Long-Run Economic Consequences of Firm Avoidance in the Threshold Regulation for Financial Reporting

Yu Cao and Lisa Yao Liu*

December 2024

Abstract

We estimate the long-term economic consequences of firm avoidance of Section 404 of the Sarbanes-Oxley Act (SOX 404), a significant threshold regulation. Firms distort R&D investments to avoid regulatory costs. However, SOX 404's information benefits, such as improved internal controls, mitigate this distortion. Extending an investment model with information frictions, we quantify both the regulatory cost and the information benefit channels. We estimate that, firm avoidance reduces R&D investment by 12.3%, and productivity by 1.15%. Counterfactual analyses suggest raising the threshold reduces productivity loss, and firms comply voluntarily if the benefits outweigh the costs.

Keywords: Financial reporting regulation, regulatory avoidance, intangible investment, R&D, bunching, threshold regulation, productivity

JEL Classification: E22, M41, M42, M48, O32

*Yu Cao: The World Bank (ycao5@worldbank.org). Lisa Yao Liu: Columbia Business School, Columbia University, New York, NY 10027, United States of America (yl4689@gsb.columbia.edu). This paper supersedes a structural estimation draft of the same title from Liu's Curriculum Paper at Chicago Booth. We appreciate helpful comments from Tim Baldenius, Ray Ball, Phil Berger, Jonathan Bonham, Matthias Breuer, Jung Ho Choi, Hans Christensen, Dhammika Dharmapala, Max Farrell, Pingyang Gao, Eric Gelsomin (discussant), Stephen Glaeser, Jon Glover, Chris Hansen, Erik Hurst, Christian Leuz, Irina Luneva, Mark Maffett, Valeri Nikolaev, Amedeo Pugliese (discussant), Haresh Sapra, Doug Skinner, Chad Syverson, Joseph Vavra, Thomas Winberry, David Windisch (discussant), Anastasia Zakolyukina, Frank Zhou (discussant), and workshop participants at the AAA, Accounting Summer Camp, Columbia Business School, European Winter Meeting of the Econometric Society, Fox and Haskayne Accounting Conference in collaboration with the International Association for Accounting Education and Research (IAAER), Minnesota Accounting Empirical Research Conference, the University of Chicago, the World Bank, and Workshop on Accounting and Economics. Liu gratefully acknowledges financial support from the University of Chicago Booth School of Business and Columbia Business School. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent. All errors are our own.

1 Introduction

Threshold regulations, or size-dependent policies, are increasingly used by financial reporting regulators to exempt smaller companies from compliance (e.g., Bernard et al., 2018; Breuer, 2021; Reuters, 2024; SEC, 2018). In the US, the SEC frequently uses public float as a threshold for financial reporting regulations.[1] For instance, in 2018, the SEC amended its definition of "smaller reporting companies" to include firms with a public float under $250 million. In March 2024, the SEC introduced a climate disclosure rule for companies with a public float over $75 million. Leuz and Wysocki (2016) advocate for regulatory thresholds to help with economic analysis and causal inference. However, these thresholds can also motivate firms to stay small or go private to avoid regulatory costs, a behavior known as avoidance. As R&D investment significantly affects firm valuation and market capitalization (e.g., Bloom et al., 2013; Glaeser and Lang, 2024), firms near the public float threshold may have incentives to strategically reduce their R&D investment. Such reduction aims to manipulate market expectations about their long-term values, allowing them to remain below the regulatory threshold to avoid compliance. Therefore, understanding such avoidance behavior and its long-term economic impact is crucial for assessing the economic costs of financial reporting regulation.

To examine this question, we use SOX 404, an important threshold regulation in financial reporting that aims to improve the quality of a firm's financial reporting and internal control environment. The cost of compliance with SOX 404 increases significantly when a firm's public float exceeds $75 million, with the impact being particularly pronounced for smaller firms whose public float is just above this threshold (e.g., Risk et al., 2002; SEC, 2011).[2] Specifically, SOX 404 has a direct impact on firms' compliance costs (e.g., audit fees), generating a wedge between firms of different sizes (Iliev, 2010).

We ask the following question in this paper: How does avoidance (i.e., public float manage-

[1] Public float is defined as the proportion of a company's publicly traded common stock that is not held by insiders such as management, firm officers, directors, or controlling-interest investors.

[2] Section 404(a) of SOX requires that management assesses the effectiveness of the internal controls over financial reporting, while 404(b) of SOX requires that firms have an independent auditor to attest to and report on management's assessment. Our sample period mainly focuses on the time when non-accelerated filers (i.e., firms with a public float below the threshold) are exempt from both SOX 404(a) and SOX 404(b).

ment) influence a firm's R&D investments, productivity, and aggregate productivity growth? We develop a theoretical model to demonstrate firms' avoidance incentives. This model guides our estimation and allows us to study the long-run economic consequences of SOX 404 on aggregate R&D investment, productivity, and investor welfare. Importantly, we quantify key economic channels, including regulatory costs and information benefits. Finally, we evaluate counterfactual scenarios, such as the impact of different regulatory thresholds.

We focus on R&D investment because it has long-term economic implications and is a critical driver of a firm's productivity growth and market value (Bloom et al., 2013; Glaeser and Lang, 2024). Increased R&D spending signals a firm's commitment to technological advancement and product innovation, which investors often interpret as a positive indicator of future growth and profitability (e.g., Bai et al., 2016; Lev and Sougiannis, 1996). Firms with higher R&D intensity tend to experience greater market capitalization due to their potential for innovation and long-term competitive advantage (e.g., Hirschey and Weygandt, 1985). Conversely, reductions in R&D expenditures can negatively impact market capitalization, as they may signal to investors a lack of commitment to innovation and long-term value creation (e.g., Chan et al., 2001). Consequently, declines in R&D investment can lead to diminished market expectations and a corresponding decrease in public float. In the context of SOX 404, Gao et al. (2009) find that firms, particularly those with stronger incentives to stay below the regulatory threshold, reduce investment to remain small. They identify reduced investments (including the R&D investment) as one of the most significant financial and real tools firms use to manage size.[3]

Following the R&D investment literature, we develop an investment model that endogenizes productivity. Our model extends prior R&D investment frameworks by incorporating imperfect information into the investment decision-making process. Specifically, firms make investments with incomplete information about future idiosyncratic productivity shocks. These information frictions reduce the marginal returns on R&D investments, leading to underinvestment. This underinvestment impedes individual firm growth and, consequently, aggregate productivity.

[3]In our estimation, we account for other channels that may influence public float and isolate the impact of the R&D investment channel that we emphasize.

Incorporating information frictions into the theoretical model addresses a key aspect of financial reporting regulation. SOX 404 provides potential information benefits. For example, managers and independent auditors' examinations and tests of internal controls can provide managers with more precise signals about the effectiveness of firms' internal information systems (e.g., Alexander et al., 2013; Berger et al., 2023; Feng et al., 2009; Wagner and Dittmar, 2006). By complying with SOX 404, these informational benefits can contribute to managers' learning and resource allocation within the firm (Berger et al., 2023; Bloom et al., 2013).

Hence, our model integrates an information benefit channel that captures increases in managers' information precision. We demonstrate that when regulatory costs exceed information benefits, marginal firms—those indifferent between complying and not complying—are above the threshold. Firms above the threshold but below the marginal firm point have incentives to reduce their R&D investments to bunch below the threshold and avoid regulation. Higher regulatory costs shift the marginal firm point further from the threshold, causing more firms to bunch below it and creating more distortions. This avoidance can create a "density hole" in the size distribution.

However, the improved information environment resulting from the regulation can move the marginal firm closer to the threshold, reducing the "density hole" above the threshold. Consequently, the overall distortion from avoidance can be reduced. Under threshold regulation, a firm's investment decision trades off regulatory costs, information benefits from compliance, and future productivity growth. When regulatory costs are greater, firms are incentivized to stay below the threshold and remain small.

To quantify the economic impact of firm avoidance on investment and productivity, we estimate the model using a simulated method of moments (SMM) approach. The two key parameters—regulatory costs and increases in information precision—are structurally estimated by analyzing how firms adjust their public floats and R&D-to-sales ratios around the regulatory threshold. These parameters influence a firm's behavior in distinct but complementary ways, which we leverage for identification. Specifically, bunching behavior, observed as missing mass above the threshold and excess mass below it, reflects higher regulatory costs. Meanwhile, significant adjustments in R&D-to-sales ratios are driven by changes in information precision, as greater precision enhances managers'

understanding of how R&D investments influence outcomes such as productivity, allowing firms to better align inputs (such as capital and labor) with productivity shocks.

The joint estimation of these parameters relies on the interplay between bunching behavior and R&D-to-sales ratio adjustments. Regulatory costs primarily drive bunching behavior. In contrast, changes in the R&D-to-sales ratio are particularly sensitive to information precision. By jointly analyzing the size of bunching and R&D-to-sales ratio adjustments, we can disentangle and estimate the relative contributions of regulatory costs and information precision.

With the estimated structural parameters, we find that firm avoidance leads to an average reduction of 61.38% in R&D investment, an average 1.15% decline in firm productivity, and a 1.10 percentage point drop in productivity growth among firms in the bunching region. Compliance with SOX 404 reduces the marginal return on R&D investment, leading to an average reduction of 4.01% in R&D investment. Furthermore, as only 0.21% of all firms in the entire economy bunch below the threshold, this suggests that the observed reductions in R&D investment and productivity are primarily driven by firms' compliance with the regulations. We estimate that regulatory costs reduce aggregate R&D investment by 7% and productivity by 1.5%, while information benefits increase R&D investment and productivity by 4.19% and 0.91% for the entire economy. The regulatory cost is equivalent to a 10.6% implicit corporate tax rate, whereas the information benefit is equivalent to a 7.06% implicit corporate subsidy.[4]

Using the economic model and estimated parameters, we conduct counterfactual analyses to assess the impact of alternative threshold regulations. Eliminating regulatory thresholds under mandatory compliance would remove firms' avoidance behavior but significantly reduce aggregate R&D investment and productivity due to decreased marginal returns on R&D. Increasing the regulatory threshold slightly raises the number of firms bunching just below the new threshold but substantially reduces the average drop in R&D investment among these firms. For example, with a $250 million threshold, 0.40% of firms bunch, and their average reduction in R&D investment falls to 49.15%. This indicates that the costs of exceeding the threshold often outweigh the benefits, in-

[4]The "implicit corporate tax rate" represents the cost of regulation as a percentage reduction in revenues, similar to a direct tax, while the "implicit corporate subsidy" reflects the benefit of regulation—improved information precision—as a percentage boost to revenues, akin to a direct subsidy.

centivizing firms to remain below it. However, higher thresholds reduce the number of firms subject to compliance, mitigating the negative effects on overall R&D investment and productivity. These counterfactual results provide useful insights for ongoing discussions about raising the public float threshold in financial reporting regulations.

Our framework also allows us to conduct a counterfactual analysis where the informational benefits outweigh the compliance costs—a scenario of voluntary compliance. In this scenario, the significant information benefit incentivizes managers to comply with SOX 404, leading to increased R&D investment and a higher rate of productivity growth. We find that the improved information environment could increase aggregate R&D investment by 9.22% and productivity by 1.9%. Overall, our counterfactual analyses inform regulators on the policy design to mitigate the distortions due to firm avoidance. Regulators could consider (1) raising the public float threshold further or (2) substantially enhancing the informational benefits associated with compliance with financial reporting regulations.

Our study makes several contributions. It quantifies the magnitudes and long-term economic implications of firm avoidance in an important financial reporting threshold regulation, SOX 404 (Leuz and Wysocki, 2016). While a large body of prior research has studied the economic consequences of SOX 404 (detailed in Section 2), our paper has two distinct features. First, we explicitly study the endogenous avoidance decisions made by companies around regulatory thresholds. Previous reduced-form research often uses a size instrument to address endogeneity concerns related to regulatory thresholds, as firms may adjust their sizes in response to these regulations. In contrast, our study directly exploits firms' endogenous avoidance incentives, contributing to the literature by modeling their economic incentives, exploring the long-term economic consequences of avoidance strategies, and distinguishing between the regulatory cost and information benefit channels of SOX 404. Second, with our economic model and structural parameters, we can vary regulatory thresholds to explore hypothetical scenarios and answer "what if" questions to inform regulatory design.

Moreover, we advance the bunching literature (e.g., Ewens et al., 2024) by studying bunching variables beyond those directly regulated, specifically, variables other than public float (Diamond

and Persson, 2016).[5] By examining firms' endogenous R&D investments, we also address the call for research on the relation between management controls and innovation, highlighting regulatory avoidance as a new channel between mandated financial reporting and innovation (Glaeser and Lang, 2024). Our framework applies not only to SOX but also to a wide range of threshold-based financial reporting regulations that induce both regulatory costs and information benefits. Many regulations are threshold-based, and this trend is increasing. Therefore, our framework has broad implications for regulatory economics.

Finally, our paper contributes to the literature on size-dependent regulation, focusing on the benefits and costs of threshold regulation. Studies on size-dependent policies highlight misallocations and distortions compared to a market-solution benchmark (e.g., Guner et al., 2008; Garicano et al., 2016) and spark international debates on such regulations. Ewens et al. (2024) use bunching estimation to study firms' compliance costs with the threshold regulation and estimate the impact on firms' listing and delisting decisions. Our paper incorporates both the compliance costs and the information benefits of the threshold regulation into the economic model, quantifying these channels. In particular, we assess the reduced distortion attributable to information benefits compared to a scenario that considers only regulatory costs. Recognizing the information benefit is important, as overlooking it may obscure the overall effects of financial reporting regulations. For instance, improved information environments can reduce distortions. In particular, our counterfactual analysis shows that size-dependent policies can have overall improvements if the information benefits induced by regulation exceed the regulatory costs.

2 Institutional Background and Literature Review

On June 5, 2003, the SEC announced that firms with public floats larger than $75 million in 2002 or thereafter must comply with Section 404 (SEC release 33-8238).[6] SOX mandates a series of changes to corporate financial reporting and corporate governance for public companies. Under SOX

[5]Several studies have observed regulatory avoidance behavior in various contexts, such as Ballew et al. (2022) on banking regulation, Bernard et al. (2018) on the disclosure practices of private firms, and Garicano et al. (2016) along with Amirapu and Gechter (2020) on labor regulation. McClure (2023) quantifies the costs of tax avoidance.

[6]The $75 million public float threshold originated from Form S-3, a simplified securities registration form, which was amended in 1992. Issuers with public floats greater than $75 million could use Form S-3, allowing them to disclose less information in their registration statements. In contrast, smaller firms were required to disclose more information when issuing securities because less information was publicly available about them. This $75 million threshold has not been adjusted for inflation over time when used in other threshold regulations.

404, disclosure of any material weaknesses detected by management and a firm's auditor must be detailed in the annual report. Under SOX 404(a), management is required to test the effectiveness of the firm's internal control structure and procedures and to disclose its assessment in the annual report. Under SOX 404(b), auditors evaluate and report on a company's internal controls over financial reporting. They assess the design and effectiveness of these controls, conduct testing, report findings, and may provide recommendations for improvement.

Since 2005, once a firm's public float crosses $75 million or the revenue crosses $50 million (if the public float information is unavailable), it remains subject to SOX 404 until its public float falls below $50 million. Before 2005, the rule was stricter, requiring that the public float fall below $25 million and that certain other conditions be met to become a non-accelerated filer. The public float is determined on the last business day of the second fiscal quarter. To calculate public float, the aggregate worldwide number of shares of voting and nonvoting common equity held by non-affiliates is multiplied by the price at which the common equity was last sold or the average of the bid and ask prices of common equity in the principal market for the common equity. The difference between market value and public float is held by insiders (or "affiliates"). An affiliate is "a person that directly, or indirectly through one or more intermediaries, controls, or is controlled by, or is under common control with" the issuer (17 CFR 240.12b-2). Insiders typically include managers, officers, directors, and large blockholders. If the calculated public float is not available, annual revenues of less than $50 million during the most recently completed fiscal year (for which audited financial statements are available) should be considered.

Following SOX's enactment, the SEC extended the section 404 compliance date for smaller reporting firms seven times, largely out of concern for the costs and the uncertainty of benefits. In 2010, the Dodd-Frank Act permanently exempted smaller public companies from compliance with the 2002 version of SOX 404 (b). In so doing, Dodd-Frank determined that "non-accelerated" filers, with a public float of $75 million or less, would be subject to 404(a) but not to 404(b).

Prior research has examined the economic consequences of SOX 404. On the intensive margin, Gao et al. (2009), Iliev (2010), and Nondorf et al. (2012) document that small reporting companies near the threshold manipulate their float to remain small and avoid compliance. On the extensive

margin, Leuz et al. (2008) show that firms choose to go private (or "go dark") to avoid costly regulation. Ge et al. (2017) investigate the benefits (i.e., aggregate audit fee savings) and costs (i.e., internal control misreporting) of exemption by applying the coefficients of non-exempt firms to small firms in their prediction models. Dharmapala (2022) use the bunching method to uncover counterfactual size distributions. Regarding SOX and risk-taking behavior, Bargeron et al. (2010) and Cohen et al. (2009) find that U.S. firms reduce risk-taking behavior after SOX. However, Albuquerque and Zhu (2019) do not find a correlation between SOX 404 compliance and a decline in investment and other risk-taking behaviors.

Our paper differs from this stream of literature because we study firms' endogenous responses to avoid compliance after SOX, while prior literature treats SOX 404 as a plausibly exogenous shock. Additionally, using our economic model and structural parameters, we can vary regulatory thresholds to analyze hypothetical scenarios and address "what if" questions, providing useful and relevant insights for regulatory design.

3 Model

3.1 Setup

Based on the R&D investment literature (e.g., Aw et al., 2008; Chen et al., 2021), our model assumes that R&D investment could contribute to firms' total factor productivity (TFP) growth with uncertainty. Consider a monopolistically competitive market where each firm i produces a single good using the production technology: $y_{it} = e^{\phi_{it}} k_{it}^{\alpha} l_{it}^{1-\alpha}$. The production function exhibits constant returns to scale in labor and capital. ϕ_{it} is the log of TFP following the law of motion:

$$\phi_{it} = \rho\phi_{it-1} + \varepsilon \ln(I_{it-1} + 1) + \mu_{it} \tag{1}$$

Here, $I_{it-1} > 0$ represents R&D investment, and $\mu_{it} \overset{i.i.d.}{\sim} N(0, \sigma^2)$ denotes the unobserved productivity shock with a variance of σ^2. The parameter ε measures the elasticity of R&D investment to firm i's productivity, while ρ captures the persistence of TFP over time. Our specification of the productivity process also implies that productivity follows an AR(1) process in the absence of R&D investment, consistent with the empirical fact that a fraction of firms never engage in R&D

investment. Suppose firm i faces a downward-sloping demand curve given by $y_{it} = p_{it}^{-\theta}$, where $\theta > 1$ represents the elasticity of product substitution.[7] The firm's revenue at the end of each period t is given by $\pi_{it} = y_{it}^{1-\frac{1}{\theta}}$.

Production under incomplete information At each time t, firm i makes its production decision before the productivity shock μ_{it} is realized. Following prior literature (e.g., David et al., 2016; Breuer and Windisch, 2019; Choi, 2021), we assume that at the beginning of time t, each firm observes a noisy private signal of productivity μ_{it}: $s_{it} = \mu_{it} + e_{it}$. The firm then chooses capital (k_{it}) and labor (l_{it}) based on this signal. However, such a signal is not observed by the firm when making the R&D investment, which is made one period before (e.g., Chen et al., 2021). The timeline for investment, signal, and production is exhibited in Figure 5. We can then summarize the firm's information set when choosing k_{it} and l_{it} as $\mathcal{I}_{it} = \{\{\phi_{it-j}\}_{j=0}^{t-1}, s_{it}\}$. Assume e_{it} and μ_{it} are independently distributed, and $e_{it} \overset{i.i.d.}{\sim} N(0, \sigma_e^2)$. Given the information structure, the expected productivity shocks can be written as:

$$\mathbb{E}_{it}[e^{(1-\frac{1}{\theta})\mu_{it}}|\mathcal{I}_{it}] = e^{\frac{1}{2}(1-\frac{1}{\theta})^2\mathbb{V}+(1-\frac{1}{\theta})\frac{\mathbb{V}}{\sigma_e^2}s_{it}}$$

with $\mathbb{V} = \left(\frac{1}{\sigma^2} + \frac{1}{\sigma_e^2}\right)^{-1}$ measuring the precision of the signal. Then the log of the expected productivity shock equals the weighted sum of the signal precision and the signal itself. Given the information structure and the timeline of the investment, the optimization problem of firm i yields the expected profits before the realization of s_{it} and μ_{it} as below (see Appendix A.1 for detailed derivations):

$$\mathbb{E}_{it}[\pi_{it}] \equiv \mathbb{E}_{it}[p_{it}y_{it} - w_t l_{it} - r_t k_{it}] = g(\sigma^2, \sigma_e^2)\Phi_{it}(1 + I_{it-1})^{\tilde{\theta}\varepsilon} \tag{2}$$

Here, $\tilde{\theta} = \theta - 1$, and $\Phi_{it} \equiv \left(\frac{\tilde{\theta}(1-\alpha)}{\theta w_t}\right)^{\tilde{\theta}(1-\alpha)} \left(\frac{\tilde{\theta}\alpha}{\theta r_t}\right)^{\tilde{\theta}\alpha} e^{\tilde{\theta}\rho\phi_{it-1}}$, which increases with the firm's initial productivity. The function $g(\sigma^2, \sigma_e^2) = \frac{1}{\theta} e^{\frac{1}{2}(1-\frac{1}{\theta})\tilde{\theta}\sigma^2 + \frac{1}{2}(1-\frac{1}{\theta})\tilde{\theta}^2\sigma^2\frac{V}{\sigma_e^2}}$ represents the manager's expectation over future signal and productivity shock, which is the same across firm i. Consider the extreme case where $\sigma_e^2 \to \infty$, and define $h(\sigma^2) \equiv g(\sigma^2, \infty) = \frac{1}{\theta} e^{\frac{1}{2}(1-\frac{1}{\theta})\tilde{\theta}\sigma^2}$. In this case, the signal

[7] Assume there is a final goods producer that uses each firm i's output as an input to produce final goods. The production technology is given by $Y_t = \frac{\theta}{\theta-1}\int_0^1 y_{it}^{1-\frac{1}{\theta}} di$. We normalize the price of the final goods to 1. The optimization problem of the final goods producer yields the inverse demand function for each product i as $p_{it} = y_{it}^{-\frac{1}{\theta}}$.

is too noisy and will be ignored by the manager. If the signal perfectly reveals the productivity shock, i.e. $\sigma_e^2 = 0$, then $g(\sigma^2, 0) = \frac{1}{\tilde{\theta}} e^{\frac{1}{2}(1-\frac{1}{\tilde{\theta}})\tilde{\theta}\sigma^2 + \frac{1}{2}(1-\frac{1}{\tilde{\theta}})\tilde{\theta}^2\sigma^2} \geq g(\sigma^2, \sigma_e^2)$. $g(\sigma^2, \sigma_e^2)$ increases in the precision of the signal, or decreases in σ_e^2. The increase in the precision of the signal could increase expected profits by reducing the probability of either under-investment or over-investment. Given the information structure, the expected profit increases with the firm's initial productivity level ϕ_{it-1} and the investment I_{it-1}. The profit elasticity of R&D investment is approximated by $\tilde{\theta}\varepsilon$.

3.2 Benchmark (No Regulation)

Firm i makes its R&D investment decision at the end of period t before any production shock μ_{it+1} and signal s_{it+1} are realized. The unit cost of R&D investment for firm i is $I_{it} + \frac{b}{2} I_{it}^2$ unit of final goods, where $\frac{b}{2} I_{it}^2$ represents the quadratic adjustment cost. For simplicity and analytical tractability, we assume $b = 1$ across all firms. To focus on firms' tradeoffs, for simplicity, we model each firm's investment decision as an intertemporal problem below:[8]

$$\max_{I_{it}>0} \pi_{it} - I_{it} - \frac{1}{2} I_{it}^2 + \beta g(\sigma^2, \sigma_e^2) \Phi_{it+1} (1 + I_{it})^{\tilde{\theta}\varepsilon}$$

Here, β is the discount factor, and $g(\sigma^2, \sigma_e^2)$ is a function related to productivity uncertainty. The optimal choice of I_{it} is given by:[9]

$$I_{it}^* = \begin{cases} [\tilde{\theta}\varepsilon\beta g(\sigma^2, \sigma_e^2)\Phi_{it+1}]^{\frac{1}{2-\tilde{\theta}\varepsilon}} - 1 & \text{if} \quad \phi_{it} > \tilde{\phi}_{it} \\ 0 & \text{if} \quad \phi_{it} \leq \tilde{\phi}_{it} \end{cases} \tag{3}$$

Here, $\tilde{\phi}_{it}$ is determined through the equation $\Phi_{it+1}(\phi_{it}) = (\tilde{\theta}\varepsilon\beta g(\sigma^2, \sigma_e^2))^{-1}$. The optimal investment increases with the initial productivity level ϕ_{it} at a decreasing rate. There exists a threshold $\tilde{\phi}_{it}$ such that firms find it optimal not to engage in R&D investment if their initial productivity level is below the threshold. The threshold $\tilde{\phi}_{it}$ increases with the signal noise σ_e^2, and

[8] We model the manager's R&D investment decision as an intertemporal problem for two reasons. First, firms can be viewed as repeatedly solving the two-period intertemporal problem throughout the sample period. Second, focusing on the manager's intertemporal decision instead of using a dynamic programming model simplifies the analysis and provides an analytical solution for R&D investment. This simplification allows us to isolate the direct effects of information frictions and regulatory costs on firms' R&D investments.

[9] The second order condition requires that $\tilde{\theta}\varepsilon < 1$.

decreases with the elasticity of R&D, ε.

Given the optimal investment I^*_{it}, the expected profits for the firm at time $t+1$, denoted as $\mathbb{E}_{it}[\pi_{it+1}]$, are calculated at time t as $\mathbb{E}_{it}[\pi_{it+1}] = \frac{1}{\theta\varepsilon}(1+I^*_{it})^2$. The profit in the period $t+1$ is positively related to the optimal investment. The firm's public float at the end of period t can be linear in its expected second-period profit under certain conditions, that is, $PF_{it} = \lambda\mathbb{E}^{investor}[\pi_{it+1}]$.[10] We assume that investors do not observe the signals received by managers about future productivity (McClure and Zakolyukina, 2024). Therefore, the public float can be written as $PF_{it} = \lambda h(\sigma^2)\Phi_{it+1}(1+I_{it})^{\tilde{\theta}\varepsilon}$. Given the value λ, there's a one-to-one mapping between the investment I^*_{it} and the public float in the time t, which is a function of expected profit in the $t+1$. Firms can adjust their public float by changing the R&D investment. Empirically, factors beyond R&D investment can influence a firm's public float, and managers might not perfectly estimate how reductions in R&D investment affect the public float. To account for these possibilities, we introduce a random component into the firm's public float to capture influences beyond R&D investment. Specifically, we assume that λ follows a log-normal distribution, which allows the price-to-expected profit ratio to vary among firms.

3.3 With the Threshold Regulation

Regulatory Cost Channel In this section, we evaluate the costs and benefits of SOX 404. SOX 404 imposes regulatory costs on firms exceeding the public float threshold, $\overline{PF}$. It requires management to assess and report on internal controls and mandates the hiring of external auditors to test these controls over financial reporting. We consider the regulatory costs for firms, modeling them as a function of their public float. Firms with a public float exceeding the threshold $\overline{PF}$ must forgo a fraction c of their sales to comply with SOX 404 requirements. The profits, net of regulatory cost, of the firms that comply with SOX can then be written as $\pi_{it} - c\theta\pi_{it} = (1-c\theta)\pi_{it}$. The regulatory cost can be viewed as an "implicit" tax on firm profits. Define $\tau^c = c\theta$; the regulatory

[10] Suppose there exists a risk-neutral investor that buys s firm share in the period t in order to get s% of the firm's profit. The investor's Euler equation can be written as $P_{it} = \mathbb{E}^{investor}[\pi_{it+1} + P_{it+1}]$. Iterating on P_{it+1}, we can get $P_{it} = \lambda\mathbb{E}^{investor}[\pi_{it+1}]$. See Appendix A.2 for the derivation. If we incorporate the stock market's information structure and investors' participation constraints into our model, these constraints would likely be relaxed following the implementation of SOX 404. Consequently, the public float might increase after the regulation, making firms further reduce their R&D investment.

cost in the $t+1$ has the following structure:

$$\tau = \begin{cases} 0 & \text{if} \quad \lambda \mathbb{E}_{it}[\pi_{it+1}] < \overline{PF} \\ \tau^c = c\theta > 0 & \text{if} \quad \lambda \mathbb{E}_{it}[\pi_{it+1}] \geq \overline{PF} \end{cases}$$

Information Benefit Channel One difference from previous threshold regulations studied in the economics literature is that compliance with financial reporting regulations (e.g., SOX 404) can provide decision-makers with more precise internal control information within firms. Within a firm, managers are required to test and disclose the effectiveness of internal controls. We refer to the information benefits of SOX 404 as the "information benefit channel." Better internal controls with independent auditors' attestations should yield more reliable internal data, such as inventories, payables, and performance measures, thus leading to better internal decision-making. Institutionally, Alexander et al. (2013) use the SEC's survey data to show that, from an insider (manager) perspective, SOX 404 compliance helps improve the quality of the firms' information environments. Specifically, 80% of respondents ascribe some benefits to SOX 404 compliance. Large fractions of respondents cite a positive impact of compliance on the quality of their firms' internal control structures (73%), audit committee confidence in the ICFR (71%), quality of financial reporting (48%), and the ability to prevent and detect fraud.

Several other empirical research studies support this notion. For instance, Feng et al. (2009) find that management forecasts are less accurate in the presence of ineffective ICFR, and become more accurate when weaknesses are remediated. Wagner and Dittmar (2006) also suggest that compliance can help strengthen the control environment, standardize processes, reduce complexity, and generate other administrative efficiencies. Additionally, Berger et al. (2023) use SOX 404 compliance as a regulatory shock for the improvement in managerial information systems and study how the enhanced internal information environment impacts investment efficiency.

We model the information benefit channel as an increase in the precision of the private signal firms receive during production. The variance of the noise contained in the signal has the following

structure:

$$\sigma_e^2 = \begin{cases} \sigma_e^2 & \text{if} \quad \lambda \mathbb{E}_{it}[\pi_{it+1}] < \overline{PF} \\ \sigma_e^2(1-\eta) < \sigma_e^2 & \text{if} \quad \lambda \mathbb{E}_{it}[\pi_{it+1}] \geq \overline{PF} \end{cases}$$

Here, $\eta \in (0,1)$ is the reduction in noise.[11] The marginal benefits of making an extra investment can be written as $\tilde{\theta}\varepsilon(1-\tau^c)g(\sigma^2,(1-\eta)\sigma_e^2)\Phi_{it+1}(1+I_{it})^{\tilde{\theta}\varepsilon-1}$ if the firm's public float at the time t is above the threshold. As the function g increases with the reduction in noise η, the increase in signal precision could increase the firm's investment and, thus, future productivity. We define productivity increases resulting from a reduction in noise as the information benefit. The benefit from more precise information could induce firms to comply with SOX 404. Therefore, whether the implementation of SOX 404 increases or decreases the marginal benefits of investment depends on the relative magnitude of τ^c and η.

Bunching below the threshold If the regulation cost channel outweighs the information benefit channel, that is $(1-\tau^c)g(\sigma^2,(1-\eta)\sigma_e^2) < g(\sigma^2,\sigma_e^2)$, compliance to SOX 404 could decrease the marginal benefit of R&D investment when the firm's public float is above the threshold $\overline{PF}$. Therefore, firms that previously invested above $\overline{PF}$ have the incentive to make sub-optimal investments and keep their public float below the threshold. Considering a firm with an optimal firm size above the threshold in the absence of regulation, the firm decides whether to bunch by comparing the current value of bunching (to avoid regulation) with the current value of operating at the optimal size (while complying with regulation).[12] The value of the firm conditional on bunching is given by:

$$v(I_{it}^b|\tau) = \pi_{it} - \frac{1}{2}(I_{it}^b+1)^2 + \frac{1}{2} + \beta(1-\tau)g(\sigma^2,\sigma_e^2)\Phi_{it+1}(1+I_{it}^b)^{\tilde{\theta}\varepsilon}$$

where I_{it}^b is given by:

$$I_{it}^b = \left(\frac{\overline{PF}}{\lambda h(\sigma^2)\Phi_{it+1}}\right)^{\frac{1}{\tilde{\theta}\varepsilon}} - 1 \tag{4}$$

[11] Another interpretation of the signal noise reduction, η, involves two components. First, firms previously unwilling to implement internal systems are now mandated to do so, improving the internal information environment, which they internalize in their compliance decisions. Second, SOX can generate an aggregate information benefit, including information spillovers from other firms, which firms may not internalize in their compliance decisions.

[12] Recall that the current value is defined as the expected profit in the $t+1$ minus the cost of investment.

The value of firms at the interior optimal I_{it}^c is:

$$v(I_{it}^c|\tau^c,(1-\eta)\sigma_e^2) = \pi_{it} - \frac{1}{2}(I_{it}^c+1)^2 + \frac{1}{2} + \frac{1}{\tilde{\theta}\varepsilon}(1+I_{it}^c)^2$$

where I_{it}^c is given by:

$$I_{it}^c = [\tilde{\theta}\varepsilon\beta(1-\tau^c)g(\sigma^2,(1-\eta)\sigma_e^2)\Phi_{it+1}]^{\frac{1}{2-\tilde{\theta}\varepsilon}} - 1 \tag{5}$$

Firms that previously (before SOX) chose $I_{it}^* > I_{it}^b$ before SOX will bunch at the threshold if $v(I_{it}^b|\tau,\sigma_e^2) \geq v(I_{it}^c|\tau^c,(1-\eta)\sigma_e^2)$. By substituting I_{it}^* and I_{it}^b in $v(I_{it}^c|\tau^c)$ and $v(I_{it}^b|\tau)$ using equations (5) and (4), and applying $PF^c = \lambda h(\sigma^2)\Phi_{it+1}(1+I_{it}^c)^{\tilde{\theta}\varepsilon}$, we can express the firm's decision to bunch or not as:

$$-\frac{1}{2}\left(\frac{PF^c}{\overline{PF}}\right)^{1-\frac{2}{\tilde{\theta}\varepsilon}}\tilde{\theta}\varepsilon + \frac{1-\tau}{1-\tau^c}\frac{g(\sigma^2,\sigma_e^2)}{g(\sigma^2,(1-\eta)\sigma_e^2)} \geq (1-\frac{\tilde{\theta}\varepsilon}{2})\frac{PF^c}{\overline{PF}} \tag{6}$$

Equality holds for firms that are indifferent between bunching and not bunching (i.e., marginal firms). If the LHS of equation (6) is greater than the RHS, firms decide to bunch. On the one hand, bunching below the threshold reduces a firm's investment, future productivity gains, and the informational benefits derived from compliance. On the other hand, it increases the firm's value by lowering implicit corporate tax (i.e., regulatory costs).

The first term on the LHS of equation (6) represents the cost of bunching (i.e., future productivity reduction), while the second term captures the net gain from bunching through two channels: (1) savings on the "implicit" corporate tax, minus (2) information benefits from increased signal precision. The RHS term represents the benefit of not bunching (i.e., future productivity gain).

When there is no information benefit (i.e., $\eta = 0$) and firms have an optimal public float close to the threshold ($PF^c \approx \overline{PF}$), the relative productivity loss from bunching is small. However, the increased "implicit" corporate tax incentivizes bunching. Conversely, when there is an information benefit (i.e., $\eta > 0$), which is $g(\sigma^2,(1-\eta)\sigma_e^2) \geq g(\sigma^2,\sigma_e^2)$, the inequality is less likely to hold for $PF^c \approx \overline{PF}$. This implies that firms are less likely to bunch when regulatory compliance provides

positive information benefits.[13]

For firms where the optimal public float significantly exceeds the threshold ($PF^c >> \overline{PF}$), the relative benefits of tax savings are unlikely to outweigh the costs of future productivity losses and the loss of information benefits. Thus, these firms will choose to comply with the regulation.

Fixing the value of λ, a firm's public float in the period $t+1$ increases with time t productivity ϕ_{it}. There exists a threshold ϕ_{it}^c such that firms with $\phi_{it} \in [\bar{\phi}_{it}, \phi_{it}^c]$ will choose to bunch. As noted earlier, with information benefits, the inequality is less likely to hold, resulting in a much smaller bunching region compared to the case without information benefits. The cutoff productivity, ϕ_{it}^c, decreases with information benefits (η) and increases with regulatory costs (τ^c). The optimal investment under regulation can be expressed as follows:

$$I_{it}^{**} = \begin{cases} [\tilde{\theta}\varepsilon\beta(1-\tau^c)g(\sigma^2,(1-\eta)\sigma_e^2)\Phi_{it+1}]^{\frac{1}{2-\tilde{\theta}\varepsilon}} - 1 & \text{if} \quad \phi_{it} > \phi_{it}^c \\ \left(\frac{\overline{PF}}{\lambda h(\sigma^2)\Phi_{it+1}}\right)^{\frac{1}{\tilde{\theta}\varepsilon}} - 1 & \text{if} \quad \bar{\phi}_{it} < \phi_{it} \leq \phi_{it}^c \\ [\tilde{\theta}\varepsilon\beta(1-\tau)g(\sigma^2,\sigma_e^2)\Phi_{it+1}]^{\frac{1}{2-\tilde{\theta}\varepsilon}} - 1 & \text{if} \quad \tilde{\phi}_{it} < \phi_{it} \leq \bar{\phi}_{it} \\ 0 & \text{if} \quad \phi_{it} \leq \tilde{\phi}_{it} \end{cases} \tag{7}$$

where ϕ_{it}^c is the time t productivity level of the marginal firm, which is determined through equation (6) when equality holds. The largest initial productivity level at which a firm's optimal public float remains just below the threshold is denoted by $\bar{\phi}_{it}$. This value, $\bar{\phi}_{it}$, is determined by the equation $\overline{PF} = \lambda h(\sigma^2)\Phi_{it+1}(\bar{\phi}_{it})\left[\tilde{\theta}\varepsilon\beta(1-\tau)g(\sigma^2,\sigma_e^2)\Phi_{it+1}(\bar{\phi})\right]^{\frac{\tilde{\theta}\varepsilon}{2-\tilde{\theta}\varepsilon}}$. In other words, $\bar{\phi}_{it}$ represents the firm whose optimal size is just below the threshold. The difference between $\bar{\phi}_{it}$ and ϕ_{it}^c defines the bunching region. Under positive compliance costs, the size of the bunching region decreases with increasing η and increases with the regulatory cost τ^c.

[13]Suppose the marginal firms with $PF^c > \overline{PF}$ and where the equality holds:

$$-\left(\frac{PF^c}{\overline{PF}}\right)^{1-\frac{2}{\tilde{\theta}\varepsilon}}\tilde{\theta}\varepsilon + \frac{1-\tau}{1-\tau^c} = (1-\frac{\tilde{\theta}\varepsilon}{2})\frac{PF^c}{\overline{PF}}$$

Substituting into the bunching condition, we have:

$$-\frac{1-\tau}{1-\tau^c} + \frac{1-\tau}{1-\tau^c}\frac{g(\sigma^2,\sigma_e^2)}{g(\sigma^2,(1-\eta)\sigma_e^2)} = \frac{1-\tau}{1-\tau^c}\frac{g(\sigma^2,\sigma_e^2) - g(\sigma^2,(1-\eta)\sigma_e^2)}{g(\sigma^2,(1-\eta)\sigma_e^2)} \leq 0$$

The equality holds when $\sigma_e^2 \to \infty$ or $\eta = 0$. That is, with information benefit (i.e., $\eta > 0$ and σ_e^2 is finite), some of the previously bunched firms will not bunch.

Within the bunching region (i.e., $\phi_{it} \in [\bar{\phi}_{it}, \phi^c_{it}]$), the optimal investment decreases as the initial productivity ϕ_{it} increases. Conversely, in the two non-bunching regions—$[\tilde{\phi}_{it}, \bar{\phi}_{it}]$ and $(\phi^c_{it}, \infty]$—the optimal investment increases with the initial productivity level. In the first region, the firm's initial productivity is too low to reach the threshold, even with R&D investment. In the second region, the firm's initial productivity is high enough that the productivity and information benefits outweigh the profit loss from the implicit corporate tax after compliance.

If a firm's initial productivity falls within the interval $(\bar{\phi}_{it}, \phi^c_{it}]$, its optimal R&D investment decreases as initial productivity increases. In this region, when the cost of the implicit corporate tax outweighs future productivity and information benefits, the firm will choose to bunch. As productivity increases, the firm must further reduce investment to remain below the threshold. Compared to the optimal investment under regulatory costs without considering information benefits, the marginal firm's investment point is lower when information benefits are taken into account, resulting in more firms bunching due to regulatory costs alone.

In some cases, where the information benefit outweighs the regulatory costs, i.e., $(1-\tau^c)g(\sigma^2, (1-\eta)\sigma^2_e) > (1-\tau)g(\sigma^2, \sigma^2_e)$, a firm with an optimal public float just below the threshold may find it optimal to bunch above the threshold. In such cases, the firm's investment under regulation would be higher than it would be without regulation. We discuss this in detail in Appendix A.3.

3.4 Discussions on Major Model Assumptions

The intuition of our model is that the regulation introduces a discontinuity in the form of a regulatory tax, and firms bunch just below the threshold as a response to this discontinuity in their incentives. The intuition is demonstrated in the section of Illustrative Examples from Figures 1 to 4. In this section, we discuss major assumptions.

Why do we focus on R&D investment and not other channels? This paper mainly focuses on real effects: how firms' avoidance affects R&D investment and productivity and the long-run implications of that avoidance. Changing R&D investment is also a feasible way to adjust the public float. R&D investments boost market capitalization by signaling future innovation and growth, enhancing investor confidence and firm valuation. Conversely, reducing R&D spending can lower market capitalization by suggesting a reduced commitment to innovation, weakening investor

expectations and firm value. Gao et al. (2009) find that a key strategy firms use to remain small is to reduce investment (including the R&D investment), especially when they have strong incentives to stay below the threshold. Among financial and real strategies, they identify investment reduction as one of the most significant. While our paper focuses on manipulating public float by cutting R&D investment, this reduction leads to decreased capital investment as well.[14]

R&D investment might not be the only way to engage in avoidance. Therefore, in the estimation, we account for other factors that may impact the public float to isolate the importance of the R&D investment channel we focus on. As R&D investment impacts long-term economic efficiency, using this approach to avoid regulation can have significant efficiency implications in the long run.

Does our model incorporate possible positive externality? We model the information benefit, which primarily arises from improvements in firms' internal information environments (i.e., private information signals for managers, consistent with the work of David et al. (2016)); however, this could also be interpreted as an aggregate information benefit that incorporates information spillovers (e.g., Bernard et al., 2018; Breuer, 2021). Improving the internal information environment can also generate benefits such as better access to capital and reduced agency problems. Compared to the bunching literature in economics, we expect the investment and productivity distortions to be smaller because financial reporting regulation also provides information spillover gains.

We note that, while we address the information benefits of the regulation, the model does not explicitly justify regulatory intervention, which is often motivated by externalities where private parties fail to internalize broader economic impacts. Our focus, instead, is on quantifying the firm-level costs and benefits of threshold regulation. In the estimation, we aggregate reductions in firm-level R&D investment to assess their macro-level effects on productivity and investor welfare.

However, our model emphasizes firm-level trade-offs and abstracts from externalities associated with R&D investments, which drive broader economic growth and aggregate productivity growth (e.g., Romer, 1990). For example, in the framework of Klette and Kortum (2004), a firm's

[14]Other financial and disclosure tools include making more cash payouts to shareholders, reducing the number of shares held by non-affiliates, making more bad news disclosures, and reporting lower earnings than control firms. Firms may also face financial constraints when engaging in repurchasing. Moreover, it may be more costly for firms to go private than to adjust their investment (Engel et al., 2007; Leuz et al., 2008). Therefore, firms could adjust their R&D investments in response to the regulation before deciding to go private.

R&D investments can advance the productivity frontier through creative destruction, displacing incumbent firms while generating positive spillovers that benefit other firms in subsequent periods. In this paper, we mainly focus on incumbent firms and do not consider the effects of creative destruction. In addition, we consider firms' responses to regulation only at the intensive margin (e.g., adjusting investment) and do not model extensive margin decisions, such as firm entry or exit. Furthermore, the fixed number of intermediate goods in our model limits its ability to capture productivity gains from expanded product variety, which could influence long-term market outcomes.

Should we model the board of directors? In the benchmark model, no agency friction exists, because we assume managers are firm owners. This modeling choice is consistent with the findings of Alexander et al. (2013) that governance structure is not a major determinant of perceived compliance effects, but firm complexity is. However, in the extension with the evasion response, we test for this possibility by studying whether we can uncover a discontinuity in other variables (e.g., the difference between market value and public float) around the threshold. Specifically, we split the sample into several subgroups in which the agency issues are presumed to be higher. We use the difference between market value and public float to measure how much managers care about firm value. We provide suggestive evidence of an evasion response, but we cannot provide a precise estimate of how much firms' avoidance and the consequential productivity implications arise from the evasion response incentive.[15]

What role does concurrent regulatory guidance play in the model? The regulatory guidance (e.g., the introduction of the SEC's management guidance and the PCAOB's issuance of Auditing Standard No. 5) would help strengthen the information benefit channel, which is consistent with survey evidence of Alexander et al. (2013).[16] To the extent that AS5 improves audit efficiency and, therefore, reduces audit fees, it could also help reduce compliance costs.

[15] In Appendix C, we examine the bunching differences when firms' market value is \$10 million (13.3% of the threshold) and \$15 million (20% of the threshold) larger than the public float. The graphical evidence also suggests the bunching is greater when the difference between market value and public float is larger.

[16] AS2 required extensive ICFR audit procedures at the process level and, in the absence of guidance for applying 404(a), management typically followed the process level focus of AS2; AS5 was intended to promote a top-down, risk-based approach for auditors, and SEC guidance was intended to do the same for management.

4 Data and Descriptive Evidence

4.1 Data and Sample Construction

Our main data sources on firms' investment and audit fees are from Compustat and Audit Analytics. We exclude firms from regulated utilities and those classified as public services, international affairs, or non-operating establishments. We also exclude firms with missing or negative values on revenue, physical capital, and book value of assets. We collect the public float data from firms' annual 10-K filings, accessed through the SEC's EDGAR system. We trim public float, revenue, and capital stock at 1% and 99% levels to remove outliers. We exclude firms with negative R&D expenditures or those with R&D expenses exceeding their revenue. The sample is an unbalanced panel data set with approximately 17,985 firm-year observations spanning the fiscal year 2000 through 2007.[17] Our main variable of interest is the firm's expenditure on productivity-enhancing activities. The firm's spending on product development, patents, and software are all considered productivity-enhancing and are recorded in its R&D expenditures.

During our sample period, 39% of firm-year observations do not report R&D expenditure. The absence of reported R&D does not necessarily imply a lack of R&D activities. Firms may choose not to disclose their R&D expenditure or may allocate these expenses to other budget categories while still engaging in innovation activities, such as patent applications (e.g., Koh and Reeb, 2015). Treating unreported R&D expenditure as zero could introduce significant bias into our estimates. Following the previous studies (e.g., Hall and Oriani, 2006; Chen et al., 2021), we restrict our estimation to a sub-sample of firms that report R&D expenditures.[18] However, we exclude firms that report zero R&D expenditure consistently throughout the entire sample period. The key variables used in our structural estimation are summarized in Table 1.

[17] Our sample period ends in 2007 for two reasons. First, it concludes before the financial crisis to minimize the confounding impact of both the crisis and other financial reporting threshold regulations. Second, it includes both SOX 404a and SOX 404b for regulatory compliance above the threshold, requiring *managers* and external auditors to test the effectiveness of internal controls over financial reporting, thereby strengthening the internal information benefit channel.

[18] As a robustness check, we also use the method suggested by Koh and Reeb (2015) to impute missing R&D values using either industry averages or historical firm-specific averages. We find similar estimates.

4.2 Descriptive Evidence

Figures 6 and 7 plot the distribution of firms' public float around the $75 million threshold, both before and after the implementation of SOX 404. If firms manipulate their public float to avoid regulation, we would expect to observe bunching in the distribution of public float following the implementation of the threshold regulation. Figure 7 shows a sharp decline in the density of firms' public float just above the $75 million cutoff in the post-regulation period. These bunching patterns suggest that the costs associated with SOX 404 compliance may be substantial enough to incentivize firms to remain below the threshold to avoid these compliance costs.

We first provide preliminary evidence on the magnitude of compliance costs following the implementation of SOX 404. While compliance costs are not directly observed in our sample, we follow Iliev (2010) and use total audit fees as a proxy for these costs. Since SOX 404 requires external auditors to issue an attestation report on management's assessment, we expect an increase in audit fees for firms' compliance with the regulation. To estimate the regulation-induced changes in firms' total audit fees, we estimate the following regression:

$$\ln(AuditFee_{it}) = \alpha \times Above_{it} + \beta \times Above_{it} \times Post_t + X_{it} + \epsilon_{it}$$

The dependent variable is the log of total audit fees, including audit-related fees.[19] $Above_{it}$ is an indicator variable equal to one if the firm's public float exceeds $75 million. $Post_t$ is an indicator variable equal to one for years after 2004 and zero otherwise. X_{it} represents a set of controls, including the log of the firm's public float, year fixed effect, and firm fixed effect. Table 2 shows that for firms with a public float above $75 million, total audit fees increased by an average of 51.2% (54.7% among firms with positive R&D expenditure) after the regulation.

To examine whether firms manipulate their public float by adjusting R&D investments, we compare the R&D-to-sales ratios of firms located just above and just below the regulatory threshold. Our model predicts that firms near the threshold may bunch by manipulating their public float through cutting R&D investment. Therefore, firms just below the threshold should exhibit a lower R&D-to-sales ratio than those just above it. We test this idea by estimating the following

[19] We adjust it by the GDP deflator to be consistent with estimates in macro studies.

specification within a narrow window around the threshold, focusing on firms with public floats between $50 and $150 million.

$$RD_{it} = \alpha \times Below_{it} + \beta \times Below_{it} \times Post_t + X_{it} + \epsilon_{it}$$

The dependent variable RD_{it} is a firm's R&D to sales ratio. $Below_{it}$ is an indicator variable that equals one if firm i's public float is less than $75 million. $Post_t$ is an indicator variable that equals one for years after 2004 and zero otherwise. X_{it} is a set of controls including sector fixed effect, year fixed effect, and the lagged R&D to sales ratio. Following Ewens et al. (2024), we instrument the lagged dependent variable with the twice-lagged level of the dependent variable.

Table 3 presents the regression results. Columns (2) and (4) focus on firms with positive R&D expenditures, reflecting changes in R&D expenditure at the intensive margin. The results indicate that firms adjust their R&D investments around the threshold, likely to avoid regulation. For firms just below the threshold, regulation is associated with a decrease in R&D investment, particularly at the intensive margin.

5 Estimating Structural Parameters

In Section 3, we build a model to explain the observed bunching in our empirical setting. The SOX 404 affects firms' investments and bunching decisions through two economic channels: the increase in regulatory cost and the information benefit. In this section, we estimate the structural parameters of the model to decompose the impact from each channel.

5.1 Moments Selection

First, we assign values to production and demand parameters in our model. Following the literature (e.g., David et al., 2016), we set $\theta = 5$, which governs the elasticity of product substitution. We assume the production technology is characterized by constant returns to scale and set the production function parameters $\alpha_1 = 0.4$ (and $\alpha_2 = 0.6$) to match the average shares of labor (and capital) compensation in GDP during the sample period. The discount factor β is set to $\beta = 0.926$, implying an 8% annual discount rate. We construct our model as a partial equilibrium model where the cost of inputs is held constant over time. We assume the cost of capital equals the risk-free rate

at $r = 1/\beta - 1$. The firm optimization condition and the aggregate of labor and capital input imply the relationship between the capital-labor ratio and the input cost: $\frac{w}{r} = \frac{1-\alpha}{\alpha}\frac{K}{L}$. Hence, the wage rate is calibrated to match the average capital-labor ratio during the sample period.[20]

In our benchmark model, the public float is a linear function of expected profit in the next period, i.e., $PF_t = \lambda \mathbb{E}_{it}[\pi_{it+1}]$. With constant λ, the only force in the model that can affect the distribution of public float at time t is through adjusting a firm's R&D investment. Our model implies $var(\log(PF)) = \frac{4}{(2-2\tilde{\theta}\varepsilon)^2} var(\phi_{it})$ among non-bunching firms. However, public float can be affected by forces other than R&D investment, which are not modeled in our theoretical framework. In the estimation, we allow for channels that can affect public float unrelated to managers' R&D investment decisions. Specifically, we assume $PF = \bar{\lambda} e^{\hat{\lambda}} \mathbb{E}_{it}[\pi_{it+1}]$. $\bar{\lambda}$ is a constant number and computed using the model-implied relationship between public float and profits without regulation (See Appendix A.2 for details). That is, $e^{\hat{\lambda}}$ summarizes the other forces that affect firms' public float. We assume $\hat{\lambda}$ follows a normal distribution with mean 0 and variance σ^2_{λ}, which will be calibrated to match the variation of the public float.

The evolution of productivity is governed by ρ, σ^2, and ε. We calibrate ρ and σ^2 using the TFP for firms without any R&D investment during our sample period. Following David et al. (2016), the model assumes that both capital and labor are chosen under incomplete information. In line with David et al. (2016), we construct the log of the TFP level for each firm as $va_{it} - \left(1 - \frac{1}{\theta}\right) k_{it}$. Here, va_{it} represents the logarithm of value added, and k_{it} represents capital stock. Following Imrohoroglu and Tuzel (2014), we approximate the value added as revenue minus expenses on materials (i.e., total expenditure minus labor expenses). Capital stock is measured using Gross Property, Plant, and Equipment (PPEGT).

We restrict our sample to the initial years following the implementation of SOX 404, specifically from 2004 to 2007, and trim the 1% and 99% levels of public float, revenues, and PPEGT to limit the influence of outliers. To eliminate the common time-specific and industry-specific components that affect a firm's TFP, we regress TFP on time-industry fixed effect to obtain the residual

[20] Data on capital-labor ratio and the share of labor compensation in GDP are computed using Penn World Table 9.1. The capital-labor ratio is computed as capital stock in millions (chained value) per employee (adjusted by hours worked).

from the regression. We then estimate ρ by employing autoregressions implied by our model setup (1). σ^2 is estimated using the residuals of the regression. To avoid any influence on SOX 404 and investment bunching behavior, we estimate ρ and σ^2 on a pre-policy sub-sample (i.e., from the year 1994 to the year 2002) and of firms without any R&D investment during that period.

We make several adjustments to our benchmark model for the estimation. First, we add a fixed compliance cost f_c for firms with a public float above the threshold. As documented in SEC (2011) and Risk et al. (2002), the ratio of total compliance costs to total assets decreases with firms' total assets, implying a fixed component in firms' compliance costs. These costs may include one-time setup expenses for establishing internal control systems or recurring fixed costs like fees for independent auditors and maintenance of these systems. Such costs disproportionately burden smaller companies, incentivizing them to remain small.

However, based on our theoretical model, the inclusion of fixed compliance costs does not affect the firm's first-order condition for R&D investment, nor does it influence the firm's R&D investment choice outside the bunching region. It does, however, alter the firm's decision on bunching. Specifically, the inclusion of fixed compliance costs modifies the left-hand side of Equation (6) to $(1-\frac{\tilde{\theta}_\varepsilon}{2})\frac{PF^c}{PF} - f_c \times \frac{\lambda h(\sigma^2)}{\tilde{\theta}_\varepsilon \beta(1-\tau^c) g(\sigma^2,\sigma_\epsilon^2)\frac{PF}{PF}}$, reducing the benefits of not bunching and making the inequality more likely to be satisfied. Since f_c does not enter the firm's first-order condition for R&D investment, we calibrate it using the empirical estimates in Table 2. The estimates imply that for firms with public float above the threshold, their fixed audit fee increased by 54.7%. Given the minimum fixed audit fee before the regulation is \$0.045 million, we set $f_c = 0.547 \times 0.045 = 0.024$ million.[21]

Second, we assume that a positive fraction p_{nb} of firms will not bunch. Theoretically, firms with a public float just above the threshold and below that of marginal firms would find it optimal to bunch, leading to a "missing mass" in the public float distribution (refer to Figure 3 for illustration). However, potential frictions such as lack of awareness, switching costs, and other frictions may prevent these firms from bunching (Ewens et al., 2024). This results in a positive fraction of firms

[21] In our robustness check, we also set the fixed compliance cost, f_c, equal to the estimated increase multiplied by the median audit fee prior to the regulation, resulting in $f_c = 0.131$. Our structural estimates remain similar.

between the threshold and the public float of marginal firms, which fits the empirical distribution of the public floats. We calibrate p_{nb} using model implied moments described below.

The set of rest parameters $\Theta = \{\varepsilon, \tau^c, \sigma_e^2, \eta, \sigma_\lambda^2, p_{nb}\}$ is estimated using the Simulated Method of Moments. Specifically, we simulate initial productivity,[22] productivity shocks, and received signals over 20,000 firms and then search for Θ that minimizes the following distance function using model-implied moments from simulation and data-generated moments:

$$\hat{\Theta} = argmin_{\Theta} \sum_{k=1}^{6} \frac{||model(\Theta)_k - data_k||}{\frac{1}{2}||model(\Theta)_k|| + \frac{1}{2}||data_k||}$$

Previous studies (e.g., Dharmapala, 2022) examine the distribution of firms' public floats and the excess mass of firms below the threshold. However, these studies are insufficient to determine whether the excess density arises from high regulatory costs or reduced information precision. In our structural estimation, we use both the R&D-to-sales ratio and the share of firms around the threshold to quantify the impact of regulatory costs and the reduction in information precision. Our model shows that changes in the R&D-to-sales ratio are more sensitive to variations in information precision than to equivalent changes in regulatory costs. This sensitivity arises because more precise information allows firms to better align their capital and labor inputs with true productivity shocks, thereby influencing sales more directly than reductions in regulatory costs, which primarily affect sales indirectly through R&D investment. Hence, we use the following set of moments (M1-M6) to jointly identify the six parameters in our model. As we do not have closed-form solutions for most of these moments, we illustrate the sensitivity of each parameter to its respective selected moments in Appendix B.

M1: Average sales growth. The optimal R&D investment and productivity evolution imply that average sales growth is given by $\tilde{\theta}[(\rho - 1)\phi_{it-1} + \varepsilon \log(I^* + 1)]$, and it increases with ε.

M2: Average R&D investment-to-sales ratio for firms with a public float between \$45 million and \$75 million (excluding \$75 million). The optimal R&D investment for non-bunching firm is $I_{it}^* = \left[\tilde{\theta}\varepsilon\beta g(\sigma^2, \sigma_e^2)\Phi_{it+1}\right]^{\frac{1}{1-\tilde{\theta}\varepsilon}} - 1$, increasing with elasticity ε. Given ε, the average R&D

[22] The initial productivity is simulated to match the distribution of public float in the pre-policy period: i.e., 1994 to 2002.

investment-to-sales ratio for these non-bunching firms decreases with increasing information noise σ_e^2, as $g(\sigma^2, \sigma_e^2)$ monotonically decreases with σ_e^2.

M3: Average R&D investment-to-sales ratio for firms with a public float between \$75 million and \$105 million. For firms above this threshold, the optimal investment, $I_{it}^{**} = [\tilde{\theta}\varepsilon(1-\tau^c)\beta g(\sigma^2, (1-\eta)\sigma_e^2)\Phi_{it+1}]^{\frac{1}{1-\tilde{\theta}\varepsilon}} - 1$, and the R&D to sales ratio decreases with the regulatory cost τ^c and increases with the information benefit η.

M4: Fraction of firms with public float between \$70 million and \$75 million (excluding \$75 million). Equation (6) indicates that a higher τ^c or a lower η results in more firms bunching. Thus, the share of firms in the [70,75) range increases with regulatory cost τ^c and decreases with the information benefit η.

M5: Fraction of firms with public float between \$75 million and \$80 million (excluding \$80 million). If all firms engage in bunching, the distribution of the firm's public float would exhibit a "missing mass" just above the threshold. However, if only a fraction of firms engage in bunching while others do not, we would still observe some firms located just above the threshold. A higher fraction of non-bunching firms p_{nb} increases the share of firms in the [75,80) range.

M6: Standard deviations of public float. The variation in a firm's public float can be approximated as $var(\log(PF)) = \frac{4}{(2-2\tilde{\theta}\varepsilon)^2} var(\phi_{it}) + \sigma_\lambda^2$. Given estimates of ε, σ_λ^2 directly increase the standard deviation of public float.

The moment conditions M1 through M5 jointly pin down the parameters ε, σ_e, τ^c, η, and p_{nb}. We use moment condition M6 to back out the parameter σ_λ. To evaluate our model's fit to the data, we compute several non-targeted moments and compare the simulated moments to those observed in our data sample. Specifically, we consider the share of firms in equally divided bins below and above the threshold. We calculate the share of firms in two additional intervals below the threshold: $[60, 65)$ and $[65, 70)$; and in four intervals above the threshold: $[80, 85)$, $[85, 90)$, $[90, 95)$, and $[95, 100)$. The value of ε could also affect aggregate TFP growth. Hence, we compare the simulated and data-implied TFP growth rates.

5.2 Estimation Results and Implication

Panel A of Table 4 summarizes the parameters calibrated externally. Panel B of Table 4 presents the estimates and standard errors of our structural parameters $\Theta = \{\varepsilon, \tau^c, \sigma_e, \eta, \sigma_\lambda^2, p_{nb}\}$. Our estimates for ρ and σ are consistent with those reported by David et al. (2016). The estimated value of $\varepsilon = 0.032$ falls within the range in previous empirical studies on R&D elasticities, which varies between 0.02 and 0.17 (Hall et al., 2010). The estimate suggests that a 1% increase in R&D investment could raise the firm's TFP level by approximately 3.2 percent. The fraction of non-bunching firms, p_{nb}, is estimated at around 80%, slightly higher than Ewens et al. (2024)'s 77%. σ_λ is estimated at around 0.722. With a public float standard deviation of 1.81, the estimates suggest that our model explains 84% of the public float variation.

Our estimates of the regulatory cost τ^c suggest that, in the absence of an information benefit channel, compliance with SOX 404 is equivalent to imposing a 10.6% implicit corporate tax rate on firms. This increase in the "implicit" tax rate discourages R&D investment by reducing its marginal returns. The estimates for the standard deviation σ_e indicate that the posterior variance of the signal is $\mathcal{V} = \left(\frac{1}{\sigma_e^2} + \frac{1}{\sigma^2}\right)^{-1} = 0.052$.

The information benefit estimate η is 0.259, which corresponds to a 20.4% reduction in signal noise. This suggests that compliance with SOX 404 could enhance information precision and reduce the "implicit" tax rate firms need to pay. In the absence of regulatory costs, the information benefit can be calculated as $\frac{g(\sigma^2,\sigma_e^2(1-\eta))}{g(\sigma^2,\sigma_e^2)} - 1 = 7.06\%$. In other words, a 20.4% increase in information precision could result in a 7.06% increase in expected profits, which could then lead to increased R&D investment. The information benefit also reduces the "implicit" tax rate associated with SOX 404 compliance. The total "implicit" tax rate can be computed as $1 - (1 - \tau^c)\frac{g(\sigma^2,\sigma_e^2(1-\eta))}{g(\sigma^2,\sigma_e^2)} = 4.30\%$. Overall, our estimates show that compliance with SOX 404 is equivalent to a 4.30% increase in the corporate tax rate.

Panel A of Table 5 demonstrates that with the estimated parameters, the model-implied moments closely align with the data-generated moments, albeit slightly underestimating the standard deviation of public float. In Panel B, we present a set of non-targeted moments. The estimation results reveal that the model's predicted distribution of public float matches the empirical distribu-

tion reasonably well within the range of $60 million to $100 million. However, our model slightly overestimates average productivity growth. Additionally, it slightly overestimates the proportion of firms in the public float range of $80 million to $90 million and slightly underestimates those in the range of $95 million to $100 million.

Figure 9 plots the distribution of simulated R&D investment and public float for firms with public float between $50 million and $150 million, both with and without threshold regulation. The R&D investment distribution shifts slightly left due to reduced investment following the regulation. The public float distribution exhibits an excess density around the $75 million threshold. Our simulations suggest that firms with an optimal public float below $82.3 million in the absence of regulation would bunch below the threshold post-regulation.[23] This bunching behavior reduces the mass in the public float distribution. Figure 10 illustrates the relationship between R&D investment and initial public float, assuming no exogenous changes in public float ($\sigma_\lambda^2 = 0$) and 0% non-bunching firms. In the bunching region, R&D investment decreases with initial public float, consistent with the model prediction.

Column (1) of Table 6 shows that 1.11% of firms are in the bunching region under the current regulation threshold, with around 20% of these firms (0.21% of all firms in the entire economy) choosing to bunch below the threshold. Bunching firms in the bunching region reduce R&D investment by an average of 61.38%, resulting in a 1.15% decrease in TFP level and a 1.1 percentage point drop in the TFP growth rate. This reduction in R&D is due to firms' avoidance behavior. In our estimation, only 20% of firms in the bunching region bunch to avoid the regulation. Given the estimates, the total reduction of R&D investment in the bunching region is 12.27% (i.e., $20\% \times 61.38\%$) due to firm avoidance.

The second row of Table 6 shows the average effect on compliance firms with public float above the threshold. With a relatively low threshold, 79.31% of firms are above it. For these firms, R&D investment decreases by 4.01% on average, leading to a 0.07% drop in TFP level and a 0.07

[23]The marginal firm's public float estimate is slightly lower than the $94.5 million reported by Ewens et al. (2024) and Alvero and Xiao (2020). Figure 8 shows the counterfactual distribution of public float from 2004 to 2007, within the range of $50 million to $150 million. The counterfactual density is estimated using the model and methods outlined in Sections 3 and 5, and is similar to the counterfactual density distribution obtained using the method of Ewens et al. (2024).

percentage point decrease in TFP growth. As most firms in the entire economy above the threshold are also above the marginal firm point, they comply with the regulation. Consequently, the reduction in aggregate R&D investment and productivity is primarily due to regulatory compliance.

The decrease in R&D investment also leads firms to adjust their capital stock. On average, bunching firms reduce capital by 4.45%, while non-bunching firms above the threshold reduce it by 1.13% due to increased compliance costs. According to our theoretical model, the optimal capital stock for firms is given by: $\ln k_{it} = \ln \Phi + \left(1 - \frac{1}{\theta}\right) \rho \phi_{it} + \varepsilon \ln(1 + I_{it-1}) + \frac{1}{2}\left(1 - \frac{1}{\theta}\right)^2 \mathcal{V} + \left(1 - \frac{1}{\theta}\right) \frac{\mathcal{V}}{(1-\eta)\sigma_e^2}) s_{it}$. The average changes in capital stock can be expressed as the sum of changes in R&D investment and the changes in the firm's expectation of future productivity shocks after receiving the signal s_{it}:

$$\Delta \ln k_{it} = \varepsilon \Delta \ln(1 + I_{it-1}) + \frac{\eta \sigma_e^2 \sigma^2}{[\sigma_e^2(1-\eta) + \sigma^2][\sigma_e^2 + \sigma^2]} \left[s_{it} - \frac{\sigma^2}{2}(1 - \frac{1}{\theta}) \right]$$

Threshold regulation could affect firms' capital stock in two opposite ways. First, a reduction in R&D investment can directly decrease current productivity and marginal returns on investment. The first term in the above equation indicates that a 1% decrease in R&D investment could result in a $\varepsilon = 0.032\%$ drop in capital stock. However, compliance increases a firm's expected profits by enhancing the informativeness of the signals received. The second term on the above equation reflects changes in the firm's expectations regarding future productivity shocks. For firms receiving sufficiently positive signals, i.e., $s_{it} > \frac{\sigma^2}{2}(1 - \frac{1}{\theta})$, compliance reduces uncertainty about future productivity shocks and incentivizes increased capital stock.

By estimating the regulatory costs and informational benefits of SOX through firms' bunching behavior, we aggregate the long-term impact of SOX 404 on the entire economy. Column (1) of Table 7 presents the annual changes in aggregate R&D investment, capital stock, aggregate TFP, and aggregate TFP growth. We find that implementing SOX 404 could reduce the aggregate R&D investment by 2.81%, resulting in a 0.61% loss in aggregate TFP and a 0.06 percentage point decrease in annual TFP growth, as well as a 0.29% decrease in capital stock. Given that the literature estimates aggregate annual TFP growth at around 1.7% to 2.02% (e.g., Acemoglu et al.,

2018; Cavenaile and Roldan-Blanco, 2021; De Ridder, 2024), our results suggest that the reduction in aggregate R&D investments due to SOX 404 is equivalent to a 3% decrease in documented productivity growth.

Columns (2) and (3) of Table 7 decompose the aggregate changes into two channels: regulatory costs and information benefits. The reductions in R&D investments, capital stock, TFP, and TFP growth would be more economically significant if compliance did not provide any information benefits. Our results indicate that the losses in investment and TFP due to regulatory costs outweigh the gains from improved information precision.

6 Counterfactual Analyses

In the final section, we conduct counterfactual analyses to inform policy debates about increasing public float thresholds in financial reporting regulations. We first use our estimated model to simulate the effect of alternative thresholds and quantify their implications on R&D investment and productivity. We then discuss the situation under which the gain from increasing information precision outweighs the loss from regulatory costs.

6.1 Different Regulatory Thresholds

We first analyze alternative regulation thresholds. Table 8 presents the effects of changing the regulation threshold to $100 million, $250 million, and $700 million, as well as having no regulatory thresholds ("mandatory compliance").

The threshold of a $100 million public float is slightly higher than the existing threshold (i.e., a $75 million public float). The $250 million threshold is based on the criterion for becoming a "smaller reporting company." We evaluate the economic impact of aligning the threshold for "accelerated filer" status with that of a "smaller reporting company." Next, we assess the economic impact under the scenario where the only threshold is for "large accelerated filers" with a threshold of $700 million public float, which applies to significantly larger firms. We finally examine the case of having no regulatory threshold, with all firms complying with the regulation. Much of the previous regulatory debate has focused on exempting smaller companies from compliance. Our results indicate that such a threshold may create incentives for firms to avoid compliance, with potential long-term implications. Therefore, we explore a scenario where firms have no option for

exemption, assessing the overall balance between regulatory costs and economic losses.

Tables 8 and 9 present the counterfactual results with different regulatory thresholds. Our results reveal three interesting findings. First, the gap between the public float threshold and the marginal firm point widens with higher regulatory thresholds (see column (1) of Table 9). However, the proportion of bunching firms does not monotonically increase with the regulatory threshold value. We observe more bunching firms under alternative thresholds. The share of bunching firms depends on both the size of the bunching region (i.e., the difference between the threshold and marginal firms' public float) and the proportion of firms within that region. For instance, if more firms have a public float between \$100 million and \$109 million than between \$75 million and \$82 million in the absence of regulation, setting the threshold at \$100 million could lead to more bunching compared to a threshold of \$75 million. In the case of mandatory compliance, no firms have an incentive to bunch.

Second, we find that the avoidance-induced reduction in R&D investment and productivity losses among bunching firms decreases significantly with a higher regulatory threshold. Figure 11 plots the density distribution of public float under different regulatory thresholds. Third, we find that with a higher regulatory threshold, the share of firms subject to compliance costs decreases, as does the average compliance-induced reduction in their R&D investment and productivity losses.

In general, the reduction in R&D investment and TFP decreases as the regulatory threshold increases. Column (1) in Table 9 presents the results. For example, compared to the benchmark case, increasing the threshold to \$700 million would reduce aggregate R&D investment losses by 0.55 percentage points and TFP losses by 0.2 percentage points. The relatively small "implicit" tax rate (i.e., due to the low fixed compliance cost to sales ratio) via regulation results in a minor change in R&D investment and productivity for firms above the threshold. In addition, as thresholds increase, bunching firms are less negatively affected. Given their magnitude, the aggregate impact is primarily driven by compliance costs rather than firms' avoidance behavior. Mandatory compliance results in higher R&D investment reduction and more productivity losses than the current regulation threshold.

To sum up, our counterfactual analyses with different regulatory thresholds indicate that

mandating all firms to comply with regulations increases compliance costs for all firms. Raising regulatory thresholds may decrease R&D reduction induced by avoidance, suggesting that the benefits of growth may outweigh avoidance costs for firms. Additionally, the overall reduction in investment remains relatively consistent across different thresholds, possibly due to the increasing information benefits associated with higher compliance at these thresholds.

6.2 Higher Information Benefits from Regulatory Compliance

In section 3.3, we briefly discuss that if the information benefit outweighs the increased regulatory costs, firms with public float just below the threshold will find it optimal to bunch above the threshold. The information benefit due to compliance could encourage firms to make higher R&D investments and thus increase productivity. The condition for net positive compliance benefit is given by: $(1-\tau^c)g(\sigma^2,(1-\eta)\sigma_e^2) > (1-\tau)g(\sigma^2,\sigma_e^2)$, creating a scenario of voluntary compliance. Given our estimated parameters, we can find a minimum value of $\eta = 0.404$ that makes the inequality hold. Given these parameter restrictions, we will now use our estimated model to evaluate the aggregate effect when the information benefit exceeds the regulatory cost.

Let $\eta' = 0.8$; that is, compliance will reduce the variance of the noise component by 80 percent. We choose 80 percent for illustrative purposes. The "implicit" tax rate under the new η would be: $1-(1-\tau^c)\frac{g(\sigma^2,\sigma_e^2(1-2\eta))}{g(\sigma^2,\sigma_e^2)} = -14.92\%$. In this situation, compliance with SOX 404 is equivalent to a subsidy rate of 14.92%. The optimal investment for firms that bunch above the threshold could be written as $I_{it}^b = \left(\frac{\overline{PF}}{\lambda h(\sigma^2)\Phi_{it+1}}\right)^{\frac{1}{\tilde{\theta}\varepsilon}} - 1$. The firm would find it optimal to bunch if the following inequality is satisfied:

$$-\left(\frac{PF^*}{\overline{PF}}\right)^{1-\frac{1}{\tilde{\theta}\varepsilon}}\tilde{\theta}\varepsilon + \frac{1-\tau^c}{1-\tau}\frac{g(\sigma^2,(1-\eta')\sigma_e^2)}{g(\sigma^2,\sigma_e^2)} \geq (1-\tilde{\theta}\varepsilon)\frac{PF^*}{\overline{PF}} \tag{8}$$

The higher the value $(1-\tau^c)g(\sigma^2,\sigma_e^2(1-\eta'))$, the more likely a firm would find it optimal to bunch above the threshold. We simulate the model under new η. Figures 12 and 13 depict the simulated results. We find that firms with an initial public float between \$45 million and \$50 million overinvest in R&D to bunch their next period public float to be above the regulatory threshold. As a result, there is a reduced mass in the public float distribution from \$70.16 million (i.e., the counterfactual public float of marginal firms) to \$74.99 million and an excess mass around the \$75 million threshold.

Table 10 presents the aggregate impact of compliance when the information benefit is larger than the regulatory cost. Bunching firms in the bunching region increase their R&D investment by 316.92% on average, leading to an increase in their annual TFP growth rate of around 1.38 percentage points. At the aggregate level, R&D investment increases by 9.22%, while annual productivity growth rises by 0.22 percentage points.

The results suggest that the overall aggregate effect will turn positive when the information benefit outweighs the regulatory cost. Another interpretation of the result is that if SOX aims to address the underprovision of information disclosures, the regulation's impact can be measured by allowing firms to internalize these information benefits. This result implies that the information benefit channel in financial reporting regulations should not be overlooked in future studies assessing the effectiveness of such regulations.

7 Conclusion

We study firms' avoidance behavior by reducing R&D investment in response to a significant financial reporting threshold regulation and estimate its impact on productivity. We extend an investment model by incorporating both information learning benefits and regulatory costs, and then conduct structural estimation. Our findings indicate that firms' choices to stay below the threshold represent an equilibrium outcome. The estimates show significant reductions in R&D investment and productivity, primarily due to the reduced marginal returns on investment arising from regulatory costs. The distortions observed in the aggregate are due to firm avoidance and high overall net regulatory costs, with the latter likely being the primary reason for the distortions. In addition, information benefits could significantly mitigate the distortions caused by regulatory costs.

We further conduct counterfactual analyses to inform policy designs. First, the counterfactual analyses reveal that the benefits of raising the regulatory threshold arise from two channels. Raising the threshold significantly mitigates the decrease in R&D investment caused by the avoidance behavior of firms bunching below the threshold. Additionally, increasing the threshold substantially reduces the number of firms required to comply with regulations, thereby decreasing overall R&D declines due to the net regulatory cost. Together, these changes resulting from raising the

regulatory threshold mitigate overall R&D declines caused by regulatory distortions and reduce aggregate productivity losses. Moreover, in a counterfactual scenario where the informational benefits of compliance outweigh the regulatory costs, we observe significant improvements in investments and productivity. Therefore, our study provides insights for regulators on designing policies to mitigate regulatory distortions by (1) further raising the public float threshold, or (2) enhancing the informational benefits associated with regulatory compliance.

Our study is subject to several limitations. First, while the model provides a transparent underlying mechanism through which SOX 404 could affect firms' R&D investment and productivity growth, there is a risk of estimation bias if the model is misspecified. Additionally, while our estimation accounts for other potential forces that could influence public float management, we do not explicitly incorporate them into the theoretical model. Given these considerations, our findings should be regarded as a benchmark.

Despite these limitations, our study sheds light on the economic implications of firm avoidance behavior around thresholds and estimates the costs and benefits of a key threshold regulation in financial reporting. Our framework applies not only to SOX but also to a wide range of threshold-based financial reporting regulations. These regulations often impose compliance costs while providing informational benefits. The increasing prevalence of threshold-based regulations underscores a significant trend in regulatory practices, making our framework particularly relevant to the field of regulatory economics.

References

Acemoglu, D., U. Akcigit, H. Alp, N. Bloom, and W. Kerr (2018). Innovation, reallocation, and growth. *American Economic Review 108*(11): 3450–3491.

Albuquerque, A. and J. L. Zhu (2019). Has section 404 of the sarbanes–oxley act discouraged corporate investment? new evidence from a natural experiment. *Management Science 65*(7): 3423–3446.

Alexander, C. R., S. W. Bauguess, G. Bernile, Y.-H. A. Lee, and J. Marietta-Westberg (2013). Economic effects of sox section 404 compliance: A corporate insider perspective. *Journal of Accounting and Economics 56*(2): 267–290.

Alvero, A. and K. Xiao (2020). Fuzzy bunching. *Available at SSRN 3611447*.

Amirapu, A. and M. Gechter (2020). Labor regulations and the cost of corruption: Evidence from the indian firm size distribution. *Review of Economics and Statistics 102*(1): 34–48.

Aw, B. Y., M. J. Roberts, and D. Y. Xu (2008). R&d investments, exporting, and the evolution of firm productivity. *The American Economic Review 98*(2): 451–456.

Bai, J., T. Philippon, and A. Savov (2016). Have financial markets become more informative? *Journal of Financial Economics 122*(3): 625–654.

Ballew, H., M. Iselin, and A. Nicoletti (2022). Accounting-based thresholds and growth decisions in the banking industry. *Review of Accounting Studies 27*(1): 232–274.

Bargeron, L. L., K. M. Lehn, and C. J. Zutter (2010). Sarbanes-oxley and corporate risk-taking. *Journal of Accounting and Economics 49*(1): 34–52.

Berger, P. G., F. Li, L. Y. Liu, and M. Wong (2023). Within-firm information frictions and the role of managerial reporting quality in investment efficiency. *Within-Firm Information Frictions and the Role of Managerial Reporting Quality in Investment Efficiency (October 11, 2023)*.

Bernard, D., D. Burgstahler, and D. Kaya (2018). Size management by european private firms to minimize proprietary costs of disclosure. *Journal of Accounting and Economics 66*(1): 94–122.

Bloom, N., B. Eifert, A. Mahajan, D. McKenzie, and J. Roberts (2013). Does management matter? evidence from india. *The Quarterly Journal of Economics 128*(1): 1–51.

Bloom, N., M. Schankerman, and J. Van Reenen (2013). Identifying technology spillovers and product market rivalry. *Econometrica 81*(4): 1347–1393.

Breuer, M. (2021). How does financial-reporting regulation affect industry-wide resource allocation? *Journal of Accounting Research 59*(1): 59–110.

Breuer, M. and D. Windisch (2019). Investment dynamics and earnings-return properties: A structural approach. *Journal of Accounting Research 57*(3): 639–674.

Cavenaile, L. and P. Roldan-Blanco (2021). Advertising, innovation, and economic growth. *American Economic Journal: Macroeconomics 13*(3): 251–303.

Chan, L. K., J. Lakonishok, and T. Sougiannis (2001). The stock market valuation of research and development expenditures. *The Journal of finance 56*(6): 2431–2456.

Chen, Z., Z. Liu, J. C. Suárez Serrato, and D. Y. Xu (2021). Notching r&d investment with corporate income tax cuts in china. *American Economic Review 111*(7): 2065–2100.

Choi, J. H. (2021). Accrual accounting and resource allocation: A general equilibrium analysis. *Journal of Accounting Research 59*(4): 1179–1219.

Cohen, D., A. Dey, and T. Z. Lys (2009). The sarbanes-oxley act of 2002: implications for compensation contracts and managerial risk-taking. Working paper, New York University, University of Chicago and Northwestern University.

David, J. M., H. A. Hopenhayn, and V. Venkateswaran (2016). Information, misallocation, and aggregate productivity. *The Quarterly Journal of Economics 131*(2): 943–1005.

De Ridder, M. (2024). Market power and innovation in the intangible economy. *American Economic Review 114*(1): 199–251.

Dharmapala, D. (2022). Estimating firms' responses to securities regulation using a bunching approach. *American Law And Economics Review 24*(2): 449–494.

Diamond, R. and P. Persson (2016). The long-term consequences of teacher discretion in grading of high-stakes tests. National Bureau of Economic Research.

Engel, E., R. M. Hayes, and X. Wang (2007). The sarbanes-oxley act and firms' going-private decisions. *Journal of Accounting and Economics 44*(1): 116–145.

Ewens, M., K. Xiao, and T. Xu (2024). Regulatory costs of being public: Evidence from bunching estimation. *Journal of Financial Economics 153*: 103775.

Feng, M., C. Li, and S. McVay (2009). Internal control and management guidance. *Journal of Accounting and Economics 48*(2): 190–209.

Gao, F., J. S. Wu, and J. L. Zimmerman (2009). Unintended consequences of granting small firms exemptions from securities regulation: Evidence from the sarbanes-oxley act. *Journal of Accounting Research 47*(2): 459–506.

Garicano, L., C. Lelarge, and J. Van Reenen (2016). Firm size distortions and the productivity distribution: Evidence from france. *American Economic Review 106*(11): 3439–3479.

Ge, W., A. Koester, and S. McVay (2017). Benefits and costs of sarbanes-oxley section 404(b) exemption: Evidence from small firms' internal control disclosures. *Journal of Accounting and Economics 63*(2): 358–384.

Glaeser, S. and M. Lang (2024). Measuring innovation and navigating its unique information issues: A review of the accounting literature on innovation. *Journal of Accounting and Economics*: 101720.

Guner, N., G. Ventura, and Y. D. Xu (2008). Macroeconomic implications of size-dependent policies. *Review of Economic Dynamics 11*(4): 721–744.

Hall, B. H., J. Mairesse, and P. Mohnen (2010). Measuring the returns to r&d. In *Handbook of the Economics of Innovation*, Volume 2, pp. 1033–1082. Elsevier.

Hall, B. H. and R. Oriani (2006). Does the market value r&d investment by european firms? evidence from a panel of manufacturing firms in france, germany, and italy. *International Journal of Industrial Organization 24*(5): 971–993.

Hirschey, M. and J. J. Weygandt (1985). Amortization policy for advertising and research and development expenditures. *Journal of accounting Research*: 326–335.

Iliev, P. (2010). The effect of sox section 404: Costs, earnings quality, and stock prices. *The Journal of Finance 65*(3): 1163–1196.

Imrohoroglu, A. and S. Tuzel (2014). Firm-level productivity, risk, and return. *Management Science 60*(8): 2073–2090.

Klette, T. J. and S. Kortum (2004). Innovating firms and aggregate innovation. *Journal of political economy 112*(5): 986–1018.

Koh, P.-S. and D. M. Reeb (2015). Missing r&d. *Journal of accounting and economics 60*(1): 73–94.

Leuz, C., A. Triantis, and T. Y. Wang (2008). Why do firms go dark? causes and economic consequences of voluntary sec deregistrations. *Journal of Accounting and Economics 45*(2): 181–208.

Leuz, C. and P. D. Wysocki (2016). The economics of disclosure and financial reporting regulation: Evidence and suggestions for future research. *Journal of Accounting Research 54*(2): 525–622.

Lev, B. and T. Sougiannis (1996). The capitalization, amortization, and value-relevance of r&d. *Journal of accounting and economics 21*(1): 107–138.

McClure, C. G. (2023). How costly is tax avoidance? evidence from structural estimation. *Accounting Review 98*(6).

McClure, C. G. and A. A. Zakolyukina (2024). Non-gaap reporting and investment. *Accounting Review 99*(2).

Nondorf, M. E., Z. Singer, and H. You (2012). A study of firms surrounding the threshold of sarbanes-oxley section 404 compliance. *Advances in Accounting 28*(1): 96–110.

Reuters (2024). Britain Plans to Simplify Reporting Rules for Small Businesses.

Risk, S. et al. (2002). Study of the sarbanes-oxley act of 2002 section 404 internal control over financial reporting requirements. *Study of the Sarbanes-Oxley Act of*.

Romer, P. M. (1990). Endogenous technological change. *Journal of political Economy 98*(5, Part 2): S71–S102.

SEC (2011). Study and recommendations on section 404 (b) of the sarbanes-oxley act of 2002 for issuers with public float between 75and 250 million.

SEC (2018). Amendments to the Definition of Smaller Reporting Companies.

Wagner, S. and L. Dittmar (2006). The unexpected benefits of sarbanes-oxley. *Harvard Business Review 84*(4): 133.

Figures

Figure 1: Example: Density Distribution with Cutoff

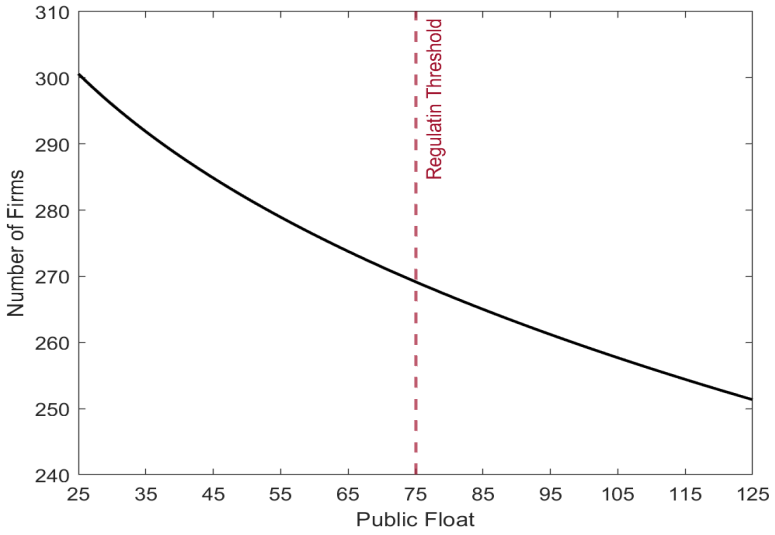


Figure 2: Example: Density Distribution with Cutoff and Marginal Firm

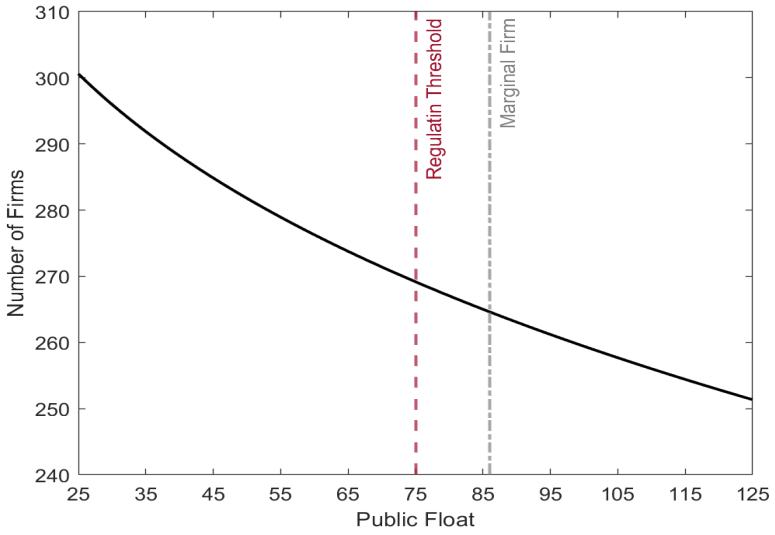


Figure 3: Example: Density Distribution with Cutoff and Marginal Firm & Bunching

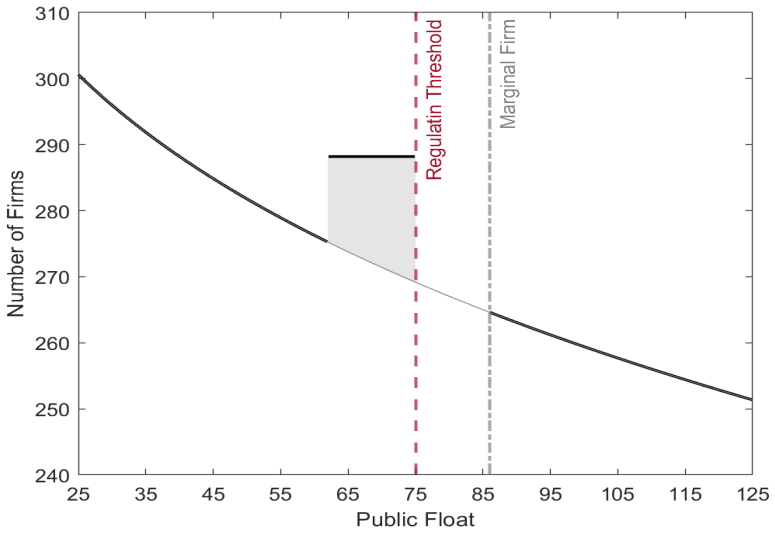
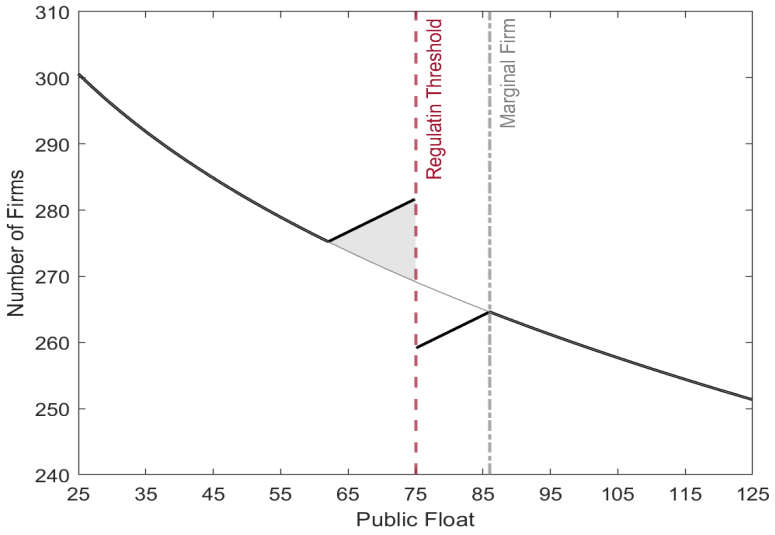
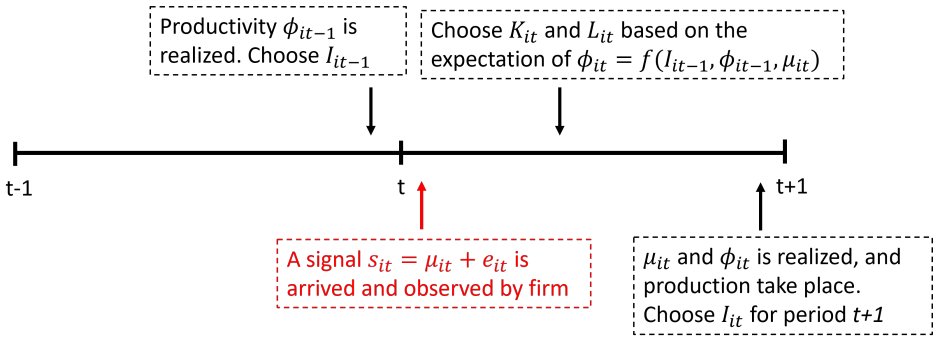


Figure 4: Example: Density Distribution with Cutoff and Marginal Firm & Bunching due to Frictions and Heterogeneity



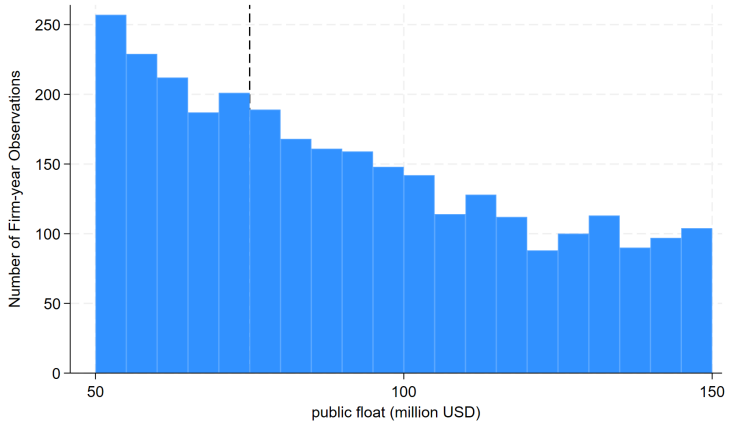
Notes: Figures 1 - 4 serve as an illustrative example to demonstrate the model intuition and empirical predictions.

Figure 5: Timeline for Investment, Signal, and Production



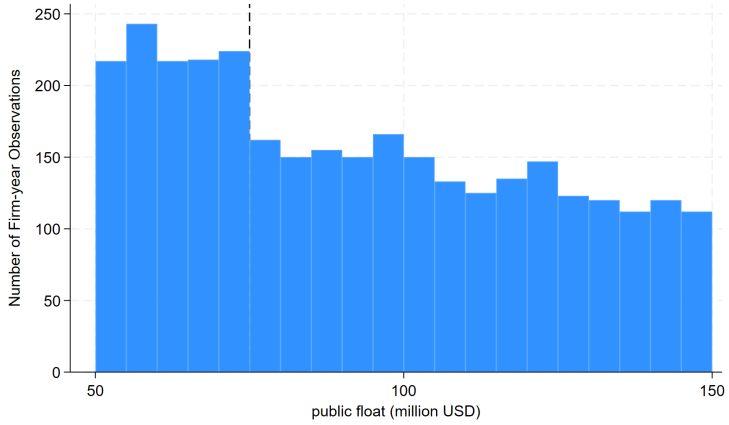
Notes: This figure shows the timeline of investment and production with the incomplete information.

Figure 6: Distribution of Public Float (2000-2003)



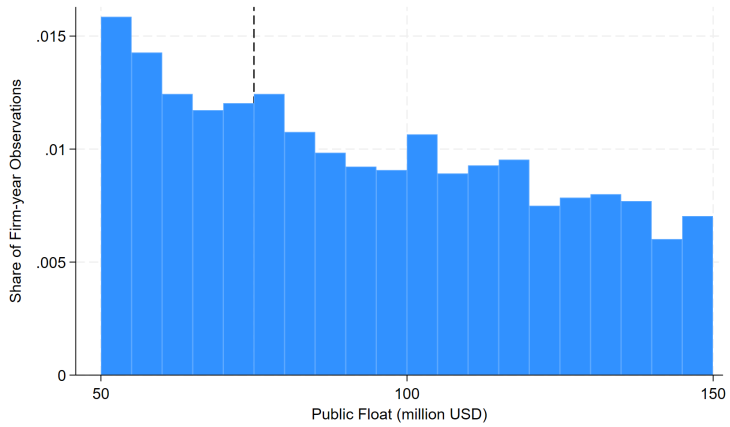
Notes: This figure represents the number of firm-year observations of the public float during 2000-2003 and within the range of $50 million to $150 million. The vertical dash line represents a public float of $75 million. The bin width is $5 million.

Figure 7: Distribution of Public Float (2004-2007)



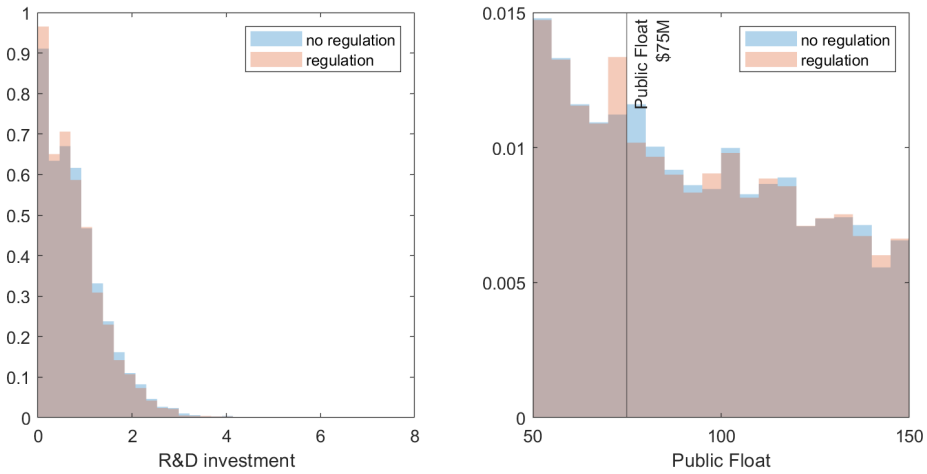
Notes: This figure represents the number of firm-year observations of the public float during 2004-2007 and within the range of $50 million to $150 million. The vertical dash line represents a public float of $75 million. The bin width is $5 million.

Figure 8: Counterfactual Distribution of Public Float (2004-2007)



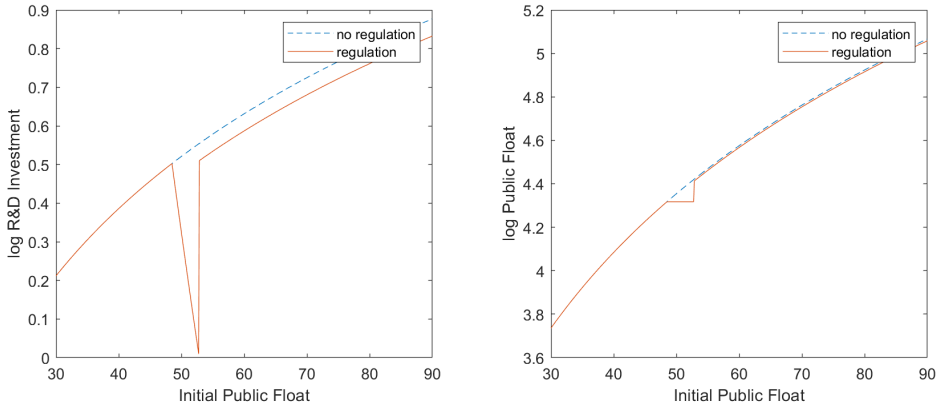
Notes: This figure represents the counterfactual distribution of public float in 2004-2007 and within the range of $50 million to $150 million. The vertical dash line represents a public float of $75 million. The bin width is $5 million. The counterfactual density is estimated using the model and estimation outlined in sections 3 and 5.

Figure 9: Distribution of Simulated Investment and Public Float



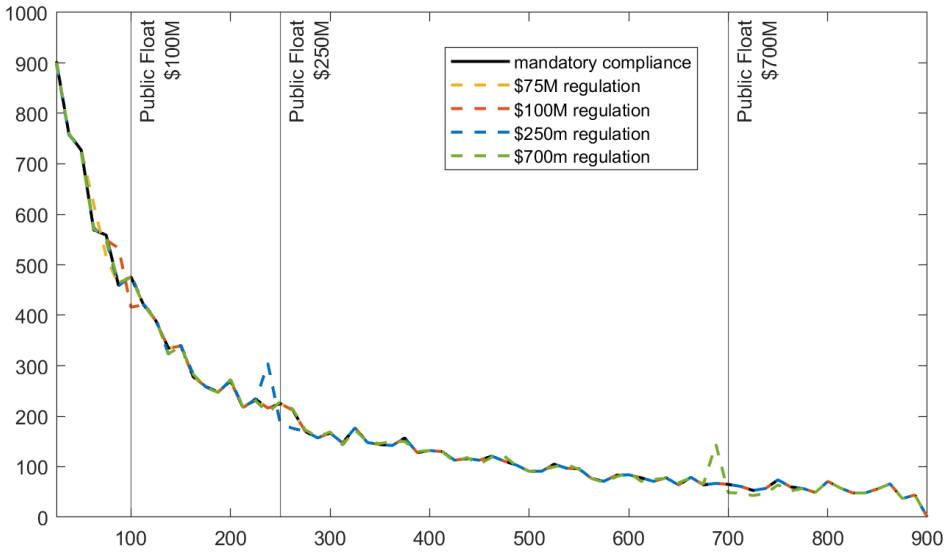
Notes: This figure represents the distribution of log R&D investment (left panel) and public float (right panel) simulated from the estimation model. The light rose bar is the estimated distribution of our benchmark model. The light blue bar is the counterfactual distribution without regulation. The right panel shows there's an excess mass on the threshold of $75 million.

Figure 10: Initial Productivity and Optimal Investment/Public Float



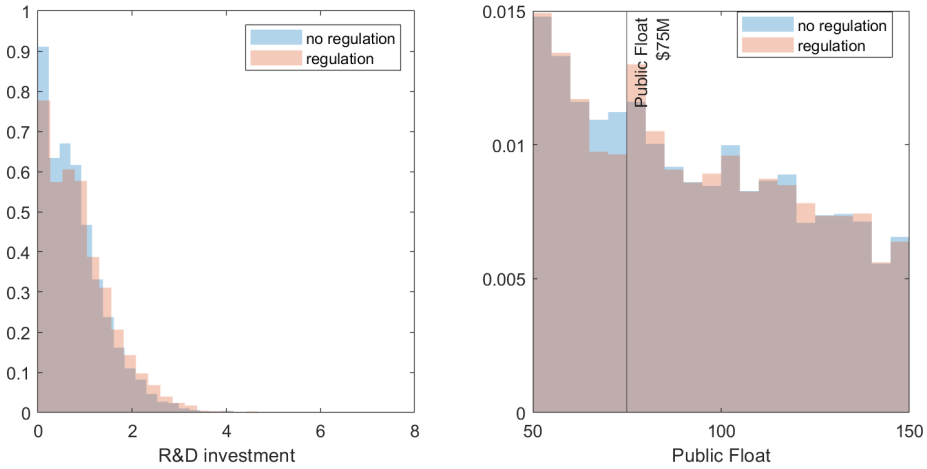
Notes: This figure compares the firm's optimal R&D investment (left panel) and public float (right panel) under the benchmark regulation threshold to the R&D investment and public float under the situation without regulation. For illustrative purposes, we set the exogenous variation of public float $\sigma_\lambda = 0$, and a fraction of non-bunching firms $p_{nb} = 0$. We plot the R&D investment and public float against the firm's prior public float. The R&D investment and public float increase with the firm's initial public float level despite a bunching region. Our estimated model shows that the bunching region contains firms with initial public float between $50 million to $55 million. For firms in the bunching region, their R&D investment decreases with initial public float to keep their public float under the regulatory threshold.

Figure 11: Distribution of Simulated Public Float under Different Regulatory Thresholds



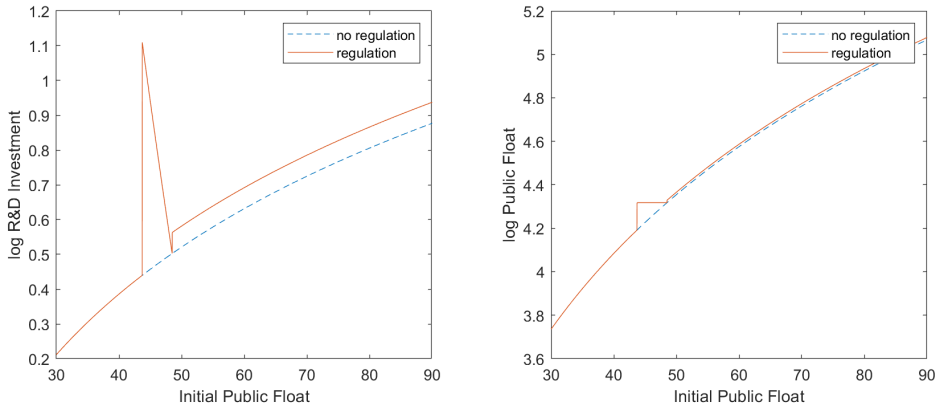
Notes: This figure plots the simulated density of Public Float distribution under different regulatory thresholds.

Figure 12: Distribution of Simulated Investment and Public Float



Notes: This figure represents the distribution of log R&D investment (left panel) and public float (right panel) simulated from the estimation model. The light rose bar is the estimated distribution of the model where the information benefit outweighs the regulatory costs. The light blue bar is the counterfactual distribution without regulation. The right panel shows there's an excess mass on the threshold of $75 million.

Figure 13: Initial Productivity and Optimal Investment/Public Float



Notes: This figure compares the firm's optimal R&D investment (left panel) and public float (right panel) under a situation without regulation to a situation where the information benefit outweighs the regulatory costs. For illustrative purposes, we set the exogenous variation of public float $\sigma_\lambda = 0$, and a fraction of non-bunching firms $p_{nb} = 0$. We plot the R&D investment and public float against the firm's prior public float. The R&D investment and public float increase with the firm's initial public float level despite a bunching region. Our estimated model shows that the bunching region contains firms with initial public float between $45 million to $50 million. For firms in the bunching region, their R&D investment decreases with initial public float to make their public float just above the regulatory threshold.

Tables

Table 1: Descriptive Statistics (year 2004 - year 2007)

	Mean	P25	P50	P75	Std	No. Obs
Audit Fee	2.489	0.543	1.232	2.656	4.426	4,442
Public Float	2763.315	120.960	448.560	1600.000	7858.825	4,442
Capital Stock	1057.578	28.246	103.679	486.942	4839.647	4,442
R&D Expenditure	107.182	4.522	17.712	55.434	428.351	4,442
Sales	2194.488	104.438	327.721	1307.339	8514.832	4,442
Sales Growth	0.093	0.004	0.082	0.178	0.184	4,442
TFP Level	-0.106	-0.552	-0.027	0.453	0.883	4,442
TFP Growth	0.029	-0.118	0.023	0.188	0.406	4,442

Notes: This table shows summary statistics of key variables used in our structural estimation. Our sample focuses only on firms with positive R&D investment and Public Float during the post-SOX 404 period (i.e., 2004-2007). Audit fees, R&D expenditure, and sales are in millions of USD deflated by the GDP deflator. Capital is the gross value of property, plant, and equipment in millions of USD, deflated by a nonresidential price deflator. TFP is computed using the method suggested by David et al. (2016).

Table 2: Auditing Fee around Threshold

	(1)	(2)	(3)	(4)
Above 75× Post	0.512***	0.547***		
	(0.029)	(0.033)		
Above 75	-0.102***	-0.128***		
	(0.024)	(0.040)		
Public Float (log)	0.049***	0.053***	0.040***	0.040***
	(0.008)	(0.014)	(0.006)	(0.011)
Public Float (log) × Post			-0.106***	-0.112***
			(0.012)	(0.017)
Public Float (log)×Above 75 × Post			0.114***	0.121***
			(0.008)	(0.009)
No. Observations	13,824	7,045	13,824	7,045
R-Squared	0.914	0.914	0.913	0.912
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes

Notes: This table examines changes in firms' auditing fee from 2000 to 2007. The dependent variable is the log of the firm's total auditing fee (including audit-related fee), deflated by the GDP deflator. We exclude the year 2003 in our regression. Column (2) and column (4) focus only on innovative firms (i.e., firms make positive R&D at least once during the sample period). *Above*75 is a dummy variable that equals one if the firm's public float is above 75. *Post* is a dummy variable that equals one if the year is beyond 2004.

Table 3: R&D Investment to Sales Ratio around Threshold

	All Firms		Exclude Public Float 75-90	
	(1)	(2)	(3)	(4)
Below 75	0.002	0.006	0.003	0.008*
	(0.002)	(0.004)	(0.002)	(0.005)
Below 75× Post	-0.004*	-0.010*	-0.006**	-0.013**
	(0.002)	(0.005)	(0.002)	(0.005)
R&D Sale Ratio (lag)	0.925***	0.893***	0.918***	0.890***
	(0.029)	(0.038)	(0.037)	(0.048)
No. Observations	1,948	989	1,621	833
R-Squared	0.824	0.764	0.795	0.727
Sector FE	Yes	Yes	Yes	Yes

Notes: This table examines firms' R&D investment to sales ratio with a public float in the range [50,150] million from 2000 to 2007. The dependent variable is the firm's R&D to sales ratio. We exclude the year 2003 in our regression. Column (2) and column (4) focus only on innovative firms (i.e., firms make positive R&D at least once during the sample period). Columns (3) and (4) exclude firms between [75,90]. *Below*75 is a dummy variable that equals one if the firm's public float is below 75. *Post* is a dummy variable that equals one if the year is beyond 2004. The lagged dependent variable is instrumented using a twice-lagged level of the dependent variable.

Table 4: Summary of Calibrated and Estimated Parameters

Panel A: calibrated outside the Model		Calibrated value	
θ	elasticity of substitution	5	
α	capital share	0.4	
β	discount rate	0.926	
ρ	persistence of TFP	0.878	
σ	standard deviation of TFP shock	0.447	
f_c	fixed cost of compliance	0.024	
Panel B: Structure Estimates		Estimates	Standard Error
ε	TFP-investment elasticity	0.032	(0.002)
τ^c	regulatory cost	0.106	(0.023)
σ_e	standard deviation in signal	0.267	(0.037)
η	information gain	0.259	(0.061)
p_{nb}	non-bunching fraction	0.800	(0.036)
σ_λ	standard deviation in price-earning (log)	0.722	(0.191)

Notes: This table summarizes the calibrated and estimated parameters we used in the model. Panel A shows the parameters that can be calibrated independently of our structure model. Panel B shows the parameters we estimated using the Simulated Method of Moments. Column (2) and column (3) lists each point estimate and the standard errors from bootstrapping, respectively.

Table 5: Data v.s. Simulated Moments

Panel A: Targeted moments	Data	Simulated
$\frac{\text{R\&D}}{\text{sales}}$ (mean, public float $\in [45, 75)$)	0.083	0.084
$\frac{\text{R\&D}}{\text{sales}}$ (mean, public float $\in [75, 105]$)	0.107	0.106
fraction of firms locate in public float $\in [70, 75)$	0.164	0.163
fraction of firms locate in public float $\in [75, 80)$	0.125	0.124
Standard Deviation of Public Float	1.893	1.819
average sales growth	0.090	0.090
Panel B: Non-Targeted moments	Data	Simulated
fraction of firms locate in public float $\in [60, 65)$	0.148	0.141
fraction of firms locate in public float $\in [65, 70)$	0.131	0.133
fraction of firms locate in public float $\in [80, 85)$	0.102	0.118
fraction of firms locate in public float $\in [85, 90]$	0.102	0.110
fraction of firms locate in public float $\in [90, 95)$	0.105	0.102
fraction of firms locate in public float $\in [95, 100)$	0.128	0.110
average productivity growth	0.029	0.034

Notes: This table compares the moments generated by our simulations of the benchmark model to those from the data. The simulation is based on 20,000 firms and the parameters we calibrated in Table 4. Panel A is the targeted moments and Panel B shows the non-targeted moments. The table shows our model does a good job of matching moments from the data.

Table 6: Average Effect of Threshold Regulation via Bunching

		Change in Investment		Change in TFP	
	Share of Firms	R&D	Capital Stock	Level	Growth
	(1)	(2)	(3)	(4)	(5)
Bunching	0.21%	-61.38%	-4.45%	-1.15%	-1.10%
above the Threshold	79.31%	-4.01%	-1.13%	-0.07%	-0.07pp
overall		-3.31%	-0.91%	-0.06%	-0.06pp

Notes: This table shows the bunching results generated by simulating our benchmark model. Column (1) shows the fraction of bunched firms and the fraction of non-bunching firms above the threshold. Columns (2) and (3) show the average changes in the firm's R&D and Capital investment in each region. Columns (4) and (5) show the average changes in the firm's productivity level and productivity growth in each region. In the last row, we compute the total decrease in R&D and capital investment in the bunching region, defined as firms located between the threshold and marginal public float using the counterfactual distribution.

Table 7: Decomposition of the Aggregate Impact

	aggregate changes	via regulatory cost	via information gain
	(1)	(2)	(3)
R&D Investment	-2.81%	-7.00%	4.19%
Capital Stock	-0.29%	-0.75%	0.46%
TFP Level	-0.61%	-1.53%	0.91%
TFP growth	-0.06%	-0.16%	0.10%
Investor's Welfare	-0.30%	-0.76%	0.46%
Public float of marginal firms	82.27	85.09	71.43

Notes: This table decomposes the aggregate changes in R&D investment, Capital investment, TFP, TFP growth, and investor's welfare into two channels. Our benchmark model shows that threshold regulation could reduce firms' R&D investment overall and cause productivity and investor welfare loss. Column (1) shows the aggregate changes in key variables in our benchmark model. Column (2) shows the reduction comes from regulatory costs, and column (3) shows that such reduction could be partially eliminated through an information benefit channel. The last row shows the marginal firm's public float (the unit is a million). The marginal firm is defined as a firm that is indifferent between bunching or not bunching. Firms between the threshold and the public float of the marginal firm will bunch.

Table 8: Simulated Counterfactual Policies: Average Effect of Different Threshold Regulation via Bunching

		Change in Investment		Change in TFP	
	Share of Firms	R&D	Capital Stock	Level	Growth
	(1)	(2)	(3)	(4)	(5)
Panel A: Alternative Threshold Public Float: $0 Million					
Bunching above the Threshold	100.00%	-5.88%	-1.19%	-0.07%	-0.07pp
Panel B: Alternative Threshold Public Float: $100 Million					
Bunching	0.32%	-54.93%	-4.19%	-1.08%	-1.04%
above the Threshold	73.82%	-3.70%	-1.12%	-0.07%	-0.07pp
Panel C: Alternative Threshold Public Float: $250 Million					
Bunching	0.40%	-49.15%	-4.91%	-1.27%	-1.21%
above the Threshold	55.37%	-3.10%	-1.07%	-0.07%	-0.07pp
Panel D: Alternative Threshold Public Float: $700 Million					
Bunching	0.36%	-38.09%	-4.33%	-1.12%	-1.09%
above the Threshold	34.27%	-2.78%	-1.06%	-0.07%	-0.06pp

Notes: This table reports the effects of different thresholds on a firm's bunching decisions. Panel A lists the results from mandatory compliance with the threshold of $0 million. Panel B-C shows the results from larger thresholds of $100 million, $250 million, and $700 million. See the notes in Table 6 for explanations of each column.

Table 9: Simulated Counterfactual policies: Decomposition of the Aggregate Impact

	aggregate changes (1)	via regulatory cost (2)	via information gain (3)
Panel A: Alternative Threshold Public Float: $0 Million			
R&D Investment	-2.90%	-7.27%	4.37%
Capital Stock	-0.30%	-0.76%	0.46%
TFP Level	-0.69%	-1.75%	1.06%
TFP Growth	-0.06pp	-0.17pp	0.11pp
Investor's Welfare	-0.30%	-0.76%	0.46%
Panel B: Alternative Threshold Public Float: $100 Million			
R&D investment	-2.77%	-6.89%	4.12%
Capital Stock	-0.30%	-0.75%	0.46%
TFP Level	-0.59%	-1.46%	0.87%
TFP Growth	-0.05pp	-0.15pp	0.10pp
Investor's Welfare	-0.30%	-0.76%	0.46%
Public float of marginal firms	109.152	118.489	92.642
Panel C: Alternative Threshold Public Float: $250 Million			
R&D investment	-2.55%	-6.33%	3.78%
Capital Stock	-0.29%	-0.74%	0.45%
TFP Level	-0.47%	-1.16%	0.69%
TFP Growth	-0.05pp	-0.13pp	0.09pp
Investor's Welfare	-0.29%	-0.75%	0.45%
Public float of marginal firms	274.039	296.666	230.750
Panel D: Alternative Threshold Public Float: $700 Million			
R&D investment	-2.26%	-5.64%	3.38%
Capital Stock	-0.28%	-0.71%	0.44%
TFP Level	-0.41%	-1.05%	0.63%
TFP Growth	-0.04pp	-0.11pp	0.08pp
Investor's Welfare	-0.29%	-0.73%	0.44%
Public float of marginal firms	768.486	828.533	646.228

Notes: This table reports the impact of different regulation thresholds on aggregate outcomes: investment, productivity, and investor welfare. Panel A lists the results from mandatory compliance with the threshold of $0 million. Panel B-C shows the results from larger thresholds of $100 million, $250 million, and $700 million. See the note in Table 7 for explanations of each column.

Table 10: Higher Information Gain: Aggregate Impact from Compliance Benefits

	Aggregate Change (1)	in Bunching Region (Average Effect) (2)	Above the Threshold (Average Effect) (3)
Share of Firms		0.32%	79.17%
Public Float of Marginal Firm	70.164		
R&D investment	9.22%	316.92%	14.14%
Capital Stock	0.93%	5.96%	0.93%
TFP Level	1.90%	1.43%	0.23%
TFP Growth	0.22pp	1.38pp	0.22pp
Investor's Welfare	0.94%		

Notes: This table shows the aggregate changes and average changes in R&D investment, TFP, TFP growth, and investor's welfare under the alternative situation where the information gain outweighs regulatory costs. The net positive compliance benefits could incentivize firms to bunch above the threshold. Column (1) shows the aggregate changes. Columns (2) and (3) shows the average changes for firms within each region. The reduction of capital investment in the bunched region is due to the simulated large negative signal received by the bunched firms.

A Appendix: Model Derivation

This appendix provides additional details on derivation equations in section 3.

A.1 Profit Maximization

input choices under incomplete information without signals Given the investment I_{it-1}, both k_{it} and l_{it} are chosen under imperfect information. That is, the value of μ_{it} is unobserved by the firms. The optimization problem for each firm is:

$$\max_{k_{it},l_{it}} \mathbb{E}_{it}[e^{(1-\frac{1}{\theta})\phi_{it}}]k_{it}^{\tilde{\alpha}_1} l_{it}^{\tilde{\alpha}_2} - w_t l_{it} - r_t k_{it}$$

The FOC yields:

$$k_{it} = \left(\frac{\tilde{\alpha}_2}{w_t}\right)^{\frac{\tilde{\alpha}_2}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}} \left(\frac{\tilde{\alpha}_1}{r_t}\right)^{\frac{1-\tilde{\alpha}_2}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}} \left(\mathbb{E}_{it}[e^{(1-\frac{1}{\theta})\phi_{it}}]\right)^{\frac{1}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}}$$
$$l_{it} = \left(\frac{\tilde{\alpha}_2}{w_t}\right)^{\frac{1-\tilde{\alpha}_1}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}} \left(\frac{\tilde{\alpha}_1}{r_t}\right)^{\frac{\tilde{\alpha}_1}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}} \left(\mathbb{E}_{it}[e^{(1-\frac{1}{\theta})\phi_{it}}]\right)^{\frac{1}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}}$$

where $\tilde{\alpha}_i = (1-\frac{1}{\theta})\alpha_i$. The actual profit can be written as:

$$\begin{aligned}
\pi_{it} =& e^{(1-\frac{1}{\theta})\phi_{it}} k_{it}^{\tilde{\alpha}_1} l_{it}^{\tilde{\alpha}_2} - w_t l_{it} - r_t k_{it} \\
=& \left(e^{(1-\frac{1}{\theta})\phi_{it}} - (\tilde{\alpha}_1+\tilde{\alpha}_2)\mathbb{E}_{it}[e^{(1-\frac{1}{\theta})\phi_{it}}]\right) \left(\mathbb{E}_{it}[e^{(1-\frac{1}{\theta})\phi_{it}}]\right)^{\frac{\tilde{\alpha}_1+\tilde{\alpha}_2}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}} \left(\frac{\tilde{\alpha}_2}{w_t}\right)^{\frac{\tilde{\alpha}_2}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}} \left(\frac{\tilde{\alpha}_1}{r_t}\right)^{\frac{\tilde{\alpha}_1}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}} \\
=& \left(e^{(1-\frac{1}{\theta})\mu_{it}} - (\tilde{\alpha}_1+\tilde{\alpha}_2)\mathbb{E}_{it}[e^{(1-\frac{1}{\theta})\mu_{it}}]\right) \left(\mathbb{E}_{it}[e^{(1-\frac{1}{\theta})\mu_{it}}]\right)^{\frac{\tilde{\alpha}_1+\tilde{\alpha}_2}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}} \left(\frac{\tilde{\alpha}_2}{w_t}\right)^{\frac{\tilde{\alpha}_2}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}} \left(\frac{\tilde{\alpha}_1}{r_t}\right)^{\frac{\tilde{\alpha}_1}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}} \\
& \times e^{\tilde{\theta}\rho\phi_{it-1}} (1+I_{it-1})^{\tilde{\theta}\varepsilon}
\end{aligned}$$

where $\tilde{\theta} = \frac{1-\frac{1}{\theta}}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}$ and $\tilde{\theta} = \theta - 1$ if $\alpha_1 + \alpha_2 = 1$. As $\mathbb{E}_{it}[e^{(1-\frac{1}{\theta})\mu_{it}}] = e^{(1-\frac{1}{\theta})^2\sigma^2}$, Let's define:

$$\begin{aligned}
h(\sigma^2) \equiv & \mathbb{E}_{it}\left[\left(e^{(1-\frac{1}{\theta})\mu_{it}} - (\tilde{\alpha}_1+\tilde{\alpha}_2)\mathbb{E}_{it}[e^{(1-\frac{1}{\theta})\mu_{it}}]\right)\left(\mathbb{E}_{it}[e^{(1-\frac{1}{\theta})\mu_{it}}]\right)^{\frac{\tilde{\alpha}_1+\tilde{\alpha}_2}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}}\right] \\
= & (1-\tilde{\alpha}_1-\tilde{\alpha}_2)\left(\mathbb{E}_{it}[e^{(1-\frac{1}{\theta})\mu_{it}}]\right)^{\frac{1}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}} \\
= & (1-\tilde{\alpha}_1-\tilde{\alpha}_2)e^{\frac{1}{2}(1-\frac{1}{\theta})\tilde{\theta}\sigma^2} \\
\Phi_{it} \equiv & \left(\frac{\tilde{\alpha}_2}{w_t}\right)^{\frac{\tilde{\alpha}_2}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}}\left(\frac{\tilde{\alpha}_1}{r_t}\right)^{\frac{\tilde{\alpha}_1}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}} e^{\tilde{\theta}\rho\phi_{it-1}}
\end{aligned}$$

Then, the expected profit before making the investment can be written as:

$$\mathbb{E}_{it}[\pi_{it}] = h(\sigma^2)\Phi_{it}(1+I_{it-1})^{\tilde{\theta}\varepsilon}$$

The expected profit increases with the productivity uncertainty σ^2 and the initial productivity ϕ_{it-1}.

input choices under incomplete information with signals Suppose each firm can observe signal $s_{it} = \mu_{it} + e_{it}$ after making the investment and before k_{it} and l_{it} is chosen. Let the information set at time t be written as: $\mathcal{I}_{it} = \{s_{it}, \phi_{it-1}\}$. Then, the optimal choice of k_{it} and l_{it} can be written as:

$$\begin{aligned}
k_{it} = & \left(\frac{\tilde{\alpha}_2}{w_t}\right)^{\frac{\tilde{\alpha}_2}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}}\left(\frac{\tilde{\alpha}_1}{r_t}\right)^{\frac{1-\tilde{\alpha}_2}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}}\left(\mathbb{E}_{it}[e^{(1-\frac{1}{\theta})\phi_{it}}|\mathcal{I}_{it}]\right)^{\frac{1}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}} \\
l_{it} = & \left(\frac{\tilde{\alpha}_2}{w_t}\right)^{\frac{1-\tilde{\alpha}_1}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}}\left(\frac{\tilde{\alpha}_1}{r_t}\right)^{\frac{\tilde{\alpha}_1}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}}\left(\mathbb{E}_{it}[e^{(1-\frac{1}{\theta})\phi_{it}}|\mathcal{I}_{it}]\right)^{\frac{1}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}}
\end{aligned}$$

with

$$\mathbb{E}_{it}[e^{(1-\frac{1}{\theta})\mu_{it}}|\mathcal{I}_{it}] = e^{\frac{1}{2}(1-\frac{1}{\theta})^2\mathbb{V}+(1-\frac{1}{\theta})\frac{\mathbb{V}}{\sigma_e^2}s_{it}}$$

with $\mathbb{V} = \left(\frac{1}{\sigma^2}+\frac{1}{\sigma_e^2}\right)^{-1}$ measures the precision of the signal. Since the signal s_{it} is unobserved when making the investment at the end of $t-1$. The expected profit before making R&D investment can be written as:

$$\begin{aligned}
\mathbb{E}_{it}[\pi_{it}] = & \mathbb{E}_{it}[e^{(1-\frac{1}{\theta})\phi_{it}}k_{it}^{\tilde{\alpha}_1}l_{it}^{\tilde{\alpha}_2} - w_t l_{it} - r_t k_{it}] \\
= & \mathbb{E}_{it}\left[\left(e^{(1-\frac{1}{\theta})\mu_{it}} - (\tilde{\alpha}_1+\tilde{\alpha}_2)\mathbb{E}_{it}[e^{(1-\frac{1}{\theta})\mu_{it}}|\mathcal{I}_{it}]\right)\left(\mathbb{E}_{it}[e^{(1-\frac{1}{\theta})\mu_{it}}|\mathcal{I}_{it}]\right)^{\frac{\tilde{\alpha}_1+\tilde{\alpha}_2}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}}\right]\Phi_{it}(1+I_{it-1})^{\tilde{\theta}\varepsilon}
\end{aligned}$$

Since:

$$\begin{aligned}
\mathbb{E}_{it}\left[e^{(1-\frac{1}{\theta})\mu_{it}}\left(\mathbb{E}_{it}[e^{(1-\frac{1}{\theta})\mu_{it}}|\mathcal{I}_{it}]\right)^{\frac{\tilde{\alpha}_1+\tilde{\alpha}_2}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}}\right] =& e^{\frac{1}{2}\frac{\tilde{\alpha}_1+\tilde{\alpha}_2}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}\left(1-\frac{1}{\theta}\right)^2\mathbb{V}} \times \mathbb{E}_{it}\left[e^{\left(1-\frac{1}{\theta}\right)\left(\mu_{it}+\frac{\mathbb{V}}{\sigma_e^2}\frac{\tilde{\alpha}_1+\tilde{\alpha}_2}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}s_{it}\right)}\right] \\
=& e^{\frac{1}{2}\left(1-\frac{1}{\theta}\right)^2\sigma^2\theta+\frac{\theta^2}{2}\left(1-\frac{1}{\theta}\right)^3\frac{\sigma^2}{\sigma_e^2}\mathbb{V}} \\
\mathbb{E}_{it}\left[\left(\mathbb{E}_{it}[e^{(1-\frac{1}{\theta})\mu_{it}}|\mathcal{I}_{it}]\right)^{\frac{1}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}}\right] =& \mathbb{E}_{it}\left[e^{\frac{1}{2}\left(1-\frac{1}{\theta}\right)^2\frac{\mathbb{V}}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}+\left(1-\frac{1}{\theta}\right)\frac{1}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}\frac{\mathbb{V}}{\sigma_e^2}s_{it}}\right] \\
=& e^{\frac{1}{2}\left(1-\frac{1}{\theta}\right)^2\sigma^2\theta+\frac{\theta^2}{2}\left(1-\frac{1}{\theta}\right)^3\frac{\sigma^2}{\sigma_e^2}\mathbb{V}}
\end{aligned}$$

Define the term in the bracket as:

$$\begin{aligned}
g(\sigma^2,\sigma_e^2) =& \mathbb{E}_{it}\left[\left(e^{(1-\frac{1}{\theta})\mu_{it}}-(\tilde{\alpha}_1+\tilde{\alpha}_2)\mathbb{E}_{it}[e^{(1-\frac{1}{\theta})\mu_{it}}|\mathcal{I}_{it}]\right)\left(\mathbb{E}_{it}[e^{(1-\frac{1}{\theta})\mu_{it}}|\mathcal{I}_{it}]\right)^{\frac{\tilde{\alpha}_1+\tilde{\alpha}_2}{1-\tilde{\alpha}_1-\tilde{\alpha}_2}}\right] \\
=& (1-\tilde{\alpha}_1-\tilde{\alpha}_2)e^{\frac{1}{2}(1-\frac{1}{\theta})\tilde{\theta}\sigma^2+\frac{1}{2}(1-\frac{1}{\theta})\tilde{\theta}^2\sigma^2\frac{\mathbb{V}}{\sigma_e^2}}
\end{aligned}$$

The expected profit decreases with the variance of the noise σ_e^2. Consider the case where the signal is too noisy such that $\sigma_e^2 \to +\infty$, then $g(\sigma^2,\sigma_e^2) = h(\sigma^2)$. If the signal is precise, that is $\sigma_e^2 = 0$, then $g(\sigma^2,\sigma_e^2) = h(\sigma^2)e^{\frac{1}{2}(1-\frac{1}{\theta})\tilde{\theta}^2\sigma^2} > h(\sigma^2)$. The increase in information precision makes firms choose k_{it} and l_{it} more preciously, avoiding the case of negative profits.

A.2 Investor's Problem

Suppose there exists a risk-neutral investor with an endowment e, and buys s_{it} shares of firm i in period $t-1$ and sells the share at $t+1$. The optimization problem of the investor can be written as:

$$\begin{aligned}
&\max_{\{s_{it}\}} \sum_{t=0}^{\infty} \mathbb{E}_{it}\beta^t c_{it} \\
s.t. \quad & c_{it} = e - P_{it}s_{it} \\
& c_{it+1} = e + P_{it+1}s_{it} + \hat{\pi}_{it+1}s_{it}
\end{aligned}$$

$\hat{\pi}_{it+1}$ is the next period's profits minus the investment cost (if any). The Euler Equation implies that:

$$P_{it} = \beta\mathbb{E}_{it}(\pi_{it+1} + P_{it+1}) = \sum_{s=1}^{\infty}\beta^s\mathbb{E}_{it}[\hat{\pi}_{t+s}]$$

Suppose the investor cannot observe whether the firm is bunching or not. He always assumes that an interior solution exists for firms' optimization problems. That is, $\ln(1+I_{it+1}) = \tilde{\theta}\epsilon \ln(\beta h(\sigma^2)\Phi e^{\tilde{\theta}\rho\phi_{it}})$. Where $\Phi \equiv \left(\frac{\tilde{\theta}(1-\alpha)}{\theta w}\right)^{\tilde{\theta}(1-\alpha)} \left(\frac{\tilde{\theta}\alpha}{\theta r}\right)^{\tilde{\theta}\alpha}$ is a constant number across firms. And the expected profits can be written as $\hat{\pi}_{it+1} = (1 - \frac{\tilde{\theta}_\epsilon}{2})(1 + I_{it})$. Given the evolution of productivity, we can write the series of future profits in terms of current profits:

$$\begin{aligned}
\pi_{it+1} =& (1 - \tilde{\alpha}_1 - \tilde{\alpha}_2) e^{\left(1+\frac{\tilde{\theta}_\varepsilon}{2-\tilde{\theta}_\varepsilon}\right)\tilde{\theta}\rho\phi_{it} + \tilde{\theta}\mu_{it+1} + \frac{\tilde{\theta}_\varepsilon}{2-\tilde{\theta}_\varepsilon}\ln[\tilde{\theta}_\varepsilon\beta h(\sigma^2)\Phi]} \\
\pi_{it+s} =& (1 - \tilde{\alpha}_1 - \tilde{\alpha}_2) e^{\tilde{\theta}q^s\phi_{it} + \tilde{\theta}\frac{1-q^{s-1}}{1-q}\ln[\tilde{\theta}_\varepsilon\beta h(\sigma^2)\Phi] + \tilde{\theta}\sum_{j=0}^{s-1} q^j \mu_{t+s-j}}
\end{aligned}$$

where $q \equiv \frac{2\rho}{2-2\tilde{\theta}_\varepsilon} < 1$. As $q < 1$, the expected π_{it+s} will converge to a finite number. We have:

$$\frac{E_{it}[\pi_{it+s}]}{E_{it}[\pi_{it+1}]} = exp\left[\tilde{\theta}(q^s - \rho)\phi_{it} + \frac{\tilde{\theta}_\varepsilon}{2-\tilde{\theta}_\varepsilon}\frac{q-q^{s-1}}{1-q}\ln[\tilde{\theta}_\varepsilon\beta h(\sigma^2)\Phi] + \frac{q-q^{s-1}}{1-q}\frac{\tilde{\theta}}{2}\sigma^2\right]$$

Summing over s, we can get

$$P_{it} = \lambda \mathbb{E}_{it}[\pi_{it+1}]$$

A.3 Bunching above the threshold

In some cases where the information benefit outweighs the increased regulatory costs, i.e., $(1-\tau^c)g(\sigma^2, (1-\eta')\sigma_e^2) > (1-\tau)g(\sigma^2, \sigma_e^2)$, compliance to SOX 404 could increase the marginal benefit intangible investment when the firm's public float is above the threshold $\overline{PF}$. A firm with optimal public float just below the threshold would find it optimal to bunch above the threshold. In such cases, the investment under regulation would be higher than those without regulation. The value of firms that are conditional on bunching is then:

$$v(I_{it}^a|\tau, (1-\eta)\sigma_e^2) = \pi_{it} - \frac{1}{2}(I_{it}^a + 1)^2 + \frac{1}{2} + \beta(1-\tau^c)g(\sigma^2, (1-\eta)\sigma_e^2)\Phi_{it+1}(1 + I_{it}^a)^{\tilde{\theta}_\varepsilon}$$

where I_{it}^a is given by:

$$I_{it}^a = \left(\frac{\overline{PF}}{\lambda h(\sigma^2)\Phi_{it+1}}\right)^{\frac{1}{\tilde{\theta}_\varepsilon}} - 1 \tag{9}$$

The value of firms that at the interior optimal I_{it}^f with optimal size below the threshold is:

$$v(I_{it}^f|\tau,\sigma_e^2) = \pi_{it} - \frac{1}{2}(I_{it}^f+1)^2 + \frac{1}{2} + \beta(1-\tau)g(\sigma^2,\sigma_e^2)\Phi_{it+1}(1+I_{it}^f)^{\tilde{\theta}\varepsilon}$$

where I_{it}^f is given by:

$$I_{it}^f = [\tilde{\theta}\varepsilon\beta(1-\tau)g(\sigma^2,\sigma_e^2)\Phi_{it+1}]^{\frac{1}{2-\tilde{\theta}\varepsilon}} - 1 \tag{10}$$

Firm that previously (before SOX) choose $I_{it}^* < I_{it}^a$ will choose to bunch above the threshold if $v(I_{it}^a|\tau,(1-\eta)\sigma_e^2) \geq v(I_{it}^f|\tau,\sigma_e^2)$. Substituting the out I_{it}^f and I_{it}^a in $v(I_{it}^f|\tau,\sigma_e^2)$ and $v(I_{it}^a|\tau^c,(1-\eta)\sigma_e^2)$ using equations (10) and (9), and public float $PF^f = \lambda h(\sigma^2)\Phi_{it+1}(1+I_{it+1}^f)^{\tilde{\theta}\varepsilon}$, we can express the decision to bunch or not as:

$$-\left(\frac{PF^f}{\overline{PF}}\right)^{1-\frac{2}{\tilde{\theta}\varepsilon}}\frac{\tilde{\theta}\varepsilon}{2} + \frac{1-\tau^c}{1-\tau}\frac{g(\sigma^2,(1-\eta')\sigma_e^2)}{g(\sigma^2,\sigma_e^2)} \geq \left(1-\frac{\tilde{\theta}\varepsilon}{2}\right)\frac{PF^f}{\overline{PF}} \tag{11}$$

$PF^f = \lambda h(\sigma^2)\Phi_{it+1}(1+I_{it+1}^c)^{\tilde{\theta}\varepsilon}$ is the optimal public float without any regulation. The first term on the LHS of equation (11) is the cost of bunching (i.e. extra cost in additional investment), and the second term on the LHS is the net gain of bunching above from two channels 1) information benefit from increasing prevision of the signal, minus 2) the payment of implicit corporate tax. The term on the RHS is the benefit of not bunching (i.e., saving on current investment minus the productivity gain in the future). In the absence of regulatory cost, firms with the optimal public float close to the threshold $PF^f \approx \overline{PF}$, bunching has a small relative investment cost, but the reduction in uncertainties (or increase in signal precision) incentivizes it to bunch. Let $\hat{\phi}_{it}$ be the productivity of the marginal firms that are indifferent between bunching or not, the optimal investment under regulatory cost and information benefit can be written as follows:

$$I_{it}^{**} = \begin{cases} [\tilde{\theta}\varepsilon\beta(1-\tau^c)g(\sigma^2,(1-\eta')\sigma_e^2)\Phi_{it+1}]^{\frac{1}{2-\tilde{\theta}\varepsilon}} - 1 & \text{if} \quad \phi_{it} > \phi'_{it} \\ \left(\frac{\overline{PF}}{\lambda h(\sigma^2)\Phi_{it+1}}\right)^{\frac{1}{\tilde{\theta}\varepsilon}} - 1 & \text{if} \quad \hat{\phi}_{it} < \phi_{it} \leq \phi'_{it} \\ [\tilde{\theta}\varepsilon\beta(1-\tau)g(\sigma^2,\sigma_e^2)\Phi_{it+1}]^{\frac{1}{2-\tilde{\theta}\varepsilon}} - 1 & \text{if} \quad \tilde{\phi}_{it} < \phi_{it} \leq \hat{\phi}_{it} \\ 0 & \text{if} \quad \phi_{it} \leq \tilde{\phi}_{it} \end{cases} \tag{12}$$

where ϕ'_{it} is determined by equation $\overline{PF} = \lambda h(\sigma^2)\Phi_{it+1}(\phi'_{it})[\tilde{\theta}\varepsilon\beta(1-\tau^c)g(\sigma^2,(1-\eta')\sigma_e^2)\Phi_{it+1}(\phi'_{it})]^{\frac{\tilde{\theta}\varepsilon}{1-\tilde{\theta}\varepsilon}}$. The difference between ϕ'_{it} and $\hat{\phi}_{it}$ is the bunching region. The size of the bunching region increases with the value of η' and decreases with the regulatory cost τ^c.

B Appendix: Parameter's Sensitivity to Moments

Figure A.1: Parameter's Sensitivity to Moments

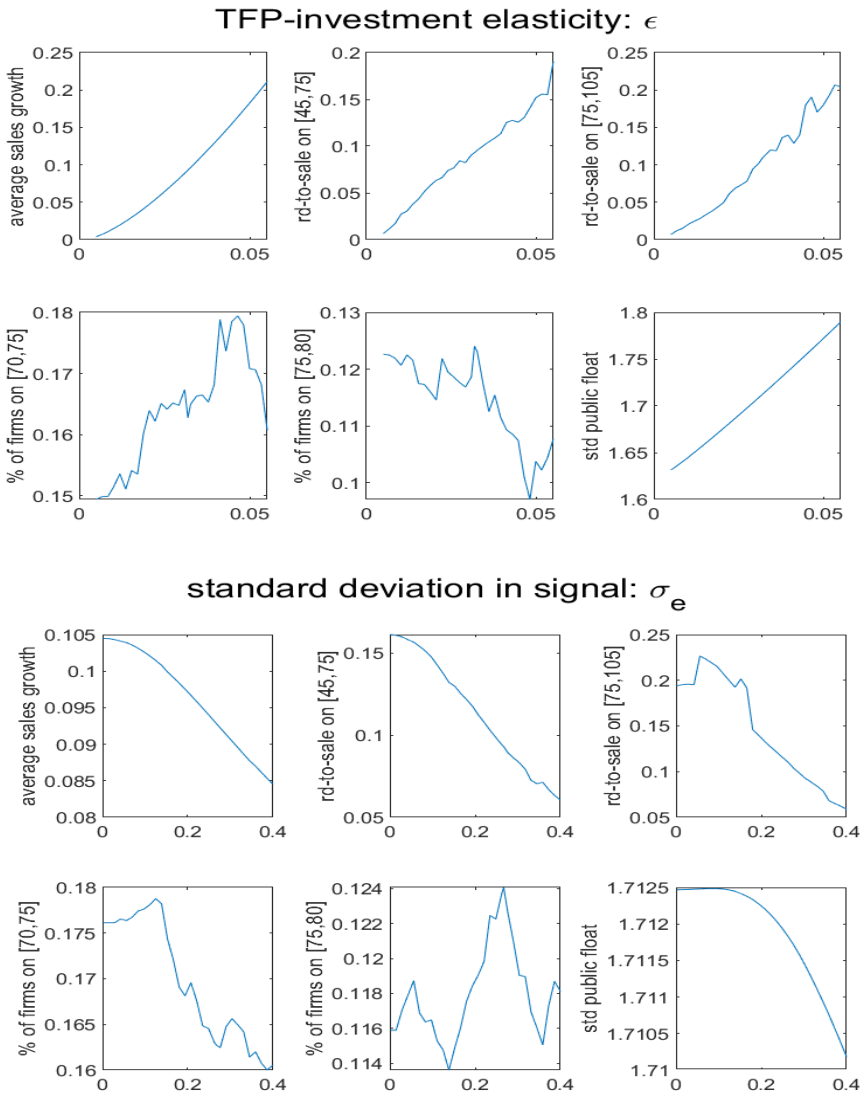


Figure A.2: Parameter's Sensitivity to Moments

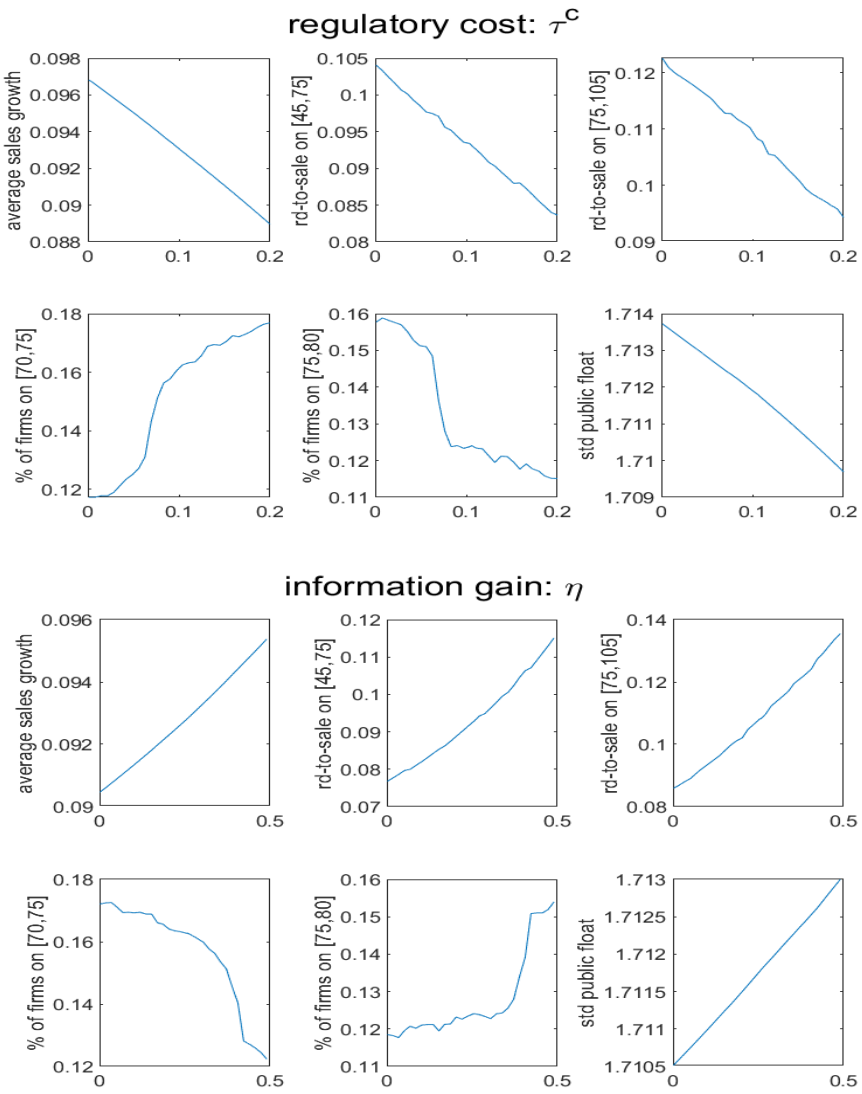
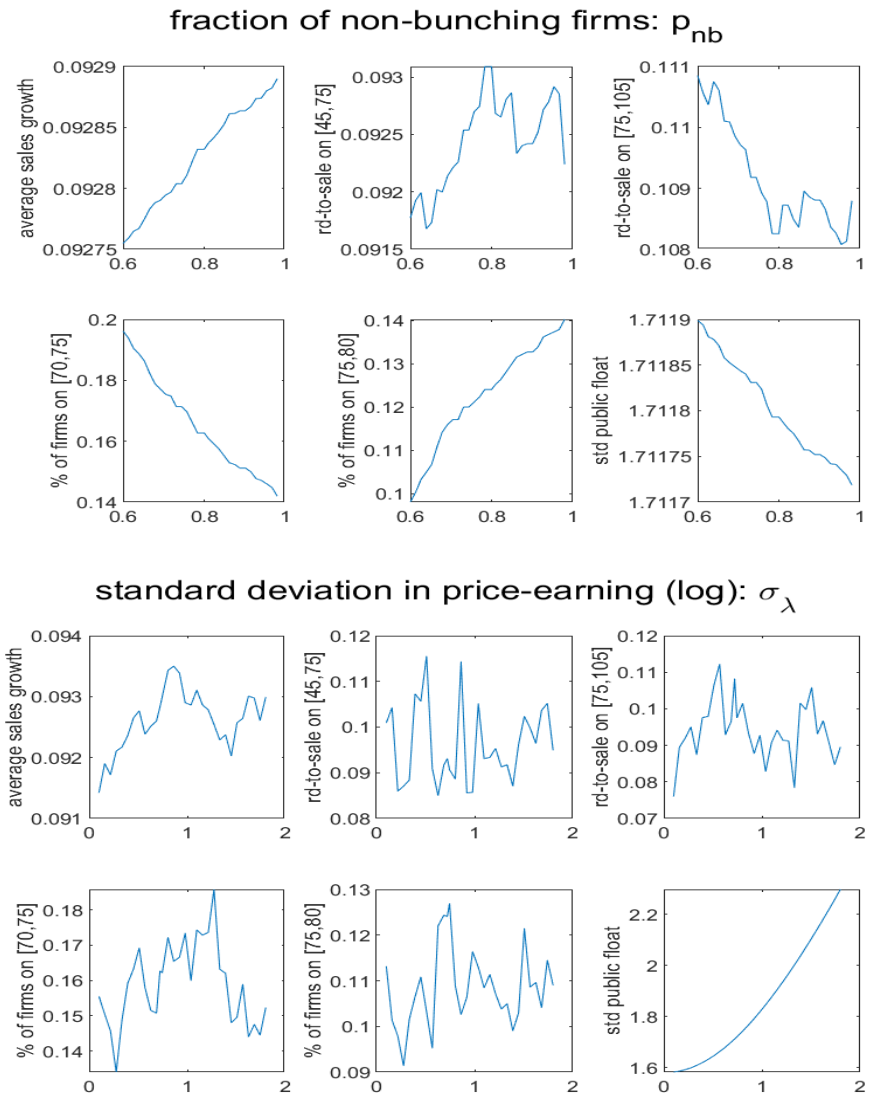
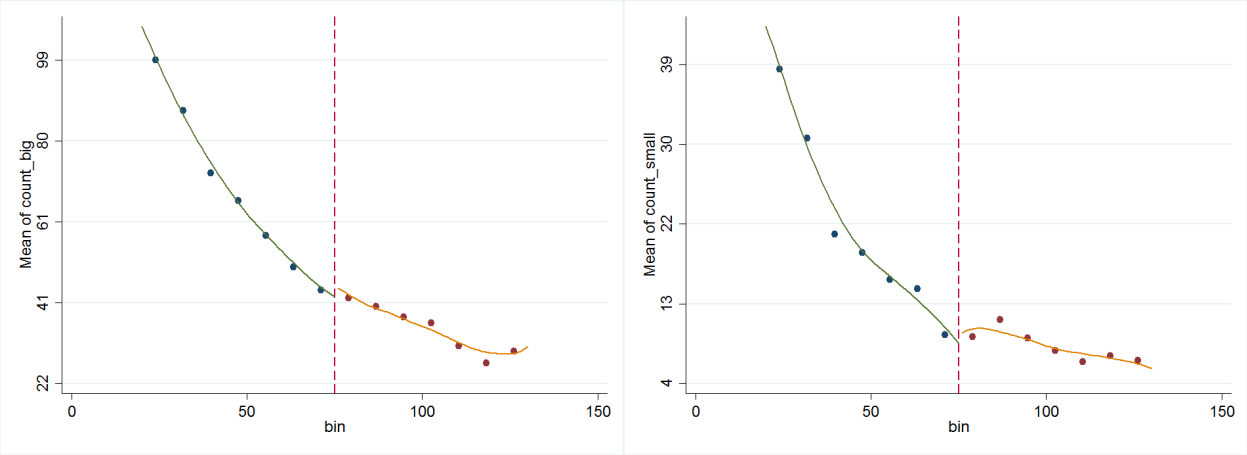


Figure A.3: Parameter's Sensitivity to Moments



C Appendix: Exploring Agency Issues

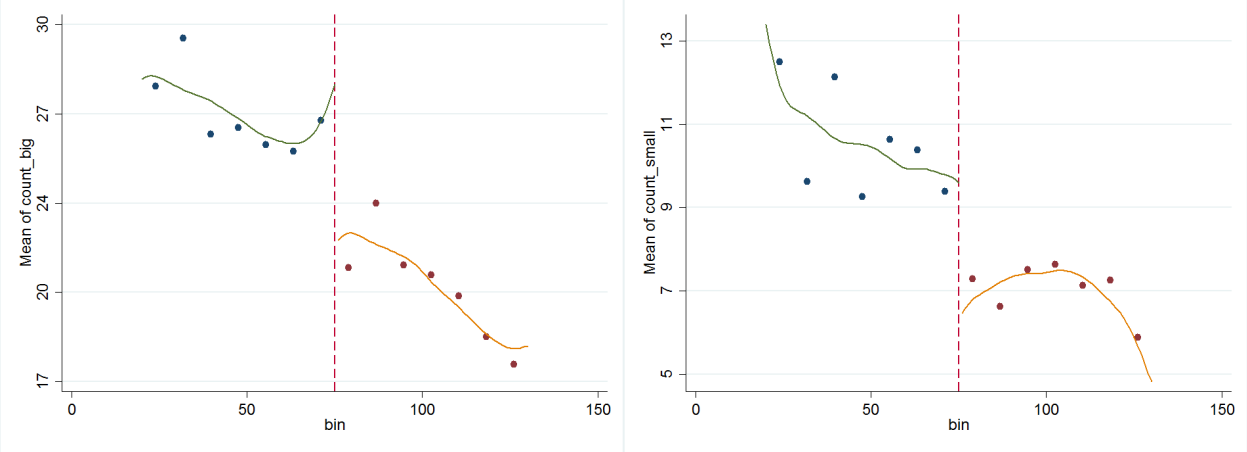
C.1 When Market Value - Public Float = $10m:



(a) Public Float Local Linear Smooth Density Distribution for more potential-agency-friction in Pre-Period

(b) Public Float Local Linear Smooth Density Distribution for less potential-agency-friction in Pre-Period

Figure A.4: Market Value - Public Float = $10m in Pre-Period

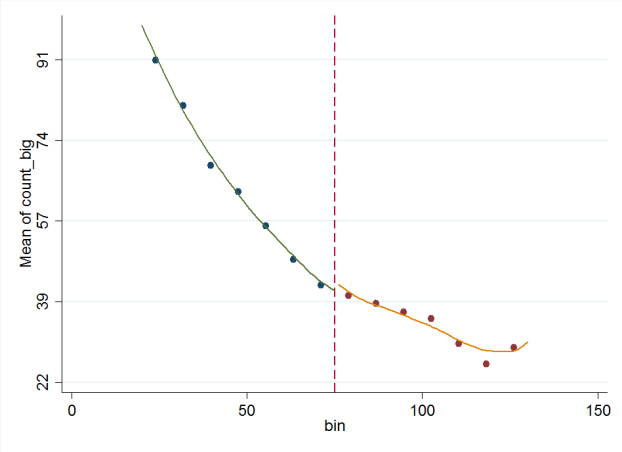


(a) Public Float Local Linear Smooth Density Distribution for more potential-agency-friction in Post-Period

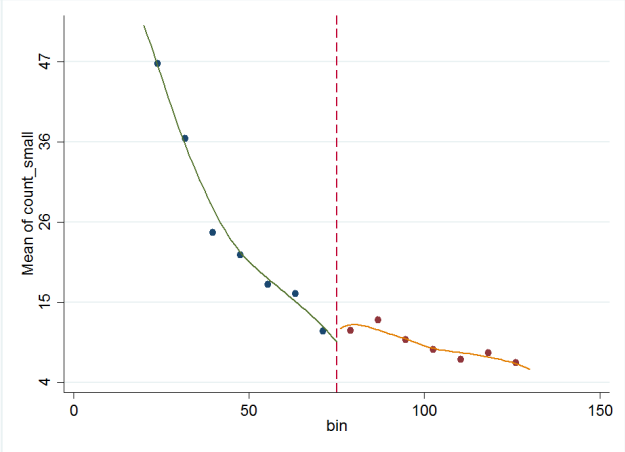
(b) Public Float Local Linear Smooth Density Distribution for less potential-agency-friction in Post-Period

Figure A.5: Market Value - Public Float = $10m in Post-Period

When Market Value - Public Float = $15m:

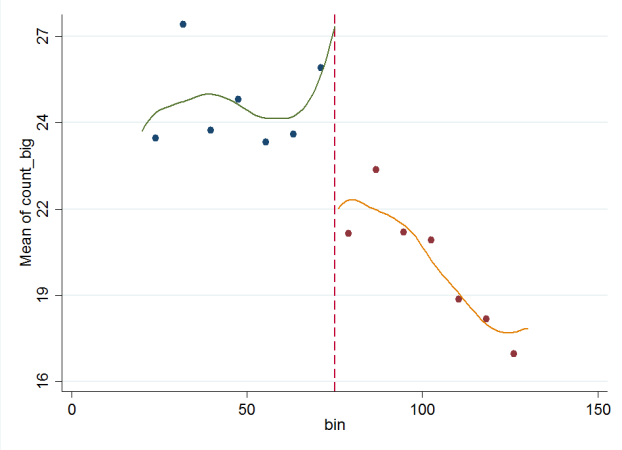


(a) Public Float Local Linear Smooth Density Distribution for more potential-agency-friction in Pre-Period

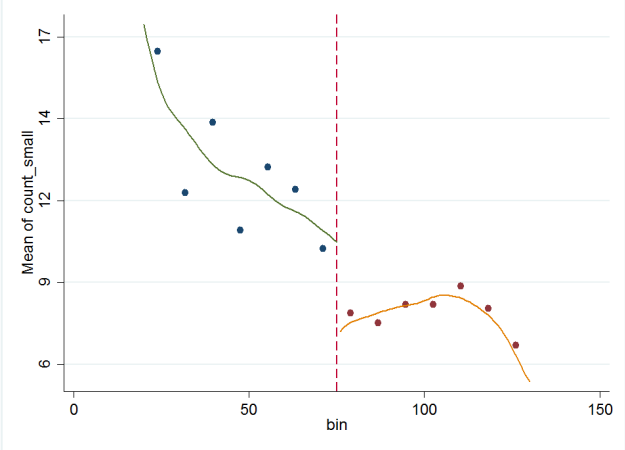


(b) Public Float Local Linear Smooth Density Distribution for less potential-agency-friction in Pre-Period

Figure A.6: Market Value - Public Float = $15m in Pre-Period



(a) Public Float Local Linear Smooth Density Distribution for more potential-agency-friction in Post-Period



(b) Public Float Local Linear Smooth Density Distribution for less potential-agency-friction in Post-Period

Figure A.7: Market Value - Public Float = $15m in Post-Period