year of the Halve-Levy program inception, up to 86.68% of firms miss the opportunity to achieve greater gains by cutting production. Subsequently, the probability of firms missing bunching decreases each year from 2014 to 2017. Possible reasons for this include increasing opportunities to learn about tax benefits and the fact that firms are more likely to respond to policies as they accumulate experience. Notably, our estimates reveal that over the five years of policy implementation observed in our data, more than 50% of the firms motivated to reduce production did not ultimately profit from the reduction.

Compare the miss bunching probability presented in Table 6 with the proportion of firms that fail to respond to the program outlined in Section 3.1 under bunching estimates. Regarding the estimation of the proportion of firms overlooking clustering opportunities in 2014, the two methods yield remarkably similar results (86.68% vs. 87%). Furthermore, the structural model parameters indicate a probability range for firms missing clustering behaviors from 2015 to 2018 (59.99-84.92%), which closely aligns with the range estimated by the bunching method (58-82%). The proximity between these two sets of estimates underscores the robustness and reliability of our structural parameter estimates from an alternative perspective.

[18] Intuitively, $(1-\tau_0)\pi > (1-\bar{\tau})\alpha_t \Rightarrow (1-25\%)\pi > (1-10\%)\alpha_t \Rightarrow \pi > 1.2\alpha_t$. Hence, $[\alpha_t, 1.2\alpha_t]$ represents the intuitive bunch bound.

[19] $1.2\alpha_t * (1-25\%) = \alpha_t * (1-10\%)$

Table 7: Bunching Parameters - Reduction Outcome Distribution

	Over Bunching Hazard h^o (%)									
	Intervals from the bunching range edge to the notch									
Year	1	2	3	4	5	6	7	8	9	10
2014	-4.38	1.68	3.48	0.58	3.68	8.38	6.96	5.05	14.26	60.32
2015	0.54	0.60	5.54	7.63	12.57	-2.04	5.74	5.80	19.60	44.02
2016	3.07	4.49	-1.07	9.51	4.24	7.04	10.93	9.83	14.85	37.12
2017	1.49	1.70	5.47	9.01	7.88	10.11	13.21	9.63	14.96	26.54
2018	7.70	8.08	7.99	3.01	5.56	11.97	10.67	9.92	7.28	27.84

Table 7 illustrates the final production reduction results for firms that decide to reduce production to the notch point. The table displays how firms that cut production from the right end of the notch point to the left end of the notch point each year are distributed in ten equal intervals from $0.9\alpha_t$ to α_t. The probability distribution of these small intervals sums to 100% for each row (per year). As our estimation results demonstrate, most firms tend to bunch into the 8th, 9th, and 10th intervals. Our non-parametric estimation results also show that although firms do not precisely bunch to the notch point, they have a higher probability of bunching to intervals closer to the notch point.

Comparing the miss bunching rate, bunching upper bound, and over-bunching probabilistic distribution through our structural estimation with the bunching excluded region and bunching friction parameters from Section 3.1, reveals strikingly similar results. This comparison further solidifies the reliability of our structural model in capturing a firm's decisions related to bunching behavior.

The structural parameter estimation results presented in this section are crucial for our counterfactual analyses. The parameter estimates for demand and supply will be utilized to predict how tax incentives affect the production size of firms. Moreover, the estimates of parameters related to bunching behavior will be used to predict the extent to which tax policies influence firms' production reductions and distortions. This will contribute to evaluating the expectation of firms' eventual production reduction outcomes.

5.3 Model Fit

In this section, we assess the adequacy of the model's fit on firms' output.

Figure 5 examines whether the model aligns with the data on firms' output (taxable income), considering both tax incentives for firm growth and bunching behavior. We begin by ranking the taxable income of each firm in each year from the real data and comparing it with the sales of each artificial firm in the model's prediction. Subsequently, we plot the log sales for each firm in the actual data against the model-predicted sales of the corresponding firm at the same rank for each year.

Figure 5: Model Fit

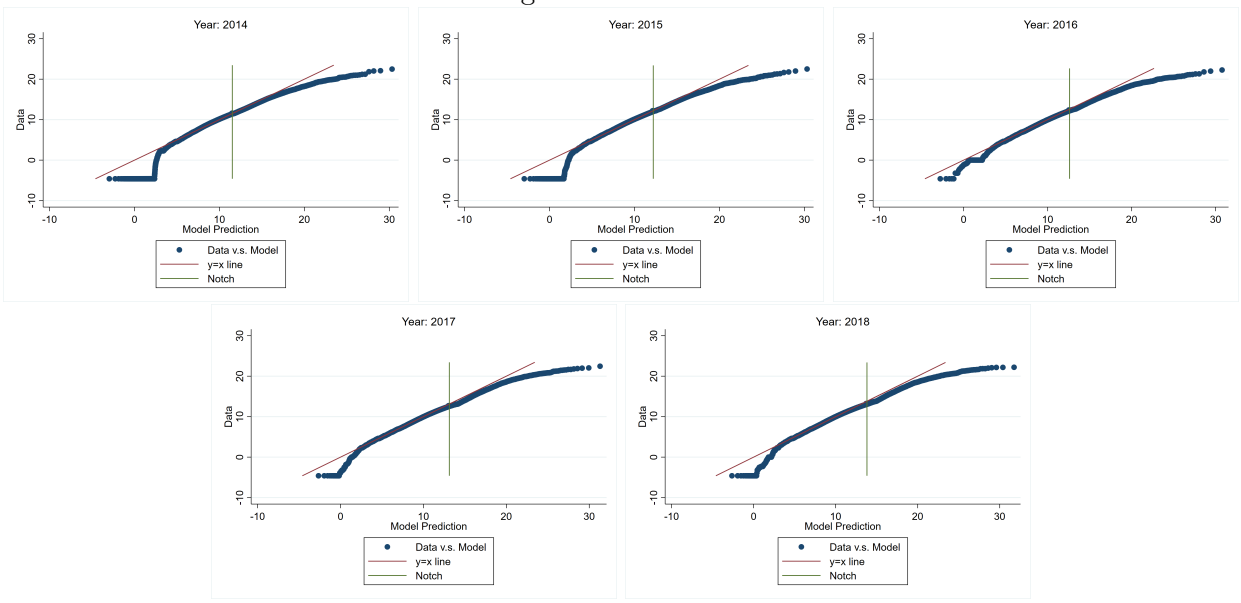


Figure 5 reveals a close fit, although discrepancies between the data and the model's simulation are observed for the largest and smallest firms. The data show that the smallest firms exhibit consistently low values (taxable income equal to 0.01 RMB), whereas our model assumes a productivity level distribution with a smoothly rising slope. However, as shown in the figure, the differences between the model and the data are mainly concentrated in firms with taxable income below $e^2 = 7.38$ RMB [20]. Since the amounts involved are minimal, the overall output and tax revenue from these firms are negligible compared to those of larger firms. Therefore, the discrepancies between the model's predictions and the data for these smallest firms will not significantly impact our subsequent policy analysis.

For the largest firms, our model's predictions are slightly higher than the actual data, primarily due to excessive variation in Shanxi firms and our estimation of a small productivity dispersion parameter. In our counterfactual analysis, we excluded the largest 0.5% of firms from the sample to mitigate the bias from overestimating their output and taxable income.

Since the maximum taxable income threshold for the 2014-2018 Halve-Levy program policy is 1 million RMB, even when extending the policy impact range to firms with taxable income from 10 RMB to 2 million RMB, Figure 5 shows that our model's fit is very close to the data for the relevant firms[21].

In Appendix A.6, we present additional model fit strategies as robustness checks for our model's fit. Figure A5 and Figure A6 further demonstrate our model's close fit to the output of the vast majority of firms affected by the Halve-Levy program.

[20] Firms with a logarithm of taxable income less than 2.
[21] $log(10) = 2.30$ and $log(2 \times 10^6) = 14.51$.

6 Counterfactual Analysis

In this section, we leverage our model and estimation results to conduct three counterfactual analyses. First, we simulate the combined effect of the existing tax preferential policy on total output and tax revenue. Further, as illustrated in Table 8, we categorize potential market entrants into four distinct groups and analyze the contribution of each category to total output and tax revenue. Second, we compare the alternative policies of kink-based tax cuts policies and output-based subsidies to the current notch-based policy. Our comparison focuses on tax revenue while ensuring equivalent economic boost and stimulation for SMEs.

Table 8: Firms by Categories of Exposure to Policy Shocks

Category	Firms
1	Exit the market under 25% tax but survive under the existing policy
2	Exist under 25% tax and enlarge their production under the existing policy
3	Have the motivation to reduce taxable income to the notch
4	Not covered by the lower tax rate and will not reduce production for profit

To facilitate these analyses, we consider the adjustment of policy concerning input prices and labor employment. A counterfactual outcome in our model setup relies on the assumption that either input prices or employment are fixed and inelastic. Notably, the economic output of SMEs covered by the Halve-Levy program constitutes less than 3% of the total provincial output in our sample year. While preferential policies for micro and small enterprises prove highly effective in promoting growth and employment for such businesses, they have limited influence on market commodity prices because of their limited output share. Consequently, to close the model, we assume that the wages and input prices utilized by firms in the model are fixed and determined by the general exogenous economic environment. Moreover, we consider that the economy in Shanxi is not operating at its maximum employment capacity.

Therefore, for our counterfactual analysis, we maintain input prices and wage levels as constants. Additionally, we incentivize firms within Shanxi Province through policies to increase hiring and expand capital, aiming to facilitate economic growth and boost production among SMEs.

6.1 Welfare analysis of the existing policy

We analyze the impact of the existing notch-based corporate income tax incentives on overall economic output. The current policy motivates more firms to enter the market through the extensive margin. Moreover, it expands the production of firms covered by the preferential policy through the intensive margin. However, due to the notch property of

the policy, some firms inefficiently curtail production. In the empirical findings section, we employ DID and bunching estimates to verify the effect of the policy on firms' expansionary production and bunching behavior independently. In this section, we utilize model simulations to quantify the collective effect of the policy on economic output. Furthermore, we determine the contributions of different types of firms, as shown in Table 8, to the overall change in output.

6.1.1 Combined effect

Table 9 illustrates the combined effect of the Halve-Levy program on aggregate economic outcomes for each year. Our focus lies on the effects of the policies on aggregate output ($\Sigma_j r_j$), aggregate log output ($\Sigma_j log(r_j)$), and tax revenues ($\Sigma_j \tau_j \pi_j^{tax}$), respectively. Further, we consider the aggregate log output to provide substantial emphasis on SMEs in the economy, achieving a balanced distribution of economic output among firms. In the table, the 25% Tax column represents economic output and tax revenue under the original 25% linear tax system, while the Policy column displays the simulation of economic outcomes under the actual policy of the Halve-Levy program. The Effect column indicates the percentage change in the economic outcome resulting from the policy compared to the initial tax system.

The simulation results presented in Panel A show that the presence of tax incentives for undersized firms boosts total output by 0.09–0.1% per year.[22] Despite the inefficiencies and distortions around the notch point, the Halve-Levy program results in a positive impact on overall economic output.

Panel B indicates that the program drives an increase of 2.37–5.18% in the output for treated small firms. [23] The magnitude of the increase tends to decrease over the years, suggesting that as time passes and the policy threshold enhances, the size of new firms supported by the policy in later years becomes larger. Consequently, larger firms receive a smaller degree of growth in output boosts from the preferential policy compared to smaller firms.

Our counterfactual simulations demonstrate that existing policies boost total log output by 7.66–7.82% and the total log output of firms covered by preferential policy by 8.91–10.29%. The effect of existing policies on the total log output of the economy is larger relative to the boost in the economy's total output. Moreover, our results here closely align with the regression results in Appendix Section A.3 of our empirical evidence, further confirming the accuracy of our model.

[22]Our analysis excludes the top 0.5% of firms from the sample to prevent bias in the prediction of production for the largest firms because of the relatively small θ we estimated. Subsequent counterfactual analyses adhere to the same approach.

[23]In Table 9, the treated SMEs are specifically characterized as Type 1 and Type 2 firms, which are referenced in Table 8. These particular types of firms represent the Halve-Levy program's intended focus for support.

Table 9: the Combined Impact of Halve-Levy program on Economic Outcomes

Panel A. Overall Economy

Year	Aggregate Output ($\Sigma_j r_j$) Unit: Billions of Yuan			Aggregate Log Output ($\Sigma_j log(r_j)$) Unit: Thousands of Log Yuan			Tax Revenues ($\Sigma_j \tau_j \pi_j^{tax}$) Unit: Billions of Yuan		
	25% Tax	Policy	Effect (%)	25% Tax	Policy	Effect (%)	25% Tax	Policy	Effect (%)
2014	45.64	45.68	0.09	679.79	731.89	7.66	11.41	11.28	-1.11
2015	44.88	44.93	0.10	677.65	729.62	7.67	11.22	11.01	-1.83
2016	65.17	65.23	0.09	761.59	820.46	7.73	16.29	15.97	-2.00
2017	97.56	97.64	0.08	898.00	967.90	7.78	24.39	23.81	-2.37
2018	132.88	133.00	0.09	1059.92	1142.81	7.82	33.22	32.12	-3.30

Panel B. Treated SMEs

Year	Aggregate Output ($\Sigma_j r_j$) Unit: Millions of Yuan			Aggregate Log Output ($\Sigma_j log(r_j)$) Unit: Thousands of Log Yuan			Tax Revenues ($\Sigma_j \tau_j \pi_j^{tax}$) Unit: Millions of Yuan		
	25% Tax	Policy	Effect (%)	25% Tax	Policy	Effect (%)	25% Tax	Policy	Effect (%)
2014	869.11	914.09	5.18	506.68	558.81	10.29	217.28	91.41	-57.93
2015	1392.39	1443.13	3.64	550.08	602.08	9.45	348.10	144.31	-58.54
2016	2185.70	2260.82	3.44	626.98	685.89	9.40	546.43	226.08	-58.63
2017	3829.82	3946.76	3.05	757.16	827.12	9.24	957.45	394.68	-58.78
2018	7268.15	7440.49	2.37	930.82	1013.75	8.91	1817.04	744.05	-59.05

Note: The "25% Tax" column represents economic output and tax revenue under the original 25% linear tax system, whereas the "Policy" column displays the simulated economic outcomes under the real policy of the Halve-Levy program. The "Effect" column indicates the percentage change in the economic outcome resulting from the policy compared to the initial tax system. The treated SMEs are specifically characterized as Type 1 and Type 2 firms, as referenced in Table 8.

Finally, we analyze the impact of the existing policy on tax revenues. While the policy stimulates more firms to remain in the market and expand their output, it lowers overall tax revenues by anywhere from 1.11 to 3.3% due to the reduction in tax rates for covered firms.

6.1.2 Decomposition

Next, we categorize all firms into four distinct types, as displayed in Table 8, and examine the contributions of each category to the overall policy effects, as outlined in Table 10. We present the output and tax-related outcomes for different types of firms under the initial 25% linear tax rate and the existing preferential policies. The "Ratio" column indicates the proportion of the value added from each category of firms to the overall change in economic outcomes. Notably, the total ratio for all four categories equals 100% for each year. As shown in Table 10, the primary increase in economic output each year predominantly stems from the second category of firms, denoting the production expansion of existing firms. Conversely, the impact of new entrants and the reduction in production due to bunching behavior within the third category of firms demonstrate a relatively modest effect on the overall change in total output.

Table 10: Impact of Tax Cuts on the Economy by Firm Types

Type	# of Firms	Aggregate Output Billions of Yuan			Aggregate Log Output Thousands of Log Yuan			Tax Revenue Millions of Yuan		
		25% Tax	Policy	Ratio	25% Tax	Policy	Ratio	25% Tax	Policy	Ratio
Year: 2014										
1	7115	0	0.001	3.52	0.00	35.48	68.10	0.00	0.15	-0.11
2	59846	0.869	0.913	105.28	506.68	523.34	31.96	217.28	91.26	99.28
3	1674	0.193	0.189	-8.80	19.51	19.48	-0.06	48.26	47.20	0.83
4	11705	44.580	44.580	0.00	153.60	153.60	0.00	11145.07	11145.07	0.00
Year: 2015										
1	7103	0	0.001	3.16	0.00	35.32	67.97	0.00	0.14	-0.07
2	63552	1.392	1.442	108.43	550.08	566.75	32.08	348.10	144.17	99.20
3	1114	0.253	0.248	-11.58	13.74	13.71	-0.04	63.33	61.54	0.87
4	8435	43.236	43.236	0.00	113.83	113.83	0.00	10809.09	10809.09	0.00
Year: 2016										
1	7746	0	0.002	3.41	0.00	40.73	69.19	0.00	0.21	-0.06
2	69885	2.186	2.259	119.69	626.98	645.16	30.88	546.43	225.87	98.44
3	1133	0.386	0.372	-23.11	14.43	14.39	-0.07	96.43	91.16	1.62
4	8694	62.594	62.594	0.00	120.18	120.18	0.00	15648.42	15648.42	0.00
Year: 2017										
1	8888	0	0.003	3.83	0.00	49.09	70.23	0.00	0.31	-0.05
2	81458	3.830	3.944	139.91	757.16	778.03	29.86	957.45	394.36	97.38
3	1128	0.638	0.602	-43.74	14.94	14.87	-0.09	159.39	143.92	2.67
4	8880	93.093	93.093	0.00	125.90	125.90	0.00	23273.19	23273.19	0.00
Year: 2018										
1	10324	0	0.004	3.45	0.00	58.66	70.78	0.00	0.42	-0.04
2	97300	7.268	7.436	136.86	930.82	955.09	29.28	1817.04	743.62	97.92
3	963	1.081	1.031	-40.32	13.41	13.37	-0.05	270.21	246.94	2.12
4	7984	124.526	124.526	0.00	115.69	115.69	0.00	31131.55	31131.55	0.00

Note: Column "25% Tax" presents the output and tax outcomes of different types of firms under the initial 25% linear tax rate, while the "Policy" column presents the output and tax outcomes of different types of firms under the existing preferential policies. The "Ratio" column indicates the proportion of the value added from each category of firms to the overall change in economic outcomes, with the total ratio for all four categories summing up to 100% for each year.

Furthermore, the influence of the bunching behavior of the third category of firms on economic output is more pronounced in the later years of the program than the earlier years. This is primarily because the lower corporate income tax rates provide less stimulus for the larger firms covered in the current years than in the earlier ones. As firms' miss-bunching probability decreases, the economic loss from bunching behavior increases. For example, in

2018, although the distorting effect of bunching behavior on economic output is only nearly one-third of the positive outcome from the expansion of Type 2 firms, the inefficiency of bunching behavior on the economy is already significant than it was in 2014.

When considering aggregate log output, the results in Table 10 indicate that Types 1 and 2 firms predominantly drive the impact, suggesting that the survival of new entrants significantly enhances the policy objectives of policymakers, particularly when they are focused on aggregate log output.

This analysis reveals the time-varying effect of the tax rate reduction policy on economic output. Over time, newly covered firms receive diminishing incentive boosts for output, while the negative effects of the notch-based policy increase as firms increasingly exploit policy rules to capitalize on bunching behavior. Our simulation results also underscore that even if the policymaker does not consider the reduction in tax revenue from tax incentives, providing wider coverage of incentives does not necessarily guarantee a substantial increase in the economy's output. Practically, Chinese policies have shown some iteration in selecting the threshold of preferential policies.

6.2 Alternative size-dependent policies

Next, we compare the efficiency of alternative policies, including a linear tax system, kink-based tax cuts, and output-based subsidies, with the benchmark results of notch-based tax cuts in terms of stimulating economic output and collecting tax revenue. Specifically, as listed in Table 11, the first scenario, "Linear," represents the absence of any tax cuts, with all firms facing a 25% corporate tax rate on their taxable income. "Notch" refers to the notch-based tax cuts as implemented in the Halve-Levy program. "Kink" denotes kink-based tax cuts with the same thresholds and tax rate levels as in the Halve-Levy program. "Subsidy" considers a scenario where the opportunity cost of operating a firm's sales can be accurately estimated, and the opportunity cost for the portion of taxable income under the threshold is subsidized by 1/6.[24] We set the subsidy ratio to 1/6 of the opportunity cost because, at this level, the mass of new entrants (and thus the mass of active firms) remains the same as in the scenarios with notch- and kink-based tax cuts.

[24]For example, in 2014, under the notch-based Halve-Levy program, firms with taxable income below 100,000 yuan pay a reduced 10% tax rate, while others pay the standard 25% tax rate. In the Kink scenario, a firm pays a reduced 10% tax rate if its earnings are just below 100,000 yuan and a 25% tax rate on any additional income above 100,000 yuan. In the Subsidy scenario, a firm with taxable income below 100,000 yuan receives a subsidy equivalent to 1/6 of its opportunity cost of operating (denoted as F_s in the model). For firms with taxable income above 100,000 yuan, the subsidy is equivalent to 1/6 of the opportunity cost of operating a firm with exactly 100,000 yuan in taxable income.

Table 11: Size-Dependent Policies for Comparison

Policy	Policy Threshold	Preferential Policy
Linear	N.A.	N.A., 25% for all the taxable income
Notch	As real policy	10% for all the taxable income if it is smaller than the threshold
Kink	As real policy	10% for the part of taxable income under the threshold
Subsidy	As real policy	subsidy of 1/6 opportunity cost for taxable income under the threshold

6.2.1 Tax cuts and subsidies with the same threshold

Table 12 presents the impact of different size-dependent policies on total economic output, total log output, and tax revenue. These policies include versions of notch-based tax cuts, kink-based tax cuts, and output-based subsidies, all set at the same policy thresholds and with the same mass of active firms.

Panels A and B illustrate the effects of different policies on the overall economic output as well as the economic output of smaller businesses that benefit from preferential policies.[25] Panel C provides further insights into the effects of different policies on tax revenues. For our counterfactual analysis, we consider the long-run impact of policies and assume that firms are fully informed about policy changes. Specifically, we assume that all firms with an incentive to reduce output profitably from notch-based preferences enact such reductions precisely to the point of the notch.

The results presented in Panel A demonstrate that the notch policy can increase total economic output by approximately 0.04–0.05% per year. Further, for firms coverred by preferential policies, the notch policy can lead to an output growth of 0.66–1.83%. Conversely, kink and subsidy policies can increase total economic output by approximately 0.1–0.13% per year while enabling SMEs coverred by the policy to achieve an output growth of 2.07–4.24%. Our simulation results indicate that kink and subsidy policies can significantly amplify the policy effect on economic output and SME production by 2–3 times more than notch-based tax cuts. This enhancement is primarily because of the absence of distortions caused by firms profiting from production cuts.

In Panel B, the results show that kink and subsidy policies do not increase the overall log of production much more than notch-based policies. As indicated by our previous results and analysis in Table 10, the effect of tax policy on overall log output is significantly driven by new entrants from the extensive margin. As the three types of policies are set to provide equivalent stimulus effects in the extensive margin for comparison, it is logical to arrive at these conclusions.

[25] In Table 12, the covered SMEs are precisely identified as Type 1, Type 2, and Type 3 firms, which can be cross-referenced in Table 8. These specific types of firms contribute to driving the overall changes in production and tax revenue within the Halve-Levy program.

Table 12: Aggregate Impact of Different Support Policies on the Economy

Panel A. Aggregate Production

	Total Economy				Covered SMEs			
	Original	Change Rate			Original	Change Rate		
Year	25% Tax	Notch	Kink	Subsidy	25% Tax	Notch	Kink	Subsidy
Unit:	Millions		%		Millions		%	
2014	45642.4	0.04	0.10	0.10	1062.2	1.83	4.24	4.10
2015	44882.1	0.05	0.11	0.11	1645.7	1.24	3.09	3.03
2016	65165.1	0.05	0.12	0.11	2571.4	1.15	2.93	2.88
2017	97560.1	0.05	0.12	0.12	4467.4	0.98	2.63	2.59
2018	132875.2	0.04	0.13	0.13	8349.0	0.66	2.07	2.05

Panel B. Aggregate Log Production

	Total Economy				Covered SMEs			
	Original	Change Rate			Original	Change Rate		
Year	25% Tax	Notch	Kink	Subsidy	25% Tax	Notch	Kink	Subsidy
Unit:	Thousands		%		Thousands		%	
2014	679.8	7.63	7.67	7.67	526.2	9.86	9.91	9.90
2015	677.7	7.65	7.67	7.67	563.8	9.20	9.22	9.22
2016	761.6	7.72	7.74	7.73	641.4	9.16	9.18	9.18
2017	898.0	7.78	7.79	7.79	772.1	9.04	9.06	9.06
2018	1059.9	7.81	7.82	7.82	944.2	8.77	8.78	8.78

Panel C. Tax Revenue

	Total Economy				Covered SMEs			
	Original	Reduction			Original	Reduction		
Year	25% Tax	Notch	Kink	Subsidy	25% Tax	Notch	Kink	Subsidy
Unit:		Millions				Millions		
2014	11410.6	157.38	320.53	30.17	265.5	157.38	150.97	30.17
2015	11220.5	244.82	478.21	37.95	411.4	244.82	237.19	37.95
2016	16291.3	382.75	742.86	57.13	642.9	382.75	371.30	57.13
2017	24390.0	665.71	1275.75	91.83	1116.8	665.71	647.33	91.83
2018	33218.8	1246.81	2327.65	143.96	2087.2	1246.81	1217.35	143.96

Note: The "Original 25% Tax" column represents economic output and tax revenue under the original 25% linear tax system. The "Notch," "Kink," and "Subsidy" columns display the percentage change in simulated economic outcomes under the actual notch-based tax cuts, counterfactual kink-based tax cuts, and counterfactual output-based subsidies to the 25% tax scenario. The covered SMEs are identified as Type 1, Type 2, and Type 3 firms, which can be cross-referenced in Table 8.

Finally, Panel C focuses on the impact of different policies on tax revenues. The panel first displays the corporate income tax revenues of the overall economy under the 25% linear tax scenario and those of the SMEs covered by the actual tax incentives. Subsequently, it reveals the ultimate reduction in tax revenues under the different tax incentives. Although the supportive policy leads to the growth of economic output, it also results in a reduction in tax revenue. Kink-based tax cuts notably mitigate the distortionary behavior of firms

that reduce production and significantly promote output. However, the annual reduction in tax revenue is generally approximately twice as substantial as that resulting from the notch-based tax cuts. This is primarily due to the tax break for the part of income below the threshold for all firms. This highlights that while kink-based policy eases distortionary effects, it concurrently exerts greater pressure on fiscal balance.

Conversely, compared to notch and kink-based corporate income tax reductions, an output-based subsidy policy stimulates economic output and minimizes the reduction in tax revenues. This is because tax cuts require foregoing a significant percentage of taxable income, whereas an output-based subsidy only requires compensating a portion of the opportunity cost of business operations. Nevertheless, implementing such an efficient subsidy policy presupposes the ability to find a one-to-one correspondence between the subsidy amount and the operational profit of a specific firm. Therefore, our exploration of this subsidy policy is framed within a more idealized context.

6.2.2 Tax cuts and subsidies with the same stimulating effect

In this subsection, we compare notch-based tax cuts, kink-based tax cuts, and output-based subsidies, focusing on achieving the same stimulus effect on economic output but with varying extents of tax revenue reduction.

Our counterfactual analysis follows these steps:

1. We calculate the extent to which the existing notch-based tax incentive increases total output and total log output compared to a 25% linear tax rate in a scenario where all firms are rationally and optimally informed about the government's tax policies and make a bunching decision.

2. We adjust the kink-based tax incentives and output-based subsidies introduced in the previous subsection. Maintaining the same tax rates and subsidy ratios, we move the threshold location for beneficiary firms to a point that enables the incentive to achieve the same stimulus effect on economic output as the notch-based tax cuts under the new location.

3. We compare the differences in the reduction of tax revenue across the different types of size-dependent policies to determine their efficiency.

Table 13 presents the four types of policies we compare, with flexible benefit thresholds for kink-based tax cuts and output-based subsidies to achieve the same stimulus effect.

Table 13: Alternative Support Policies to Achieve the Same Objective

Policy	Policy Threshold	Preferential Policy
Linear Tax	N.A.	N.A., 25% for all the taxable income
Notch	As real policy	10% for all the taxable income if it is smaller than the threshold
Kink	**New threshold**	10% of the part of taxable income under the threshold
Subsidy	**New threshold**	subsidy of 1/6 opportunity cost for taxable income under the threshold

The results of this analysis are presented in Table 14. The 25% Tax refers to the tax revenue under the initial 25% linear tax rate, and the "Aggregate Output" column displays the amount of tax revenue reduction required for Notch, Kink, and Subsidy policies to achieve the same effect as the existing Halve-Levy policies if the government's optimal goal is to optimize total output. The "Aggregate log output" column represents the scenario where the government treats total log output as the policy objective. As the actual policy is the notch-based tax cuts, the values in this column are similar to the ones in the previous table, serving as a benchmark for comparison.

Table 14: Effectiveness of Different Types of Support Policies

Objective		Aggregate Output			Aggregate Log Output		
Year	Original Tax	Tax Reduction			Tax Reduction		
	Unit: Millions of Yuan				Thousands		
	25% Tax	Notch	Kink	Subsidy	Notch	Kink	Subsidy
2014	11,411	157.4	53.8	9.4	157.4	148.6	19.5
2015	11,221	244.8	58.9	10.0	244.8	208.8	24.3
2016	16,291	382.8	85.6	14.6	382.8	320.7	36.6
2017	24,390	665.7	126.2	21.6	665.7	537.0	58.7
2018	33,219	1246.8	151.1	26.5	1246.8	921.4	91.8

Note: The "25% Tax" column presents tax revenue under the original 25% linear tax system. The "Notch," "Kink," and "Subsidy" columns display the amount of tax revenue reduction compared to the 25% linear tax scenario when the actual notch-based tax cuts, the counterfactual kink-based tax cuts, and the counterfactual output-based subsidies are applied. The "Aggregate Output" column displays the amount of tax revenue reduction required for Notch, Kink, and Subsidy policies to achieve the same effect as the existing Halve-Levy policies if the government's optimal goal is to optimize total output. The "Aggregate Log Output" column presents the scenario where the government treats total log output as the policy objective.

Our simulations demonstrate that if the government aims to optimize total output, kink-based tax cuts achieve the same policy effect with only 1/8–1/3 of the tax reduction required for notch-based tax cuts. As the government raises the income threshold for preferential policies, notch-based tax cuts become more inefficient, and kink-based policies can save even more fiscal expenditure. If the government aims to optimize total log output, a kink-based policy remains more effective than a notch-based policy; nonetheless, the difference is not substantial. This is primarily because of the significant weight of all the promotion effects on the micro new entrants.

Furthermore, when comparing a policy of output-based subsidies with notch-based tax cuts, the fiscal reduction of the subsidy policy can be even less, possibly up to as lower as 1/10 of that under the notch tax cuts.

Our simulations reveal that while notch-based tax relief, a prevalent approach in many countries, including China, may not require precise prediction of firm costs or reduction of tax revenues for all firms (including large ones), the financial commitment associated with such policies is significant than other types of size-dependent policies. This discrepancy is due to the inefficiencies related to firm production reductions at the notch.

7 Conclusion

Governments place significant emphasis on fostering the growth and stability of SMEs. To shed light on the impact of China's cascading notch-based tax cuts for SMEs during 2014–2018, we employ a combination of parsimonious empirical tests and structural counterfactual analysis. For the theoretical foundation and counterfactual framework, we integrate tax incentives into Arkolakis (2010) to model SME's behaviors. Our model proposes that notch-based tax cuts motivate qualified SMEs to boost production while also prompting some above-notch firms to decrease production for arbitrage opportunities. This approach enables a comprehensive analysis of the effects of these policies, facilitating a comparison of the advantages and disadvantages of different policy types.

In our theoretical framework, firms' optimal output results from balancing post-tax marginal benefits with marginal costs due to expansion. Consequently, corporate income tax cuts stimulate economic output by enabling firms to achieve higher levels of marginal returns at lower corporate income tax rates, ultimately leading to increased production. Owing to the model's simplicity and low data requirements for more detailed information, our quantitative framework can be extended to study size-dependent policies. Meanwhile, in the real economy, lower corporate income tax rates also boost firms' output through various mechanisms, including increased R&D activity and export investment under higher budget constraints. While our model focuses on macro-level impact mechanisms, there is potential for further development by incorporating more detailed data to analyze how firms expand their financing budget constraints through reduced corporate income tax incentives and dynamically grow through R&D and export activities. This avenue of research offers opportunities to refine the model and better understand how changes in tax incentives affect firms' production decisions through specific mechanisms of action. Hence, this topic has significant potential for exploration in conjunction with fiscal and tax policy.

In conclusion, our study comprehensively analyzes the impacts of China's notch-based tax cuts on SMEs, facilitating a thorough comparison of different policy types. Our findings

suggest that utilizing kink-based tax cuts and output-based subsidies may be more effective than notch-based policies in stimulating SMEs, and provide specific quantitative assessments. These insights contribute valuable information for policymakers and offer empirical evidence from China's experience regarding the nuanced effects of tax incentives on firms' production decisions and economic output.

References

[1] Almunia, Miguel, and David Lopez-Rodriguez. "Under the Radar: The Effects of Monitoring Firms on Tax Compliance." *American Economic Journal: Economic Policy*, Vol. 10, No. 1 (2018): 1-38.

[2] Arkolakis, Costas. "Market penetration costs and the new consumers margin in international trade." *Journal of Political Economy*, Vol. 118, No. 6 (2010): 1151-1199.

[3] Bachas, Pierre, and Mauricio Soto. "Corporate Taxation under Weak Enforcement." *American Economic Journal: Economic Policy*, Vol. 13, No. 4 (2021): 36-71.

[4] Boonzaaier, Wian, Jarkko Harju, Tuomas Matikka, and Jukka Pirttilä. "How Do Small Firms Respond to Tax Schedule Discontinuities? Evidence From South African Tax Registers." *International Tax and Public Finance*, Vol. 26, No. 5 (2019): 1104-36.

[5] Chaney, Thomas. "Distorted gravity: the intensive and extensive margins of international trade." *American Economic Review*, Vol. 98, No. 4 (2008): 1707-21.

[6] Chen, Qiaoyi, Zhao Chen, Zhikuo Liu, Juan Carlos Suárez Serrato, and Daniel Xu. "Regulating Conglomerates in China: Evidence from an Energy Conservation Program." *National Bureau of Economic Research*, No. w29066 (2021).

[7] Chen, Zhao, Zhikuo Liu, Juan Carlos Suárez Serrato, and Daniel Yi Xu. "Notching R&D investment with corporate income tax cuts in China." *American Economic Review*, Vol. 111, No. 7 (2021): 2065-2100.

[8] David, Joel M., and Venky Venkateswaran. "The Sources of Capital Misallocation." *American Economic Review*, Vol. 109, No. 7 (2019): 2531-67.

[9] Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda. "The Role of Entrepreneurship in Us Job Creation and Economic Dynamism." *Journal of Economic Perspectives*, Vol. 28, No. 3 (2014): 3-24.

[10] Devereux, Michael P., Li Liu, and Simon Loretz. "The Elasticity of Corporate Taxable Income: New Evidence from Uk Tax Records." *American Economic Journal: Economic Policy*, Vol. 6, No. 2 (2014): 19-53.

[11] Diamond, Rebecca, and Petra Persson. "The long-term consequences of teacher discretion in grading of high-stakes tests." *National Bureau of Economic Research*, No. w22207 (2016).

[12] Dixit, Avinash K., and Joseph E. Stiglitz. "Monopolistic competition and optimum product diversity." *The American Economic Review* Vol. 67, No. 3 (1977): 297-308.

[13] Dwenger, Nadja, and Viktor Steiner. "Profit Taxation and the Elasticity of the Corporate Income Tax Base: Evidence from German Corporate Tax Return Data." *National Tax Journal*, Vol. 65, No. 1 (2012): 117-50.

[14] Eaton, Jonathan, Samuel Kortum, and Francis Kramarz. "An anatomy of international trade: Evidence from French firms." *Econometrica*, Vol. 79, No. 5 (2011): 1453-1498.

[15] Fuest, Clemens, Andreas Peichl, and Sebastian Siegloch. "Do higher corporate taxes reduce wages? Micro evidence from Germany." *American Economic Review*, Vol. 108, No. 2 (2018): 393-418.

[16] Gabaix, Xavier. "Zipf's law for cities: an explanation." *Quarterly Journal of Economics*, Vol. 114, No. 3 (1999): 739-767.

[17] Garicano, Luis, Claire Lelarge, and John Van Reenen. "Firm size distortions and the productivity distribution: Evidence from France." *American Economic Review*, Vol. 106, No. 11 (2016): 3439-79.

[18] Gruber, Jonathan. *How elastic is the corporate income tax base?*. Ross School of Business, Office of Tax Policy Research, 2005.

[19] Guner, Nezih, Gustavo Ventura, and Yi Xu. "Macroeconomic Implications of Size-Dependent Policies." *Review of Economic Dynamics*, Vol. 11, No. 4 (2008): 721-44.

[20] Haltiwanger, John, Ron S. Jarmin, and Javier Miranda. "Who Creates Jobs? Small Versus Large Versus Young." *Review of Economics and Statistics*, Vol. 95, No. 2 (2013): 347-61.

[21] Harju, Jarkko, Aliisa Koivisto, and Tuomas Matikka. "The Effects of Corporate Taxes on Small Firms." *Journal of Public Economics*, Vol. 212, No. 8 (2022): 104704.

[22] Hassett, Kevin A., and R. Glenn Hubbard."Tax policy and business investment." In *Handbook of Public Economics*, Vol. 3, pp. 1293-1343. Elsevier, 2002.

[23] He, Daixin, Langchuan Peng, and Xiaxin Wang. "Understanding the Elasticity of Taxable Income: A Tale of Two Approaches." *Journal of Public Economics*, Vol. 197, No. 5 (2021): 104375.

[24] Helpman, Elhanan, Marc J. Melitz, and Stephen R. Yeaple. "Export Versus FDI with Heterogeneous Firms." *American Economic Review*, Vol. 94, No. 1 (2004): 300-16.

[25] Hopenhayn, Hugo A. "Entry, exit, and firm dynamics in long run equilibrium." *Econometrica: Journal of the Econometric Society* (1992): 1127-1150.

[26] Kleven, Henrik J., and Mazhar Waseem. "Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan." *Quarterly Journal of Economics*, Vol. 128, No. 2 (2013): 669-723.

[27] Kostøl, Andreas R., and Andreas S. Myhre. "Labor Supply Responses to Learning the Tax and Benefit Schedule." *American Economic Review*, Vol. 111, No. 11 (2021): 3733-66.

[28] Luttmer, Erzo GJ. "Selection, growth, and the size distribution of firms." *Quarterly Journal of Economics*, Vol. 122, No. 3 (2007): 1103-1144.

[29] Martin, Leslie A., Shanthi Nataraj, and Ann E. Harrison. "In with the Big, out with the Small: Removing Small-Scale Reservations in India." *American Economic Review*, Vol. 107, No. 2 (2017): 354-86.

[30] Melitz, Marc J. "The impact of trade on intra-industry reallocations and aggregate industry productivity." *Econometrica*, Vol. 71, No. 6 (2003): 1695-1725.

[31] Melitz, Marc J., and Stephen J. Redding. "New trade models, new welfare implications." *American Economic Review*, Vol. 105, No. 3 (2015): 1105-1146.

[32] Notowidigdo, Matthew J. "The Incidence of Local Labor Demand Shocks." *Journal of Labor Economics*, Vol. 38, No. 3 (2020): 687-725.

[33] OECD. *Taxation of SMEs in OECD and G20 Countries,* OECD Tax Policy Studies. Paris: OECD Publishing, 2015.

[34] Patel, Elena, Nathan Seegert, and Matthew Smith. "At a Loss: The Real and Reporting Elasticity of Corporate Taxable Income." Available at SSRN 2608166 (2017).

[35] Zwick, Eric, and James Mahon. "Tax policy and heterogeneous investment behavior." *American Economic Review* 107, No. 1 (2017): 217-248.

A Appendix (For Online Publication)

A.1 Bunching Estimates

A.1.1 Cross-Validation of p and $[TI_l, TI_u]$ in Bunching Analysis

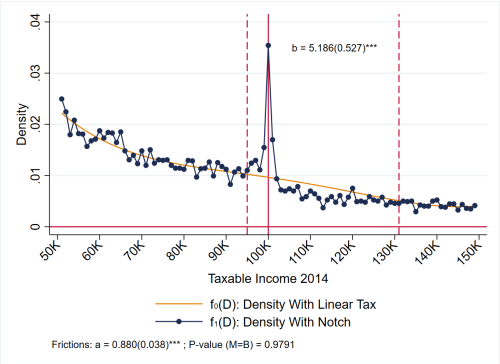
We follow Diamond and Persson (2016) and Chen et al. (2021) and use a data-based approach to select the width of the exclusion region (i.e., $[TI_l, TI_u]$) and the order of the polynomial, p. In particular, we use K-fold (K = 5) cross-validation to search over a triple (p, TI_l, TI_u) that determines the excluded region. We first test the hypothesis that $f_0(\cdot)$ and $f_1(\cdot)$ have equal mass over the exclusion restriction and keep the combination of parameters that do not reject the test of equality at the 10% level. Subsequently, we compute the sum of squared errors across the test subsamples and select a triple (p, TI_l, TI_u) that minimizes the sum of squared errors. Finally, we bootstrap the entire procedure to calculate standard errors.

A.1.2 Robustness of Bunching Estimates

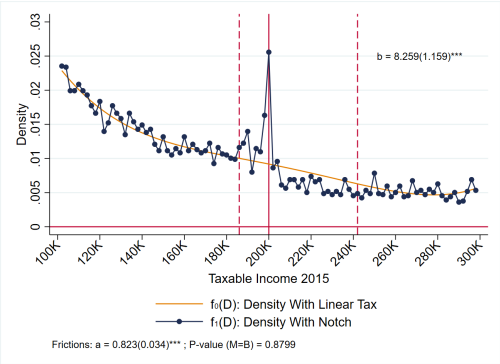
This section explores the robustness of our bunching estimates. First, Figure A1 shows that restricting (p, TI_l, TI_u) to the second-best estimate results in extremely similar estimates with the baseline estimates.

Second, as our sample data contains information on the number of employees reported by enterprises, we can conduct further research on firms that meet the requirements of the number of employees—fewer than 100 employees for industrial sector and 80 employees for other sectors. Figure A2 displays the estimates of the counterfactual density of taxable income, which are qualitatively and quantitively similar with baseline estimates.

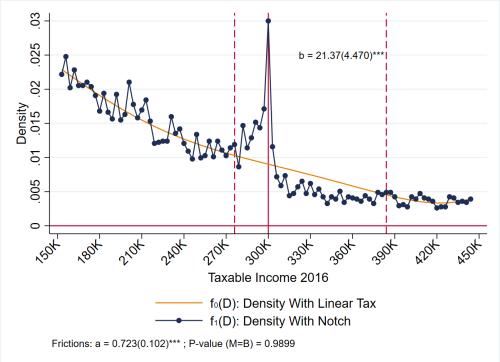
Figure A1: Robustness of Bunching Estimates to Specification of Counterfactual Density



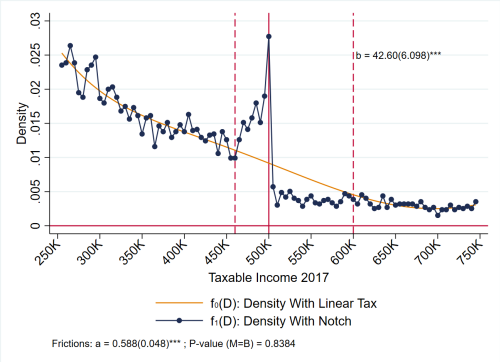
(a) Panel A. 2014



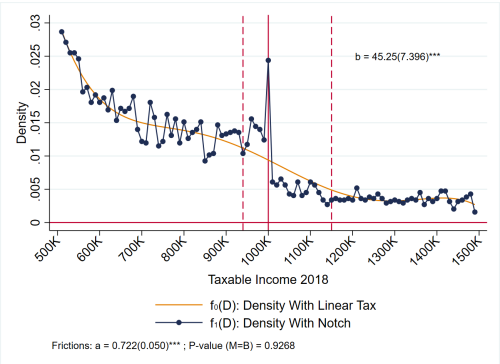
(b) Panel B. 2015



(c) Panel C. 2016



(d) Panel D. 2017

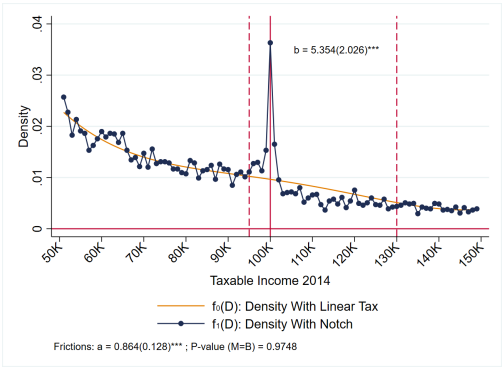


(e) Panel E. 2018

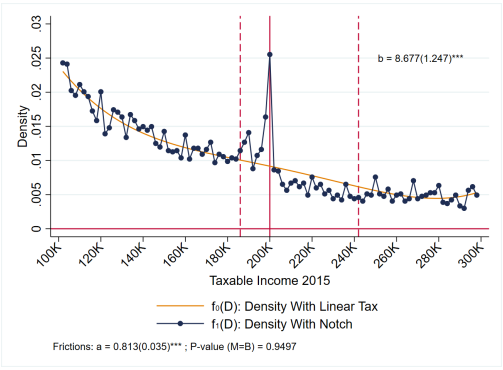
Note: This figure conducts a robustness check of the benchmark bunching analysis in Figure 2. As discussed in the appendix, we select $(p,\ TI_l, TI_u)$ via cross-validation. We use the second-best choice for the specification of $(p,\ TI_l, TI_u)$, and the graphs in Figure A1 show that our benchmark results are robust to how we specify $(p,\ TI_l, TI_u)$.

Source: Shanxi Administrative Tax Return Database.

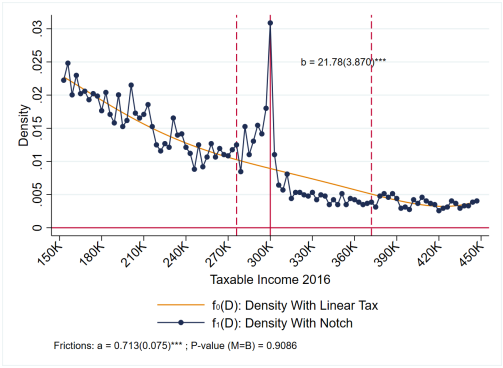
Figure A2: Robustness check: Requirements of the Number of Employees



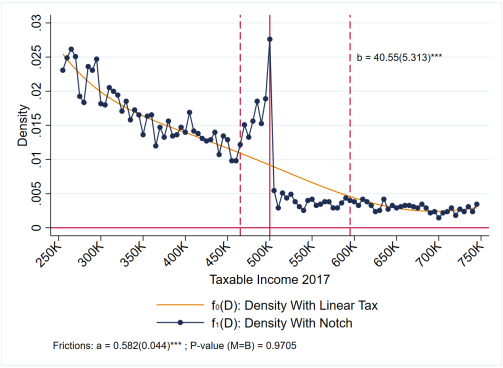
(a) Panel A. 2014



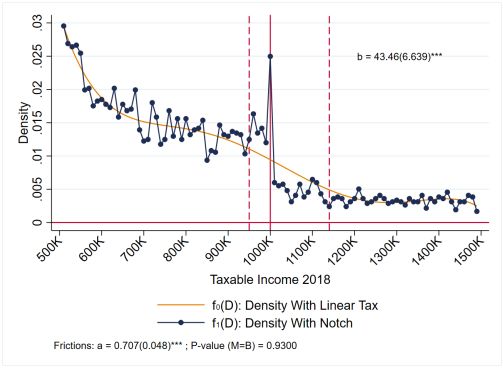
(b) Panel B. 2015



(c) Panel C. 2016



(d) Panel D. 2017



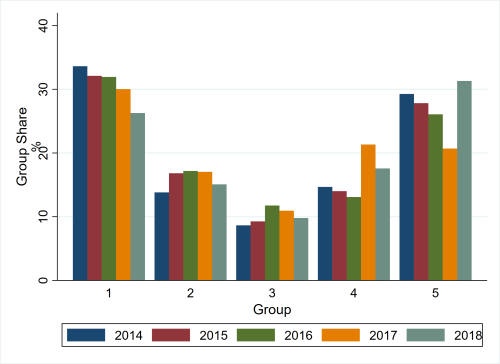
(e) Panel E. 2018

Note: This figure presents a robustness check of the benchmark bunching analysis in Figure 2. We conduct further sample screening to meet the other specific employees' requirements of the Halve-Levy program in addition to the taxable income requirements.

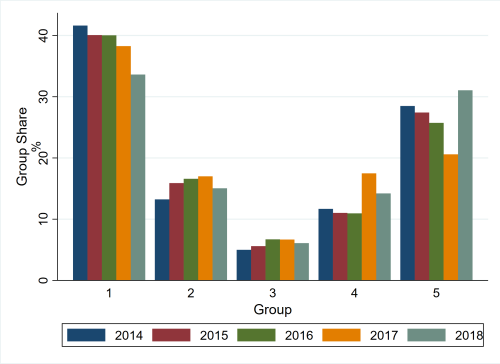
Source: Shanxi Administrative Tax Return Database.

A.2 Robustness of Fact 2

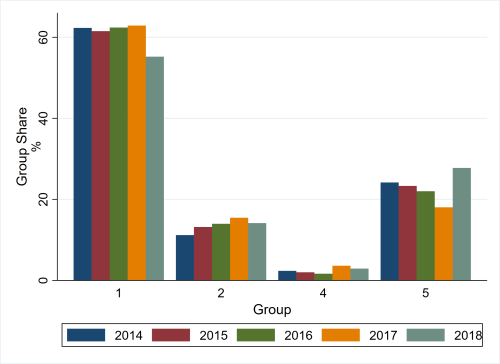
Figure A3: Robustness Check of Fact 2



(a) Bunching excluded regions: 0.95×Notch to 1.1×Notch



(b) Bunching excluded regions: 0.9×Notch to 1.2×Notch



(c) Bunching excluded regions: 0.8×Notch to 1.3×Notch

Note: Based on the tax incentives of the Halve-Levy program, we divided firms with annual taxable income below 1 million RMB into five groups: Group 1 (0 to 100 thousand), Group 2 (100 thousand to 200 thousand), Group 3 (200 thousand to 300 thousand), Group 4 (300 thousand to 500 thousand), and Group 5 (500 thousand to 1 million). We then plotted the annual share of total taxable income for each group among firms with taxable income below 1 million RMB. Panels A, B, and C show the annual share for each group with adjustments to exclude certain bunching regions to mitigate bias from firms' bunching behaviors. Specifically, Panel A's bunching excluded regions are from 0.95 times the notch value to 1.1 times the notch value, Panel B's are from 0.9 times the notch value to 1.2 times the notch value, and Panel C's are from 0.8 times the notch value to 1.3 times the notch value.

Source: Shanxi Administrative Tax Return Database.

A.3 Firms' Responses to Tax Cuts: DID results

Due to the fact that our sample observation period spans only five years and there are policy shocks each year, applying the traditional difference-in-differences (DID) estimation to arrange natural experiment control and treatment groups to test the impact of Halve-Levy policy shocks on firm production and bunching decisions poses certain challenges. Nonetheless, we still employ the DID method as a robust check to examine the effects of tax cuts on promoting the growth of eligible firms. As the qualifying threshold is adjusted annually, we conduct separate DID regressions on subsamples for two consecutive years. For instance, in the 2014–2015 sample, the threshold increased from 100 thousand in 2014 to 200 thousand in 2015. Consequently, firms with taxable income ranging from 100 thousand to 200 thousand were ineligible in 2014 but eligible in 2015. Conversely, firms with taxable income above 200 thousand remained ineligible. Therefore, the treatment group comprises firms with initial taxable income ranging from 100 thousand to 200 thousand in 2014, and the control group includes firms with initial taxable income ranging from 200 thousand to 500 thousand in 2014. The bandwidth of the control group is set as three times larger than that of the treatment group to maintain a sufficient number of control group observations. To eliminate the confounding impact of bunching near the threshold, we drop observations in the excluded bunching region (see Section 3.1). For example, in the 2014–2015 sample, we exclude firms with taxable income ranging from 184 thousand to 250 thousand in 2015.

The DID specifications are as follows:

$$log(taxable\ income)_{it} = \beta Treat_i * Post_t + (Z_i * \lambda_t)'\Theta + \gamma_i + \rho_t + u_{it},$$

where the dependent variable is the logarithm of the taxable income for firm i in year t. $Treat_i$ is a dummy variable, as previously defined.[26] $Post_t$ is defined as 1 in the year 2015 and 0 in 2014. We include the interaction term between firm-level initial characteristics Z_i [27] and year dummies λ_t in the regressions to control for the time effects of these characteristics on the outcome of interest. γ_i and ρ_t represent firm and year-fixed effects, respectively. The standard errors are clustered at the firm level. Notably, the corporate income tax rate for the treatment group decreased from 20% in 2014 to 10% in 2015, while the control group's tax rate remained unchanged. The coefficient β is expected to demonstrate a sign, indicating the impact of the reduced tax rate on the growth of firm income.

[26]The regression yields an intention-to-treat effect. Firms with initial taxable income ranging from 100 thousand to 200 thousand in 2014 may not necessarily experience treatment effects. If their taxable income grows beyond 200 thousand, they will no longer enjoy the preferential tax rate.

[27]Firm-level initial characteristics include factors such as age, number of employees, industry dummies and city dummies.

Table A1: Firms' Responses to Tax Cuts: DID results

	Dependent variable: logarithm of taxable income			
	2014–15 sample	2015–16 sample	2016–17 sample	2017–18 sample
	(1)	(2)	(3)	(4)
Tax Cuts	0.168	0.110	0.152	0.113
	(0.036)	(0.046)	(0.046)	(0.041)
Initial Characteristics * Year FE	YES	YES	YES	YES
Observations	9018	4256	4038	4198
R^2	0.126	0.150	0.178	0.133

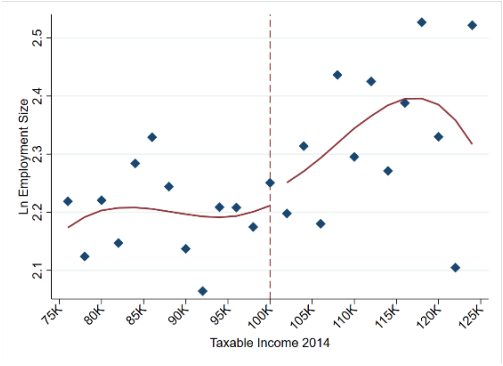
Note: The firm-level clustered standard errors are reported in parentheses. All regressions control for the interaction term of firm-level initial characteristics and year dummies. Firm-level initial characteristics include age, number of employees, industry dummies, and city dummies.

Similar DID specifications are conducted for the 2015–2016, 2016–2017, and 2017–2018 samples in Table A1. [28] In each case, the results show that tax cuts exert a positive effect on taxable income. This suggests that the reduced tax rates encourages newly eligible firms to expand their business size. For example, Column 1 demonstrates that after cutting the corporate income tax rate from 20% to 10%, firms increase their taxable income by 16.8%.

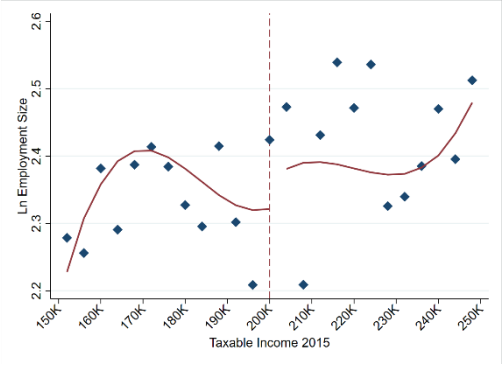
[28] In the 2015–2016 sample, firms with initial taxable income in 2015 ranging [200 thousand, 300 thousand) and [300 thousand, 600 thousand) are defined as treated and control groups, respectively. The corporate tax rate of the treated group has fallen from 20% in 2015 to 10% in 2016, while the tax rate of the control group remains at 25%. In the 2016–2017 sample, firms with initial taxable income in 2016 ranging [300 thousand, 500 thousand) and [500 thousand, 1.1 million) are defined as treated and control groups, respectively. In the 2017–2018 sample, firms with initial taxable income in 2017 ranging [500 thousand, 1 million) and [1 million, 2.5 million) are defined as treated and control groups, respectively. For the 2016–2017 and 2017–2018 samples, the corporate tax rate of the treated group has fallen from 25% to 10%, while the tax rate of the control group stays at 25%.

A.4 Misreporting Evidence

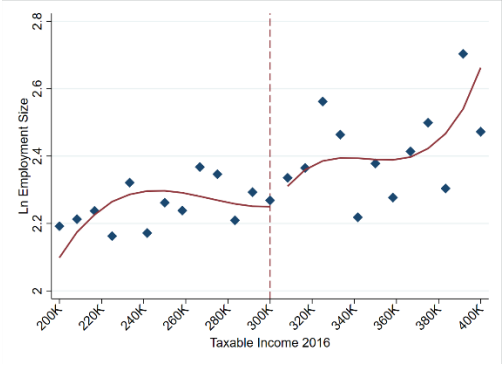
Figure A4: Alternative Measure of Firm Production at the Notch



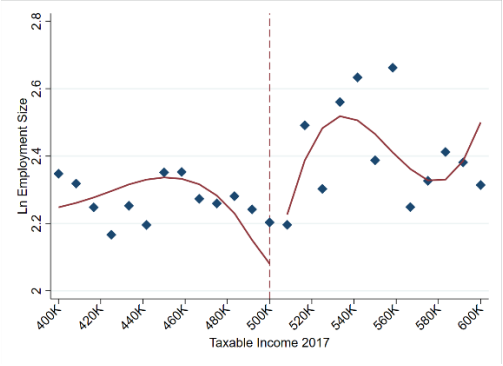
(a) Panel A. 2014



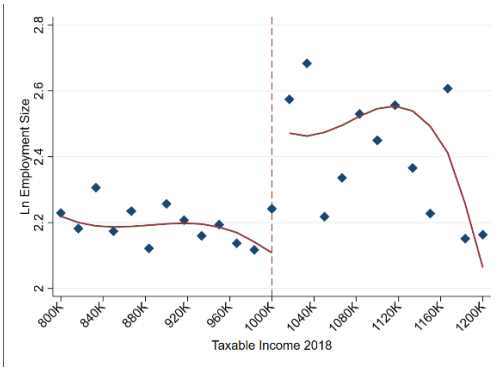
(b) Panel B. 2015



(c) Panel C. 2016



(d) Panel D. 2017



(e) Panel E. 2018

Note: This figure plots log of the number of employees at each size category, along with an estimated cubic regression of taxable income with heterogeneous coefficients above and below the notches.
Source: Shanxi Administrative Tax Return Database.

A.5 Firm Heterogeneity Distribution

According to the properties of Pareto distribution, the portion of the distribution $\frac{\varphi_{jt}}{\overline{\varphi}_t}$ that maps to the observed data (market entrants) follows a Pareto distribution with a scale parameter of 1 and shape dispersion of θ.

Proof: According to our Pareto distribution setting of firm productivities, the distribution of each firm j's relative productivity to the cutoff productivity, denoted as $\frac{\varphi_{jt}}{\overline{\varphi}_t}$, has a distribution as

$$Pr(\frac{\varphi_{jt}}{\overline{\varphi}_t} \leq y) = Pr(\varphi_{jt} \leq \overline{\varphi}_t \times y) = 1 - T_t(\overline{\varphi}_t y)^{-\theta}.$$

Consequently, the productivity distribution of market entrants observed in the data is the conditional distribution $Pr(\frac{\varphi_{jt}}{\overline{\varphi}_t} \leq y | \varphi_{jt} \geq \overline{\varphi}_t)$ and we have

$$\begin{aligned}
Pr(\frac{\varphi_{jt}}{\overline{\varphi}_t} \leq y | \varphi_{jt} \geq \overline{\varphi}_t) &= Pr(\frac{\varphi_{jt}}{\overline{\varphi}_t} \leq y | \frac{\varphi_{jt}}{\overline{\varphi}_t} \geq 1) \\
&= Pr(1 \leq \frac{\varphi_{jt}}{\overline{\varphi}_t} \leq y) / Pr(\frac{\varphi_{jt}}{\overline{\varphi}_t} \geq 1) \\
&= [Pr(\frac{\varphi_{jt}}{\overline{\varphi}_t} \leq y) - Pr(\frac{\varphi_{jt}}{\overline{\varphi}_t} \leq 1)] / Pr(\frac{\varphi_{jt}}{\overline{\varphi}_t} \geq 1) \\
&= [T_t(\overline{\varphi}_t)^{-\theta} - T_t(\overline{\varphi}_t y)^{-\theta}] / [T_t(\overline{\varphi}_t)^{-\theta}] \\
&= 1 - y^{-\theta}
\end{aligned}$$

Consequently, the relative productivities of the firms we observe from the data follows a Pareto distribution of scale 1 and shape dispersion θ.

A.6 Model Fit

Figure A5: Model Fit 2

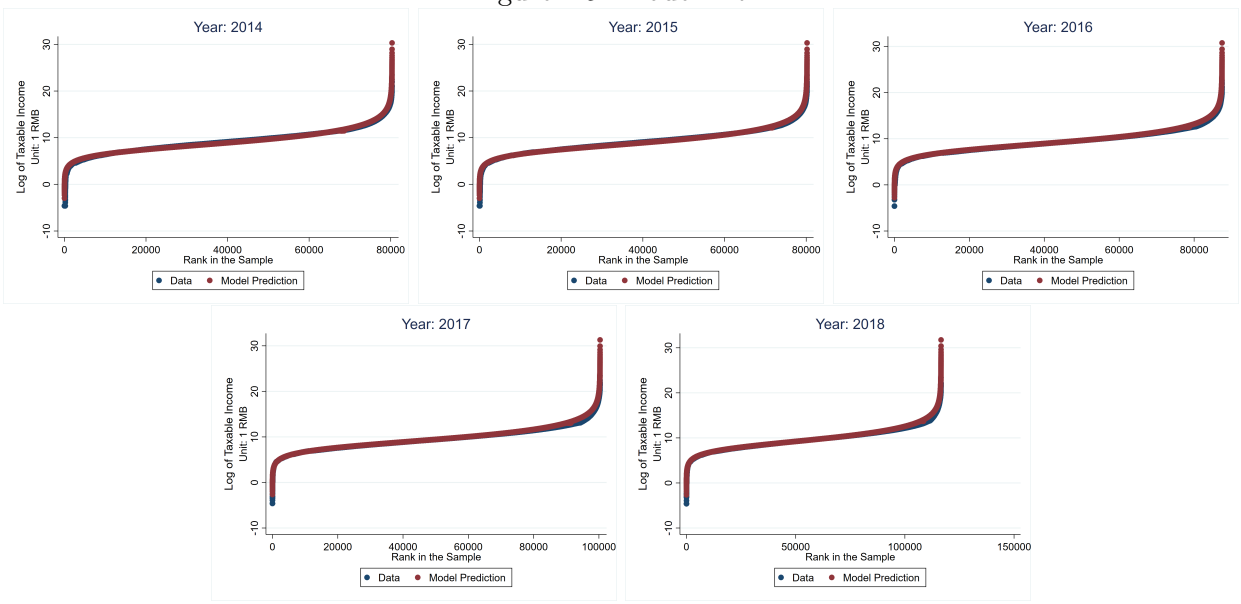
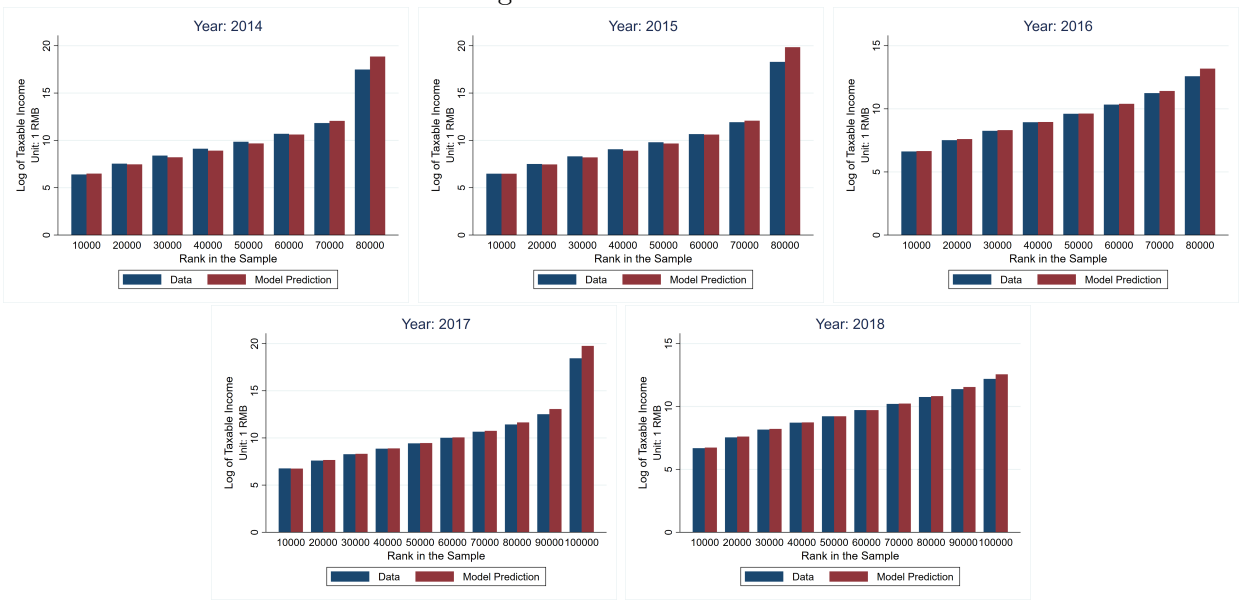


Figure A6: Model Fit 3



Similar to the methodology in Figure 5, we rank the sales of each firm in each year from the actual data and compare them with the sales of each artificial firm in the model's prediction.

In Figure A5, we plotted the logarithm of taxable income for firms at each taxable income rank using blue lines for the actual data and red lines for model predictions. Apart from

very small and very large firms, our model closely matches the data for the main range of firms affected by the Halve-Levy policy.

In Figure A6, we plot the log of taxable income for the 10,000th, 20,000th, 30,000th, 40,000th, 50,000th, 60,000th, 70,000th, 80,000th, 90,000th, and 100,000th firms in the actual data using the blue bars and in the model's prediction using the red bars. The bars demonstrate that our model exhibits a considerably close fit to the data. In particular, in the non-largest 10% of companies per year, the model predictions closely align with the actual data at each stage.