# Not(ch) Your Average Tax System: Corporate Taxation Under Weak Enforcement

Pierre Bachas (UC Berkeley) & Mauricio Soto (Banco Central de Costa Rica) *

December 14, 2015

JOB MARKET PAPER. LATEST VERSION AVAILABLE HERE.

**Abstract**

Boosting domestic tax collection is a major policy challenge for low and middle-income countries, however there remains considerable debate regarding the appropriate rate and breadth of the tax base. We use the unusual design of the corporate income tax in Costa Rica, administrative data on the universe of formal firms, and a novel methodology, which combines bunching at tax notches with a discontinuity approach, to estimate important parameters for the design of optimal tax policy. First, the elasticity of reported profits with respect to the tax rate is very large at roughly four. As a result the highest possible optimal tax rate is substantially lower than in rich economies. Second, we separate the profit elasticity into changes in reported revenue versus reported costs, and show that the cost elasticity is larger than the revenue elasticity. Firms' apparent ease to understate profits by over-reporting costs rationalizes the use of broad tax bases with few deductions. Third, we provide evidence that tax evasion is a key driver, while real effects appear limited. Taken together, the data suggests that Costa Rican firms evade taxes on up to 75% of their profits when faced with a 30% tax rate, and that the revenue maximizing rate is between 18-26%. This implies that the optimal design of the corporate income tax in lower-income countries might be substantially different from that of OECD countries.

*Corresponding author: Pierre Bachas, Economics Department UC Berkeley (bachas@berkeley.edu).
We gratefully acknowledge the assistance of Oscar Fonseca Villalobos and Jorge Richard Munoz Nunez in helping us with the data. PB thanks his advisors Emmanuel Saez and Edward Miguel for their guidance and we have benefited from discussions with Miguel Almunia, Juan Pablo Atal, Alan Auerbach, Youssef Benzarti, Michael Best, Anne Brockmeyer, Natalie Cox, Paul Gertler, Anders Jensen, Etienne Lehman, Joana Naritomi, Alvaro Ramos Chavez, Andres Rodriguez Clare, Danny Yagan and seminar participants at UC Berkeley, the Universidad de Costa Rica, the ZEW Public Finance conference at Mannheim, the Oxford Center for Business Taxation doctoral meeting, Paris II and the Institute for Fiscal Studies. Financial support from the Burch Center for Tax Policy and the Center for Equitable Growth is gratefully acknowledged. The findings and conclusions are entirely those of the authors and do not necessarily represent the views of the Banco Central de Costa Rica or any other institution.

## 1 Introduction

Taxation provides low-income countries with the revenue needed to alleviate poverty and invest in public goods, while strengthening state capacity and the social contract between government and citizens. It also relaxes aid dependence: in the last decade tax revenue increases have dwarfed foreign aid flows, and it is estimated that even in Sub-Saharan Africa, governments collect $10 in own-revenue for every dollar in foreign aid (World Bank 2013). Even though domestic tax collection has recently grown (United Nations 2014, IMF 2015), many low and middle-income countries still have tax to GDP ratios below 20%[1] compared to 35%, on average, for OECD countries (Besley and Persson 2013).

Tax revenue growth can be achieved by designing tax systems which consider the constraints and limited capacity of tax administrations in low income countries. In particular, tax administrations have difficulty monitoring transactions[2], which leads to informality and high tax evasion rates. Increases in tax revenue are partially driven by the development process and the structure of the economy, however recent research shows that efficient design of the tax system and investment in tax capacity boost revenue collection at a given income level (Best, Brockmeyer, Kleven, Spinnewijn and Waseem 2014, Pomeranz 2015, Khan, Khwaja and Olken 2015, Naritomi 2015). Two important features in the design of the corporate income tax, the object of our study, are the *tax rate* and *breadth of the tax base*. When tax evasion is a primary concern, profit elasticities with respect to the tax rate can be large, which lowers the revenue maximizing rate. In addition, the standard tax base, which allows for all costs to be deducted, might not be desirable since it provides evasion opportunities on both the revenue and the cost base.

In this paper we make four contributions. First, we show that even for a middle-income country, Costa Rica, the elasticity of profit with respect to the tax rate for small and medium firms is very large and the top of the Laffer curve appears to correspond to a tax rate below 25%.

Second, we separate the profit elasticity into changes in declared revenue versus declared cost, and show that the cost elasticity is substantially larger than the revenue elasticity. This has important implications for policy: firms' apparent ease in understating profit by misreporting cost

[1] A tax to GDP ratio of 20% is considered the minimum to achieve the UN development goals (UN 2010)

[2] Information and capacity constraints are certainly not the only reasons for differences in tax revenue across income levels: an exciting new literature studies preferences for redistribution and tax morale as other candidate explanations (Luttmer and Singhal 2011, Luttmer and Singhal 2014, Kleven 2014). Moreover, Olken and Singhal (2011) show that households in low-income countries pay substantial informal taxes, which could get substituted for formal taxes through the development path.

rationalizes the use of broad tax bases and policies determined by revenue, instead of profits. This constellation of policies, while rare in rich countries, is in fact sometimes observed in low and middle income countries[3].

Third, we provide evidence that tax evasion is likely the key driver of our results, while real effects and avoidance responses appear limited. Taken together, the data implies that Costa Rican firms evade up to a massive 75% taxes on their profits when faced with a 30% tax rate.

Fourth, we develop a new methodology which combines bunching and discontinuities at tax notches, to separate the profit response into revenue and cost responses and to deal with selection into bunching based on multiple attributes, here revenue distance to the threshold and costs.

Estimating the parameters needed to evaluate the design of tax policy in developing countries is challenging. However, new estimation techniques in public economics (Saez 2010, Chetty, Friedman, Olsen and Pistaferri 2011, Kleven and Waseem 2013), combined with improved access to large and high quality administrative datasets, are increasing researchers' capacity to address these questions. Even five years ago, many low and middle income countries had neither online tax fillings with automatic quality checks, nor integrated data systems covering the universe of taxpayers. As part of this small but growing literature, we use rich administrative data and the design of the corporate tax in Costa Rica to study small and medium enterprises' behavioral responses to taxation.

Our unique setup allows us to estimate the elasticity of corporate profits with respect to the net of tax rate and to separate the profits response into revenue and cost responses. While most corporate tax systems tax profits at a flat rate, Costa Rica imposes increasing average tax rates on profits as a function of firms' revenue. The average tax rate jumps from 10 to 20% at the first revenue threshold and from 20 to 30% at the second threshold. The change in average tax rate generates two distinct behavioral responses. First, some firms reduce their revenue below the threshold, in order to lower the tax rate they face on their whole profit base. This generates excess mass in the firm distribution just below the threshold and missing mass above it, which we use to measure the elasticity of revenue, following the notch estimation technique in Kleven and Waseem (2013). Second, firms remaining above the revenue threshold respond to the higher tax rate by sharply dropping their reported profits. This is evidenced by a large discontinuity in average profit

[3] Examples of such policies are presumptive taxes, which tax revenue instead of profits, enforcement and registration thresholds determined by revenue (e.g. Large Taxpayers Units), and corporate tax systems with different rates as a function of revenue.

margins on either side of the threshold, when plotted against revenue. The response of profits to higher average tax rates is a mix of revenue and cost responses. Using the revenue elasticity, estimated with the bunching methodology, we hold revenue responses constant, such that the remaining profits discontinuity only identifies changes in reported costs. Finally, by combining the revenue and cost responses we estimate the elasticity of profits with respect to the net of tax rate. The resulting elasticities are very large: 4.7 at the first threshold and 2.7 at the second threshold. These are an order of magnitude higher than those of small firms in OECD countries, estimated at around 0.5 (Devereux, Liu and Loretz 2014, Patel, Seegert and Smith 2015) and severely constrain the range of optimal tax rates: given the current policy and enforcement environment, rates above 25% are on the wrong side of the Laffer curve for Costa Rica.

The reduced form estimation provides a robust measure of the profit elasticity: an overestimate of the revenue responses would mechanically underestimate cost responses, leaving profit responses practically unchanged. In contrast, the relative shares of revenue and cost responses are not robust to this estimation. Under heterogeneity in revenue elasticities the bunching method recovers the response of the highest revenue elasticity firm, hence providing an upper bound on the true revenue elasticity. This mechanically implies that the estimated cost elasticity is a lower bound of the average cost elasticity. If we make no assumption about the counterfactual distribution of profit margin by revenue, then our estimates are the tightest possible bounds on the revenue and cost elasticities. However, if we assume that the counterfactual distribution of profit margin by revenue is constant around the threshold, then these elasticities are rejected by the data, as they predict substantially more bunching than observed. With this new counterfactual, we develop a model-based numerical method to estimate the revenue and cost responses, as a joint function of the revenue distance to the threshold and the costs of the firm. The model-based estimation recovers the average elasticity of revenue, while dealing with two-dimensional selection into bunching. Our preferred estimates show that cost responses account for 71% of the drop in reported profits, while revenue responses only for 29%. The relative ease to manipulate cost, compared to revenue, rationalizes the use of tax bases with few deductions and policies based on revenue instead of profits, often observed in lower-income countries.

Behavioral responses due to evasion responses, production effects or avoidance have the same impact on revenue collection, but call for different policy action. To study the mechanisms of firms' responses we draw on rich administrative datasets. In addition to the corporate tax returns, we use information on audits, the central bank's firm registry, social security data on wages and employees, and monthly sales tax receipts. We find evidence that tax evasion is a key driver of responses for bunching firms: these firms are significantly more likely to display inconsistencies with third-party reported information and adjust revenue upwards following audit threats at the industry level. Furthermore, all statistical tests for real effects and avoidance responses are rejected: the social security data shows no discontinuity in the number of employees and wage bill at the threshold, monthly sales receipts do not display evidence of revenue shifting across fiscal years and the registry of economic groups only shows very modest evidence of firms dividing themselves into smaller firms. In a literature that often remains agnostic on the mechanisms of behavioral response, our paper takes innovative steps to support tax evasion as the key mechanism. Further, if we assume that profit responses are only due to evasion, then firms facing a 30% tax rate evade taxes on as much as 75% of their profits.

Our paper contributes to the growing literature on tax design and tax enforcement in low and middle income countries. We are the first to estimate a corporate tax elasticity[4] in this context and find that it is substantially higher than in rich countries: Devereux et al. (2014) for the UK and Patel et al. (2015) for the US both estimate corporate elasticities of 0.5. These estimates are particularly relevant for comparison, since they are identified with bunching at kink points and also concern small and medium firms. To make sense of the magnitude of our results one needs to consider the weak enforcement environment - an expanding empirical literature (Pomeranz 2015, Naritomi 2015) shows that difficulties in monitoring transactions and missing third-party information lead to large evasion rates in developing countries. Even in Denmark, Kleven et al. (2011) estimate tax evasion rates as high as 40% on income not subject to third-party reporting[5]. In Costa Rica, firms facing a 30% tax rate have an implied evasion rate on profits of 75%, which is comparable to the 60% evasion rate estimated with micro-data for large Pakistani firms by Best et al. (2014) and the 65% average evasion for Costa Rican firms, estimated with aggregate data by the IMF (2012).

Our paper also contributes to a recent literature on the two-dimensional aspect of the corporate

[4]A large literature summarized in Auerbach et al. (2010) studies firms' responses to the tax code but few articles estimate corporate tax elasticities. Some exceptions are Gruber and Rauh (2007) who estimate an elasticity of 0.2 for large US corporations, using panel data and an instrument for the effective tax rate change. With a similar methodology, Dwenger and Steiner (2012) estimates a corporate tax elasticity of 0.5 in Germany

[5]Kleven et al. (2011) and Slemrod et al. (2001) use randomized audits to estimate tax evasion - the latter study also finds that tax evasion is concentrated among self-employed individuals in Minnesota

tax declaration and provides the first separation of the elasticity of profit into cost and revenue responses. The relative ease to manipulate costs, compared to revenue, complements the findings of Carrillo et al. (2014) and Slemrod et al. (2015): both studies show that following tighter enforcement on revenue by the tax administration, firms' reported revenue increases to limit the risk of an audit, but reported costs also increase, such that overall profits and tax liability are practically unchanged. Regarding tax policy, the result supports theoretical work on the desirability of "production inefficient" tax instruments under evasion (Emran and Stiglitz 2005, Gordon and Li 2009) and the empirical findings of Best et al. (2014): when evasion opportunities are large, limiting deductions or switching to a turnover tax with no deductions could be optimal.

Finally, from a methodological standpoint, we contribute to the literature using discontinuities in tax design to identify structural parameters. Saez (2010) and Chetty et al. (2011) developed the framework to recover taxable income elasticties from kink points, which was extended to notches by Kleven and Waseem (2013). In our setting, tax notches are determined by revenue, but the tax rate applies to profit: we show how to use this variation to separate revenue and cost elasticities. Revenue-dependent policies, such as registration and enforcement thresholds[6], are common, especially in low-income countries, and our methodology could be applied to these settings, with the caveat that it requires large sample sizes and regularity in the data.

The remainder of the paper is organized as follows. Section 2 introduces the tax system and provides a theoretical framework. Section 3 presents the data, methods and results. Section 4 adds structure to refine the previous results. Section 5 shows evidence of evasion as a key mechanism, while Section 6 rejects some specific real and avoidance mechanisms. Section 7 discusses policy implications and concludes.

## 2 Tax System and Theoretical Framework

### 2.1 Corporate Tax System in Costa Rica

Figure 1 presents the Costa Rican corporate tax schedule. A corporation pays an average tax rate of 10%, 20% or 30% on its profit as a function of its revenue - firms with revenue below the first threshold[7] face a 10% average tax rate, firms with revenue in between the two thresholds face a

[6]For example, Almunia and Lopez-Rodriguez (2015) study the impact of an enforcement threshold, the Large Taxpayer Unit, on Spanish firms' reporting behavior.

[7]In 2014, the revenue thresholds are 49,969,000 and 10,0513,000 Colones, corresponding to 150,000 and 300,000 USD in Purchasing Power Parity. The thresholds are indexed on inflation and therefore grow 5% yearly, on average.

20% rate and firms with revenue above the second threshold face a 30% tax rate. Importantly the revenue thresholds only determine tax liability and are not used for any other policy. Loss carry-forwards are limited to the manufacturing sector and last for three years while loss carry-backs are never allowed. The historical motivation for the system stems back from political pressure to apply preferential rates to small and medium enterprises and the current system has been unchanged since 1988.

### 2.2 Theoretical Framework: Baseline

We develop a simple theoretical framework of firm behavior, to consider simultaneous revenue and cost responses to tax changes, and allow for heterogeneity in revenue and in cost for a given revenue. A representative firm decides how much to produce and can simultaneously evade taxes by under-reporting revenue and over-reporting cost. When evading taxes, the firm incurs resource costs and risks detection. Under this simple framework, we discuss the impact of the Costa Rican corporate tax system on firm behavior. We then derive the empirical predictions of the model and extend it to discuss heterogeneity in revenue and cost elasticities.

Consider a firm that produces good $y$, subject to a convex cost function $c(y)$. The costs incurred by the firm are fully tax-deductible and therefore a flat tax rate on profit is non-distortionary[8]. The firm can under-report revenue, such that revenue evasion is $(y_i - \tilde{y}_i)$, where $\tilde{y}_i$ is declared revenue, and over-report cost, such that cost evasion is $(\tilde{c}_i - c_i)$, where $\tilde{c}_i$ is declared cost. In doing so it incurs resource costs[9] and risks detection: this generates a convex cost of evasion $R(y_i - \tilde{y}_i, \tilde{c}_i - c_i)$. The convexity captures the idea that detection is increasingly likely for large amounts evaded. Finally, the firm faces the tax rate $\tau$ that applies to declared profit, $\tilde{\pi} = \tilde{y}_i - \tilde{c}_i$. The firm's expected profits are therefore:

$$E\pi_i = y_i - c(y_i) - \tau.(\tilde{y}_i - \tilde{c}_i) - R(y_i - \tilde{y}_i, \tilde{c}_i - c_i) \quad (1)$$

To generate heterogeneity and simplify the exposition we make two more assumptions. First, we assume that the cost function takes the following form $c(y_i; \phi_i, \alpha_i) = \alpha_i + \frac{k(y_i)}{\phi_i}$, where $\alpha_i$ is the

[8] We do not pretend that corporate taxation is generally non-distortionary but make this assumption for the tractability of the model. The corporate tax is non-distortionary in a cost of capital model (Jorgenson and Hall 1967) with immediate expensing: if all costs, including returns to capital, are immediately deductible, then the corporate income tax is a tax on pure profits and does not impact production decisions.

[9] Resource costs from evasion include, forgoing business opportunities with formal firms, keeping multiple sets of accounting records and limiting interactions with the financial sector. See Chetty (2009) for a discussion.

fixed cost (equivalent to a demand shifter) and $\phi_i$ is a productivity parameter, which scales variable costs $k(y_i)$. Second, we assume that the cost of evasion function is separable in revenue and cost evasion such that $R(y_i - \tilde{y}_i, c_i - \tilde{c}_i) = h(y_i - \tilde{y}_i) + g(\tilde{c}_i - c(y_i))$. Under these conditions the firm's expected profits are:

$$E\pi_i = y_i - c(y_i; \phi_i, \alpha_i) - \tau.(\tilde{y}_i - \tilde{c}_i) - h(y_i - \tilde{y}_i) - g(\tilde{c}_i - c(y_i)) \tag{2}$$

The firm maximizes expected profits, by choosing the triple of revenue to produce, revenue to declare and costs to declare $(y_i, \tilde{y}_i, \tilde{c}_i)$. An interior optimum satisfies the following first order conditions:

$$1 = \frac{k'(y_i)}{\phi_i} \tag{3}$$

$$h'(y_i - \tilde{y}_i) = \tau \tag{4}$$

$$g'(\tilde{c}_i - c_i) = \tau \tag{5}$$

Equation (3) determines the revenue produced $y$. Since, in our model, taxation is non-distortionary, the production decision is independent of the tax rate. Equations (4) and (5) state that the marginal return to revenue and cost evasion, $\tau$, equals the marginal cost, which is a function of the amount evaded. Firm revenue is a function of its productivity draw $\phi_i$ but independent of the fixed cost draw $\alpha_i$, such that $\frac{dy_i^*}{d\phi_i} > 0$ and $\frac{dy_i^*}{d\alpha_i} = 0$. Firm costs are given by $c^*(y^*; \phi_i, \alpha_i) = \alpha_i + \frac{k(y^*)}{\phi_i}$ and depend both on the productivity draw $\phi_i$ and the fixed cost $\alpha_i$, such that $\frac{dc_i^*}{d\phi_i} > 0$ and $\frac{dc_i^*}{d\alpha_i} > 0$. Finally, we define profit margin as profit over revenue, $\pi_{margin} = \frac{y^* - c^*}{y^*}$, and is determined jointly by $\phi_i$ and $\alpha_i$.

Under a continuous and differentiable joint distribution of productivity and fixed cost parameters $f_0(\phi, \alpha)$ the distribution of revenue and cost is smooth. We assume that the cost of evasion functions $h(y_i - \tilde{y}_i)$ and $g(\tilde{c}_i - c_i)$ are continuous and differentiable and therefore the distributions of reported revenue, reported costs and reported profit margins are also smooth.

### 2.3 Theoretical framework: Impact of the tax system

A noteworthy aspect of Costa Rica's corporate schedule is that the average tax rate applied on profits increases from $\tau$ to $\tau + d\tau$ when firms declare revenue above the threshold $y^T$. Tax liability is a function of declared revenue $\tilde{y}$ and declared costs $\tilde{c}$:

$$
\begin{aligned}
&T(\tilde{y}_i - \tilde{c}_i; \tilde{y}_i) = \tau(\tilde{y}_i - \tilde{c}_i) && if \quad \tilde{y}_i \leq y^T \\
&T(\tilde{y}_i - \tilde{c}_i; \tilde{y}_i) = (\tau + d\tau)(\tilde{y}_i - \tilde{c}_i) && if \quad \tilde{y}_i > y^T \\
&T(\tilde{y}_i - \tilde{c}_i; \tilde{y}_i) = 0 && if \quad \tilde{y}_i - \tilde{c}_i \leq 0
\end{aligned}
\tag{6}
$$

We consider that the above tax system is imposed as a tax reform over a previously flat corporate tax at rate $\tau$. Since only the productivity parameter $\phi_i$ determines firm revenue and all firms face the same cost of evasion, there exists a productivity threshold $\underline{\phi}$ such that a firm with productivity $\phi_i = \underline{\phi}$ reports revenue exactly equal to the threshold $\tilde{y}_i = y^T$, and all firms with $\phi_i \leq \underline{\phi}$ declare revenue below the threshold $\tilde{y} \leq y^T$. These firms are not affected by the tax change. For firms with $\phi_i > \underline{\phi}$ there are two possible responses: (1) reduce revenue, declared or real, by an amount such that the new revenue equals the threshold (the "bunchers") or (2) stay above the threshold and face a higher tax rate. These firms then change their reporting revenue and cost such that the marginal cost of evasion equals the new tax rate.

Firms choose one of two responses depending on their *productivity* and *fixed cost* draw: for every productivity draw $\phi_i$ in an interval $[\underline{\phi}, \overline{\phi}_{max}]$ there exists a fixed cost $\alpha_i$ such that all firms within the interval $[\underline{\phi}, \overline{\phi}_\alpha]$ bunch at the threshold. $\overline{\phi}_\alpha$ is determined by the indifference condition between expected profits at the threshold and expected profits at the interior solution, $E\pi^{Threshold}(y, \tilde{y}^T, \tilde{c}|\overline{\phi}_\alpha, \alpha) = E\pi^{Interior}(y', \tilde{y}', \tilde{c}'|\overline{\phi}_\alpha, \alpha)$. Firms with $\phi_i > \overline{\phi}_\alpha$ remain above the threshold and adjust their reporting behavior.

To illustrate the effect of costs on the bunching decision we consider a firm with productivity $\phi_i > \underline{\phi}$ and fixed costs $\alpha_i$ mapping into true revenue and cost $(y_0, c_0)$ and reported revenue and cost $(\tilde{y}_0, \tilde{c}_0)$ such that $\tilde{y}_0 > y^T$ before the tax change. To reach the threshold the firm can reduce declared income with a combination of real and reporting behavior. Real income reduction is $dy$ and reported income is $d\tilde{y}$. The total change in revenue is $\Delta y = dy + d\tilde{y}$ such that $\Delta y$ is the revenue distance to the threshold, $\Delta y = \tilde{y}_0 - y^T$. We compare the firm's utility when it reports revenue at the threshold versus when it report its pre-tax change revenue - and approximate the expected gains from bunching as:

$$
E\ Gains \approx d\tau(y^T - \tilde{c}_0) + \Delta y(\tau + d\tau) - d\tilde{y}.h'(y_0 - \tilde{y}_0 + d\tilde{y}) - dy.[1 - c'(y_0 - dy)] \tag{7}
$$

Where we have used the envelope condition and ignored intensive margin changes past the threshold. The *first term* of equation (7) is a noteworthy feature of the Costa-Rican setting: it shows that the gains from lowering revenue to reach the threshold are proportional to the change in the tax rate $d\tau$ and to the firm's declared tax base at the threshold, $y^T - \tilde{c}_0$. Therefore variation in cost, due to fixed cost heterogeneity, generate different incentives to bunch for firms of equal productivity.

The other terms of equation (7) state that the firm directly gains by not paying taxes on undeclared and non-produced revenue $\Delta y$, but incurs larger resource costs, due to the additional revenue under-reporting (evasion responses) and looses profit due to its lower production level (real responses). Note that if all responses are due to revenue under-reporting, equation (7) simplifies to: $EGains \approx d\tau(y^T - \tilde{c}_0) + d\tilde{y}(\tau + d\tau) - h'(y_0 - \tilde{y}_0 + d\tilde{y})$

**Prediction 1: Bunching at the revenue thresholds**

From the distribution of productivity and fixed cost parameters $f(\phi, \alpha)$ we obtain a direct mapping into the distribution of declared revenue and declared costs $\psi_0(\tilde{y}_0, \tilde{c}_0)$ such that the total number of firms bunching at the revenue threshold is:

$$B = \int_{\tilde{c}_0} \int_{\tilde{y}_0 = y^T}^{y^T + \Delta y(\tilde{c}_0)} \psi_0(\tilde{y}, \tilde{c}) d\tilde{y}.d\tilde{c} \tag{8}$$

With knowledge of the joint distribution of revenue and cost we can estimate the elasticity of revenue $\epsilon_y$ that generates a given amount of bunching.

Absent the counterfactual cost distribution we can still estimate the revenue response of the marginal buncher, defined as the firm with the maximal revenue change. For firms with the same revenue distance to the threshold the marginal buncher is the firm with the lowest declared costs: given a support of costs $[\underline{c}_0; \bar{c}_0]$ the marginal buncher's revenue response is $\Delta y^{mb} = \Delta y(\tilde{c}_0 = \underline{c}_0)$. With knowledge of the lower support of the distribution of $c_0$, we can identify the revenue response of the marginal buncher as the maximum revenue response, which in the model corresponds to the response of the firm with the lowest costs. Under homogeneous revenue elasticities the marginal buncher's revenue elasticity and the average revenue elasticity are the same.

**Prediction 2: Missing mass above the thresholds but no strictly dominated region**

A corollary to the first prediction states that some revenue intervals past the threshold display missing density, which corresponds to the excess density at the threshold. At each revenue level past the threshold, the missing density is a function of the distance to the threshold and the cost distribution at that revenue level. In the standard notch setting (Kleven and Waseem 2013) there is a deterministic dominated revenue interval just above the threshold: firms that report revenue in that interval are making an irrational decision under any preferences, since lowering production would increase their after tax profits. Whereas, for the Costa Rican notches, only a subset of firms with sufficiently low costs are dominated. For example, a firm with zero profits has no incentives to lower its revenue since its tax liability is already null. Being dominated is not only a feature of the revenue distance to the threshold but also of costs and hence firm specific. As a consequence, even in a frictionless world, there will be firm density in revenue intervals just past the threshold.

**Prediction 3: Increased Revenue and Cost Evasion Past the Thresholds**

Infra-marginal firms do not bunch at the revenue threshold but face an increase in the marginal return to evasion, which jumps from $\tau$ to $\tau + d\tau$. They respond to the tax hike by increasing revenue and cost evasion such that the marginal resource costs of each evasion type equals the new tax rate. As a consequence firms above the threshold declare less revenue and more costs than under the counterfactual. As a consequence observed profits and profit margins by revenue, jump downwards discontinuously at the threshold.

**Prediction 4: Excess Profit at the Thresholds (Under evasion responses)**

On the one hand, firms selecting into bunching have higher profit than the average firm (Selection effect). On the other hand, by lowering declared revenue to reach the threshold, bunchers lower their profits (Evasion effect). Theoretically, the average declared profit margin of firms at the threshold could be higher or lower than that of firms below the threshold, depending on the variance of the distribution of costs, in the revenue bins above the threshold. Under homogeneous costs (no variance) then the Evasion effec dominates and the average observed profit margin of bunchers is lower than that of firms below the threshold. With sufficient heterogeneity in the cost distribution the Selection effect dominates and bunching firms display excess profit at the threshold. The domination of the selection effect is better understood from equation 7: while gains from bunching are linear in the firm's costs, the cost of bunching are convex in the firm's revenue distance to the threshold, due to the convexity of the resource cost of evasion. Therefore, for a sufficiently large revenue distance to the threshold, the revenue change is small compared to the cost difference between selected bunchers and the average firm.

## 3 Behavioral responses and tax elasticities

The goal of this section is to estimate the firms' elasticity of profit with respect to the net of tax rate and separate the profit response between changes in revenue versus changes in costs. To this end, we develop an estimation procedure that takes into account the intertwined revenue and cost responses, without assuming a functional form for the firms' utility. Elasticities are defined with respect to the net of tax rate and when discussing size and bounds we refer to their absolute value. We initially make few assumptions and recover an upper bound on the revenue elasticity and a lower bound on the cost elasticity. In section (4) we add structure and obtain tighter estimates for each of the revenue and cost elasticities, under a model-based numerical estimation. The profit elasticity is stable across the different estimation strategies. We summarize the different methodologies, elasticity estimates and assumptions in Table 1.

The estimation relies on two assumptions: absent the tax increase past the threshold, the distribution of firms by revenue would be smooth and continuous and therefore it can be approximated by a flexible polynomial[10] and average cost by revenue would be continuous. In a first step, we use bunching at the revenue thresholds to identify the change in revenue for the marginal buncher. In a second step, we use the discontinuity in average cost by revenue, on both sides of the thresholds, to estimate the cost response. The cost discontinuity is adjusted to take into account the revenue response estimated in the first step. It recovers the average increase in declared cost at the threshold caused by the tax rate increase. In a third and final step, we combine the revenue and cost responses to compute the profit response at the threshold.

### 3.1 Data

We base our study on administrative data from the Ministry of Finance (Ministerio de Hacienda) and have access to the universe of corporate tax returns over the 2008-2014 period. All registered corporations are required to submit the yearly tax declaration D101 ("Declaracion Jurada del Impuesto Sobre la Renta") and report their profits, revenue and costs. The tax declaration can be filled electronically since 2008, and in recent years a large majority of firms opted for this format. The data consists of 617,929 firm-year observations and 222,352 unique firms. As a whole, the

[10]The ability to approximate the counterfactual density with a flexible polynomial is the standard assumption in the bunching literature and we show that in our data the parametric choices (polynomial order, bunching interval limits, etc.) have little effect on estimated parameters.

corporate income tax raises slightly less than 20% of total tax revenue, which is equivalent to 3% of GDP. The firms we study are small and micro enterprises with yearly revenue below 150 million Costa Rican Colones ($450,000 in PPP). They represent 83% of the 80,000 firms filling taxes in a given year and declare 25% of total profits, which generates 15% of corporate tax revenue.

Figure 3 presents the key features of the data by revenue bins of half million CRC, pooling all years together. Panel A shows the number of firms by revenue. We observe a clear excess mass below each revenue threshold and missing mass just above, as predicted by theory. Panel B shows the average profit margin by revenue[11], where profit margin is defined as profit over revenue. Profit margin by revenue resembles a downward step function - constant within a given tax bracket and jumping down at the thresholds. Average profit margin within the first tax bracket is 16%, 7-8% in the second bracket and 4-5% in the third bracket. We also observe that firms reporting revenue at the thresholds display profit margins in excess of 22% and 9%, respectively at the first and second thresholds. As discussed in theory (Prediction 4), this could arise from the selection into bunching of firms with low costs.

The reduced form estimation uses the distributions of Figure 3 to estimate profit elasticities and separate them between revenue and cost responses. Intuitively, the excess mass at the revenue thresholds provides evidence of revenue responses while the jump in profit margin combines revenue and cost responses. Therefore, in a first step we apply the bunching methodology to the firm density to estimate revenue elasticities. In a second step, we use the discontinuities in profit margin on either sides of the thresholds to estimate the cost elasticities, holding constant revenue responses. In a third step, we combine the revenue and cost responses to obtain profit elasticities.

## 3.2 Revenue elasticity estimation

### 3.2.1 Bunching methodology

To estimate the revenue elasticity, we use the distribution of firms by revenue and the point of convergence method described in Kleven and Waseem (2013)[12]. We slice the data in half million CRC bins. To obtain the counterfactual density, we fit a flexible polynomial of degree five[13]

[11]Figure A1 shows average profits and average costs by revenue. We choose to present profit margin since it is unit free and very stable within tax bracket, in our data. It therefore highlights the large discontinuity at the threshold.

[12]The notch estimation builds upon the kink method of Saez (2010) and Chetty et al. (2011)

[13]The order of the polynomial is chosen to maximize Akaike's criteria. Table A1 shows the impact on the results of using different orders of polynomial.

$$F_j = \sum_{k=0}^{5} \beta_k.(y_j)^k + \sum_{i=y_l}^{y_u} \delta_i.\mathbb{1}(y_j = i) + \nu_j \tag{9}$$

where $F_j$ is the number of firms in revenue bin $j$, $y_j$ is the revenue midpoint of interval $j$, $[y_l, y_u]$ is the excluded region and $\delta_i$'s are dummy shifters for the excluded region. We use the estimated $\beta_k$'s to obtain the counterfactual firm distribution by revenue absent the tax change:

$$\hat{F}_j = \sum_{k=0}^{5} \hat{\beta}_k.(y_j)^k \tag{10}$$

The estimation procedure requires that the excess mass below the threshold (E) equals the missing mass past the threshold (M), defined as:

$$\hat{E} = \sum_{j=y_l}^{y^*} (F_j - \hat{F}_j) \qquad and \qquad \hat{M} = \sum_{j=y^*}^{y_u} (\hat{F}_j - F_j) \tag{11}$$

Where $y^*$ is the revenue threshold and the bounds of the excluded region $[y_l, y_u]$ are obtained as follows: the lower limit $y_l$ is *chosen* by the researchers as the revenue bin where excess density starts appearing[14]. The upper limit, $y_u = y^* + dy$, is *estimated* using the identity that the excess mass (E) has to equal the missing mass (M). Starting from $y_u$ just above the threshold, we estimate equation (9) and compute $\hat{E}$ and $\hat{M}$. For a low value of $y_u$, the excess density is much larger than the missing density ($\hat{E} > \hat{M}$). We iteratively increase $y_u$ until the excess mass equals the missing mass ($\hat{E} = \hat{M}$). The estimated upper bound, $y_u$, is the revenue of the marginal firm responding to the tax change. Under heterogeneity in revenue elasticities, this is the response of the highest elasticity individual and therefore provides an upper bound on the revenue response.

By forcing the excess mass to equal the missing mass, this estimation method generates two potential concerns[15]. First, it assumes that there are no extensive margin responses. Extensive responses could occur if firms decided to become informal when faced with higher tax rates. This would generate additional missing mass past the threshold and imply that $E < M$. In our setting extensive margin responses should play a limited role, as Costa Rica is one of Latin America's country with the least informalityILO 2012), and it is unlikely that firms with growing revenue and already registered decide to reverse back to informality, after increasing revenue past the threshold.

[14]We show in table A1 that changing $y_l$ has negligible impacts on the results.
[15]Those limitations are also noted by Kleven and Waseem (2013).

In terms of results, if extensive margin responses are present then the true revenue elasticity is smaller than the estimated one. Second, the standard bunching method ignores intensive margin revenue responses past the threshold. Intensive responses imply that the revenue of firms above the threshold is lower than the counterfactual, estimated from the polynomial fit. In our estimation, we take into account this second order effect: we shift the counterfactual distribution above the threshold by the factor implied from the estimated revenue elasticity. The intensive margin adjustment occurs simultaneously with the point of convergence method and the iterative process of determining the upper bound on revenue $y_u$. In our setting where elasticities are substantial, this adjustment does have a modest impact on the results, reducing slightly the estimated revenue response.

In the case of a notch, and in particular of a notch with two-dimensional incentives, revenue distance to threshold and costs, obtaining the change in the marginal tax rate is less straightforward than with a kink. Given the tax liability $T(y-c;y)$, we define the implicit marginal tax rate $\tau^*$, for an increase in revenue $dy$, as the change in tax liability over the change in revenue:

$$\tau^* = \frac{T(y^*+dy)-T(y^*)}{dy} = \frac{(\tau_0+d\tau)(y^*+dy-c)-\tau_0(y^*-c)}{dy}$$

$$\tau^* = (\tau_0+d\tau) + \frac{d\tau(y^*-c)}{dy} \tag{12}$$

Where $\tau_0$, $d\tau$ and $y^*$ are known parameters and $dy$ is estimated with the bunching point of convergence method. However, the cost of the marginal buncher $c$ is unknown. From the theory section we know that the marginal buncher is the firm with the lowest cost, within its revenue bin. Therefore, the marginal buncher should have costs in the 1st percentile of the cost distribution for its revenue bin. To ensure that we obtain an upper bound on the revenue elasticity we assume the cost of the marginal buncher are at the 10th percentile. With this assumption, we obtain $c$ and can compute the implicit marginal tax rate $\tau^*$, given the estimate of the revenue response $dy$. The revenue elasticity is then defined as:

$$\epsilon_{y,1-t} = \frac{\%change\ revenue}{\%change\ (net\ of\ tax\ rate)} = \frac{dy}{y^*}.\frac{(1-t_0)}{(t^*-t_0)} \tag{13}$$

### 3.2.2 Bunching results

Figure 4 shows the firm distribution by revenue and the estimated counterfactual density, obtained from the polynomial fit around each threshold[16]. The estimated parameters are displayed in the top right corner of each panel. For the first threshold (Panel A), the excess mass is 2.3 times the counterfactual, meaning that there is 3.3 times the density that should be expected. In the absence of the notch, the marginal buncher would have an income of 58.3 million CRC, 16% higher than the threshold. For the second threshold (Panel B), the excess mass is 1.1 times the counterfactual and the marginal buncher has revenue of 107.7 M CRC, 7.6% higher than the threshold.

Given our estimates of revenue responses, we compute the implicit marginal tax rate faced by the marginal buncher, with equation (12) and obtain a revenue elasticity with respect to the net of tax rate of 0.33. This implies that for a 10% reduction in the net of tax rate, firms respond by reducing reported revenue by 3.3%. At the second threshold, the elasticity of revenue is 0.08. Table 3 reports the parameters used for the revenue estimation and the resulting revenue elasticities. Standard errors are estimated from a 1,000 bootstrap iterations from the residuals of the polynomial fit[17].

Despite being graphically compelling, the large behavioral responses to the revenue notches produce moderate revenue elasticities. Three points are worth mentioning. First, on a small profit base a modest change in revenue could generate a large profit elasticity, holding costs constant. Second, notches differ from kinks in that they generate sizable changes in implicit marginal tax rates and therefore large behavioral responses are consistent with moderate elasticities. Third, firms can also reduce their tax liability by increasing reported costs and therefore changing revenue is only one of two possible margins of response to an increase in the tax rate. We investigate the latter point in the next section.

## 3.3 Cost discontinuity

In Figure 3, Panel B, we showed the step-like pattern of average profit margins by revenue. Profit margins by revenue are visually attractive since unit free and, in our data, very stable within tax

[16] Due to the intensive margin adjustment above the threshold, the counterfactual does not exactly fit the observed density for revenue bins above $y_u$.

[17] Since the data contains the universe of corporate tax returns, there is no sampling error in the firm distribution. The source of uncertainty arises from the functional form of the polynomial. When running the bootstrap iterations we therefore draw from the sample of residuals of equation 9 and obtain new firm densities, with which we repeat the point of convergence method

brackets. However, to quantify the jump in costs at the threshold, caused by the tax rate increase, we turn to the relation between reported cost and revenue. Figure 5 plots the average reported costs by revenue at the first threshold. Importantly, some firms have selected into the revenue range around the threshold, as a function of their costs. From the bunching analysis, we know that selection occurs precisely in the revenue bins corresponding to the excess and missing mass intervals, $[y_l, y_u]$. Therefore, when analyzing the cost discontinuity, we exclude these intervals, with dummy variables for the excess and missing mass areas. We measure the discontinuity in cost at the threshold with the following specification:

$$cost_j = \alpha + \delta.\mathbb{1}(\tilde{y}_j > 0) + \beta_1.\tilde{y}_j + \beta_2.\tilde{y}_j \mathbb{1}(\tilde{y}_j > 0) + \sum_{j=\tilde{y}_l}^{\tilde{y}_u} \gamma_j \mathbb{1}(\tilde{y}_j = j) + \epsilon_j \tag{14}$$

where $cost_j$ is firms' average cost in bin $j$, $\tilde{y}_j = y_j - y^*$ is the revenue distance to the threshold and $\gamma_j$ are dummy shifters for firms with revenue in the excluded excess and missing mass intervals. $\beta_1$ provides the slope of average cost on revenue below the threshold and $\beta_1 + \beta_2$ the slope past the threshold. The parameter of interest is $\delta$, the discontinuity in costs at the threshold. The specification directly provides the percentage change in cost at the threshold as $dc = \frac{\delta}{\alpha}$.

Our objective is to measure the discontinuity in costs, holding revenue responses constant. However, the cost discontinuity estimated from equation (14) could entirely be due to intensive margin responses of revenue. To understand this, note that the "running" variable is revenue, which is also distorted by the change in the tax rate: absent the tax change, firms in the upper tax bracket would have declared larger revenue. However, since we have estimated the revenue elasticity in Section (3.2), we can adjust for the intensive margin revenue change. To illustrate the revenue adjustment, we consider firms belonging to revenue bin $j$, with revenue midpoint $y_j$. Absent the tax change their counterfactual revenue would be:

$$\begin{aligned} y_j^{adj} &= y_j & if \quad y_j \leq y^* \\ y_j^{adj} &= y_j + \epsilon_{y,1-t}.y.\frac{dt}{1-t} = y_j + 0.33 * y_j * \frac{0.1}{0.9} = y_j * 1.037 & if \quad y_j > y^* \end{aligned} \tag{15}$$

We clarify three aspects of the revenue adjustment. First, the adjustment only applies to firms with revenue above the threshold. Second, for firms with revenue sufficiently above the threshold the increase in the average tax rate is equivalent to an increase in the marginal tax rate. Since we exclude firms from the missing mass interval, this holds for the vast majority of firms. Third, the revenue adjustment assumes that the revenue elasticity is the average revenue elasticity. Since under heterogeneity in revenue responses the revenue elasticity is an upper bound of the average elasticity, the revenue adjustment is an upper bound of the true adjustment. We return to this point when interpreting the cost elasticity.

We apply the revenue adjustment, that is we replace $\tilde{y} = y_j - y^*$ with $\tilde{y}_j^{adj} = y_j^{adj} - y^*$, and then estimate equation (14). $\delta$ now measures the increase in reported costs due to the tax change, but holding revenue responses constant. The results, with and without, the revenue adjustment[18] are reported in Table 2 and displayed for the first threshold in Figure 5. Figure 5, Panel A plots average cost by revenue at the first threshold. It also shows average cost by revenue after the revenue adjustment, which shifts costs horizontally for firms past the threshold. We then fit separate lines to the right and to the left of the threshold, excluding the interval affected by bunching responses between $z_l$ and $z_u$. The linear extrapolation to the threshold on the left provides a counterfactual average cost for firms at the threshold under a 10% tax rate and absent the notch. The linear extrapolation to the right, provides the average cost for a 20% tax rate, assuming no revenue responses. The resulting discontinuity is the change in reported costs that arises at the threshold due to the change in the tax rate. In Panel B, we zoom in on the discontinuity in the predicted average cost at the first threshold. We observe a large jump in average cost of 2.39 million on a cost base of 41.9 million. Given the net of tax rate increase of 11%, the elasticity of cost is:

$$\epsilon_{c,1-t} = \frac{dc}{c^*}.\frac{(1-t_0)}{dt} = \frac{-2.39}{41.9} * \frac{0.9}{0.1} = -0.51$$

This implies that when the *net of tax rate* is reduced by 10%, firms respond by increasing their reported costs by 5.1%. At the second threshold costs jump by 1.1 Million on a 92 Million base. Together with the net of tax rate increase of 12.5% this implies a cost elasticity of -0.11.

The estimation is equivalent to a donut RD, with a local linear fit. Linearity is an important assumption to which we provide support in Appendix A. Figure A2 shows the linear and quadratic fits of average costs by revenue, above and below each threshold. The quadratic fit is indistinguishable from the linear fit. Table A2 displays the adjusted R-squared from the linear, quadratic and cubic regressions and shows that the linear model has the highest adjusted R-squared. In Table A3

[18] The revenue adjustment method shifts horizontally average costs, such that under a sufficiently large revenue elasticity, the entire cost discontinuity could arise due to intensive margin revenue responses. For this to be the case, the elasticity of revenue would have to be 0.79 at the first threshold and 0.19 at the second, slightly under three times what we estimated.

we present the results of Equation (14) using a quadratic fit: the cost discontinuity is even larger than under the linear model, and therefore if the linearity assumption introduces bias, we would be underestimating the discontinuity in cost an the cost elasticity. In addition, table A3 shows that the results are robust to variation in the revenue interval used to estimate the model and to the assumption that the revenue elasticity falls with revenue[19].

### 3.4 Profit elasticity

By combining the revenue and cost responses, we can now estimate the elasticity of profit with respect to the net of tax rate. The elasticity of profit is the central parameter to set optimal tax rates and a sufficient statistic for revenue collection under a flat tax rate. It is defined as:

$$\epsilon_{\pi,1-t} = \frac{\%\ change\ profit}{\%\ change\ (net\ of\ tax\ rate)} = \frac{\Delta\pi}{\pi} * \frac{1-\tau}{\Delta\tau} = \frac{(\Delta y - \Delta c)}{\pi} * \frac{1-\tau}{\Delta\tau} \tag{16}$$

We already estimated the change in cost at the threshold $\Delta c$ and compute the change in revenue $\Delta y$ using the revenue elasticity: $\Delta y = y * \epsilon_{y,1-t} * \frac{\Delta t}{1-t}$.

Table 3 summarizes the elasticity estimates and changes in revenue, cost and profit, at each threshold. At the first threshold, we estimate a profit elasticity with respect to the net of tax rate of 4.7, and at the second threshold an elasticity of 2.7. These are very large elasticities and imply that the revenue maximizing rate is 17% for micro firms and 27% for small firms[20]. Tax rates above these are on the wrong side of the Laffer curve and Pareto dominated, since government revenue would fall, as the base diminishes faster than the rate increases. These large elasticities are a function of the current policy environment and of evasion and avoidance opportunities, which we investigate in further detail in Sections 5 and 6. However, we highlight in Figure 6 that the estimated elasticities correspond to an interesting reporting behavior. The figure shows average tax payment as a share of revenue on revenue: despite the 10% tax rate increase at each threshold, tax liability as a share of revenue is continuous and stable, at roughly 1.5% of revenue. Faced with a tax hike, firms adjust their reported profits such that tax payments represent a near constant

[19] The revenue adjustment uses the estimated elasticity at the threshold and applies homogeneously to all firms with revenue above the threshold. Since the revenue elasticity is larger at the first than second threshold, an alternative adjustment assumes a linearly decreasing elasticity as a function of revenue, with slope proportional to the drop in the elasticity between the first and second threshold.

[20] Under a flat corporate tax, the government revenue maximizing rate is $\tau^* = \frac{1}{1+\epsilon_{\pi,1-\tau}}$

share of their revenue.

Another novel result in Table 3 is the comparison between the cost and revenue elasticities. Roughly 60% of the discontinuity in profit is due to an increase in costs and 40% from an increase in revenue[21]. The difference is not statistically significant, but the qualitative result holds: reported costs react stronger to a change in the tax rate than reported revenue, despite estimating a lower bound on the cost elasticity and an upper bound on the revenue elasticity. In Section 4 we impose additional assumptions to obtain point-estimates of the revenue and cost elasticities.

### 3.5 Robustness and heterogeneity

We discuss three important dimensions of robustness and heterogeneity: elasticity estimates, industry variation and distributional results.

The large drop in average profit margin on either side of the threshold is the key identifying variation for the total profit response, while the two step estimation decomposes this variation into revenue and cost responses. In doing so, a near perfect negative correlation mechanically arises, due to the revenue adjustment term applied to the cost discontinuity: a larger revenue elasticity implies a larger revenue adjustment, which reduces the cost elasticity. Accordingly, the profit elasticity is robust to results from the bunching estimation and hinges upon the assumption that average profit (or cost) by revenue would have been smooth around the threshold, absent the tax change. Following the same logic, the cost to revenue elasticity ratio estimate is less robust and we only note that despite estimating a loose upper bound on the revenue elasticity, and a loose lower bound on the cost elasticity, the qualitative result that the cost elasticity is larger than the revenue elasticity holds. The relative magnitudes of the revenue and cost elasticities should be taken with more caution and our preferred estimates are those of Section (4). As a robustness test, we show the resulting cost and profit elasticities for a range of revenue elasticities in appendix A.

Some of the results could be driven by industry variation. Figure 7 plots for the first threshold the within industry density (left axis) and profit margin (right axis) by revenue, separating the economy in fifteen sectors and binning at the 1M CRC level. All industries display excess mass at the threshold, however sectors with high evasion potential such as construction, real estate and legal and economic consultancies exhibit stronger bunching than retailers and manufacturers. Most

[21] In addition, costs are measured on a smaller base than revenue, therefore the cost elasticity is larger in asbolute term than the revenue elasticity at both thresholds ($\epsilon_{c,1-t} > \epsilon_{y,1-t}$).

importantly, average profit margin drops discontinuously for all industries and the proportional downward fall in margins appears homogeneous. The only sector which does not display any discontinuity is Public Administrations and Associations.

Our two step estimation methodology relies on large sample sizes such that the firm distribution[22] and average cost are smooth. As a consequence we can not obtain precise revenue and cost elasticity estimates at the industry level. Instead, to summarize the revenue response at the industry level, we turn to the excess mass of bunchers[23] at the threshold, $\hat{E}_s$ for industry $s$. The industry excess ratio $\hat{E}_s$ is robust to the polynomial fit and a provides a proxy for the size of revenue responses. We also estimate the profit margin discontinuity for each industry using equation (14). Figure 8 plots for the first threshold the excess mass on the percentage change in profit margins by industry. On the one hand, we observe a large variation in excess mass ranging from just above one for retailers and wholesalers to six for consultancies. On the other hand, the proportional change in profit margin is very homogeneous: most sectors declare margins 40 to 50% lower above the threshold. Despite large variation in revenue responses, total profit responses are fairly homogeneous.

Finally, figure 9 shows the quartiles of profit margin. The median profit margin drops from 6% below the threshold to 3% above it and we observe similar drops in proportion at the 25th and 75th percentiles. It appears that profit margin discontinuities arise from an entire downward shift of the distribution of profit margin and not only from a change in profit of a few high profitability firms.

## 4 Model-Based Estimation of Revenue and Cost Elasticities

If we make no assumption about the counterfactual distribution of cost by revenue, then the estimation of Section 3 provides the tightest possible bounds on the revenue and cost elasticity. This estimation faces several limitations: first, under heterogeneity in revenue elasticity, it only provides an upper bound on the true revenue elasticity and a lower bound on the cost elasticity. Second, it does not take into account selection into bunching as a function of costs. Third, it does not consider the feedback effect of cost responses on revenue responses.

[22] The main limitation to apply the two step estimation method for each industry arises from the point of convergence method, which iteratively fits the polynomial to equate the excess mass with the missing mass. With small sample sizes the iterative process can fail to converge and multiple revenue bins can satisfy the equality condition.

[23] We fit a new polynomial for each industry but do not re-estimate, $y_u$ the upper bound of the excluded interval.

To address these limitations, we assume a specific counterfactual distribution of profit margins: absent the tax change, the distribution of profit margins by revenue remains constant within the revenue intervals around the threshold. We then combine the counterfactual firm density with the new counterfactual *profit margin distribution*, to determine the revenue elasticity at which the number of numerically estimated bunching firms corresponds to the observed excess mass. The additional structure allows us to model responses to the notch as a joint function of the firms' revenue distance to the threshold and costs, as suggested by theory. We also consider the impact of cost responses on the bunching decision: we assume that firms' cost responses, if not bunching, would equal the average cost response, estimated from infra-marginal firms with the discontinuity approach. The model-based estimation of bunching improves our estimates of the revenue and cost elasticities. However, it does not impact the profit elasticity, which is identified by the large profit margins discontinuity.

### 4.1 Main Estimation

To improve upon our estimation of Section 3, we impose the restriction that absent the tax change, the entire distribution of profit margin by revenue would stay constant in an interval around the threshold. Under this restriction, we use the profit margin distribution of firms with revenue below the threshold, and away from the bunching zone, as the counterfactual for firms with revenue just below or just above the threshold. Specifically, this assumes that the profit margin distribution of firms with 45 million CRC in revenue, would apply to firms with revenue 10 to 20% larger, absent the tax change. We test if this assumption holds away from the first threshold. Figure 10, Panel A, plots several distributions of profit margin for revenue intervals below the threshold. The distributions of profit margin appear extremely stable across different revenue intervals 10 to 20% lower than the threshold, and we can never reject the Kolmgorov-Smirnov tests that profit margins are generated from identical distributions. Figure 10, Panel B, also shows that within revenue intervals 20 to 30% higher than the threshold, the new distributions of profit margins are stable. Firms in these revenue intervals are not impacted by selection into bunching, therefore the stability of profit margin distributions across these intervals gives us confidence that locally constant counterfactual profit margin distributions is a reasonable assumption. Therefore, we use the distributions in Panel A as the counterfactual profit margin distribution for firms in the bunching impacted zone, under a flat 10% corporate tax rate.

By combining the constant counterfactual profit margin distribution within each revenue bin, with the distribution of firms by revenue from the polynomial fit, we obtain a joint counterfactual distribution of revenue and costs. This allows us to model selection into bunching as a common function of the firms' costs and revenue distance to the threshold. For each revenue bin past the threshold, and given an elasticity of revenue, we can compute the cost threshold at which the firm is indifferent between bunching and remaining above the threshold. To compute the cost threshold, we first return to the expression of the implicit marginal tax rate, $t^*$, and model cost responses, $dc$:

$$\begin{aligned} \tau_i^* &= \frac{T(y^*+dy_i)-T(y^*)}{dy_i} = \frac{(\tau_0+d\tau)(y^*+dy_i-c_i-dc_i)-\tau_0(y^*-c_i)}{dy_i} \\ \tau_i^* &= (\tau_0+d\tau)+\frac{d\tau(y^*-c_i)-(\tau_0+d\tau).dc_i}{dy_i} \end{aligned} \tag{17}$$

The above equation states that the implicit marginal tax rate $t_i^*$, faced by firm $i$, is a function of the firm's revenue distance to the threshold $dy_i$, costs $c_i$, and the change in reported costs conditional on facing the higher tax rate $dc_i$. To understand this later term, consider a firm that can easily evade costs: facing the higher average tax rate hardly increases its implicit marginal tax rate, since it can freely adjust its tax liability by reporting higher costs (large $dc$). On the contrary, a firm facing large resource costs of evasion on costs, can not adjust it tax liability by over-reporting costs, and therefore faces a large increase in its implicit marginal tax rate (low $dc$). With the explicit marginal tax rate we can now express the revenue elasticity as:

$$\begin{aligned} \epsilon_{y,1-\tau} &= \frac{dy}{y}.\frac{1-\tau_0}{d\tau} = \frac{dy}{y}.\frac{1-\tau_0}{\tau^*-\tau_0} \\ \epsilon_{y,1-\tau} &= \frac{(dy)^2}{y}.\frac{(1-\tau_0)}{d\tau.dy+d\tau(y^*-c)-(\tau_0+d\tau).dc} \end{aligned} \tag{18}$$

With knowledge of $dc$, then for a given elasticity of revenue, $\epsilon_{y,1-\tau}$, and distance to the threshold, $dy$, we can measure *the cost threshold,* $\tilde{c}$, such that all firms with cost lower than $\tilde{c}$ bunch. $dc$ is the change in reported costs of bunchers, had they remained above the threshold and faced the higher tax rate. In practice $dc$ is unknown: our preferred estimation assumes that the cost response of bunchers would have equaled the average cost response ($dc = d\bar{c}$). This is equivalent to say that bunchers, would have had the same cost response as infra-marginal firms, had they not selected into bunching[24]. The average cost response, $d\bar{c}$, is estimated from the cost discontinuity, adjusted

[24] In our model, under heterogeneous cost elasticities, this assumption provides an upper bound on the elasticity

for revenue responses, following the methodology of section 3.3.

We rewrite equation 18 such that the cost threshold is a function of all parameters:

$$\tilde{c}_j = y^* + dy_j - \frac{(dy_j)^2.(1-\tau_0)}{d\tau.\epsilon_{y,1-\tau}.(y^*+dy_j)} - \frac{(\tau_0+d\tau).d\bar{c}}{d\tau} \tag{19}$$

Equation 19 states that firm $i$ in revenue bin $y_j$, with distance $dy_j$ to the threshold, will bunch under revenue elasticity $\epsilon_{y,1-\tau}$, if its costs are below the cost threshold, $c_{ij} < \tilde{c}_j$. With the counterfactual revenue distribution, we know the number of firms that would have declared revenue in bin $y_j$, absent the tax change. With the counterfactual profit margin distribution, we know the distribution of costs within each revenue bin. Therefore we can numerically estimate the number of bunching firms for a given elasticity of revenue.

The estimation is an iterative procedure: the initial values of the revenue elasticity, $\epsilon_{y,1-t}^{Step1}$, and of the cost response, $d\bar{c}^{Step1}$, are taken from our estimation in Section 3. This combination of revenue and cost responses predict substantially more bunching than observed. We estimate the new revenue elasticity $\epsilon_{y,1-t}^{Step2}$, such that the number of numerically estimated bunchers equates the excess mass at the threshold. With the resulting revenue elasticity, $\epsilon_{y,1-t}^{Step2}$, we measure the average cost response from the cost discontinuity equation (Equation 14), adjusted for the newly estimated revenue elasticity, and obtain a new average cost response, $d\bar{c}^{Step2}$. $d\bar{c}^{Step2}$ is then used as the average cost response to measure a new revenue elasticity $\epsilon_{y,1-t}^{Step3}$. We iterate this process[25] until we converge to the fixed point $(\hat{\epsilon}_{y,1-t}, \hat{\epsilon}_{c,1-t})$, where the revenue and cost elasticities are consistent with each other.

We report the iteration steps in table 4 and represents graphically the last iteration in Figure 11. In Panel A, The number of bunchers are represented by the area between the elasticity curve and the counterfactual density. We show the elasticity curves for three values of the revenue elasticities. On the one hand, even under a very small revenue elasticity, some firms with revenue just above the threshold bunch, since their implicit marginal tax rate is above one. On the other hand, even with a very large revenue elasticity some firms do not bunch since they have very large costs.

of revenue: bunchers select into bunching because their revenue elasticity is large and because their cost elasticity is low. Therefore, assuming the average cost response over-estimates bunchers potential response, and under-estimates bunching for a given elasticity. However, under a non-separable resource cost of evasion function the estimation is not necessarily an upper bound. For example, if substitution across revenue and cost evasion is easy, then firms with large cost evasion can substitute for revenue evasion to reach the threshold.

[25] At each iteration we also re-estimate the counterfactual density distribution of firms, since we applied a correction for intensive margin responses. This only has second order effect

As the revenue elasticity increase firms further away from the threshold bunch. Panel B shows the number of bunching firms by revenue bins, for the equlibrium revenue and cost elasticities.

At the first threshold, the resulting revenue elasticity is 0.22, while the cost elasticity is -0.65 and the profit response is practically unchanged at 4.72. This implies that 71% of the total profit responses are due to an increase in reported costs and only 29% to a decrease in reported revenue. The share of cost to revenue responses is substantially larger than the one estimated from the point of convergence bunching method. This result was expected since under heterogeneity in revenue elasticity, the point of convergence method estimated the revenue elasticity of the highest revenue elasticity firm. Instead the model-based estimation measures the average revenue elasticity, in a frictionless environment, which mechanically gives us the average cost elasticity via the discontinuity method. This comes at the price of two additional assumptions: first the profit margin counterfactual distribution by revenue is stable around the threshold and second the counterfactual cost responses of bunchers, had they not bunched, would equal the average cost response.

## 5 Mechanisms: Evasion Responses

Behavioral responses generated by evasion, real and avoidance responses have similar consequences for government revenue collection, but entail different policy responses. It has proven difficult to separate the different mechanisms in the literature: evasion responses are by nature secretive and measuring real responses requires very precise production data linked to tax records. Using additional dimensions of the data and new data sources we perform a series of empirical test of the mechanisms. Even though those tests are far from exhaustive, we believe that taken together they provide a convincing picture that evasion is the key driver while real effects and avoidance responses appear limited.

In this section we study two direct channels to uncover evasion behavior. First we look at the percentage of firms being flagged in the internal data cross-checks and show that firms at the threshold are significantly more likely to be selected for discrepancies. Second we look at variation in audit intensity by sectors across years. We find evidence that bunching firms under higher scrutiny increase their declared revenue and move past the threshold but also increase cost by a similar amount. Those empirical tests both support the idea that reporting responses are strong drivers of revenue bunching.

### 5.1 Audits and tax corrections in Costa Rica

Costa Rica undertakes five hundred in depth taxpayer audits every year and three hundred of those are targeted to firms. As a consequence, only 0.4 % of all firms are audited in a given year. Taxpayers are selected following a risk based analysis which incorporates information from third-parties, deviation from industry averages and the taxpayers' history. Figure 15 shows the number of audits performed and the percentage of firms audited for broad revenue bins for 2009 and 2010. The small and medium firms we study have revenue to the very left of the figure and are rarely audited: just above 100 audits over two years and a 0.2 percentage chance of being audited. The percentage of audited firms increases with firm revenue to reach over 3% for large firms.

The low capacity to conduct in depth audits is partly mitigated by the extensive automatic warning system: auditors send notification letters to firms raising "red flags" in the internal data inteligence process. Specifically, anytime a computer operated system observes discrepancies between self-declared revenue and revenue based on third-parties reports, it generates a correction letter. A large part of third-party information is collected through the D151 informative tax form, which requires all individuals and firms to declare purchases and sales to the same entity when the value within the tax year is above two million Colones ($6,000 in PPP) and any commission, professional fee or rental agreement above fifty thousand Colones ($150). Other third-party information such as sales tax retentions, credit card payments and insurance policies also enter the database.

The letters sent by tax auditors ask for a correction or justification of the tax declaration in order to match the amount assessed by the tax administration. Importantly, this process does not treat bunching firms differently. For legal reasons this information can not be linked to the individual tax records, however we obtained the number of correction letters sent by revenue bins for the year 2012. Figure 16 shows the proportion of firms receiving a letter by revenue bins of two Million CRC. The green line shows the linear fit excluding the revenue intervals around the thresholds. Around a third of small Costa Rican firms receive correction letters, which highlights that tax declarations are often incomplete and that the environment is prone to evasion. Bunchers face large and significant increase in the probability of receiving correction letters compared to their expected probability: they are 8.3% more likely to receive correction letters at the first threshold and 11.5% at the second. This result highlights that bunching and revenue changes are most likely the result of tax evasion and that bunchers get noticed for inconsistencies at a higher rate but are willing to incur the expected costs[26].

Two other results are worth noticing. First, firms declaring revenue just above the threshold (potentially dominated firms) are less likely to get flagged and the joint F-test shows that the difference is significant at both thresholds. This might indicate that firms that do not adjust their revenue to the threshold do so partly because of honesty. Second the proportion of correction letters by revenue is fairly constant on either side of the threshold contrarily to profit margins. However we saw that the increase in declared costs explains a large share of the discontinuity. A possible explanation is that third-party information on costs is not sufficient to establish evasion as it only provides a lower bound on the true cost. On the contrary, third-party information on revenue provides an upper bound on true revenue, hence a clear signal of tax evasion. Carillo et al also observe this asymmetry and document it in the case of Ecuador.

### 5.2 Sectors of special audit attention

A second test of tax evasion uses the variation in audit probability generated at the industry level by the program of "Special audit attention". In 2012 the tax agency determined during the first semester of the calendar year a list of industries assigned to special audit attention, which was posted on the ministry of finance website. In practical terms it implied that the selected industries are assigned a dedicated group of auditors and that their risk of an audit increased. Industries are not randomly selected but determined by the underlying evasion risk and the industry's growth rate compared to its tax revenue growth. The twelve sectors selected in 2012 were real estate, private education, hotels and tour agencies, transport of merchandise, sale of vehicles, sports, production of pineapple, yucca, flowers and plants, casinos and betting, performances and recycling[27]. The difference in difference analysis of firms within the audit sectors versus other sectors shows significant growth in reported profits following the assignment of the sector to "special audit attention". However it is difficult to establish causality due to the endogenous selection mechanism.

Instead we use a triple difference strategy to study firms' evasion behavior at the threshold: we compare the change in revenue, costs and profits reported by bunchers in the sectors of special

[26] Measuring precisely the expected costs from tax evasion is a difficult task and from our discussions it is unclear how many firms actually make significant adjustments to their tax payment following a correction letter. Firms can certainly revise their tax payment at minimal cost and do not get systematically prosecuted. However failure to comply increases the risk of an in depth audit which is considered to be very costly for the firm.

[27] The Sectors selected in 2013 were quasi-identical and therefore we do not perform a sector event study for each year of the variation

audit attention with bunchers in other sectors and non-bunchers in the same sectors. Our hypothesis is that bunchers are evading more revenue compared to smaller firms and dominated firms. Therefore, when faced with a higher audit probability, bunching firms should lower revenue evasion and increase reported revenue by more than non-bunchers. We estimate the following equation:

$$y_{ist} = \alpha_i + \beta * Bunch_i * Audit_j * Post_t + \gamma * Bunch_i * Post_t + \delta * Audit_s * Post_t + Post_t + \epsilon_{ijt} \quad (20)$$

Where depending on the specification $y_{ijt}$ is revenue, costs or profits of firm $i$ in sector $j$ at time $t$. *Bunch* is a firm level dummy equal to one if the firm declared revenue in the two Million revenue interval below the threshold in 2011 and zero otherwise. *Audit* is a sector level dummy that equals one if the firm belongs to the sectors of special audit attention and zero otherwise. *Post* is a time dummy equal to one in 2012 and 2013 and zero in 2011.

Table 8 presents two sets of results: one using firms with revenue in 2011 below the bunching interval (declared revenue 4 to 8 Million CRC below the threshold) as a control group in columns (1)-(3) and one using dominated firms (revenue 0 to 3 M Colones above the threshold) as a control group in columns (4)-(6). Columns (1) and (4) presents the main result: bunchers in special audit sectors increase revenue by 10% or more of their initial size compared to non-bunchers. However they simultaneously increase their reported costs by a large amount in columns (2) and (5), leading to a statistical significant decrease in declared profits of above 1 M Colones column (3) and (6). Placebo treatment effects for previous years are not significant on any of the three margin.

Those results support the idea that bunching firms are evading revenue since bunchers belonging to audit sectors increase their declared revenue substantially more than firms in the control groups following an increase in their audit probability. Interestingly, reported cost also increase by a large amount, which indicates that away from the bunching segment firms lower their tax liability by increasing cost. Firms with revenue just above the threshold have strong incentives for revenue evasion in order to decrease both their tax base and tax rate, while for firms with revenue further past the threshold it is equivalent to under-report revenue or over-report costs which only decreases the tax base.

## 6 Mechanisms: dynamic, real and avoidance responses

In this section we investigate four other dimensions of firm behavior: dynamic responses, firm division, labor input usage and time-shifting of monthly revenue.

First, we show that growing firms weakly increase their profit margin conditional on staying in the same tax bracket. However when changing tax bracket, firms display donward jumps in profit margin mirroring the cross-sectional results. Second, we investigate whether larger firms divide themselves and create subsidiaries with revenue below the threshold on which to offload their profits. We show that this is unlikely to be an important mechanism. Few firms repeatedly bunch which would be a prediction of this model. More importantly a dataset of economic groups collected by the central bank shows very limited excess profit of subsidiaries when compared to non-subsidiaries and no excess mass of subsidiaries at the threshold. Third, we explore responses to employment and wage bill using data from social security, which is considered to be well-reported information. Neither employment nor wage bill show any discontinuity at either threshold. Finally, looking at monthly revenue declared for sales tax we do not observe reduced economic activity in the last month of the fiscal yea nor any time-shifting to the first month of the next fiscal year.

### 6.1 Dynamic Responses

Does the cross-sectional result of discontinuous profit margins also apply to firms dynamic reporting behavior? To answer this question we use the panel dimension of the data and look at the difference in reported profit margin as a function of the firm's tax bracket in a given year. Figure 12 shows the average profit margin difference between years t+1 and t, conditional on firms tax bracket in those years. Firms remaining within the same bracket in consecutive years do not change on average their profit margins. Whereas, growing firms jumping to a higher tax bracket declare lower profit margins and symmetrically, shrinking firms falling to a lower bracket declare higher profit margins.

A fair critique is that profit margins could decrease with revenue for structural reasons[28]. To investigate this claim, we regress firm revenue on profit margins controlling for the tax bracket:

$$margin_{it} = \alpha_i + \gamma_t + \beta y_{it} + \delta \mathbb{1}(\tau_{it} = \tau + d\tau) + \sum_{j=\tilde{y}_l}^{\tilde{y}_u} \psi_j \mathbb{1}(\tilde{y}_j = j) + \epsilon_{it}$$

Where $\alpha_i$ and $\gamma_t$ are respectively firm and year fixed effects, $y_{it}$ is the revenue of firm $i$ at time $t$, $\tau_{it}$ is the average tax rate faced by firm $i$ at time $t$ and the dummy variables are shifters for the revenue intervals impacted by bunching. Table (6) presents the results from the above regression. The coefficients on revenue $\hat{\beta}$ shows that conditional on staying within the same tax

[28] In the model this corresponds to a positive correlation between changes in productivity $\phi$ and changes in fixed costs $\alpha$.

bracket, growing firms increase their declared profit margins rejecting the possibility that profit margin should be falling for a growing firms. The dummy coefficients for changing tax bracket measures the discontinuity that occurs at the threshold and are significant and negative at each threshold. A firm growing past the first threshold decreases its profit margin by 3.06% and a firms growing past the second threshold decreases its profit margin by 0.86%. For these results to be consistent with real responses under distortionary taxation firms would need to have increasing returns to scale such that lowering production would also lower profit margin.

### 6.2 Firm Division

Given the design of the corporate tax system, it seems attractive for a large "mother" firm to create a subsidiary firm on which to "offload" its profits: the small subsidiary then declares high profits, taxed at a 10% rate, while the larger firm declares low profits, taxed at a 30% rate. If there are large profits to shift, then the subsidiary should locate its revenue just below the threshold. Therefore, sufficient firm division and profit shifting could explain the revenue bunching and the discontinuities in profit margins. To test for firm division we use the registry of economic groups, a unique dataset, compiled by the Central Bank exclusively for statistical purposes. It links corporate groups and their subsidiaries by combining the registry of corporate ownership, the census and direct visits and calls to firms offices[29].

Firms operating under common ownership are defined as forming part of an economic group. On the one hand, shared ownership structure of firms can exist for structural reasons and the existence of subsidiaries does not provide in itself evidence of tax avoidance. On the other hand, if avoidance motivations are important, the following hypothesis should hold: First, firms in the 10% tax bracket should be more likely to be subsidiaries compared to firms in the 20% tax bracket since they represent better tax instruments, while subsidiaries in the 30% tax bracket do not serve any tax goal. Second, there should be an excess number of subsidiaries with revenue in the bunching interval and few subsidiaries with revenue just above the threshold: if subsidiaries are tax-related vehicles, then changing their declared revenue should produce minimal resource costs, while generating large tax gains. Third, the profitability of subsidiaries should be large since these are profit-shifting vehicles while the profitability of mother firms should be low.

[29]This data project, named REVEC, was motivated by the re-estimation of the input-output matrix for Costa Rica in 2012 and to obtain an accurate view of corporate ownership structure. We provide additional information on this dataset in Appendix B.

Figure 14 shows the share of subsidiaries by revenue bin and fits a linear relation on both sides of the threshold, excluding revenue bins in the bunching and dominated intervals. On average 4 to 5% of firms are subsidiaries of larger firms, however the relation appears rather continuous on either sides of the threshold: The estimated drops in the number of subsidiaries at the first threshold of 0.39% and 0.42% at the second threshold are not significant. Qualitatively this still represents a 10% decrease in the probability of being a subsidiary past the threshold, which indicates that the strategy might exist but is marginal in explaining our results. Bunchers do not exhibit a significant particularly large response. Finally, when comparing subsidiaries to non-subsidiary firms we only find very modest evidence of excess profitability.

To summarize, we only find modest evidence of firms dividing themselves and gaming the tax system. The results can not explain the excess bunching nor the large difference in profitability on each side of the thresholds. The absence of firm division could be a combination of several factors. The monetary and non-monetary costs of setting up a corporation and keeping it active are not trivial: In addition to cumbersome administrative work, Costa Rica has a registration fee and a yearly stamp duty payment. related to the above, the relative ease of evading taxes by inflating costs might make this avoidance strategy suboptimal. Finally, the data might not reveal the full extent of firm division. Economic groups might have hard to detect dilution ownership strategies, and the Central Bank dataset is a (large) subset of the universe of tax filling corporations (around 80% of firms filling a tax declaration).

### 6.3 Employment and wage bill

The break down of costs into the five categories reported on the tax return only brings limitted insights. The two main categories,"Administrative and Distributional Costs" and "Material and Production Costs" each explain around half of the cost discontinuity. Interest deductions, depreciation and other costs do not display a discontinuity and each only represents 5% of total costs[30]. More interestingly, we study employment and wage bill data from social security records. There are reasons to believe that labor inputs are better reported than other deductible costs. First, employees have incentives for correct reporting of their wages as the benefits and social security provided depend on it and are generous in Costa Rica. Second estimated evasion on payroll taxes and on the personal income tax of wage earners is much lower than evasion on other margins echoing survey

[30] We report those results in appendix B.

results that labor informality in Costa Rica is among the lowest in Latin America[31]. Finally, the sum of personal income taxes and payroll taxes is larger than the corporate income tax at all tax brackets. Therefore a firm evading on its wage bill has incentives to under-report labor and not over-report it.

Figure 13 plots the average number of employees and average wage bill by revenue at each threshold. We show the linear fit of employment on revenue on each side of the threshold, excluding revenue bins in the bunching and dominated intervals. We run the following regression:

$$E_j = \alpha + \delta.\mathbb{1}(y_j > 0) + \sum_{j=y_l}^{y_u} \gamma_j \mathbb{1}(y_j = j)\beta_1.y_j + \beta_2.y_j\mathbb{1}(y_j > 0) + \epsilon_j \tag{21}$$

Where $E_j$ is average employment or average wage bill in revenue bin $j$, $y_j$ is the midpoint of revenue bin $j$ and $[y_l, y_u]$ is the excluded revenue interval around the threshold. We report the results for the threshold discontinuity and the dummy for bunchers in Figure 13. The discontinuity in average employment and wage bill is never significant. The only possible indication of a change in labor inputs is the significantly lower wage bill of bunchers at the second threshold[32]. Overall the absence of a dip in labor inputs for bunchers and of a downward discontinuity past the threshold suggests that firms are not distorting their production nor evading on labor inputs.

### 6.4 Production and timing responses

Another potential explanation for bunching is to limit production or to shift revenue across time. For example firms could limit their operations in September, the last month of the fiscal year and/or date September revenue in October, in order for their revenue to remain below the threshold. We use the subsample of firms liable for sales taxes[33] to obtain a monthly measure of revenue for the years 2008 to 2013. We run the following specification:

$$y_{imt} = \beta_1 \mathbb{1}(m = Sept) * Bunch_{it} + \beta_2 \mathbb{1}(m = Oct) * Bunch_{it} + \delta * Bunch_{it} + \alpha_m + \gamma_t + \epsilon_{imt} \tag{22}$$

where $\beta_1$ measures the differential monthly revenue of bunchers in September and $\beta_2$ in October. Under limitation of production at the end of the fiscal year, we should observe $\beta_1$ to be negative

[31] IMF 2012 study, MAU doc

[32] We are not currently controlling for average profitability which could minimize the significant lower tax bill estimated at the second threshold.

[33] Costa Rica only has a final consumer sales tax that only applies to certain industries. More details in Appendix B

and significant and under time-shifting we should observe $\beta_2$ to be positive and significant. Table 7 reports the results for the first threshold for all firms in columns (1)-(4) and for firms who's corporate income tax revenue equals the sum of the monthly sales tax in column (5)-(8) and who therefore have no revenue not subjected to sales tax. The regressions reject evidence of either responses: bunchers do not report revenue differently in September and October compared to other months and other non-bunching firms. The only significant coefficient shows a positive revenue increase of bunchers revenue in September - possibly a sign of reverse time-shifting: as firm realize their revenue fall in the lower tax bracket they decide to bring revenue forward due to uncertainty about next year's revenue. A caveat to keep in mind when interpreting this non-result is that sales tax liable firms belong to sectors with relatively less bunching (e.g.: retail, restaurants and hotels, car dealers). Nonetheless those results support the limitting role of production and time-shifting responses in the bunching decision.

## 7 Conclusion

A conventional view regarding taxation in low-income countries states that monitoring the income of individuals and self-employed is difficult and therefore developing countries should resort to taxing capital, which is supposedly easier to observe. In this paper, we used the unique design of Costa Rica's corporate income tax, administrative data on the universe of formal firms and a novel estimation method to estimate new parameters relevant for optimal tax design. We showed that even in a middle-income country, the capacity to tax small and medium enterprises is limited, particularly when the tax base is broad and allows for many deduction possibilities. As a result, low and middle-income countries might be constrained in their ability to generate additional revenue from the corporate income tax and the optimal design could be substantially different from that of OECD countries.

Two key dimensions should be considered for the external validity of the results: the size of firms in our study and Costa Rica's institutional environment. First, our elasticity estimates concern small and medium firms and might not apply to large firms. Based on only two observations, one at each threshold, the elasticity of profits with respect to the net of tax rate appears to decrease with firm revenue. However large firms could have access to more elaborate evasion and avoidance schemes and it is therefore hard to conclude that the elasticity of profit is necessarily falling with firm revenue. Second, tax elasticities are always dependent on the institutional and policy environment. On the one hand, Costa Rica's institutions are solid given its income: for example, the country ranks above its income level on Transparency International's corruption perception index. In weaker institutional environments, profit elasticities could be even larger. On the other hand, Costa Rica's tax structure is complex and fragmented[34], which contributes to the large estimated profit elasticity and the ease of over-reporting costs.

With these caveats about interpretation in mind, how should low and middle-income countries design their corporate income tax? We discuss four types of corporate tax design, and their potential implications for tax collection and efficiency. Across countries, the most common corporate schedule taxes profits at a flat rate and permits the deduction of most production costs. A direct consequence of the large estimated elasticities is that flat rates above 18-26% are on the wrong side of the Laffer curve and should be excluded[35]. Some countries have increasing marginal tax rates on profits, which reduce bunching incentives but generate a loss in tax revenue, as infra-marginal firms with large profits would see their tax bill decrease on the initial part of their earnings. These concerns are particularly important if the profit elasticity falls with firm revenue, such that large firms with a lower profit elasticity, obtain a reduction of their tax liability without substantially increasing reported profits. In light of this plausible assumption, Costa Rica's tax system permits to tag[36] firms based on revenue, which we show is relatively hard to manipulate, and to assign increasing tax rates, potentially satisfying an inverse elasticity rule. In addition, the low initial tax rate for small firms might contribute to firm registration and formalization: once firms are registered it could be difficult to return to informality and at this stage increasing enforcement could yield revenue gains. However, the current system does not satisfy horizontal equity (i.e. quasi-identical firms face vastly different tax treatment) and only imperfectly deals with the large cost over-reporting. To this end the introduction of presumptive tax schemes could be beneficial. In such schemes tax liability is the maximum amount of a low rate applied on revenue and a higher rate on profits. Best et al. (2014) show that in Pakistan, the revenue gains from a presumptive scheme outweigh the production distortion it generates. The large ratio of cost to revenue elasticity we estimate,

[34] In addition to the corporate and personal income taxes, Costa Rica has a self-employed regime and a micro-sellers regime, which applies to firms with revenue substantially below the firms we study. In addition it does not have a VAT with self-enforcement incentives (Pomeranz 2015), but only a final sales tax.

[35] Gorodnichenko et al. 2009 and Kopczuk 2012 both find that flat tax reforms in Eastern Europe, which decreased substantially the rate and simplified the tax code, led to large increases in reported income.

[36] See Ito and Sallee 2014 for a model of attribute based regulation and enforcement.

further supports their empirical results on the desirability of using revenue as the tax base. More generally, a tax design which limits deductions and/or requires tangible evidence of incurred costs could generate a substantial increase in tax revenue.

To conclude, we show in figure 17 the types of corporate tax systems used worldwide, separating countries by their per-capita income quintile. Flat and increasing marginal tax systems are the most common and are ubiquitous in the top-income quintile. However, within low and middle-income countries, we observe large variation in corporate income tax policies: presumptive schemes are common, especially in the bottom quintile, while systems with revenue dependent rates, like Costa Rica's, are sometimes used in middle-income countries. The observed heterogeneity in corporate tax systems highlights that developing countries do use non-standard tax instruments, potentially to deal with the specific constraints they face and to enhance their revenue collection.

## References


**Abramovksy, Laura, Pierre Bachas, and Anders Jensen**, "Tax Structures and Size Based Tax Policies: A Cross Country Database and Analysis," *Working Paper*, 2015.

**Almunia, Miguel and David Lopez-Rodriguez**, "Heterogeneous Responses to Effective Tax Enforcement: Evidence from Spanish Firms," *Working Paper, Warwick University*, 2015.

**Auerbach, Alan, Michael Devereux, and Helen Simpson**, "Taxing Corporate Income," *Chapter 9 of: Dimension of Tax Design: the Mirrlees Review*, 2010, *Oxford University Press.*

**Besley, Timothy and Torsten Persson**, "Chapter 2 - Taxation and Development," in Martin Feldstein Alan J. Auerbach, Raj Chetty and Emmanuel Saez, eds., *handbook of public economics, vol. 5*, Vol. 5 of *Handbook of Public Economics*, Elsevier, 2013, pp. 51 – 110.

**Best, Michael, Anne Brockmeyer, Henrik Kleven, Johannes Spinnewijn, and Mazhar Waseem**, "Production vs Revenue Efficiency With Limited Tax Capacity: Theory and Evidence From Pakistan," CEPR Discussion Papers 9717, C.E.P.R. Discussion Papers November 2014.

**Carrillo, Paul, Dina Pomeranz, and Monica Singhal**, "Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement," Harvard Business School Working Papers 15-026, Harvard Business School October 2014.

**Chetty, Raj**, "Is the Taxable Income Elasticity Sufficient to Calculate Deadweight Loss? The Implications of Evasion and Avoidance," *American Economic Journal: Economic Policy*, 2009, *1* (2), 31–52.

**_ , John N. Friedman, Tore Olsen, and Luigi Pistaferri**, "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records," *The Quarterly Journal of Economics*, 2011, *126* (2), 749–804.

**Devereux, Michael P., Li Liu, and Simon Loretz**, "The Elasticity of Corporate Taxable Income: New Evidence from UK Tax Records," *American Economic Journal: Economic Policy*, May 2014, *6* (2), 19–53.

**Dwenger, Nadja and Viktor Steiner**, "Profit Taxation And The Elasticity Of The Corporate Income Tax Base: Evidence From German Corporate Tax Return Data," *National Tax Journal*, March 2012, *65* (1), 118–50.

**Emran, M. Shahe and Joseph E. Stiglitz**, "On selective indirect tax reform in developing countries," *Journal of Public Economics*, April 2005, *89* (4), 599–623.

**Gordon, Roger and Wei Li**, "Tax structures in developing countries: Many puzzles and a possible explanation," *Journal of Public Economics*, 2009, *93* (7–8), 855 – 866.

**Gorodnichenko, Yuriy, Jorge Martinez-Vazquez, and Klara Sabirianova Peter**, "Myth and Reality of Flat Tax Reform: Micro Estimates of Tax Evasion Response and Welfare Effects in Russia.," *Journal of Political Economy.*, 2009, *117* (3).

**Gruber, Jonathan and Joshua Rauh**, "How Elastic Is the Corporate Income Tax Base?," *In Auerbach, Alan J., James R. Hines Jr., and Joel B. Slemrod (eds.), Taxing Corporate Income in the 21st Century*, 2007, *Cambridge University Press*, 140–163.

**ILO**, "http://laborsta.ilo.org/informal_economy_E.html," 2012.

**IMF**, "Medición del Incumplimiento y del Gasto Tributario en Costa Rica," 2012.

_ , "Current challenges in revenue mobilization: improving tax compliance," April 2015.

**Ito, Koichiro and James M. Sallee**, "The Economics of Attribute-Based Regulation: Theory and Evidence from Fuel-Economy Standards," Working Paper 20500, National Bureau of Economic Research September 2014.

**Jorgenson, Dale and R.E. Hall**, "Tax Policy and Investment Behavior," *American Economic Review*, 1967, *57*, 391–414. Reprinted in Bobbs-Merrill Reprint Series in Economics , Econ-130. Investment 2, ch. 1, pp 1-26.

**Khan, Adnan Q., Asim I. Khwaja, and Benjamin A. Olken**, "Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors," *The Quarterly Journal of Economics*, 2015.

**Kleven, Henrik J. and Mazhar Waseem**, "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan," *The Quarterly Journal of Economics*, 2013, *128* (2), 669–723.

**Kleven, Henrik Jacobsen**, "How Can Scandinavians Tax So Much?," *Journal of Economic Perspectives*, 2014, *28* (4), 77–98.

**Kleven, Henrik, Martin Knudsen, Klaus Kreiner, Soren Pedersen, and Emmanuel Saez**, "Unwilling or Unable to Chear: Evidence from a Tax Audit Experiment in Denmark," *Econometrica*, May 2011, *Vol. 79* (No. 3), 651–692.

**Kopczuk, Wojciech**, "The Polish Business "Flat" Tax and its Effect on Reported Incomes: a Pareto Improving Tax Reform?," *Working Paper, Columbia University*, 2012.

**Luttmer, Erzo F. P. and Monica Singhal**, "Culture, Context, and the Taste for Redistribution," *American Economic Journal: Economic Policy*, February 2011, *3* (1), 157–79.

_ **and** _ , "Tax Morale," *Journal of Economic Perspectives*, Fall 2014, *28* (4), 149–68.

**Naritomi, Joanna**, "Consumers as Tax Auditors," *Working Paper, London School of Economics*, 2015.

**Olken, Benjamin A. and Monica Singhal**, "Informal Taxation," *American Economic Journal: Applied Economics*, October 2011, *3* (4), 1–28.

**Patel, Elena, Nathan Seegert, and Matt Smith**, "At a Loss: the Elasticity of Corporate Income at the Zero Kink," *Working Paper*, 2015.

**Pomeranz, Dina**, "No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax," *American Economic Review*, 2015, *105* (8), 2539–69.

**Saez, Emmanuel**, "Do Taxpayers Bunch at Kink Points?," *American Economic Journal: Economic Policy*, 2010, *2* (3), 180–212.

**Slemrod, Joel, Brett Collins, Jeffrey Hoopes, Daniel Reck, and Michael Sebastiani**, "Does Credit-card Information Reporting Improve Small-business Tax Compliance?," NBER Working Papers 21412, National Bureau of Economic Research, Inc July 2015.

_ **, Marsha Blumenthal, and Charles Christian**, "Taxpayer response to an increased probability of audit: evidence from a controlled experiment in Minnesota," *Journal of Public Economics*, March 2001, *79* (3), 455–483.

**United Nations**, "Report of the Intergovernmental Committee of Experts on Sustainable Development Financing," 2014.

**World Bank**, "Financing for Development Post-2015," October 2013.

Figure 1: Costa Rica's Corporate Tax Schedule

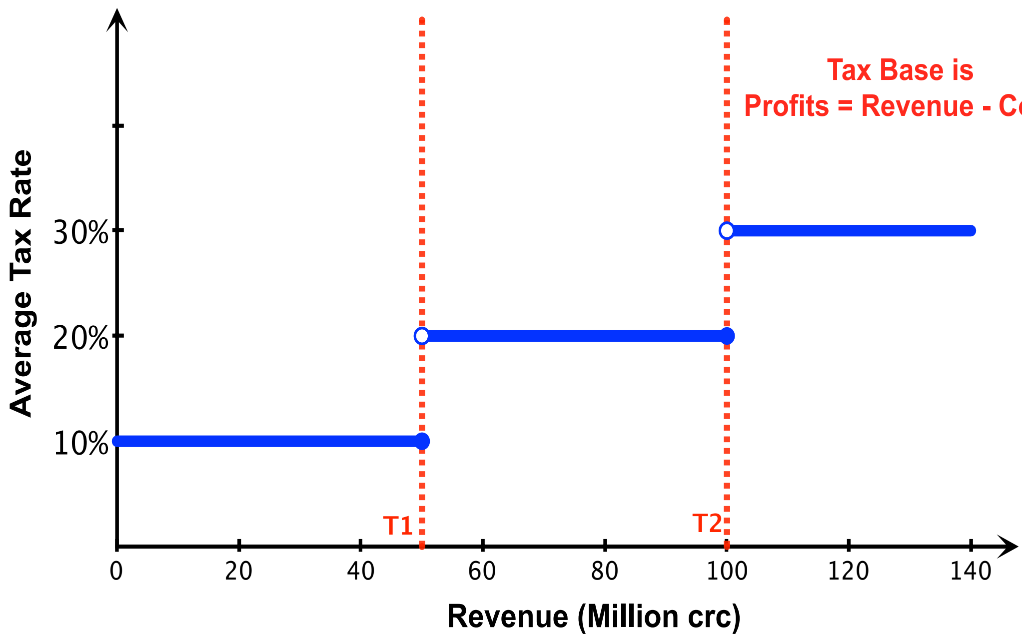



Figure 1 shows the design of the corporate income tax in Costa Rica. Firms pay increasing average tax rates on their profits as a function of their revenue. When revenue exceeds the first threshold, the average tax rate jumps from 10% to 20% and from 20% to 30% past the second threshold.

Figure 2: Bunching Theory

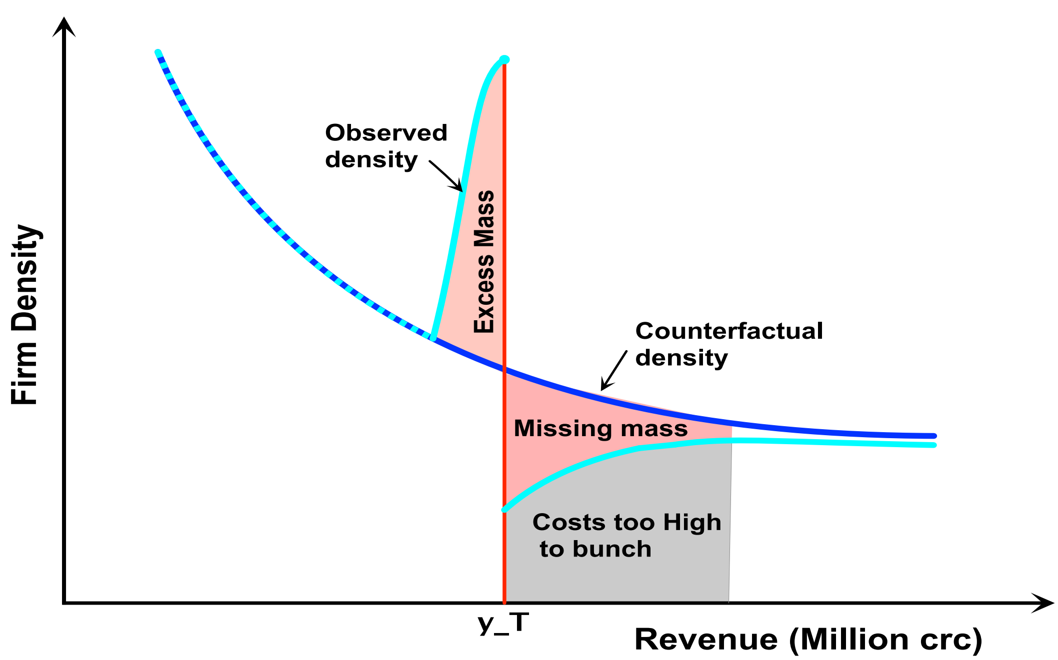
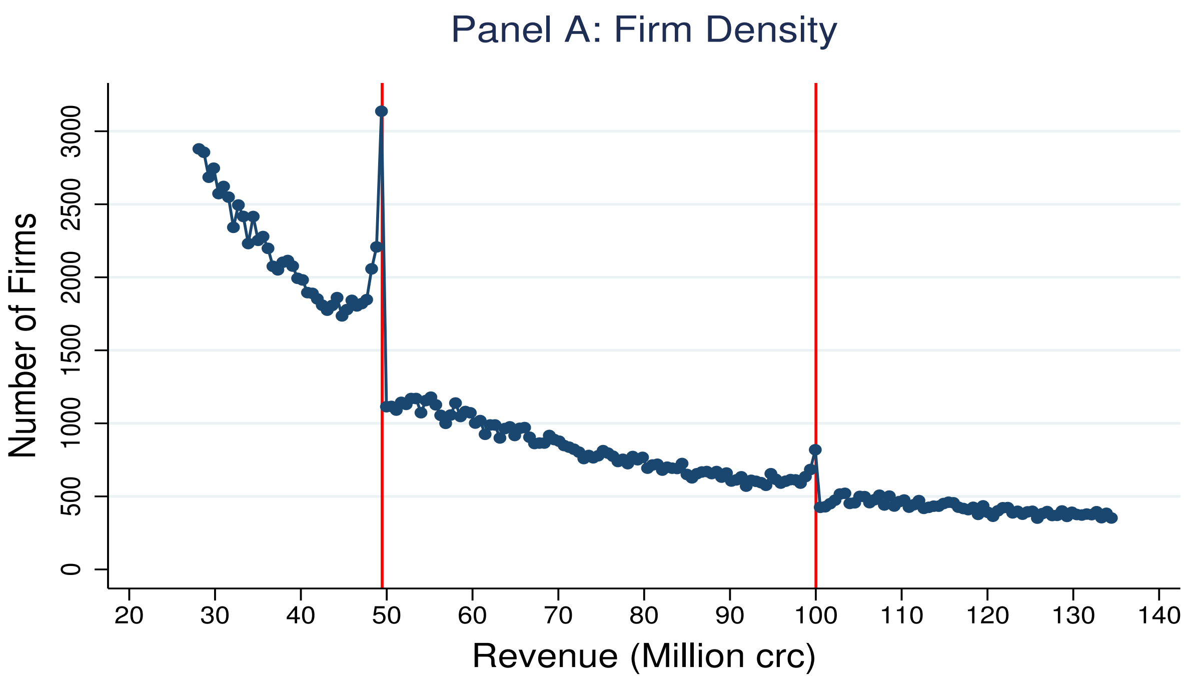



Figure 2 displays the density distributions. Under a flat 10% tax rate the counterfactual firm density follows a smooth distribution. The notch induces some firms, with counterfactual revenue above the threshold, to reduce their revenue and bunch just below the threshold. The bunching decision is a joint function of the firm's revenue distance to the threshold and costs, such that at each revenue bin, only firms with sufficiently low costs bunch.

Figure 3: Firm Density and Average Profit Margin



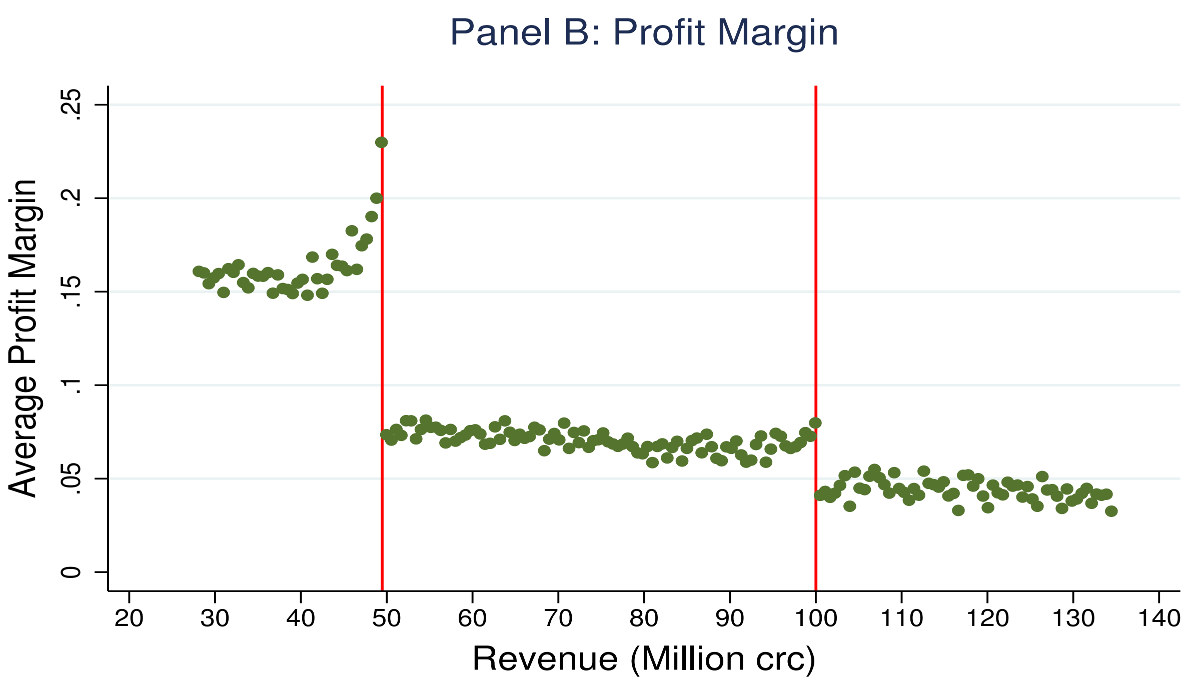
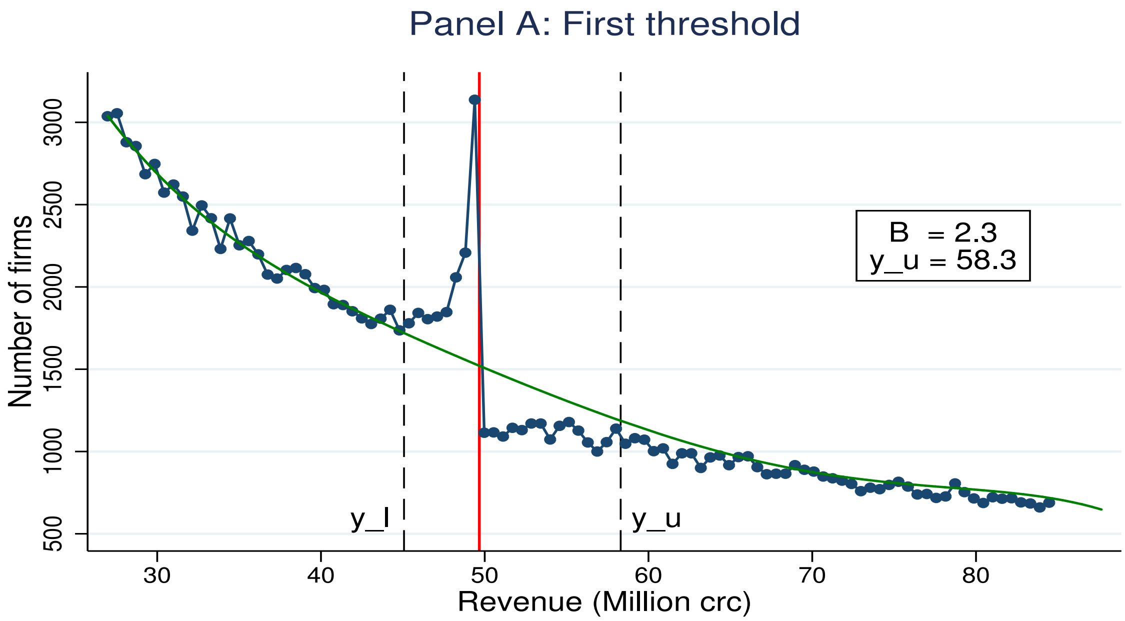




Figure 3 presents the key patterns of the corporate tax returns, pulling together the years 2008 to 2014. Panel A shows the density of firms by revenue. Panel B displays the average profit margin by revenue, where profit margin is defined as profits over revenue. The size of the revenue bins is 575,000 CRC.

Figure 4: Revenue Bunching Estimation



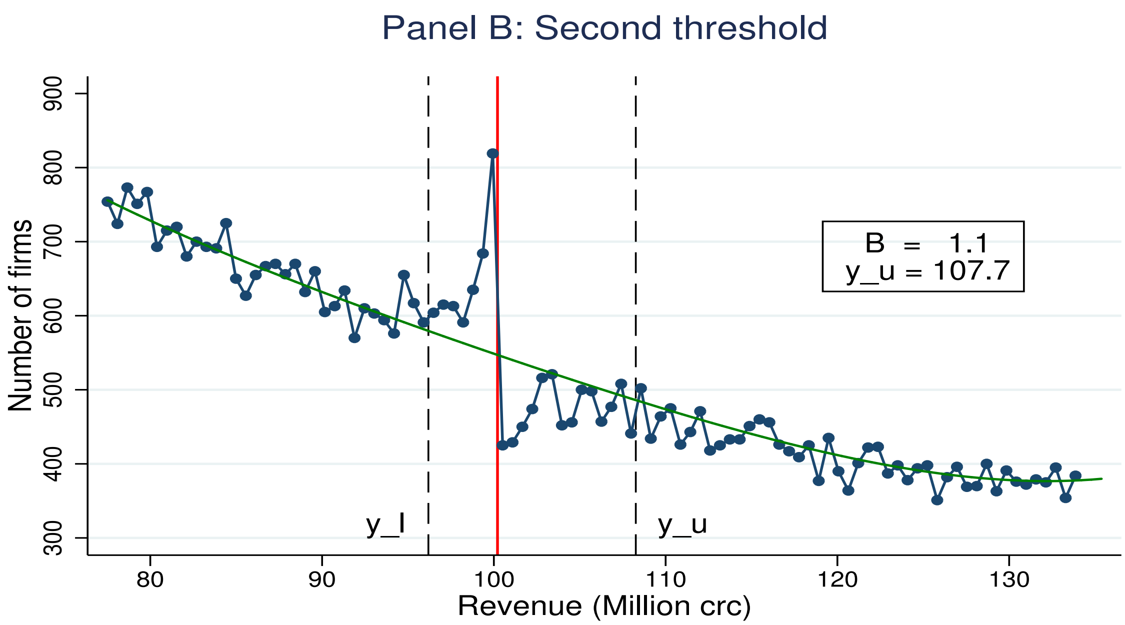




Figure 4 displays the density of firms by revenue and fits the counterfactual distribution for the first and second thresholds. In the boxes on the top right, B is the excess mass as a share of the counterfactual, and $y_u$ the revenue of the marginal buncher, obtained with the point of convergence method. The counterfactual is obtained from the regression of a polynomial of degree 5 (maximizes Akaike criteria), on all data points outside the $[y_l, y_u]$ interval. The lower bound $y_l$ is chosen by the researchers as the revenue bin which starts exhibiting excess. The upper bound $y_u$ is estimated from an iterative process: starting from $y_u$ close to the threshold, we obtain the counterfactual and estimate the excess mass (B) below the threshold and missing mass (M) above the threshold. For low $y_u$, the excess mass is larger than the missing mass, $B >> M$. We increase $y_u$ until the two masses are equal, $B = M$.

Figure 5: Cost Discontinuity

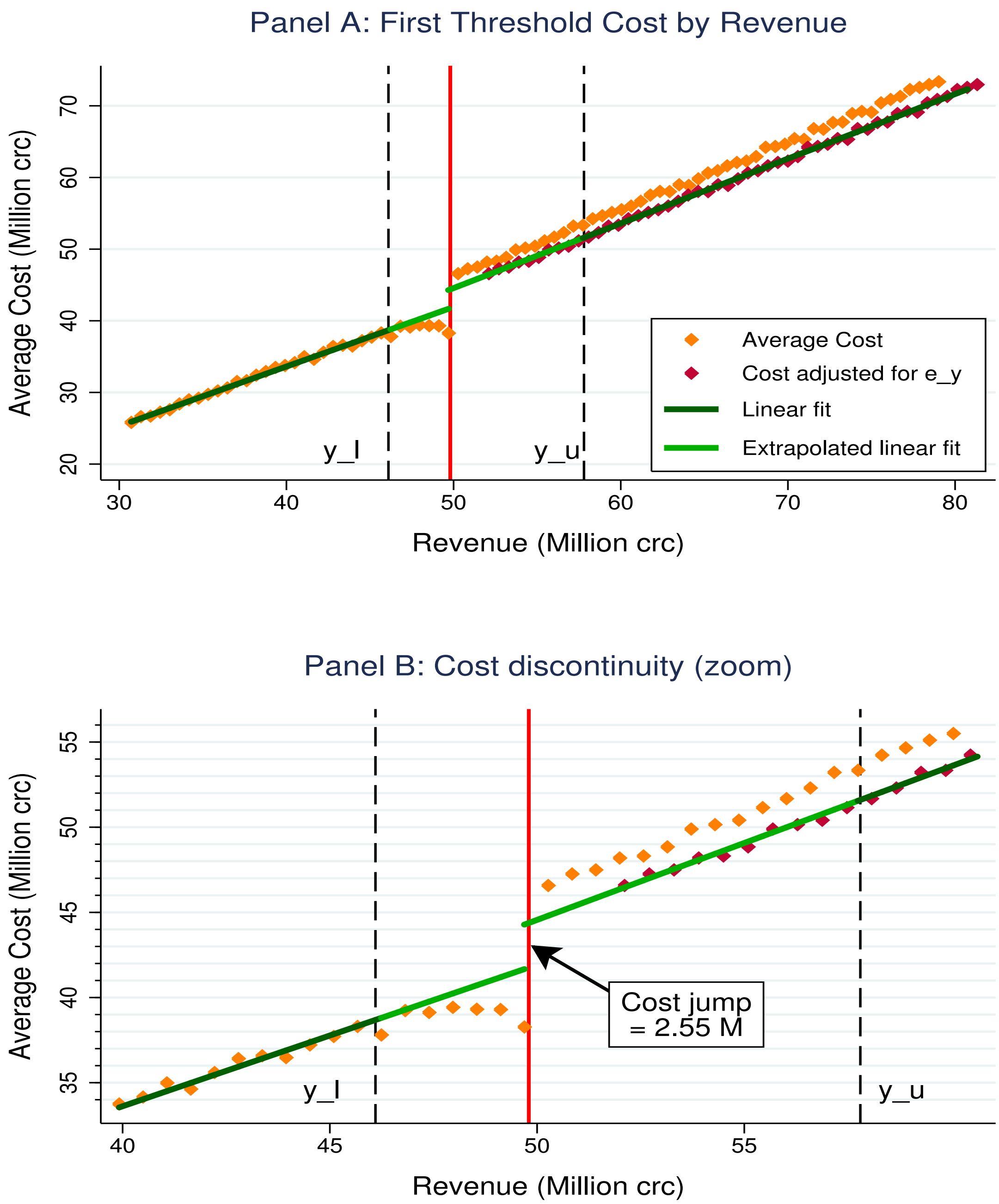



Figure 5 displays the average declared cost for each revenue bin around the first threshold. To estimate the cost discontinuity at the threshold, absent revenue responses, we adjust for intensive margin revenue responses: firms declaring revenue above the threshold reduced their declared revenue, due to the the tax rate increase. To take intensive responses into account, we horizontally shift firms' costs proportionally to the elasticity of revenue, estimated from bunching. For example, given an elasticity of revenue of 0.25 and a firm with revenue of 60M: $revenue_{counter} = 60 + \epsilon_{y,1-t}.y.\frac{dt}{1-t} = 60 + 0.25 * 60 * \frac{0.1}{0.9} \approx 61.6$. We linearly fit costs by revenue, below and above the threshold. On either side, we exclude revenue bins impacted by bunching behavior. We then extrapolate the linear fits to the threshold. The resulting cost discontinuity represents the average increase in declared costs, for a firm at the threshold, due to an increase in the tax rate from 10 to 20%.

Figure 6: Effective Tax Rate on Revenue


Figure 6 plots the average tax payment as a share of revenue by revenue bins of half million CRC. The large drop in declared profits past the thresholds implies that even though the tax rate increases by 10% past each threshold taxes paid as a share of revenue are close to constant.

Figure 7: Industry Heterogeneity

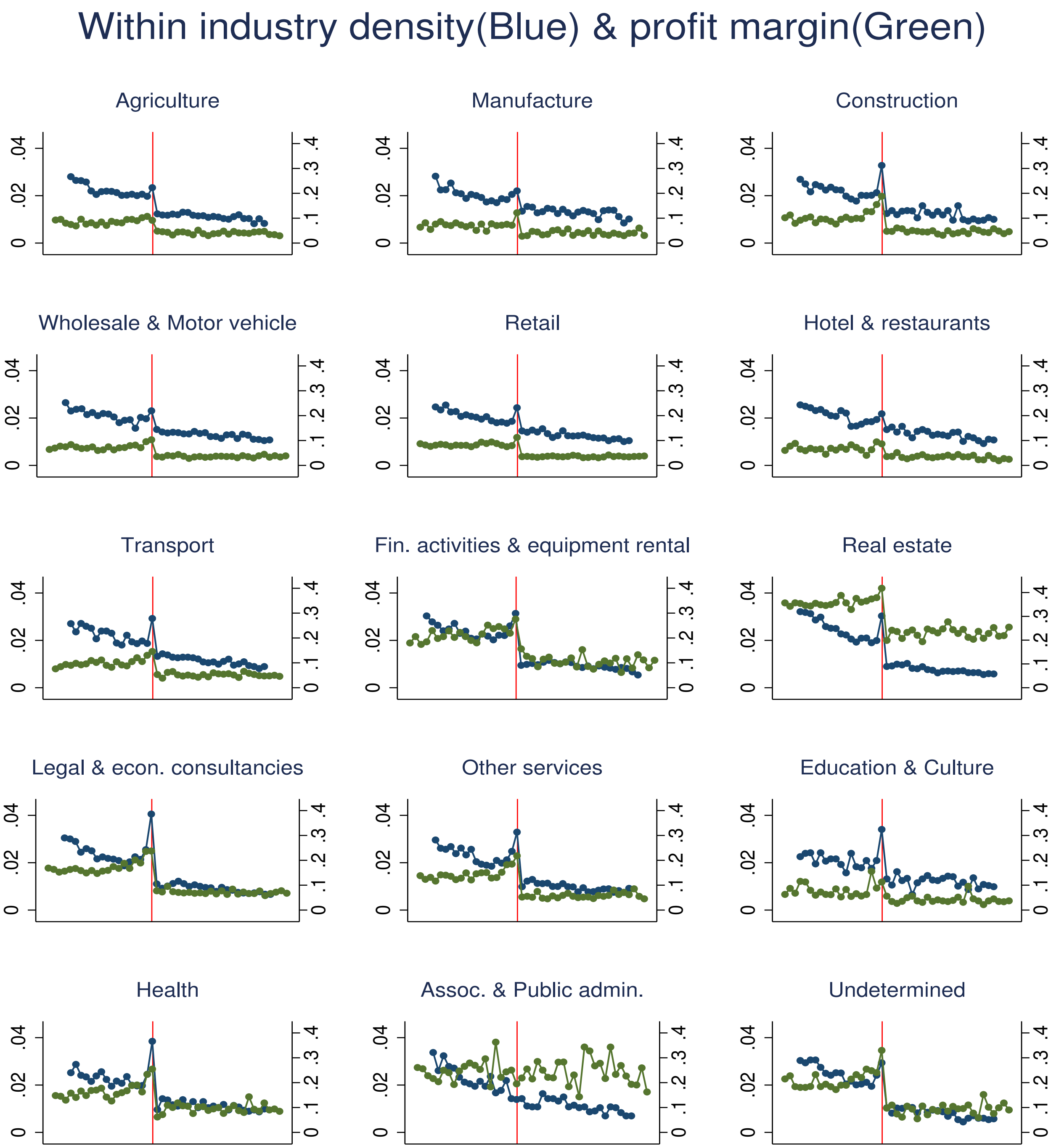



Figure 7 presents the firm density and profit margin by revenue, separating the economy in fifteen industries. In blue, the within industry firm density by revenue and in green the average profit margin by revenue.

## Figure 8: Excess Mass and Profit Discontinuity by Industry

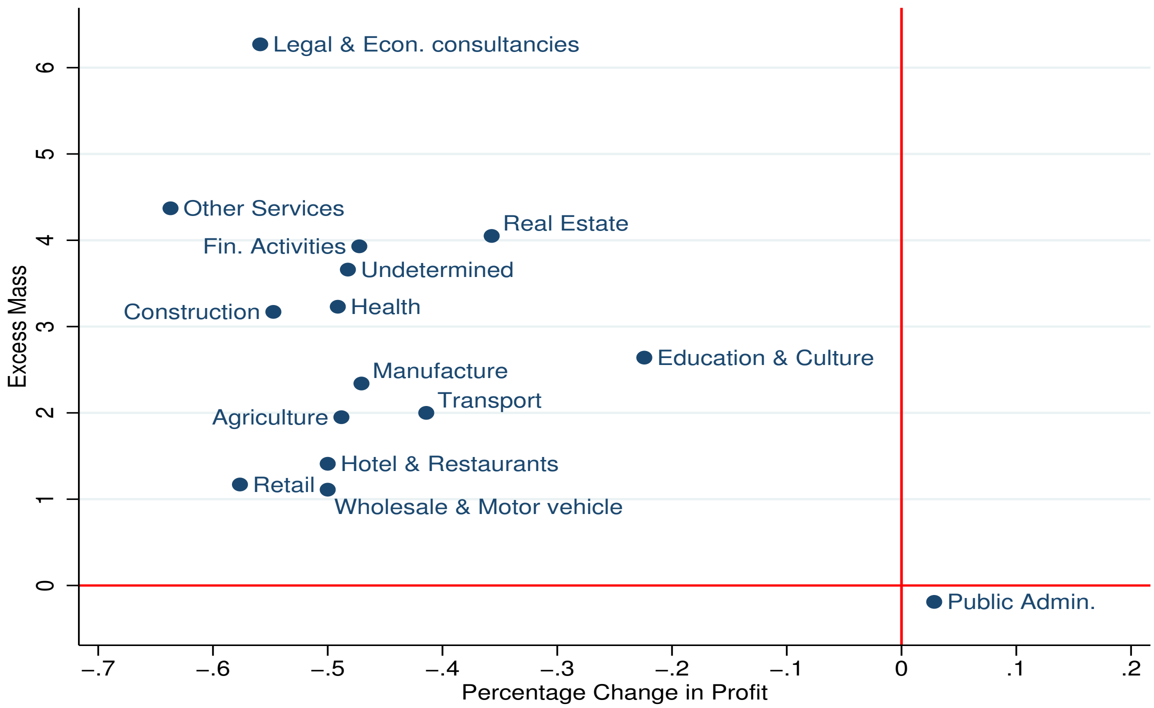
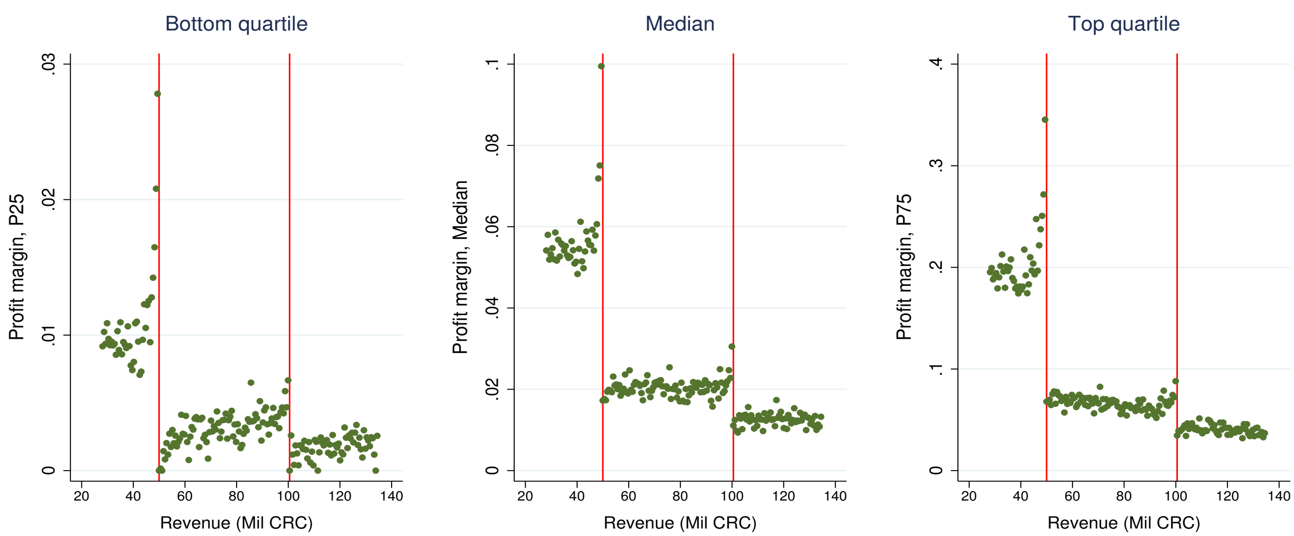



Figure 8 shows the relation between the profit margin discontinuity and the excess mass by industry, at the first threshold.

## Figure 9: Quartiles of Profit Margin by Revenue




Note: The y−axis scale is not constant across figures

Figure 10: Structural Assumption - Stable Profit Margin Distributions

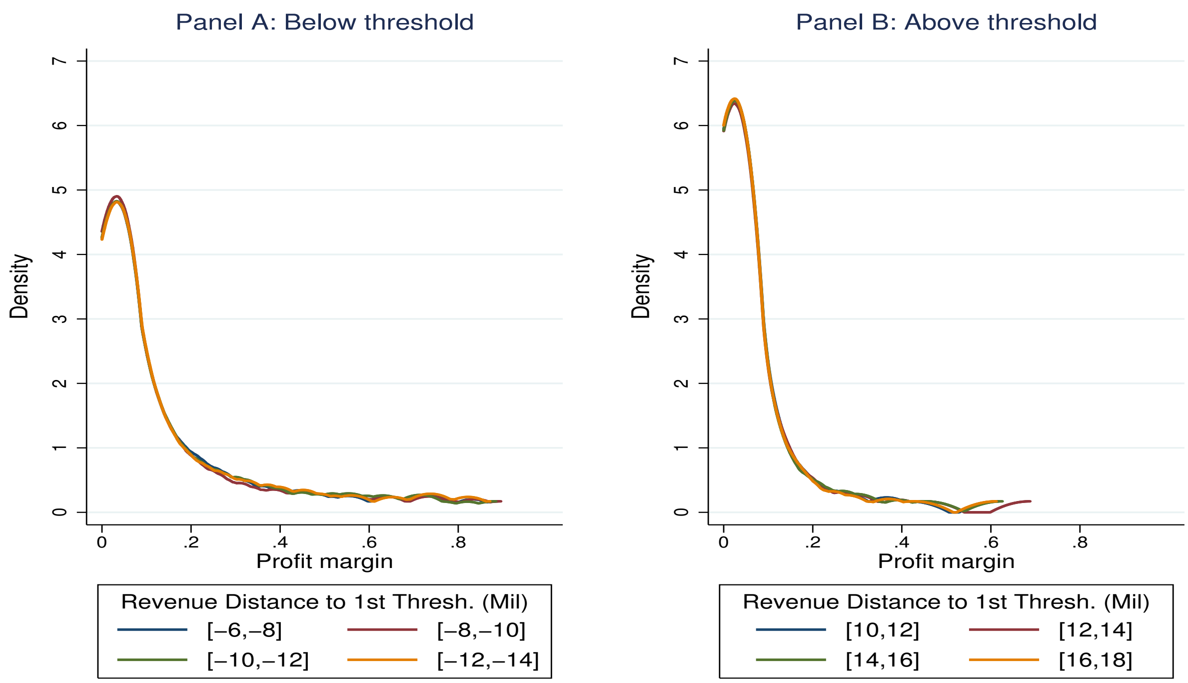



Figure 10 shows that the distribution of profit margin is stable across revenue intervals 10 to 20% below the threshold (Panel A), and revenue intervals 20 to 30% above the threshold (Panel B). The revenue distance to the threshold indicates the revenue intervals considered. We use an Epanechnikov kernel with bandwith of 0.04 across all distributions. Within each panel, we never reject the Kolmgorov-Smirnov tests, that profit margins are sampled from populations with identical distributions across all pairs of revenue intervals.

Figure 11: Numerical Estimation of Bunching Behavior

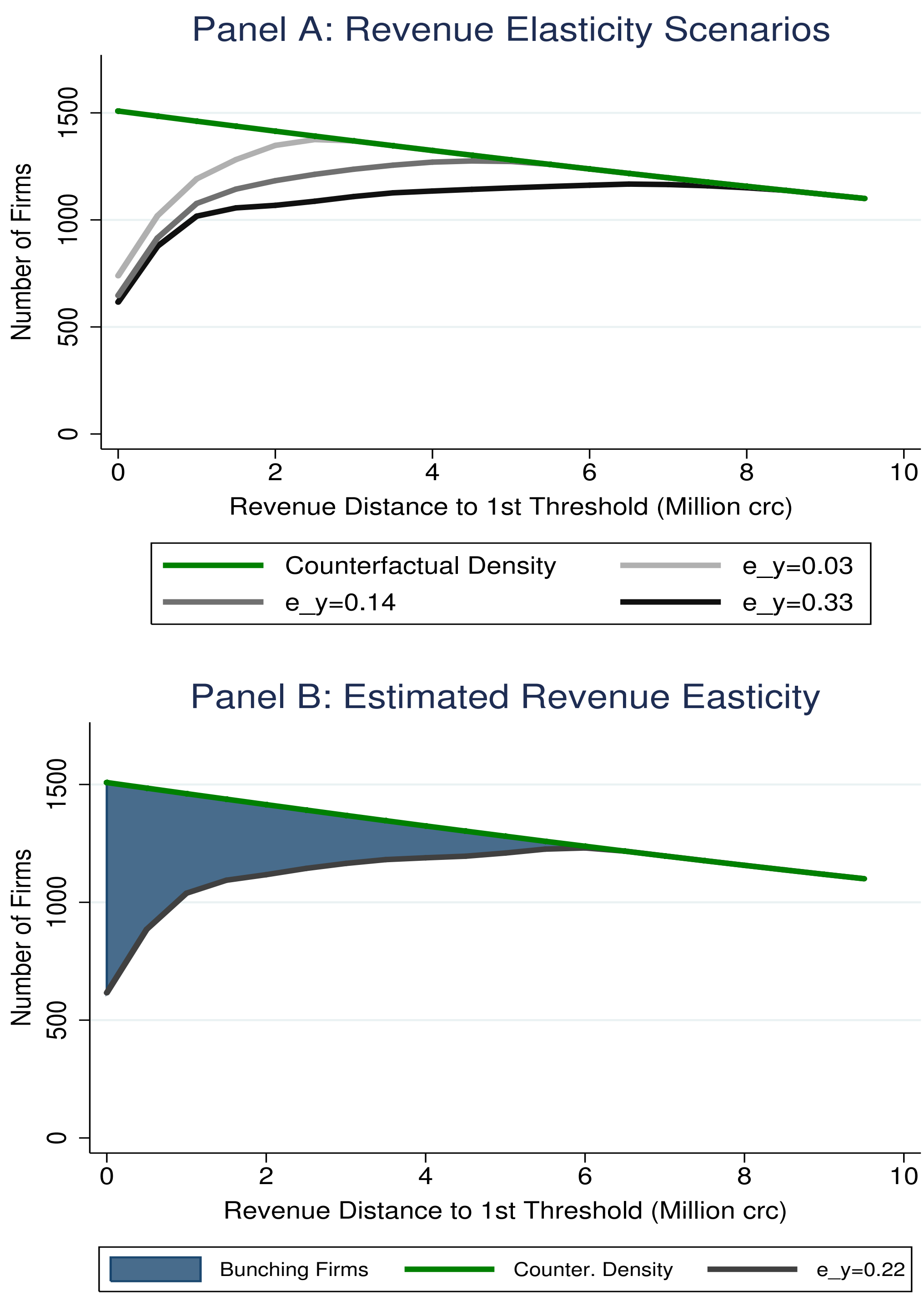



Figure 11 displays the results from the numerical model-based estimation, when assuming a joint counterfactual distribution of revenue and costs and that bunchers cost response correspond to the average cost response. For a given revenue elasticity, $e_y$, the area between the counterfactual density and the curves represents the number of bunching firms. In panel A, we display the profile of these curves for several values of the revenue elasticity. Panel B, displays the result for the last iteration: the number of estimated bunchers equals the observed bunchers and the revenue and cost elasticity are in equilibrium.

Figure 12: Dynamic firm behavior by tax bracket

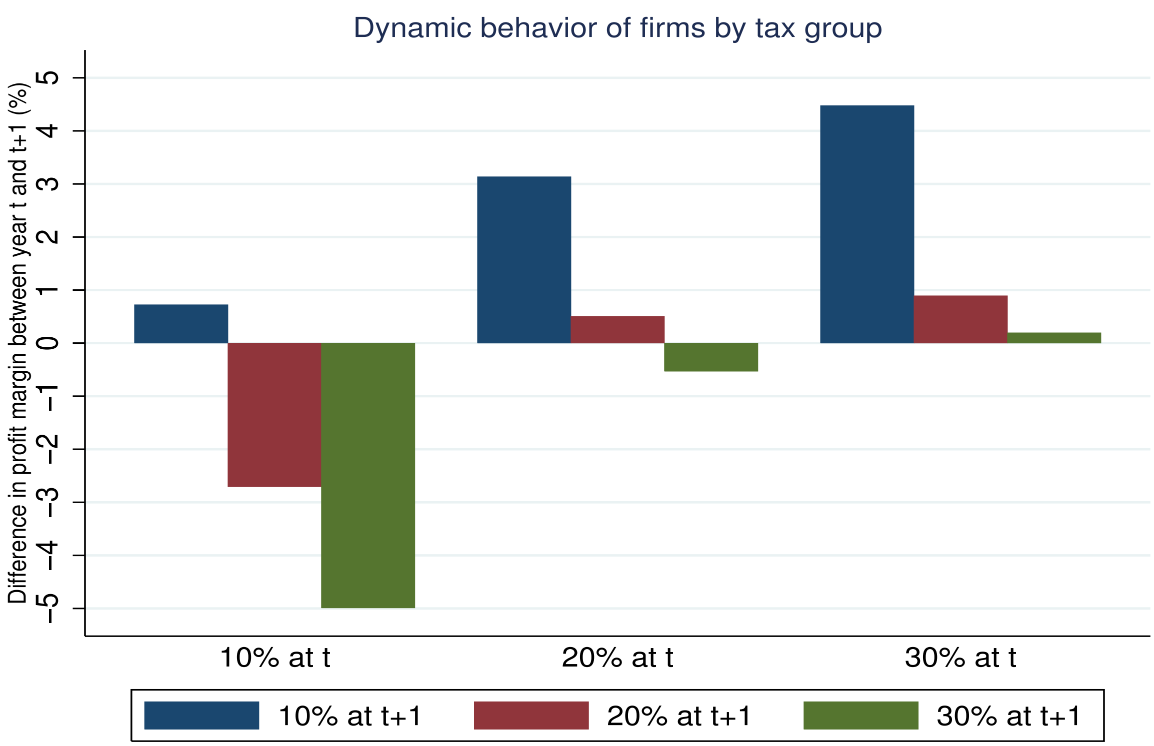



Figure 12 shows the average change in firms' profit margins between year t and t+1 as a function of their tax brackets in year t and t+1. On the one hand, firms remaining within their initial tax bracket hardly change reported profits. On the other hand, firms jumping to higher tax brackets drop their profit margins and symmetrically firms falling to lower tax bracket increase their profit margins.

## Figure 13: Employment

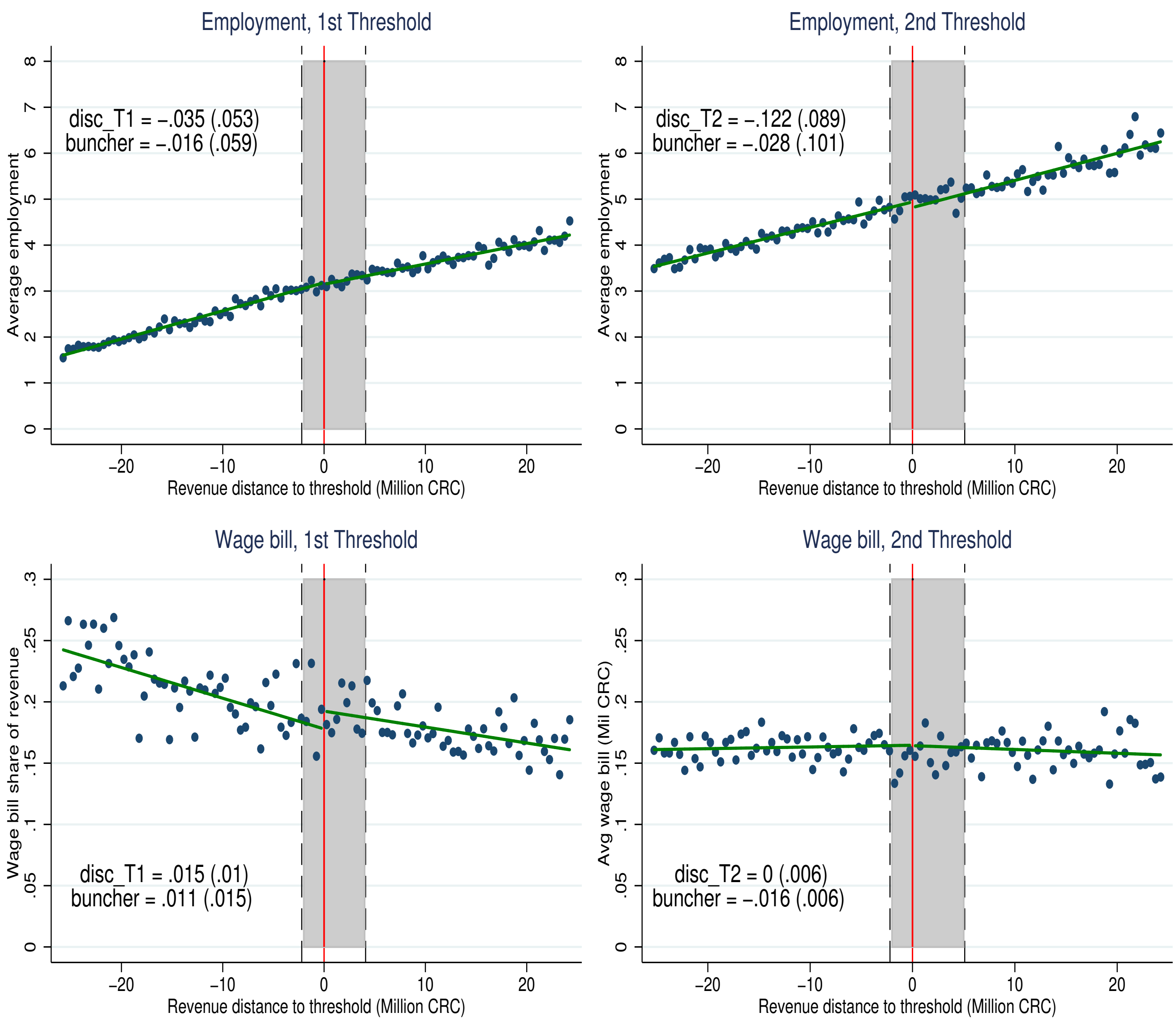



Figure 13 shows the average number of employees and wage bill by revenue around the first and second thresholds. The data is obtained from the social security. We display the coefficient and standard errors from the discontinuity regression at the threshold and the dummy coefficient for firms in the bunching interval.

Figure 14: Share of Subsidiaries by Revenue

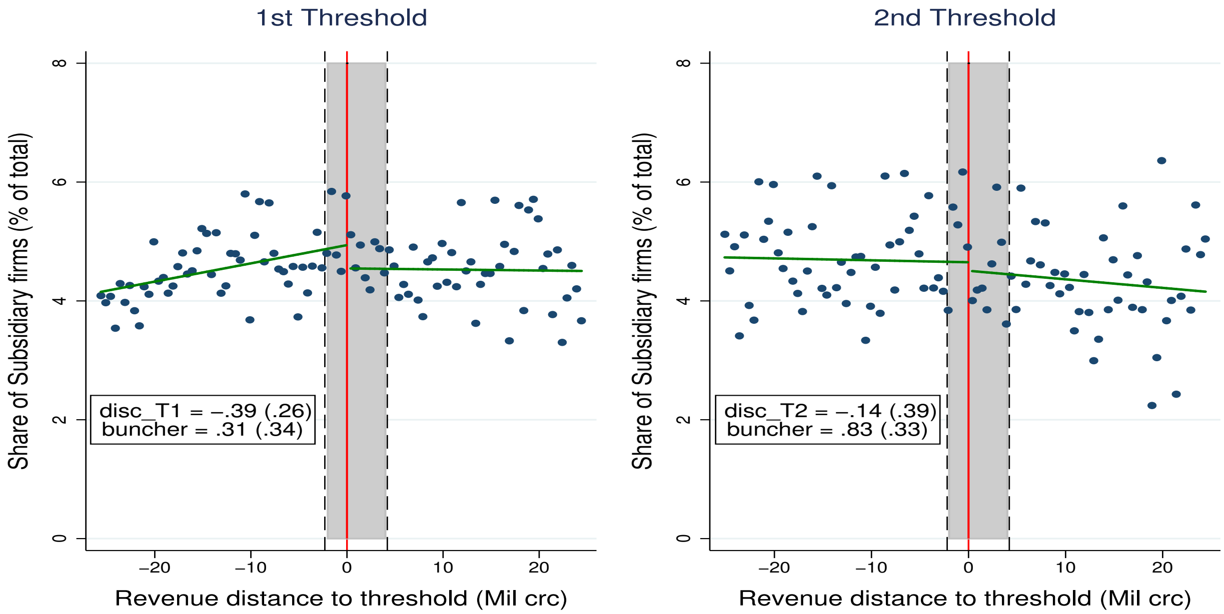



Figure 15: Audits by Revenue (2009 & 2010)

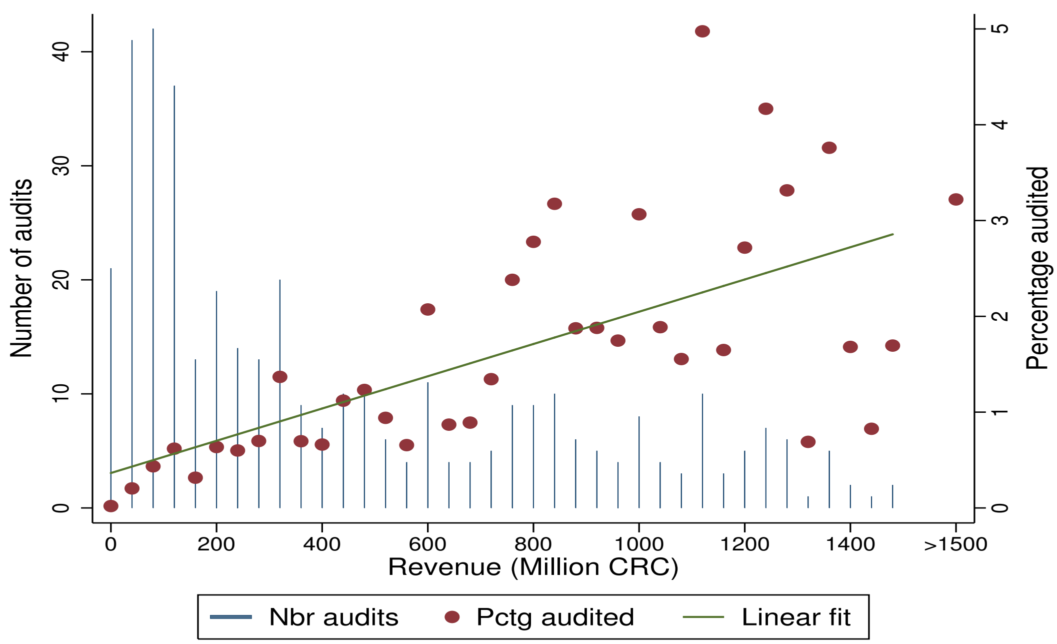



Figure 15 shows the number of in depth audits and the percentage of firms audited by revenue for the years 2009 and 2010. Revenue bins represent 40 Million CRC. The firms we study are in the initial four bins and face very low audit rates.

Figure 16: Correction Letters by Revenue (2012)

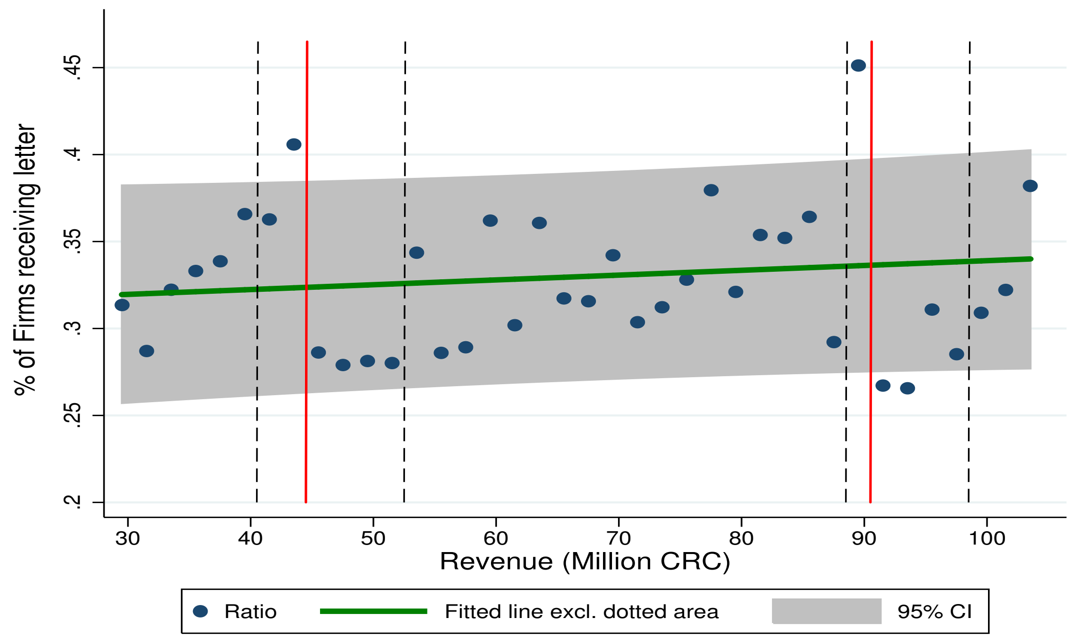



Figure 16 displays the percentage of firms receiving correction letters by revenue due to an inconsistency in their tax declaration or a discrepancy with third party information. Revenue bins represent 2 Million CRC. The fitted line excludes the revenue intervals impacted by the bunching selection.

Figure 17: Worldwide Corporate Tax Systems by Income Quintiles

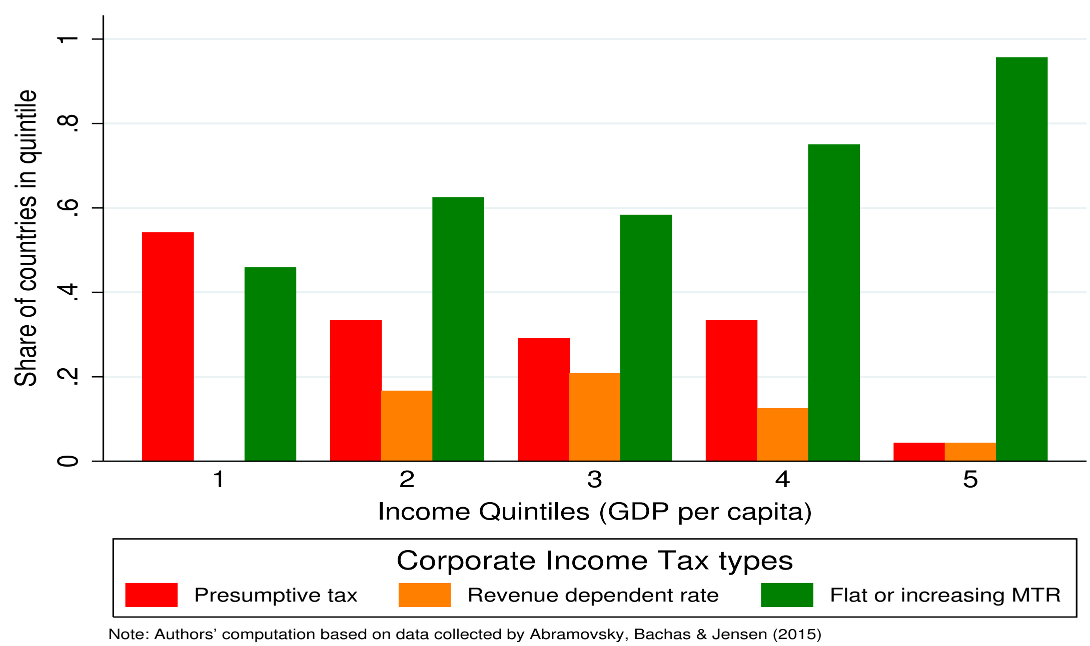



Figure 17 shows the worldwide distribution of corporate income tax types, by countries' per-capita income quintile. The sample contains 120 countries, with data on corporate incomes tax types collected from international tax guides by Abramovksy et al. 2015 and per-capita income data from the Penn Tables. Presumptive tax systems are very common in the lowest quintile, while CIT with revenue dependent rate, such as Costa Rica's, are sometimes used in middle-income countries. In the top quintile, corporate income is almost always taxed at a flat or increasing marginal rate.

Table 1: Summary of estimates and assumptions

| Method | Estimation strategy | Elasticity | Bound | Assumptions |
|---|---|---|---|---|
| **Reduced form** | | | | |
| Step 1 $\rightarrow \epsilon_y$ | Bunching: empirical point of convergence | $\epsilon_{y,1-\tau} = 0.33$ | Upper bound: $\epsilon_y$ of marginal buncher | Counterfactual firm density by revenue is regular |
| $\Downarrow$ | | | | |
| Step 2 $\rightarrow \epsilon_c(\epsilon_y)$ | Cost discontinuity | $\epsilon_{c,1-\tau} = -0.51$ | Lower bound: adjustment for $\hat{\epsilon}_y \rightarrow$ $\hat{\epsilon}_c$ negative relation with $\hat{\epsilon}_y$ | Linear cost extrapolation is correct |
| $\Downarrow$ | | | | |
| Step 3 $\rightarrow \epsilon_\pi(\epsilon_y, \epsilon_c)$ | $d\pi = dy - dc$ | $\epsilon_{\pi,1-\tau} = 4.73$ | | |
| **Added structure** | Profit margin distribution of firms with revenue 10% below the threshold would apply to firms with revenue just above the threshold absent the tax change | | | |
| **With** $dc = d\bar{c}$ | | | | |
| Step 1 $\rightarrow \epsilon_y(\epsilon_c)$ | Bunching: numerical elasticity $\epsilon_y$ | $\epsilon_{y,1-\tau} = 0.22$ | | Bunchers would have $dc = d\bar{c}$ if not bunching |
| $\Updownarrow$ Iterate | | | | |
| Step 2 $\rightarrow \epsilon_c(\epsilon_y)$ | Cost discontinuity | $\epsilon_{c,1-\tau} = -0.65$ | | |
| $\Downarrow$ | | | | |
| Step 3 $\rightarrow \epsilon_\pi(\epsilon_y, \epsilon_c)$ | $d\pi = dy - dc$ | $\epsilon_{\pi,1-\tau} = 4.72$ | | |
| **With** $dc = 0$ | | | | |
| Step 1 $\rightarrow \epsilon_y$ | Bunching: numerical elasticity $\epsilon_y$ | $\epsilon_{y,1-\tau} = 0.095$ | Lower bound: Assumption dc=0 | Bunchers would have $dc = 0$ if not bunching |
| $\Downarrow$ | | | | |
| Step 2 $\rightarrow \epsilon_c$ | Cost discontinuity | $\epsilon_{c,1-\tau} = -0.79$ | Upper bound: $\hat{\epsilon}_c$ negative relation with $\hat{\epsilon}_y$ | |
| $\Downarrow$ | | | | |
| Step 3 $\rightarrow \epsilon_\pi(\epsilon_y, \epsilon_c)$ | $d\pi = dy - dc$ | $\epsilon_{\pi,1-\tau} = 4.73$ | | |

Table 2: Cost on Revenue Distance to Threshold Relation

| | 1st Threshold | | 2nd Threshold | |
|---|---|---|---|---|
| | Cost | Cost(revenue adjust.) | Cost | Cost(revenue adjust.) |
| Jump in cost $\delta$ | 4.203** | **2.548**** | 2.223** | **1.277**** |
| | (0.212) | (0.226) | (0.416) | (0.432) |
| Slope below T. $\beta_1$ | 0.834** | 0.834** | 0.933** | 0.933** |
| | (0.010) | (0.010) | (0.015) | (0.015) |
| Slope change above T. $\beta_2$ | 0.103** | 0.069** | 0.017 | 0.008 |
| | (0.014) | (0.014) | (0.026) | (0.026) |
| Intercept $\alpha$ | 41.971 | 41.971 | 93.863 | 93.863 |
| Observations | 80 | 80 | 80 | 80 |
| % Jump in Cost $\frac{\delta}{\alpha}$ | +10.01% | +6.07% | +2.37% | +1.36% |

Table (2) shows the results from the regression of average costs by revenue on revenue distance to the threshold, estimated from equation (14): $cost_j = \alpha + \delta.\mathbb{1}(\tilde{y}_j > 0) + \beta_1.\tilde{y}_j + \beta_2.\tilde{y}_j\mathbb{1}(\tilde{y}_j > 0) + \sum_{j=\tilde{y}_l}^{\tilde{y}_u} \gamma_j \mathbb{1}(\tilde{y}_j = j) + \epsilon_j$
For each threshold we report the discontinuity in cost $\delta$ with and without the revenue adjustment. The revenue adjustment holds revenue responses above the threshold constant, using the revenue elasticity estimated with bunching, such that the discontinuity in the cost to revenue relation, after adjustment, only identifies cost responses: the results from column (2) & (4) are our main estimate of the cost discontinuity. An observation is a revenue bin of 0.575 Million Colones. Standard errors are shown in parentheses and stars indicate statistical significance level. *=5% level, **=1% level.

Table 3: Elasticity Estimates - Point of Convergence

| Parameters | | | | **Elasticity** | | | Threshold jump | | |
|---|---|---|---|---|---|---|---|---|---|
| $y^*$ | $dy^*$ | $1-\tau_0$ | $\tau^*$ | Revenue | Cost | Profit | $\Delta y$ | $\Delta c$ | $\Delta \pi$ |
| 50 | 8.3 | 0.9 | 0.55 | 0.33 | -0.51 | **4.73** | -1.84 | 2.39 | -4.23 |
| | (1.3)** | | | (0.06)** | (0.06)** | (0.27)** | | | |
| 100.5 | 7.2 | 0.8 | 0.77 | 0.08 | -0.11 | **2.72** | -0.99 | 1.11 | -2.09 |
| | (1.8)** | | | (0.04)* | (0.05)* | (0.70)** | | | |

Table (3) shows the elasticity estimates when the point of convergence method is used to estimate the revenue elasticities. Cost elasticities are estimated with the threshold discontinuity, holding constant revenue responses. The profit elasticity then combines the revenue and cost responses. Standard errors are obtained through 1,000 bootstrap iterations. $y^*$ is the revenue threshold in Million CRC and $dy^*$ is the revenue response of the marginal buncher estimated with bunching. $1-\tau_0$ is the average tax rate below each threshold and $\tau^*$ is the implicit marginal tax rate faced by the marginal buncher. This is estimated with equation (12) and assumes that the cost of the marginal buncher corresponds to the 10th percentile of the cost distribution of this revenue bin. Standard errors are shown in parentheses and stars indicate statistical significance level. *=5% level, **=1% level.

Table 4: Model-Based Numerical Estimation: Iteration Steps

| Iteration Step | Revenue Elasticity $\epsilon_{y,1-t}$ | Cost Jump $dc$ | Cost Elasticity $\epsilon_{c,1-t}$ |
|---|---|---|---|
| 1 | 0.33 | 2.43 | -0.52 |
| 2 | 0.14 | 3.42 | -0.73 |
| 3 | 0.29 | 2.63 | -0.56 |
| 4 | 0.17 | 3.27 | -0.70 |
| 5 | 0.26 | 2.83 | -0.61 |
| 6 | 0.19 | 3.17 | -0.68 |
| 7 | 0.24 | 2.93 | -0.63 |
| 8 | 0.21 | 3.07 | -0.66 |
| 9 | 0.23 | 2.98 | -0.64 |
| **Final** | **0.22** | **3.02** | **-0.65** |

Table 4 shows the iteration steps of the model-based numerical bunching estimation. With a counterfactual firm density by revenue and a counterfactual profit margin distribution, the method numerically estimates the number of bunching firms as a joint function of their revenue distance to the threshold and costs. Step 1 uses as initial values the revenue and cost elasticity from section 3. Under those parameters the revenue elasticity of Step 2 is sufficient such that the number of bunchers equal that of the excess mass. With this new revenue elasticity we re-estimate the cost elasticity using the discontinuity method of 3.3. We iterate this procedure until we find the fixed point at which the revenue and cost elasticities are consistent with each other.

## Table 5: Sector Level Results

| Sector | Profit Margin (%) Drop | Base | % Drop | Bunching Excess Mass | # Firms Total | % Below T1 |
|---|---|---|---|---|---|---|
| Agriculture | -4.1 | 8.4 | -48.8 | 1.95 | 33,095 | 59.5 |
| Manufacture | -3.2 | 6.8 | -47.1 | 2.34 | 34,799 | 45.4 |
| Construction | -5.2 | 9.5 | -54.7 | 3.17 | 26,410 | 51 |
| Wholesale & Motor Vehicle | -3.5 | 7 | -50 | 1.11 | 63,544 | 45.1 |
| Retail | -4.9 | 8.5 | -57.6 | 1.17 | 100,552 | 47.9 |
| Hotel & Restaurants | -3.5 | 7 | -50 | 1.41 | 21,483 | 49 |
| Transport | -4.1 | 9.9 | -41.4 | 2 | 36,294 | 54.7 |
| Financial Activities | -10.3 | 21.8 | -47.2 | 3.93 | 26,366 | 71.9 |
| Real Estate | -13 | 36.4 | -35.7 | 4.05 | 91,525 | 85.1 |
| Legal & Econ. Consultants | -9.5 | 17 | -55.9 | 6.27 | 64,617 | 73.3 |
| Other Services | -9.3 | 14.6 | -63.7 | 4.37 | 37,091 | 69.3 |
| Education & Culture | -1.3 | 5.8 | -22.4 | 2.64 | 14,228 | 56.8 |
| Health | -8.4 | 17.1 | -49.1 | 3.23 | 19,611 | 65.2 |
| NGO & Public Admin. | .8 | 28.1 | 2.8 | -.19 | 10,608 | 68.8 |
| Undetermined | -9.6 | 19.9 | -48.2 | 3.66 | 36,044 | 80.8 |

A

## Table 6: Dynamic Firm Behavior

| | Profit Margin 1st Threshold | Profit Margin 2nd Threshold |
|---|---|---|
| Revenue (Million CRC) | **0.0115**** | **0.0071*** |
| | (0.0039) | (0.0029) |
| Higher Tax Bracket | **-3.06**** | **-0.86**** |
| | (0.17) | (0.14) |
| Buncher (Narrow) | 1.56** | 0.46** |
| | (0.27) | (0.14) |
| Bunching (Broad) | 0.84** | 0.60* |
| | (0.20) | (0.27) |
| Above threshold (Narrow) | -0.33 | -0.10 |
| | (0.18) | (0.17) |
| Above threshold (Broad) | -0.12 | 0.02 |
| | (0.11) | (0.10) |
| Constant (Avg across years) | 14.63 | 5.62 |
| Firm & Year fixed effects | YES | YES |
| Observations | 289,744 | 88,493 |

All firms with revenue in a 70 Million CRC window centered around the thresholds are included in the sample. Profit margin is defined as profit over revenue. The Bunching and above threshold are dummies for declaring revenue in the intervals around the threshold. Bunching narrow is defined as the having revenue in the half Million interval below the threshold. Bunching wide as having revenue between 4 and 0.5 Million below the threshold. Above threshold narrow is defined as having revenue between 0 to 3 Million above the threshold and wide as having revenue 3 to 9M above threshold. Standard errors are shown in parentheses and stars indicate statistical significance level. *=5% level, **=1% level.

Table 7: Revenue Shifting at End of Fiscal Year

| | Dependent Variable: Monthly Revenue (Million CRC) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | All firms | | | | CIT rev. = sales tax rev. | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Buncher*Sept | 0.10 | 0.05 | | | -0.17 | 0.30** | | |
| | (0.12) | (0.09) | | | (0.32) | (0.07) | | |
| Buncher*Oct | | | -0.02 | -0.21 | | | -0.62 | -0.12 |
| | | | (0.12) | (0.13) | | | (0.48) | (0.10) |
| Firm FE | NO | YES | NO | YES | NO | YES | NO | YES |
| Observations | 596,705 | 596,705 | 596,705 | 596,705 | 115,649 | 115,649 | 115,649 | 115,649 |

Table 7 tests for revenue shifting at the end of the fiscal year using the revenue declared for the monthly sales tax payment. Standard errors are shown in parentheses and stars indicate statistical significance level. *=5% level, **=1% level.

Table 8: Threat of Audit Impact at the Industry Level

| | Control 1: firms too small to bunch | | | Control 2: Dominated firms | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Outcome (Million CRC) | Revenue | Costs | Profit | Revenue | Costs | Profit |
| Bunch*Audit*Post | **4.87** | **6.43** | **-1.23** | 6.30 | 9.76 | -1,68 |
| | (2.39)** | (2,24)*** | (0.33)*** | (3.99) | (3.63)*** | (0.57)*** |
| Bunch*Post | -0.22 | -0.48 | -0.42 | 0.88 | 1.35 | -1.50 |
| | (2.13) | (2.00) | (0.31) | (3.20) | (2.81) | (0.37)*** |
| Audit*Post | 5.55 | 5.88 | -1.25 | 6.98 | 9.21 | -1.70 |
| | (3.10)* | (3.01)* | (0.29)*** | (2.89)** | (3.29)*** | (0.65)** |
| Firm FE | YES | YES | YES | YES | YES | YES |
| Observations | 7,203 | 7,203 | 7,203 | 4,672 | 4,672 | 4,672 |

Table 8 shows the results of the triple difference regression estimated from equation 20. The coefficient of interest is the triple interaction $Bunch * Audit * Post$ which shows the change in reported revenue, costs and profits of bunchers following an increase in their audit risk at the industry level. Standard errors, clustered at the industry level are shown in parentheses. Stars indicate statistical significance level. *=10% level, **=5% level, ***=1% level.

## Appendix A Additional Figures and Robustness

Figure A1: Average Profit and Costs by Revenue

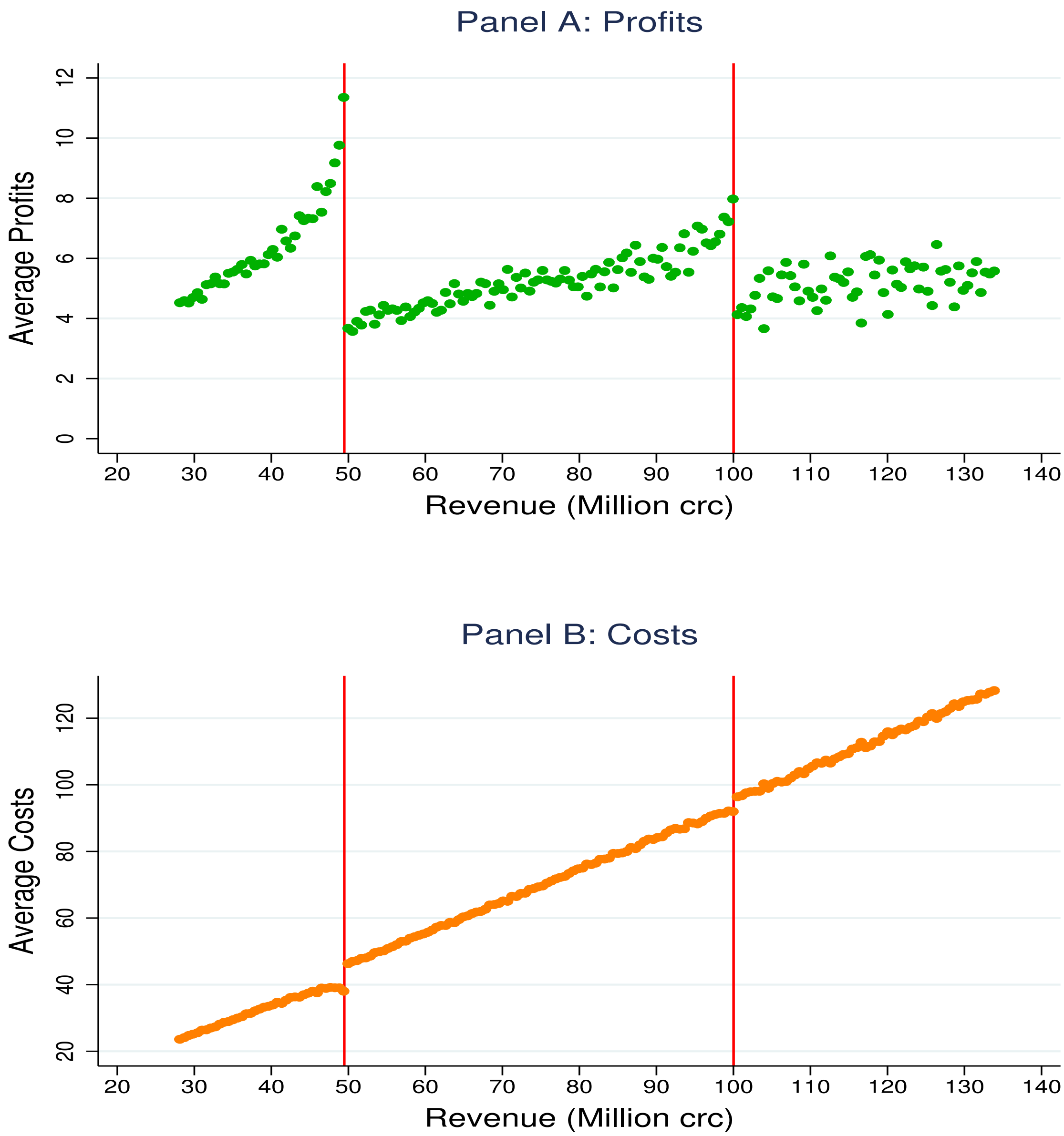



Figure A1 shows average profits (Panel A) and average costs (Panel B) by revenue, pulling together the years 2008 to 2014. The size of the revenue bins is 575,000 CRC.

Figure A2: Linear Relation of Average Costs by Revenue

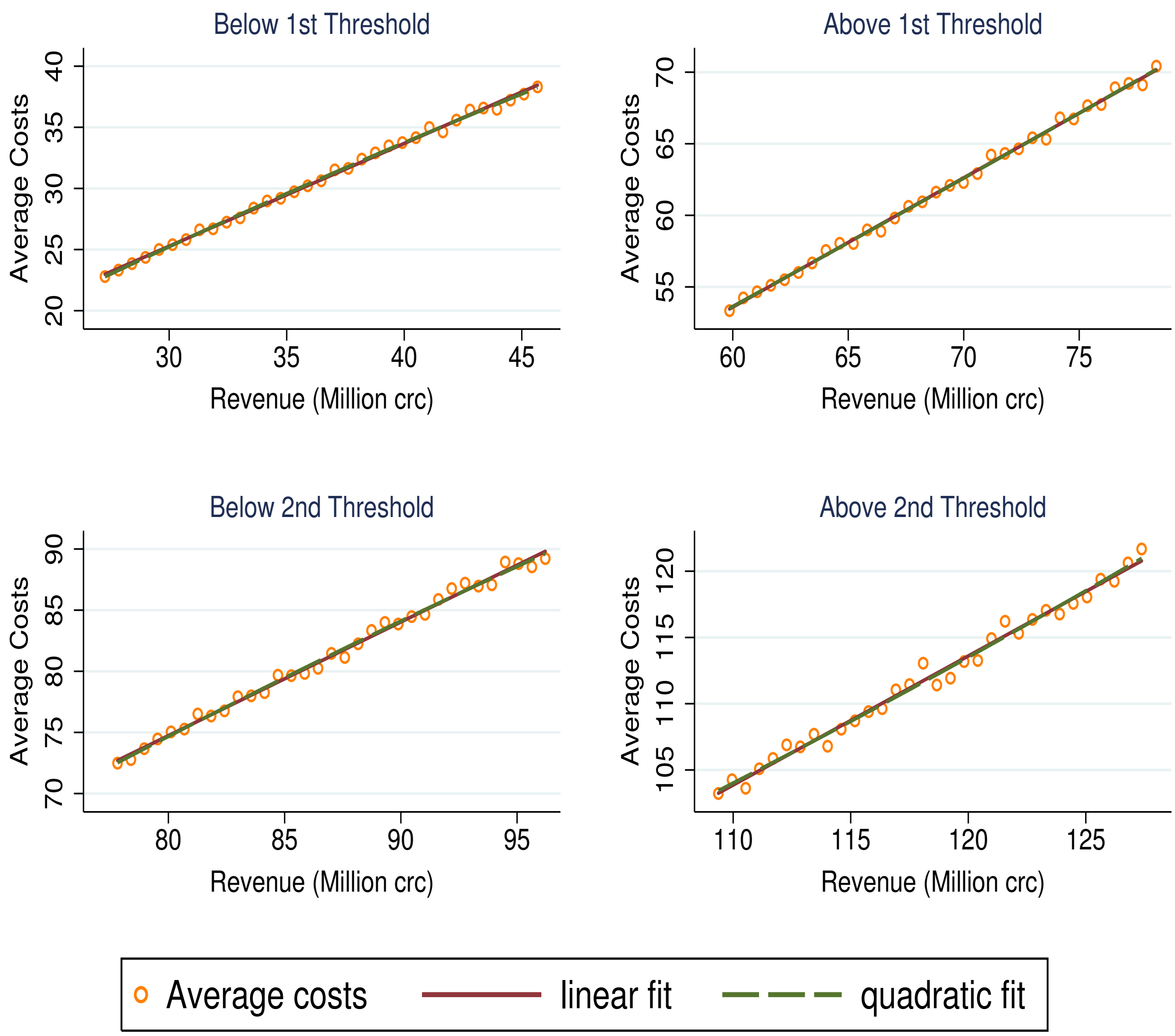



Figure A2 shows the linear relation of average costs by revenue. For each of four revenue intervals, below and above the first and second threshold, the linear and quadratic fit of the data. Quadratic fits are practically indistinguishable from Linear fits.

## Table A1: Robustness of Bunching Estimates

| **Panel A: Varying the order of the Polynomial** | | | |
|---|---|---|---|
| Order of Polynomial | 4 | 5 | 6 |
| First Threshold | | | |
| B | 2.4 | 2.2 | 2.2 |
| $y_u$ | 60 | 58.3 | 59.4 |
| $\epsilon_{y,1-\tau}$ | 0.46 | 0.33 | .41 |
| Second Threshold | | | |
| B | 1.1 | 1.1 | 1.1 |
| $y_u$ | 108.3 | 107.7 | 107.7 |
| $\epsilon_{y,1-\tau}$ | 0.10 | 0.08 | 0.08 |

| **Panel B: Varying the excluded zone, $y_l$** | | | |
|---|---|---|---|
| Number of excluded bins | 6 | 7 | 8 |
| First Threshold | | | |
| B | 2.0 | 2.2 | 2.3 |
| $y_u$ | 57.1 | 58.3 | 58.3 |
| $\epsilon_{y,1-\tau}$ | 0.25 | 0.33 | 0.33 |
| Second Threshold | | | |
| B | 1.1 | 1.1 | 1.0 |
| $y_u$ | 107.1 | 107.7 | 106.6 |
| $\epsilon_{y,1-\tau}$ | 0.07 | 0.08 | 0.06 |

Table A1 shows under different scenarios the estimates of the excess mass B, the revenue of the marginal buncher $y_u$ and the resulting revenue elasticity $\epsilon_{y,1-\tau}$. Panel A varies the order of the polynomial and Panel B the number of excluded bins on the lower side, which corresponds to $y_u$.

## Table A2: Adjusted R-squared of Average Costs on Revenue Regression

| Variable: **Adj. R-squared** | Order of Polynomial | | |
|---|---|---|---|
| Revenue Interval | Linear | Quadratic | Cubic |
| Below 1st Threshold | .9977 | **.9981** | .9981 |
| Above 1st Threshold | **.9971** | .9970 | .9969 |
| Below 2nd Threshold | **.9933** | .9933 | .9932 |
| Above 2nd Threshold | **.9872** | .9871 | .9870 |

Table A2 shows the model fit for different specifications of the regression of average costs on revenue. The simple linear model fits the data well and higher order terms are superfluous based on the adjusted R-squared. Only below the first threshold could the quadratic fit be preferred.

## Table A3: Alternative Models for Cost Discontinuity by Revenue

| | 1st Threshold | | | | 2nd Threshold | | | |
|---|---|---|---|---|---|---|---|---|
| Model Specification | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | Narrow Window | Wide Window | Falling $\epsilon_y$ | Quadratic | Narrow Window | Wide Window | Falling $\epsilon_y$ | Quadratic |
| **Jump in cost** $\delta$ | 2.688** | 2.465** | 2.326** | 3.804** | 1.389* | .85* | 1.277** | 3.113** |
| | (.249) | (.204) | (.246) | (.583) | (.591) | (.392) | (.432) | (.726) |
| Slope below T. | .823** | .841** | .834** | .603** | .924** | .944** | .933** | .821** |
| | (.012) | (.007) | (.009) | (.089) | (.02) | (.012) | (.015) | (.078) |
| $\Delta$ Slope above T. | .079** | .063** | .107** | .297** | .014 | .018 | .008 | -.052 |
| | (.017) | (.011) | (.014) | (.092) | (.042) | (.021) | (.026) | (.098) |
| Quadratic below T. | | | | -.009** | | | | -.004 |
| | | | | (.003) | | | | (.003) |
| $\Delta$ Quadratic above T. | | | | .009** | | | | .01** |
| | | | | (.003) | | | | (.003) |
| Intercept, $\alpha$ | 41.86 | 42.046 | 41.971 | 40.682 | 93.777 | 93.99 | 93.863 | 93.217 |
| Observations | 70 | 90 | 80 | 80 | 70 | 90 | 80 | 80 |
| % Jump in Cost $\frac{\delta}{\alpha}$ | +6.42% | +5.86% | +5.54% | +9.34% | +1.48% | +0.91% | +1.36% | 3.36% |

Table A3 shows the regressions of average costs by revenue on revenue for different model specifications. The parameter of interest is the jump in declared costs at the threshold, $\delta$, from Equation (14). Compared to the main specification of Table (2), Rows (1)-(2) & (5)-(6) vary the revenue interval over which the line is fitted. Rows (3) & (7) assume that the revenue elasticity is falling with revenue, at the speed estimated between the first and second threshold. Rows (4) & (8) assume a quadratic fit instead of a linear fit. An observation is a revenue bin of 0.575 Million Colones. Standard errors are shown in parentheses and stars indicate statistical significance level. *=5% level, **=1% level.