Optimal Dual-Regime Business Tax Systems*

Rishi R. Sharma
Department of Economics,
Colgate University

Joel Slemrod
Department of Economics,
University of Michigan

Michael Stimmelmayr
Department of Economics,
University of Bath

John D. Wilson
Department of Economics,
Michigan State University

Peter Choi
Department of Economics,
University of Michigan

This version: 15 September 2025

Abstract

Dual-regime business tax systems typically subject smaller firms to an output (turnover) tax and larger firms to a profit (corporate) tax. Despite their prevalence, there is little formal analysis of their optimal design. This paper addresses this gap by developing a theoretical framework to analyze the optimal tax parameters and the relative performance of two types of dual-regime systems: threshold and minimum tax systems. We show that either type of dual regime system can yield lower social costs than a single regime system. Using parameter values from recent empirical studies, we also show that a generalized minimum tax system we propose would outperform other dual regime systems under most parameter values. These findings carry important policy implications, particularly as many countries currently employ either threshold or minimum tax systems, but none have yet implemented a generalized minimum tax.

JEL-Classification: H25, H21

Keywords: dual-regime tax system, output tax, profit tax, bunching

*We thank Xinyu Chen, Michael Hartt, and Paul Kindsgrab for able research assistance. We also thank Marko Koethenbuerger and William Dodds for helpful comments as well as participants at the 2nd World Bank Tax Conference on Taxing Personal Income and Wealth in Developing Countries, the 65th annual meeting of the Standing Field Committee Public Economics of the German Economic Association, and the 79th Annual Congress of the International Institute of Public Finance.

1 Introduction

Many countries, especially developing ones, use two taxes to collect revenue from firms, a turnover tax for small firms and a profit tax for large firms. Often, the stated motivation of this type of tax system is to reduce the administrative and compliance costs incurred by small establishments, motivated by the belief (and credible evidence) that for income taxes these costs relative to measures of size are larger for small firms. To achieve this goal, firms below some size are subject to a simpler tax, usually with a flat rate and a relatively easily monitorable base such as output (sales). Thus, these firms are freed from the necessity to measure the costs of doing business. In addition, turnover taxes are presumed to be less susceptible to evasion, as costs are relatively difficult for most tax administrations to monitor; thus, such a dual-regime system arguably makes efficient use of administrative resources. On the other hand, a tax on output inefficiently distorts production, while a profit tax, or at least a pure profit tax with full cost deductibility and no evasion, is non-distortionary.

Dual-regime business tax systems that use a threshold to separate firms subject to the simpler turnover tax from firms subject to the profit tax are common in African countries such as South Africa, Egypt, and Kenya. Table 2 in Appendix A provides a non-exhaustive list of countries that, as of May 2024, apply such a dual-regime tax system. In most African countries, the threshold is set at about USD 50,000.[1] Firms with annual turnover below the threshold face a tax liability between 1 and 4 percent of turnover, while those firms exceeding the threshold are subject to the standard corporate tax regime. Several European countries also use a threshold tax system, with thresholds around USD 100,000 and turnover tax rates between 1 and 3 percent.[2]

In addition to threshold systems, several countries in Africa and Latin America use a minimum tax version of the dual-regime system. Under the minimum tax system, it is not a threshold that determines whether a firm faces the output tax or the profit tax, but instead each firm must pay the higher of the two taxes, given their chosen output level. Table 3 in the Appendix A presents a non-exhaustive list of countries that, as of May 2024, operate a minimum tax system. The design of minimum tax systems is very similar globally with the exception of India and South Korea. Countries levy tax rates between 1

[1] With the exception of Egypt and Kenya. The latter two countries have much higher thresholds of around USD 200,000.

[2] In some European countries, including France and Italy, the regime targets sole proprietors.

and 2 percent on turnover and corporate tax rates that range between 25 and 35 percent.[3]

Surprisingly, given the pervasiveness of such business tax systems, there is little formal analysis of even the basic optimal design questions.[4] In this paper, we begin such an analysis by investigating the optimal parameters of such a dual-regime system and what aspects of the environment influence the optimal structure. We investigate two kinds of dual-regime systems: (1) a threshold system where a firm with less than a given output is subject to an output tax, while larger firms face a profit tax; and (2) a minimum tax system where the firm owes the maximum of its notional output tax liability and its profit tax liability. We also introduce and analyze a variant of a minimum tax, which we label a generalized minimum tax. Under this variant, the government chooses the maximum percentage amount by which the profit (output) tax liability can exceed the output (profit) tax liability without triggering the requirement that the firm owes the profit (output) tax.

We investigate these questions using a specific-factors model where firms vary in their productivity levels. Firms can choose to operate in either an untaxed, constant-returns-to-scale (informal) sector or a formal sector that is subject to taxation. The entry of firms into the taxed sector is endogenous, and the prices of goods are fixed at the global market level. Depending on the dual-regime tax system in place, either an output threshold or the relative size of the tax liabilities determines whether a firm is subject to an output tax or a profit tax. Production costs are deductible under the profit tax and, additionally, firms can reduce tax liability by overreporting cost. The government faces a fixed revenue requirement and incurs a fixed per-firm administrative cost for each firm subject to the profit tax.

Calibrating the model, our findings show that, despite their differences, the threshold tax, the minimum tax, and the generalized minimum tax lead to very similar tax rate structures. For the assumed productivity distribution, the optimal output tax is roughly 8 percent, and the optimal profit tax rate is approximately 29 percent. However, the differences in the design of the threshold tax system vs. the minimum tax system imply that the nature of the output distortions and welfare effects vary considerably.

The existence of an output threshold causes firms to reduce output and bunch just below the threshold, making the threshold tax system the least favorable in terms of output distortion. Under the minimum tax system, firms cluster at a profit rate cutoff

[3]India and South Korea apply higher tax rates of around 15 percent on an adjusted base measure instead of gross turnover.

[4]The one exception is unpublished work by Wei and Wen (2019, 2023), which conducts simulations for different countries to quantify the optimal turnover threshold and tax rate for a given profit tax rate.

instead of an output threshold, leading to a continuous increase in the output of firms near this cutoff. The optimal generalized minimum tax system allows the profit tax liability to exceed the output tax liability by roughly 4 percent, and firms that bunch overproduce, raising their outputs above the level they would produce without taxation, in order to reduce profits and thus qualify for the output tax.

We determine the welfare effects of the different dual-regime systems by measuring the social costs of taxation. These costs include the loss from the tax-distorted entry decision of firms, the distortion in firms' output when operating under the output tax and of firms that bunch, the private costs of tax evasion due to cost overstatement under a profits tax, and the administrative costs of the government. Among the systems analyzed, the generalized minimum tax exhibits the lowest social costs, followed by the minimum tax and the threshold tax system, when applying parameter values from recent empirical literature. This finding is particularly relevant, given that several countries operate either a threshold or minimum tax system, but none have implemented a generalized minimum tax. Our results suggest that adopting a generalized minimum tax could significantly improve welfare of these countries.

The next section reviews the related literature, followed by Section 3, which introduces the basic model framework. Section 4 characterizes firm behavior under a threshold tax system and establishes the first-order conditions for the optimal system. Section 5 presents a minimum tax system and a generalized version of it, and characterizes the optimal parameters of such a system. Section 6 examines which of the systems is best, and the comparative static analysis is Section 7 identifies what country characteristics matter for the ranking. Section 8 concludes.

2 Literature Review

Although in this paper we address taxes on profits and output, a related literature has analyzed size-based adjustments to a value-added tax.[5] Notably, Keen and Mintz (2004) analyze the optimal threshold for a value-added tax system using a model where firms with heterogeneous productivity can choose to produce in either a taxed sector or a constant-returns-to-scale untaxed sector, and firms in the taxed sector incur administrative costs

[5]Exemptions in value-added tax systems are widespread. For example, the EU-wide simplified regime allows eligible small businesses to be exempt from VAT on domestic and cross-border supplies, with reduced compliance obligations like single registration, quarterly reporting, and simplified invoices.

for the government. The optimal threshold balances the gain in production efficiency with the revenue effects net of administrative costs when marginally adjusting the threshold.[6] Notably, compliance costs incurred by firms operating in the taxed sector do not affect the optimal choice of the threshold. This is because, for the marginal firm that bunches at the threshold, the savings in compliance costs equal the loss in profits from the necessary reduction in sales to stay just below the threshold.

Dharmapala et al. (2009) use a different model of heterogeneous firms, based on Melitz (2003), to investigate the optimal threshold, above which firms face an output tax and a fixed fee, i.e., a linear tax system. Again, administrative cost may justify not taxing sufficiently small firms, but a cutoff is found to be welfare-improving only if the number of untaxed firms is sufficiently large. Then, bunching of firms at the optimal cutoff creates a "missing middle" in the distribution of equilibrium outputs–a common phenomenon in developing countries.

Kanbur and Keen (2014) build on Keen and Mintz (2004) and analyze the optimal use of a threshold to partition compliant and non-compliant taxpayers. They find that the optimal threshold is set sufficiently high, i.e., above the optimal level defined in Keen and Mintz (2004) to disincentivize taxpayers from falsely declaring income just below the threshold.

Wei and Wen (2019) consider a model similar to Keen and Mintz (2004), but assume that firms below the threshold face a turnover tax while firms above the threshold face a profit tax, which is the setting we consider below. They derive conditions for the optimal turnover threshold and turnover tax rate, and conduct simulations for different exogenously given values of the profit tax rate. Similarly, Wei and Wen (2023) conduct simulations akin to those in their 2019 working paper, but use a simplified model with output heterogeneity instead of cost heterogeneity.

Best et al. (2015) study the consequences of a minimum tax regime in Pakistan, under which firms owe the higher of a tax on profits and a tax on turnover. Based on administrative data from the Federal Bureau of Revenue covering 2006-2010, they find "large and sharp" bunching of profit rates at the profit tax rate threshold, and that the extent of bunching varies with the corporate tax rate. They conclude that the extent of bunching, if it were caused only by real changes in output, would imply unrealistically ("phenomenally") large real output elasticities, due to the fairly small differences in the

[6]The gain in production efficiency originates from firms that bunch at the threshold and can increase their output when the threshold is marginally raised.

tax on output under the two taxes. They thus argue that most of the behavioral response is due to evasion in the form of overstated costs. Best et al. (2015) also solve an optimal tax problem where the government chooses the profit tax rate and the fraction of costs that can be deducted (which limits the gain from overstating costs), but do so in model with a representative firm, in contrast to the model with heterogeneous firms we develop below.

Bachas and Soto (2021) study the Costa Rican corporate tax system, which is not a dual-regime system of the type we focus on in this paper, but under which tax liability depends both on profits and revenue. In particular, the corporation owes an average tax rate of 10 percent, 20 percent, or 30 percent on its profit as a function of its revenue. This system provides incentives to bunch, and the magnitude of the bunching allows the authors to estimate the tax rate elasticity of reported profits, which they find to be substantial, between three and five, and largely due to unjustified overstatement of costs.

Mosberger (2016) studies the consequences of the introduction of a corporate minimum tax in Hungary in mid-2007, under which tax liability was the greater of the standard profit tax liability and two percent of revenues.[7] An analysis of tax administration data reveals sharp bunching in profit rates within a half year of implementation whose magnitude is positively related to the profit tax rate. Firms with more opportunity to over-report costs, such as small companies and companies in the construction and manufacturing sectors, exhibited larger bunching, while multinational companies exhibited less, arguably due to the presence of targeted audits on this sector. Contrary to theoretical expectations, they find that companies bunch above the kink point, which may be due to the perception that firms just below the cutoff would be perceived as evading firms and therefore subject to greater audit attention.

Alejos (2018) studies firm responses to the Guatemalan minimum tax system, under which firms owe the greater of a 31% tax on profits and a 0.93% tax on turnover, with one twist –the presence of exemptions from the turnover tax based on the firm's cost structure, depending on whether the reported profit rate is below or above an exogenous value. Analysis of the universe of corporate tax returns for the years 2006 to 2012 reveals clear evidence of bunching at the threshold profit margin of 3% (0.93/31*100) as well as at the exemption threshold, behavior that is consistent (only) with gross margin overreporting. As do Best et al. (2015), Alejos (2018) concludes that the extent and nature of bunching

[7]Notably, firms could continue to be subject to the profit tax if they were to submit detailed information on their cost structure and income items and be subject to a tax audit with high probability.

is consistent only with substantial cost overreporting, with estimates of the average ratio of reported-to-actual profits ranging between 0.42 and 0.91.

Lobel, Scot, and Zúniga (2024) examine the consequences of the introduction of a minimum tax in Honduras that was instituted in 2014. This tax was a hybrid of the tax regimes we study in this paper. Below a gross revenue threshold of L10 million (about $400,000), which increased to L300 million in 2018, firms owed 25% of declared taxable income. Above the threshold, they owed the maximum of the liability under a 25% profit tax and 1.5% of declared gross revenue. This system created a notch at the gross revenue threshold, and generated the incentives outlined in this paper for firms that report gross revenue above the threshold. The paper documents substantial bunching at the gross revenue threshold, and use the extent of bunching to estimate the elasticity of reported revenue with respect to the net-of-tax rate to be between 0.35 and 1.0. They argue that misreporting revenue drives part of this response, based on the fact that firms for which the tax authority has limited ability to independently assess declared revenue are more likely to declare revenue just below the exemption threshold. They also estimate that, when subject to profit taxation, firms increased their reported costs by 13-17%, with overstatement much less likely for costs that generate a paper trail, such as labor or financial costs.

3 Model Setup

Our basic setup is a modified version of the specific-factors model from international trade, amended to accommodate an endogenous number of heterogeneous firms. In this model, one sector utilizes a mobile factor, "labor", and a sector-specific factor, "entrepreneurs", to produce output Y. The other sector relies solely on mobile labor to produce output Z, using a constant-returns technology.[8] Both goods Y and Z are traded at fixed world prices, and markets are perfectly competitive. The wage received by a worker adjusts to satisfy the zero-profit condition in sector Z. We choose units of Z such that the equilibrium wage equals one, and units of Y such that the world price of Y equals one. Moreover, firms in sector Y are subject to either an output tax or a profit tax, but no tax is collected from

[8]This specification is used, for example, by Grossman and Helpman (1994), although their model has multiple sectors, each utilizing a sector-specific factor.

workers in sector Z.[9] Entrepreneurs differ in their productivities, and they enter sector Y until the marginal entrepreneur earns zero profits.

Without taxes, the total value of the outputs of sectors Y and Z, measured at world prices, is maximized in a competitive equilibrium. Under the requirement of balanced trade, the value of exports equals the value of imports. We consider only equilibria where the country produces both Y and Z; in other words, the economy does not specialize in the production of a single good.

When we introduce taxes and compare the dual-regime tax systems described in the introduction, we assume the government uses an output tax for smaller firms and a profit tax for larger firms. In developing countries, the usual motivation for implementing different taxes is that the output tax is easier to administer and avoid excessive administrative cost.[10] To model the tradeoff between profit taxation and output taxation, we assume that the government incurs a fixed administrative cost, denoted a, for each firm subject to the profit tax. This administrative cost reflects the relative difficulty of monitoring this tax, while, for simplicity, we assume that collecting the output tax incurs no such cost.

Following common practice in developing economies, as modeled by Gordon and Li (2009), we consider source-based taxes on output. These taxes lower the producer prices received by firms without affecting the prices paid by consumers, which are the world prices normalized to equal one.

Firms facing the output tax earn before-tax profits, $y - c\,(y, \varphi)$, with output denoted by y and total cost by $c\,(y, \varphi)$. This cost increases with output and decreases with firm productivity φ, i.e., $\frac{\partial c}{\partial y} > 0$ and $\frac{\partial c}{\partial \varphi} < 0$. We also assume that $\frac{\partial^2 c}{\partial y^2} > 0$, reflecting increasing marginal costs. The productivity parameter φ is distributed over the interval $[\varphi_m, \infty]$, $\varphi_m > 0$, with $H(\varphi)$ and $h(\varphi)$ denoting the cumulative distribution and density function, respectively. In addition, there is a fixed cost, $F = c\,(0, \varphi) > 0$, which can be the same for all φ, or increasing in φ.

Under the output tax, levied at rate t_o, after-tax profits are

$$\pi_o^A = (1 - t_o)y - c(y, \varphi), \tag{1}$$

[9] In developing countries, sector Z can be interpreted as the "informal" sector. For developed countries, we could extend the model to include a tax on wage income in both sectors.

[10] In Dharmapala et al. (2009), the government uses only an output tax but, due to a per-firm fixed cost for tax administration, it is optimal to tax only firms producing outputs above a certain threshold. In contrast, our model assumes that all firms in the Y sector are subject to one of the two taxes.

and profit maximization gives

$$1 - t_o = c', \tag{2}$$

where a prime denotes the derivative with respect to output, y. The first-order condition in (2) determines the supply function, $y_o(t_o, \varphi)$, and the after-tax profit function, $\pi_o^A(t_o, \varphi)$, for a firm operating under the output tax.

The output tax not only distorts the output decisions of existing firms, but it also causes some firms to exit the taxed sector of the economy. In the case where low-productivity firms face the output tax, the minimum productivity for firms that produce in this sector, $\varphi^l(t_o)$, is determined by the following zero-profit condition:[11]

$$\pi_o^A(t_o, \varphi^l) = 0. \tag{3}$$

Firms facing the profit tax owe a tax on their taxable income. We assume that all reported costs are tax deductible and that firms are able to overreport their costs, which incurs a private cost. With cost overreporting, actual profits, $y-c$, exceed reported profits, $(y-c)(1-\gamma)$, where γ represents the fraction of actual profits that are evaded. The private cost of this evasion behavior is proportional to actual profits. Specifically, $e(\gamma)$ denotes the cost per dollar of actual profits, in which case the total cost of evasion is $(y-c)e(\gamma)$.

The function $e(\gamma)$ satisfies $e\,(0) = e'(0) = 0$, and for $0 < \gamma < 1$, $e(\gamma) > 0$, $e'(\gamma) > 0$ and $e''(\gamma) > 0$. This reflects the idea that cost overreporting becomes increasingly difficult to conceal as the discrepancy between reported profits and actual profits increases. Using this notation, after-tax profits can be written as follows:[12]

$$\pi_p^A = (y-c)(1-e) - t_p(y-c)(1-\gamma). \tag{4}$$

In this model, all firms subject to the profit tax will choose the same profit-maximizing γ, which satisfies $e'(\gamma) = t_p$. This condition allows us to define the function, $\gamma(t_p)$, which is the same for all firm productivities. This property greatly simplifies simulations, where

[11]The minimum productivity $\varphi^l(t_o)$ is increasing in the output tax t_o because firms require a higher productivity to meet the zero-profit condition if the output tax increase.

[12]Alternatively, overreporting costs can be modeled as a function increasing in the percentage amount by which costs are inflated: reported costs are $\beta\, c$ for $\beta > 1$, and the cost of this overreporting is given by $e(\beta)\, c$, where $e(\beta)$ is an increasing convex function defined for $\beta \geq 1$, with $e(1) = 0$. However, if overreporting costs are sufficiently low, then some firms will not only reduce reported profits to zero, but will also report negative profits and receive a subsidy, if allowed. The current specification avoids this issue.

we assume $e(\gamma) = \beta\gamma^{\varsigma}$, with $\varsigma > 1$ and $\beta > 0$.[13]

The ratio of the marginal cost of evasion, $MEC = e'(\gamma)$, to the average cost, $AEC = e/\gamma$, satisfies $MEC/AEC = \varsigma$. A higher MEC reduces evasion, whereas a lower AEC implies fewer resources "wasted" on evasion activities. Thus, increasing ς while decreasing β to keep $\beta\varsigma$ fixed tends to make profit taxation more favorable compared to output taxation.

When a firm faces the profit tax, the first-order condition for its profit-maximizing output is

$$1 = c'. \tag{5}$$

Thus, output is provided efficiently, implying a marginal tax on output of zero. Consequently, output can be defined as a function of productivity alone: $y_p(\varphi)$. We assume that profits $y_p(\varphi) - c(y_p(\varphi), \varphi)$ approach infinity as productivity increases to infinity.

Under any dual-regime tax system, the government chooses taxes to ensure that tax revenue net of administrative costs meets the government's exogenous business tax revenue requirement, denoted G. To formulate the objective function, recall that prices are determined exogenously on world markets and that the total supply of labor across the formal Y and informal Z sectors is fixed. Additionally, the wage, which is fixed by the free entry condition for the Z sector, is normalized to one. In this setting, welfare maximization is achieved by maximizing after-tax profits.

The general logic of an optimal dual-regime tax system is as follows: the government desires to impose the profit tax on relatively large firms because this tax does not distort output and administrative costs per unit of tax revenue are relatively low. Similarly, it avoids excessive administrative costs by imposing the output tax on small firms, despite the distorting effect of the tax on output. However, some of the large firms will respond by reducing their outputs sufficiently to just qualify for the output tax. This firm behavior leads to bunching, either at a specific output level, as under a threshold tax system, or at the cutoff profit rate, as under a minimum tax system.

In equilibrium, three types of firms can be identified: First, firms with outputs up to a productivity level φ^{**}, which are subject to the output tax and do not bunch. Second, firms with productivity levels between φ^{**} and φ^{*} decide to "bunch" in order to qualify for the relatively lower output tax. The outputs chosen by these firms, along with the value of

[13]Under this specification, $\beta\varsigma\gamma^{\varsigma-1} = t_p$, can be solved to obtain a closed-form solution for γ, showing that an increase in t_p will lead to a larger γ.

φ^*, will depend on the type of dual regime system in place. Moreover, the government may be able to control the extent of bunching through the use of a non-tax policy variable that affects both φ^* and outputs. For the threshold model, this variable is simply the threshold output level, y^*, above which firms are subject to the profit tax. For firms that face the output tax, including firms that bunch, total output and total cost can be written:

$$Y_o(t_o, y^*, \varphi^{**}) = \int_{\varphi^l(t_o)}^{\varphi^{**}} y_o(t_o,\ \varphi)\, h(\varphi)\, d\varphi + y^*[H(\varphi^*) - H(\varphi^{**});$$

$$C_o(t_o, y^*, \varphi^{**}) = \int_{\varphi^l(t_o)}^{\varphi^{**}} c(y_o(t_o,\ \varphi),\ \varphi)\, h(\varphi)\, d\varphi + \int_{\varphi^{**}}^{\varphi^*} c(y^*,\ \varphi) h(\varphi)\, d\varphi.$$

In the case of the minimum tax or the generalized minimum tax, the behavior of bunching firms is more complicated, and we defer it to a later section. Third, firms above the productivity level φ^* produce the efficient levels of outputs, and total output and total cost of these firms are as follows:

$$Y_p(\varphi^*) = \int_{\varphi^*}^{\infty} y_p(\varphi)\, h(\varphi)\, d\varphi;$$

$$C_p(\varphi^*) = \int_{\varphi^*}^{\infty} c(y_p(\varphi), \varphi) h(\varphi)\, d\varphi.$$

Once these outputs and costs are determined, the government's problem is to choose the tax rates and the tax instrument variable, b, that is, the output threshold y^* in the case of the threshold tax, or the multiplier δ in the case of the generalized minimum tax, to maximize after-tax profits subject to the government's budget constraint:

$$\begin{aligned} \max_{t_o, t_p, b} \quad & (1 - t_o)Y_o - C_o + [Y_p - C_p][(1 - e(t_p)) - t_p(1 - \gamma(t_p))] \\ s.t. \quad & t_o Y_o + t_p(Y_p - C_p)(1 - \gamma(t_p)) - [1 - H(\varphi^*)]a = G. \end{aligned} \tag{6}$$

In the following two sections, we derive the solution to this problem for the threshold system and the minimum tax and generalized minimum tax systems.

4 The Optimal Threshold Tax System

In this section, we analyze a dual-regime tax system that uses an output threshold y^* to separate firms that face the profit tax from those facing the output tax. A firm that is indifferent between facing the two taxes has the cutoff productivity $\varphi^*(t_o, t_p, y^*)$, which depends on both tax rates and the output threshold, so that:

$$(y_p(\varphi^*) - c(y_p(\varphi^*), \varphi^*))[(1-e) - t_p(1-\gamma)] = (1-t_o)y^* - c_p(y^*, \varphi^*).$$

Differentiating this indifference condition when there is bunching gives

$$\frac{\partial \varphi^*}{\partial t_0} < 0; \qquad \frac{\partial \varphi^*}{\partial t_p} > 0; \qquad \frac{\partial \varphi^*}{\partial y^*} \geq 0.$$

An increase (decrease) in φ^* represents a decrease (increase) in the fraction of firms subject to the profit tax. The derivative $\frac{\partial \varphi^*}{\partial y^*}$ equals zero when there is no bunching, because $y^* = y_o(t_o, \varphi^*)$, implying that a type-$\varphi^*$ firm's after-tax profits under the output tax are unaffected by a marginal decrease in y^*. Below, we provide a proof of the desirability of profit taxation that does not rely on bunching.

A type-φ^{**} firm strictly prefers the output tax while optimizing profits at y^* : $y^* = y_o(t_o, \varphi^*)$. This equality defines the function, $\varphi^{**}(t_o, y^*)$, and differentiating this function yields:

$$\frac{\partial \varphi^{**}}{\partial y^*} > 0; \qquad \frac{\partial \varphi^{**}}{\partial t_0} < 0.$$

According to the envelope theorem, changes in φ^{**} have no effect on after-tax profits.

In Appendix B, we prove

Proposition 1. *If administrative costs are positive but sufficiently low, and overreporting costs are sufficiently low, then a dual-regime threshold system is optimal.*

To address the issue of how to design a threshold system that provides higher welfare than an output tax system, we introduce an additional assumption: the ratio of a firm's profit tax liability to its liability under an output tax increases with productivity. This is equivalent to assuming:

$$\frac{y_p(\varphi) - c_p(y_p(\varphi), \varphi)}{y_o(t_o, \varphi)} \quad \text{increases with } \varphi. \tag{7}$$

This assumption holds for our simulations, which use a quadratic function for variable costs. The assumption corresponding to equation (7) for the minimum tax system will be that each firm's profit rate, $(y-c)/y$, is increasing with φ. The presence of the fixed cost F contributes towards satisfying both assumptions. But the formulas differ because different outputs appear in the numerator and denominator of equation (7).

Proposition 2. *Starting from an output tax system with the budget-balancing tax rate t_o^s, consider designing a threshold system that divides firms between the two taxes: firms with productivities less (greater) than φ^* pay the output tax (profit tax) and $\varphi^* > \varphi_l(t_o^s)$. If administrative costs and overreporting are positive but sufficiently small, the budget-balancing tax rates, t_o^d and t_p^d, and the threshold y^* can be chosen so that $\varphi^* > \varphi_l(t_o^d)$, resulting in a dual-regime threshold system that provides higher welfare than the output tax system.*

The proof of Proposition 2 appears in Appendix B.

Proposition 2 tells us that a dual-regime threshold system with any division of firms between the output and the profit tax can be be welfare-improving if administrative costs and overreporting are low enough. However, the proof of Proposition 1 shows that low thresholds run the risk of allowing some firms' tax payments to fail to cover administrative costs, even if these costs are relatively low. In general, we expect the fraction of firms subject to the profit tax under an optimal threshold system to fall as administrative costs rise, a result that is confirmed by our comparative static analysis in Section 7.1.

We next state and use the first-order conditions for an optimal dual-regime tax system with bunching. Let λ denote the Lagrange multiplier for the government budget constraint, which measures the social marginal utility of tax revenue and exceeds one. Differentiating the Lagrangian for the optimal tax problem with respect to t_o gives the first-order condition for the output tax rate:

$$(1-\frac{1}{\lambda})Y_o + t_o\frac{\partial Y_o}{\partial t_o} + \left[t_o y^* - (t_p\pi_p^{T*} - a)\right] h(\varphi^*)\frac{\partial \varphi^*}{\partial t_o} = 0, \tag{8}$$

where π_p^{T*} denotes taxable profits for a type-φ^* firm that chooses the profit tax. A unit rise in the output tax shifts revenue from the private sector to the government, creating a welfare gain, represented by the first term. However, it also creates the two behavioral changes. First, it lowers the outputs of firms facing the output tax, creating the marginal

deadweight loss, $ML_o = -t_o \frac{\partial Y_o}{\partial t_o}$. Second, it causes $-h(\varphi^*)\frac{\partial \varphi^*}{\partial t_o}$ firms to switch from the output tax to the profit tax. This raises tax revenue net of administrative costs if a type-φ^* firm's profit tax liability net of administrative costs exceeds its output tax liability.

The first-order condition for the profit tax may be written

$$[Y_p(\varphi^*) - C_p(\varphi^*)][(1-\gamma)(1-\frac{1}{\lambda}) - t_p\gamma'] + \left[t_o y^* - (t_p \pi_p^{T*} - a)\right] h(\varphi^*)\frac{\partial \varphi^*}{\partial t_p} = 0. \qquad (9)$$

Similar to the output tax, a unit rise in the profit tax is creating a beneficial shift of revenue from the private sector to the government, but it also causes a marginal deadweight loss through its positive effect on privately costly cost overreporting: $ML_p = t_p[Y_p(\varphi^*) - C_p(\varphi^*)]\gamma'(t_p)$. In addition, it is also causing firms to switch from the profit tax to the output tax, thereby again affecting tax revenue net of administrative costs.

The final first-order condition is for the output threshold y^*. If the government attempts to expand the number of firms subject to the output tax by increasing y^*, the envelope theorem implies that the resulting changes in φ^* and φ^{**} will not directly affect after-tax profits, as these are set at their profit-maximizing levels. But bunched firms are constrained by the requirement that their outputs must not exceed y^* for them to qualify for the output tax. A unit increase in y^* changes their profits by

$$\int_{\varphi^{**}(t_o, y^*)}^{\varphi^*(t_o, t_p, y^*)} [1 - t_o - c_y(y^*, \varphi)]\, h(\varphi)\, d\varphi > 0.$$

Next, the increase in y^* raises the social value of the tax revenue obtained from firms producing at y^*:

$$\lambda\, t_o[H(\varphi^*) - H(\varphi^{**})] > 0.$$

Finally, the increase in y^* causes more firms to bunch at y^*, implying a rise in φ^* that changes the social value of tax revenue by the amount

$$-\lambda\left[(t_p \pi_p^{T*} - a - t_o y^*\right] h(\varphi^*)\frac{\partial \varphi^*}{\partial y^*}.$$

Summing and rearranging these effects gives the first-order condition

$$\int_{\varphi^{**}(t_o,y^*)}^{\varphi^*(t_o,t_p,y^*)} [1 - c_y(y^*, \varphi)]\, h(\varphi)\, d\varphi + t_o(\lambda - 1)[H(\varphi^*) - H(\varphi^{**})]$$
$$- \left[t_p \pi_p^{T*} - a - t_o y^*\right] h(\varphi^*) \frac{\partial \varphi^*}{\partial y^*} = 0. \tag{10}$$

Because firms that bunched at y^* have reduced their outputs below their profit-maximizing levels under the profit tax, the integral in this expression is positive. Thus, we have

Proposition 3. *Under an optimal threshold tax system with bunching, type-φ^* firms face a profit tax liability that exceeds their output tax liability by an amount greater than the administrative cost a.*

Bunching occurs when $t_p \pi_p^{T*} - t_o y^* > 0$, because firms attempt to escape the profit tax by reducing their outputs to the threshold. The proposition is telling us that this difference in tax payments will exceed the administrative cost when the tax system is optimized.

We can combine all three first-order conditions to prove

Proposition 4. *Under an optimal threshold tax with bunching, $t_o \left(-\frac{\partial Y_o}{\partial t_o}\frac{1}{Y_o}\right) > t_p \left(\frac{\gamma'}{1-\gamma}\right)$.*

Proof. Divide (8) by the output tax base Y_0 and (9) by the profit tax base $(Y_P - C_P)(1-\gamma)$, and then subtract the latter from the former to obtain

$$t_o \left(-\frac{\partial Y_o}{\partial t_o}\frac{1}{Y_o}\right) - t_p \left(\frac{\gamma'}{1-\gamma}\right) = \left[t_p \pi_p^{T^*} - a - t_o y^*\right] h(\varphi^*) \Psi,$$

where

$$\Psi = \frac{\partial \varphi^*}{\partial t_0}\frac{1}{Y_o} - \frac{\partial \varphi^*}{\partial t_p}\frac{1}{(Y_p - C_p)(1-\gamma)} > 0.$$

Having shown that $t_p \pi_p^{T*} - a - t_o y^* > 0$, this equality gives the result. □

Thus, the marginal deadweight loss from the output taxation per unit of output tax base must be greater than the marginal deadweight loss from overreporting per unit of profit tax base. This condition puts a limit on the amount of overreporting that is compatible with profit taxes that are high enough to induce firms to attempt to avoid them by bunching at the threshold y^*. Specifically, the condition fails for high levels of overreporting relative to the price elasticity of output.

Simulations: Optimal Threshold Tax

In the following, we solve a discretized version of the model, using a quadratic cost function of the form, $c(\varphi) = \frac{1}{2\varphi} y^2 + f$, and a cost overreporting function given by $e(\gamma) = \beta\gamma^{\varsigma}$.[14] We set the fixed cost to $f = 0.5$, ensuring that firms with minimum productivity break even in the absence of (output) taxation and that any firm shutdown in equilibrium results from the tax system. Moreover, following the findings of Lobel, Scot, and Zúniga (2024), we set the parameter values to $\beta = 8.3$ and $\varsigma = 3$, which yields a cost overreporting rate of 10% under a 25% profit tax and ensures a plausible distribution of firms under both the output and profit taxes across all three dual-regime systems.[15]

For the simulation analysis, we draw 1,000 firm productivities from a truncated (i.e., bounded) Pareto distribution with a shape parameter of $\alpha = 1.20$, following Tybout (2014), and a normalized scale parameter (minimum productivity) of $\varphi_m = 1$. The upper bound of the productivity distribution is set to 250 times the mean productivity, resulting in a broadly reasonable revenue share of approximately 30% for the top 1% of firms in the undistorted baseline.[16] Moreover, the per-firm administrative cost under the profit tax is set to $a = 0.0168$, reflecting administrative cost equivalent to 4% of revenues, and the government revenue requirement from the business tax system is set at 10% of undistorted formal sector output.[17] For the threshold system, this calibration process results in optimal tax rates of $t_o = 0.082$ for the output tax and $t_p = 0.291$ for the profit tax, along with an optimal output threshold of $y^* = 2.52$.

Figure 1a illustrates the output of firms within the lowest 85% of the productivity distribution under a single-regime output tax (red line) or a single-regime profit tax with full cost deductibility and no cost overreporting (black line).[18,19] Because the profit tax

[14]These functions allow us to derive closed-form solutions for the firm's optimum and the bunching regions under both dual-regime tax systems. In Appendix C, we present the specific expressions and derivations, along with a discussion of the empirical literature used for calibration.

[15]In Section 7.1, we conduct a comparative static analysis using values ranging from 50% to 150% of the β value used in the baseline simulations. This range corresponds to cost overreporting rates of 9–15%, aligning with the findings of Best et al. (2015), Alejos (2018), and Lobel, Scot, and Zúniga (2024).

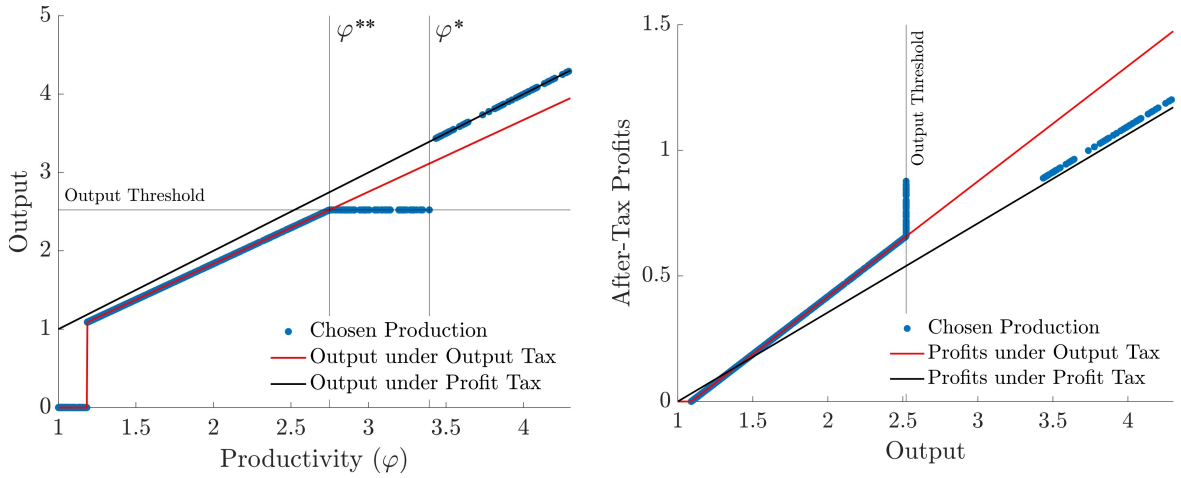
[16]We use a truncated Pareto distribution for computational efficiency, which is also commonly employed in the international trade literature to model firm productivity (e.g., Helpman et al., 2008).

[17]In Sections 7.2 and 7.3, we perform a comparative static analysis for both parameter values.

[18]All figures are restricted to firms in the lowest 85% of the productivity distribution to more clearly depict the behavior of firms near or within the bunching region.

[19]These single-regime lines represent what firm choices would be under each regime in the dual system. They do not represent choices under the optimal single regime systems. This applies to all similar figures (Figures 2, 3, and 5).

Figure 1: Threshold Tax System (lowest 85% productivity firms)



(a) Output vs. Productivity

(b) After-Tax Profit vs. Output

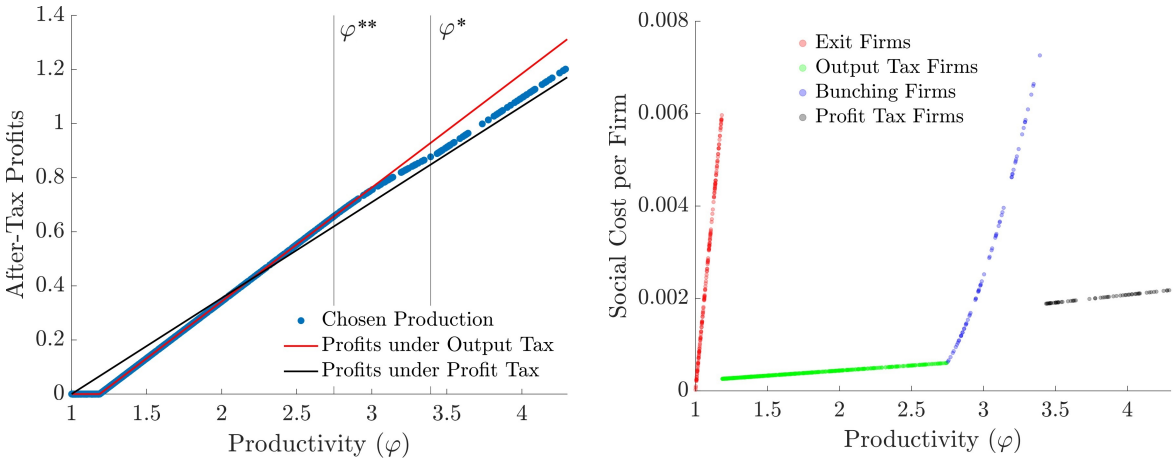
does not distort firms' output decisions and firms are able to overreport cost, output is consistently higher under this regime.

The blue dots represent the actual profit-maximizing output of firms under the optimal threshold system. Given our calibration, firms with productivity levels below $\varphi^{l} = 1.19$ operate in the informal Z sector, as their productivity is insufficient to cover the fixed cost and taxes required to begin production in the formal Y sector. Firms with productivities between φ^{l} and $\varphi^{**} = 2.75$ produce at levels consistent with a single-regime output tax. In contrast, firms with productivity exceeding $\varphi^{*} = 3.45$ produce slightly above what they would under a single-regime profit tax due to cost overreporting. For firms with productivity between φ^{**} and φ^{*}, it is optimal to reduce output to the threshold level y^{*} to remain subject to the output tax, rather than producing as they would under a single-regime profit tax.

Figure 1b displays firms' after-tax profits on the vertical axis and their output on the horizontal axis. Under the optimal tax rates of the threshold system, after-tax profit is generally higher for firms subject to the output tax (red line), except for those with relatively low productivity. Starting at a productivity level of $\varphi^{**} = 2.75$, firms produce exactly $y^{*} = 2.52$ units of output. Note, though, that firms with higher productivity incur lower costs when bunching at the output threshold, leading to higher after-tax profits. At a productivity level of $\varphi^{*} = 3.45$, a firm's output jumps from the threshold level of 2.52 to 3.17 at the cost of being subject to the profit tax.

Notably, and as illustrated in Figure 2a, the after-tax profits of firms with productivity

Figure 2: Threshold Tax System (lowest 85% productivity firms)



(a) After-Tax Profit vs. Productivity

(b) Social Costs

levels between $\varphi^{**} = 2.75$ and $\varphi^{*} = 3.45$ exhibit a continuous increase. Thus, by reducing their output levels to output threshold y^{*}, firms become subject to the lower output tax rather than the higher profit tax, enabling them to increase their after-tax profits.

Figure 2b depicts the social cost per firm under the optimal threshold system. Here, the social cost is measured by the difference in pretax profit between a no-tax regime and the optimal threshold regime as a percentage of total undistorted profits. Firms with productivity below $\varphi^{l} = 1.19$ operate in the informal Z sector because their productivity is not sufficient to offset the fixed cost and output tax necessary to start production in the formal Y sector. The social cost per firm, due to a firm not entering the formal sector, is represented by the red dots and it increases with firm productivity. The fraction of firms not entering the formal Y sector amounts to 17.9%. Firms with productivities above φ^{l} up to $\varphi^{**} = 2.75$ produce unconstrained output under the output tax. Because costs are not deductible under the output tax, taxation results in production inefficiency. The social cost from the tax-induced output distortion is depicted by the green dots which represent roughly 54.2% of all firms. Each of the blue dots represents the social cost from firms that bunch. The social cost increases with firm productivity, because the loss of output is higher, reaching its maximum for the firm with productivity $\varphi^{*} = 3.45$, which is indifferent between bunching at the output threshold $y^{*} = 2.52$ or producing unconstrained output under the profit tax. The portion of firms that bunch at the threshold amounts to 7.4%. Finally, the black dots represent the social costs generated by firms subject to the profit tax. While the full deductibility of costs under the profit tax ensures that it does not distort firms' output, the social costs involve fixed per-firm administrative cost and private

concealment cost of overreporting. The latter explains the increase in social costs under the profit tax. The share of firms operating under the profit tax amounts to 20.5%.

5 Minimum Tax Systems

As detailed in Table 3 of Appendix A, many countries implement a minimum tax version of a dual-regime business tax system. Instead of employing an output threshold to determine whether a firm should face the profit tax or the output tax, this regime mandates that each firm owes the higher of the two tax liabilities. In this section, we analyze the optimal minimum tax system and a generalized version of it. Then, in Section 6, we compare the welfare implications of the two minimum tax systems versus a threshold system.

Assuming the overreporting cost function is separable, a firm's tax liability under the minimum tax system is determined by the following expression:

$$T(y, \varphi) = \max \left(t_p(y - c\,(y, \varphi))\,(1 - \gamma\,), t_o\, y \right).$$

Implicitly, a cutoff rule, involving profit rates rather than output levels, determines the type of tax a firm will face:

$$\frac{(y - c\,(y, \varphi))\,(1 - \gamma\,)}{y} = \frac{t_o}{t_p}. \tag{11}$$

Firms with profit rates above t_o/t_p are subject to the profit tax, while those with profit rates below t_o/t_p face the output tax. Under this minimum tax system, firms bunch at the cutoff profit rate rather than at a cutoff output level.[20]

One disadvantage of the minimum tax system, compared to the threshold system, is that the government has only the two statutory tax rates to control the share of firms subject to the output tax and the profit tax. In contrast, a threshold system has three: it provides the government with an additional control variable, the threshold output level, y^*. This disadvantage can be mitigated by allowing the government to choose the maximum percentage amount by which the profit tax liability may exceed the output tax liability without triggering the requirement for the firm to become subject to the profit tax. Alternatively, it may be desirable to allow the output tax liability to exceed the profit

[20]Best et al. (2015) investigate bunching both theoretically and empirically, but do not present an explicit model of firm heterogeneity.

tax liability by a maximum percentage amount without triggering the requirement that the firm pays the output tax.

To accommodate both possibilities, we generalize the profit rate cutoff by introducing a parameter δ, which can be either greater or less than one, so the cutoff rule becomes:

$$\frac{(y - c(y, \varphi))(1 - \gamma)}{y} = \frac{\delta t_o}{t_p}. \tag{12}$$

In the case where $\delta > 1$, only firms with profit tax liabilities that exceed output tax liabilities by $(\delta - 1) * 100$ percent owe the profit tax. If $\delta < 1$, firms face the profit tax as long as their output tax liability does not exceed $(1 - \delta) * 100$ percent of their output tax liability. Our numerical simulations show that a value of δ greater than one improves welfare compared to a standard minimum tax rule, where δ equals one, and we also prove that raising δ above one raises welfare in the case of no cost overreporting. Therefore, we focus on this scenario ($\delta > 1$).

As before, the marginal cost curve declines as the productivity parameter increases, implying that firms with higher productivity produce more output under a given effective marginal tax rate on output, t^E. Additionally, we restrict the production technology to ensure that firms with higher productivity have higher profit rates under this effective marginal tax rate, given the level of overreporting. This increasing profit rate is achieved by incorporating fixed costs into the model, such that firms with higher productivity parameters have lower average fixed costs, while also imposing certain restrictions on variable costs.[21] However, there are also cost functions without fixed costs that satisfy these properties.[22]

We again consider an equilibrium in which some firms are subject to an output tax while others are subject to a profit tax. Under our specification of cost differences, firms facing the output tax are those with relatively low productivity, as their reported profit rates are below the cutoff. Conversely, high-productivity firms are subject to the profit tax because their profit rates are above the cutoff. In addition, there exists an intermediate range of firms that bunch at the profit rate cutoff, $\delta t_o / t_p$. Using previous notation, the set of productivities of firms that strictly prefer this profit rate is the interval, $(\varphi^{**}, \varphi^{*})$, and the outputs of these bunched firms increase with productivity.

[21] For example, an isoelastic variable cost function of the form y^{β}/φ can be used.

[22] For the cost function $c = (be^{dy} - b)/\varphi$, the equilibrium profit rate under a given marginal tax rises with productivity if b and d are sufficiently small.

Before stating the government's optimization problem, we must address a discontinuity in the relation between welfare and the policy instrument δ. Under the standard minimum tax system where $\delta = 1$, the government can require all bunched firms to face the output tax, as these firms are indifferent between the two taxes. However, when δ falls slightly below one, all bunched firms opt for the profit tax, because their profit tax liabilities are lower than their output tax liabilities. Consequently, total administrative costs will increase by a discrete amount as firms opt for the profit tax, causing a discontinuous drop in welfare. Conversely, welfare changes continuously as δ rises above one. Because bunched firms are already facing the output tax when $\delta = 1$, administrative costs decline continuously as δ increases above one. When $\delta > 1$, firms in the interval $(\varphi^{**}, \varphi^{*})$ strictly prefer to face the output tax, and their optimization problem can be stated as the maximization of profits net of output tax payments, subject to the constraint that their profit rate equals the cutoff rate.

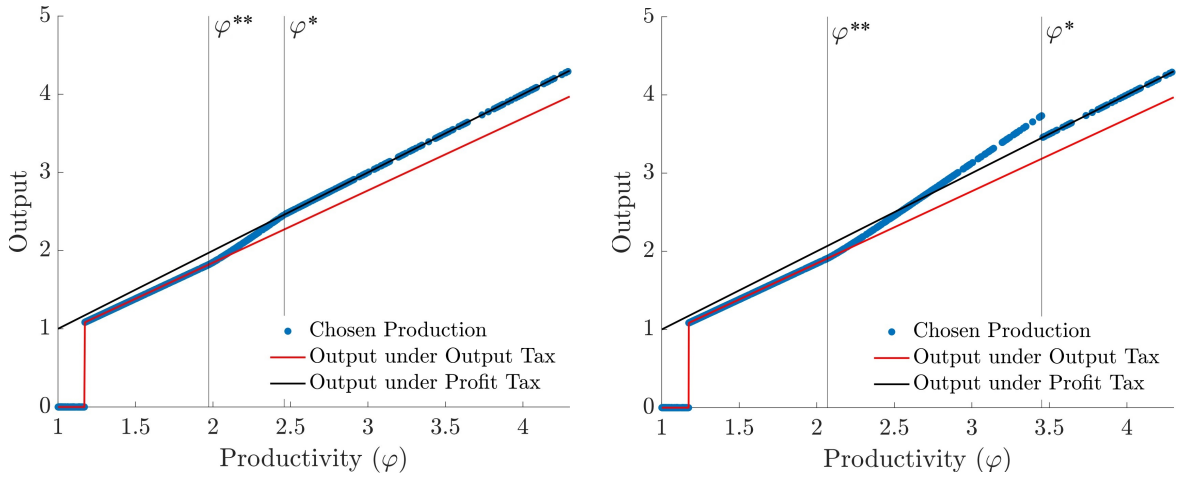
As illustrated by the simulations, when $\delta > 1$, type-φ^{*} firms are willing to set their output and overreporting levels above their unconstrained levels under the profit tax. Nevertheless, they remain indifferent about bunching because choosing not to bunch will enable them to discontinuously raise their profit rate to its unconstrained level, in which case they face the profit tax. Doing so involves discontinuously reducing both output and overreporting levels, which have been kept higher to satisfy the cutoff rule for bunching.

Simulations: Optimal Minimum and Generalized Minimum Tax

Using the same model calibration as for the threshold system (see Appendix C for details), we run simulations for the optimal minimum and optimal generalized minimum tax system. We find that the optimal tax rates under both minimum tax systems are very comparable to those under the optimal threshold system. Specifically, the optimal output tax under the minimum tax and the generalized minimum tax amounts to 0.0769 and 0.0773, respectively, while the optimal output tax rate under the threshold system is 0.0822. Similarly, the optimal profit tax under the two minimum tax systems is 0.2913 and 0.2910, respectively, whereas under the threshold system, it is 0.2907.

Figures 3a and 3b illustrate firms' optimal output choices as a function of productivity for the optimal minimum tax system, where $\delta = 1$, and the optimal generalized minimum tax system, where $\delta = 1.042$. Unlike under the threshold tax, no discontinuity in the firms' output choice emerges under the minimum tax (Figure 3a). The positive output tax distorts firms' entry decision, i.e., increases the minimum required productivity from

Figure 3: Output vs. Productivity (lowest 85% productivity firms)



(a) Minimum Tax

(b) Generalized Minimum Tax

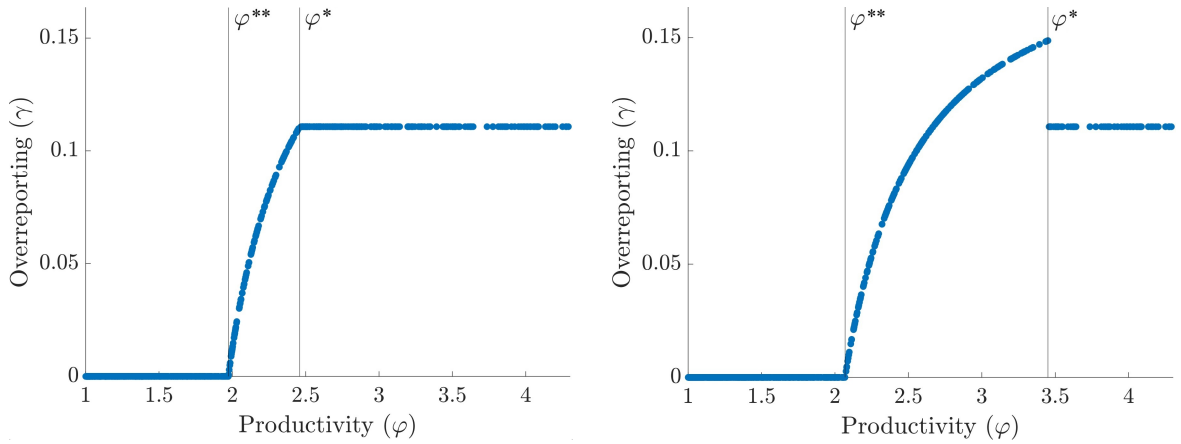
$\varphi = 1$ to $\varphi^l = 1.17$. Firms with productivity levels above φ^l up to $\varphi^{**} = 1.97$ are subject to the output tax and produce output as they would under a pure output tax (red line). Firms with productivity levels above $\varphi^* = 2.46$ are subject to the profit tax and produce output as they would produce under a pure profit tax with the possibility of cost overreporting. Firms with productivity levels between $\varphi^{**} = 1.97$ and $\varphi^* = 2.46$ are constrained in their output choice. These firms increase their output and cost overreporting to bunch at the cutoff profit rate, making them indifferent between the output tax and the profit tax.

In the case of the generalized minimum tax, shown in Figure (3b), the variable δ is an additional tax instrument of the government and is set to $\delta = 1.0422$. This value of δ implies that only firms with profit tax liabilities exceeding the output liabilities by at least 4.22% face the profit tax. As a consequence, even firms with a higher productivity level increase output and cost overreporting, as this enables them to bunch at the profit rate cutoff. Compared to the case of the minimum tax where $\delta = 1$, the productivity range where firms bunch has widened, extending from $\varphi^{**} = 2.07$ to $\varphi^* = 3.45$, and additionally an output discontinuity occurs. The bunching behavior of firms with relatively high productivity levels implies that, under the generalized minimum tax, adjustments in optimal output and cost overreporting result in excessive production.[23] This overproduction by bunching high-productivity firms introduces an output distortion under the generalized minimum tax, a distortion that does not arise under the basic

[23]Due to cost overreporting, the optimal output choices of high-productivity firms that bunch exceed the black line, which represents the optimal output choice as a function of productivity under the profit tax without cost overreporting.

minimum tax.

Figure 4: Overreporting vs. Productivity (lowest 85% productivity firms)



(a) Minimum Tax

(b) Generalized Minimum Tax

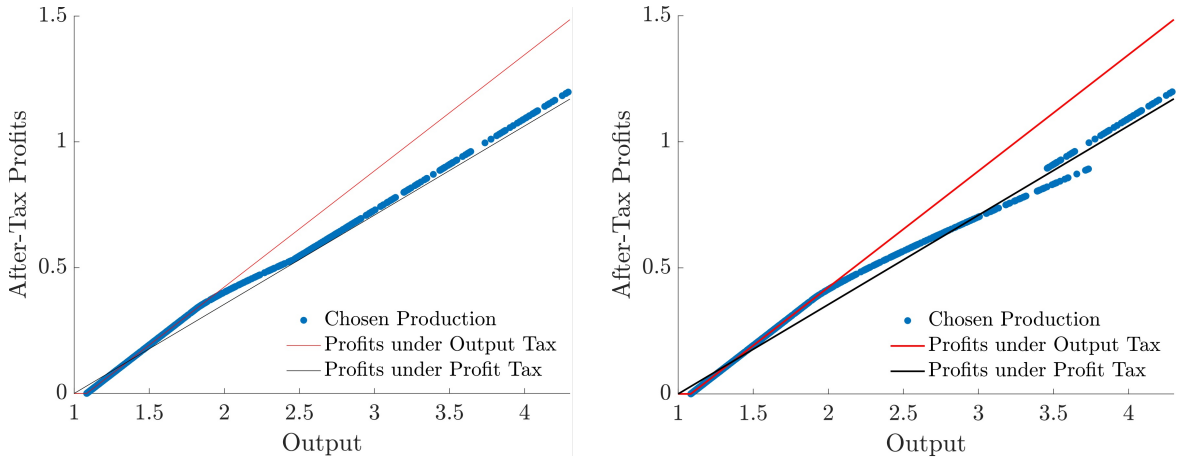
Figures 4a and 4b depict firms' optimal fraction of cost overreporting as a function of productivity under the minimum and the generalized minimum tax systems. Because costs are not deductible under the output tax, firms have no incentive for costly overreporting ($\gamma = 0$). In contrast, under the profit tax, cost overreporting ($\gamma > 0$) becomes beneficial, as it increases firms' after-tax profit.

As show in Figure 4a, firms that bunch under the standard minimum tax continuously increase cost overreporting (along with output) up to $\gamma = 0.11$, the optimal fraction of cost overreporting under a profit tax of 0.29. By comparison, under the generalized minimum tax depicted in Figure 4b, bunching firms choose cost overreporting (along with output) above the level that is optimal under the profit tax, resulting in a discontinuity in overreporting.

Firms' optimal adjustments in output and cost overreporting ensure a continuous increase in their after-tax profit as productivity increases. Figures 5a and 5b display firms' after-tax profits as a function of output. In the case of the minimum tax, after-tax profit increases continuously with firm output. However, under the generalized minimum tax, there exist firms of different productivity levels that produce the same level of output. Some of these firms increase output and cost overreporting to bunch at the profit rate cutoff, while other firms with identical outputs but higher productivity levels produce conventional output under the profit tax, yielding larger after-tax profits.

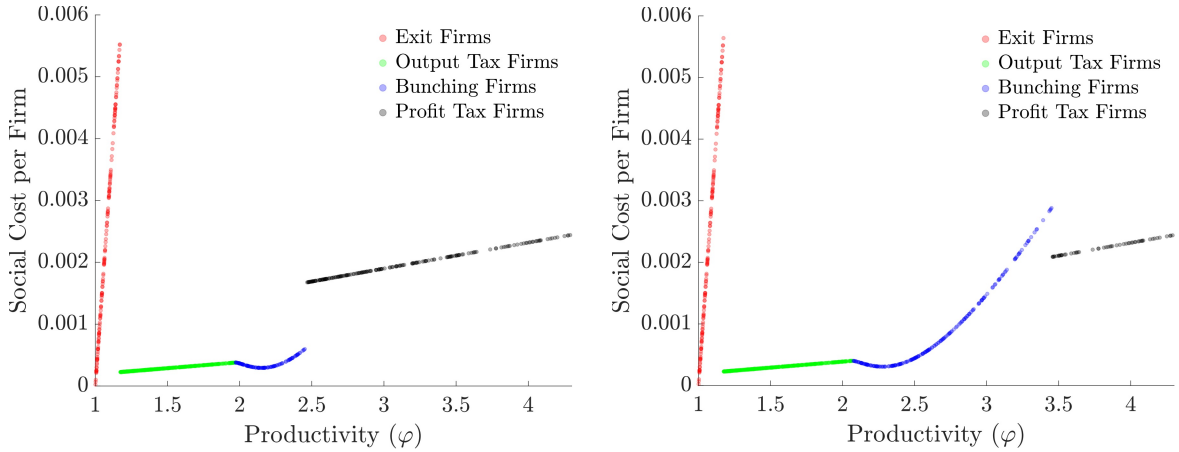
Finally, Figures 6a and 6b show the different social costs for both types of minimum tax systems measured by the difference in pretax profits between a no-tax regime and the

Figure 5: After-Tax Profit vs. Output (lowest 85% productivity firms)



(a) Minimum Tax (b) Generalized Minimum Tax

Figure 6: Social Costs (lowest 85% productivity firms)



(a) Minimum Tax (b) Generalized Minimum Tax

optimal (generalized) minimum tax regime as a percentage of total undistorted profits. Again, we distinguish four different types of social costs. First, a positive output tax rate impedes firm entry. Only firms with sufficiently high productivity levels will break even and remain in the formal (taxed) sector. The related social cost are depicted by the red dots. Second, the green dots denote the social cost of the output distortion of firms producing under the output tax. Third, the blue dots represent the social cost generated by the firms that bunch. The shape of this social cost is u-shaped because firms increase output when bunching at the profit rate cutoff, which reduces the social cost from the output distortion. However, the firms that bunch also engage in cost overreporting and

thereby incur socially wasteful concealment costs, which increase the social cost. The reason for the much larger social cost of bunching under the generalized minimum tax is due to the inefficient increase in output and cost overreporting even above the level that is optimal under the profit tax. Fourth, the black dots denote the social costs under the profit tax, which consist of fixed administrative costs per firm and concealment costs from overreporting. Significantly fewer firms incur these costs under the generalized minimum tax, offsetting to some extent the higher social costs due to the more extreme bunching.

Government Optimization Problem

Next, we state the government's optimization problem. The same notation employed in the threshold case is employed to identify aggregate functions, but now the arguments differ. In place of the output threshold y^*, we now have the instrument δ. Also, the aggregate output chosen by firms under the output tax now depends on both the tax rates as well as on δ, because all three variables determine bunching behavior. Finally, φ^* is also determined by the two tax rates and δ: $\varphi^* = \varphi^*(t_o, t_p, \delta)$.

With this setup, the government's maximization problem may be stated as follows:

$$\max_{t_o, t_p, \delta} \quad (1 - t_o)Y_o(t_o, t_p, \delta, \varphi^*) + [Y_p(\varphi^*) - C_p(\varphi^*)][(1 - e(t_p)) - t_p(1 - \gamma(t_p))]$$

$$s.t. \quad t_o Y_o(t_o, t_p, \delta, \varphi^*) + t_p[Y_p(\varphi^*) - C_p(\varphi^*)](1 - \gamma(t_p)) - [1 - H(\varphi^*)]a = G.$$

Letting Π_p^T denote taxable profits under the profit tax, the first-order conditions for the tax rates are:

$$\begin{aligned} (1 - \frac{1}{\lambda})Y_o + t_o \frac{\partial Y_o}{\partial t_o} + ah(\varphi^*)\frac{\partial \varphi^*}{\partial t_o} = 0; \\ (1 - \frac{1}{\lambda})\Pi_p^T + t_p \frac{\partial \Pi_p^T}{\partial t_p} + t_o \frac{\partial Y_o}{\partial t_p} + ah(\varphi^*)\frac{\partial \varphi^*}{\partial t_p}. \end{aligned} \tag{13}$$

In contrast to the threshold system, the difference in a type-φ^* firm's tax rates does not appear under a standard minimum tax system ($\delta = 1$), because the tax payment by a type-φ^* firm is the same under the output tax and the profit tax: $t_o y_o^* - t_p \pi_p^{T*} = 0$. Also, recall that under the threshold system, a rise in the profit tax rate reduces the number of firms that face the profit tax, and a rise in the output tax rate increases this number. In contrast, under a minimum tax system, the opposite occurs:

$$\frac{\partial \varphi^*}{\partial t_o} > 0; \qquad \frac{\partial \varphi^*}{\partial t_p} < 0.$$

This difference suggests that too many firms face the profit tax under a minimum tax system, resulting in too much administrative cost.

Another difference between the two tax systems involves the effects of tax changes on the welfare level of firms that bunch. Under the threshold system, a rise in the profit tax t_p has no effect on the output tax liability of bunched firms. However, under the standard or generalized (with $\delta > 1$) minimum tax system, a higher profit tax causes bunched firms to increase their outputs and/or overreporting levels so that their profit rates stay equal to the lower cutoff. Both of these responses reduce the after-tax profits of firms that bunch. In particular, firms are already setting outputs beyond the profit-maximizing level under the output tax, and a higher profit tax increases this discrepancy, although the higher outputs do raise additional tax revenue for the government. In contrast, firms bunching at the output level y^* under the threshold system obtain after-tax profits that are independent of the profit tax t_p.

Assume now that δ can be set optimally. In this case, it is convenient to use the profit rate cutoff, $\tau = \frac{\delta t_o}{t_p}$, as a control variable because it can now be chosen independently of the tax rates. As a result, a rise in the profit tax no longer affects the tax base of the output tax, Y_o, given φ^*, and the first-order condition for the profit tax rate t_p becomes

$$(1 - \frac{1}{\lambda})\Pi_p^T + t_p \frac{\partial \Pi_p^T}{\partial t_p} + (t_o y_o^* - t_p \pi_p^{T*} + ah(\varphi^*)\frac{\partial \varphi^*}{\partial t_p} = 0. \tag{14}$$

Note, too, that the sign of $\frac{\partial \varphi^*}{\partial t_p}$ changes to positive: there is no change in the after-tax profits obtained by bunched firms, which depend on t_o and τ, but the after-tax profits fall for firms facing the profit tax, implying that their number falls (i.e., φ^* rises). Using these results, we prove:

Proposition 5. *If there is no cost over-reporting, then $\delta > 1$ holds under the optimal generalized minimum tax.*

Proof. The assumption of no cost overreporting implies that $\frac{\partial \Pi_p^T}{\partial t_p} = 0$, in which case the first-order condition for t_p becomes

$$(1 - \frac{1}{\lambda})\Pi_p^T + (t_o y_o^* - t_p \pi_p^{T*} + ah(\varphi^*)\frac{\partial \varphi^*}{\partial t_p} = 0.$$

The first term is positive. Because $\frac{\partial \varphi^*}{\partial t_p} > 0$, a negative second term requires that $t_p \pi_p^{T*} > t_o y_o^*$, which holds only if $\delta > 1$. Not only does this condition rule out the standard minimum tax, where $t_p \pi_p^{T*} = t_o y_o^*$, but also the case where $\delta < 1$, under which $t_p \pi_p^{T*} < t_o y_o^*$, violating the first-order condition. □

6 Comparison of Tax Systems

We rely on simulations to demonstrate that switching from the optimal threshold tax system to the optimal (generalized) minimum tax system will improve welfare in most, but not all, cases.

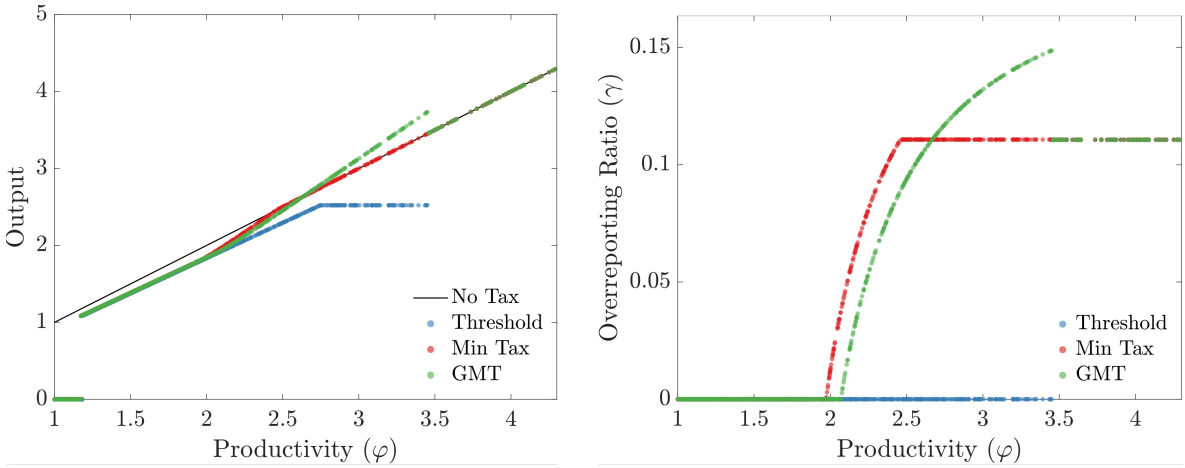
Table 1: Optimal Tax Rates and Instruments under Different Dual-Regime Systems

Dual-Regime System	Output Tax Rate	Profit Tax Rate	Tax Instrument
Threshold Tax	$t_o = 8.22\%$	$t_p = 29.1\%$	$y^* = 2.52$
Minimum Tax	$t_o = 7.69\%$	$t_p = 29.1\%$	$(\delta = 1.0)$
Generalized Minimum Tax	$t_o = 7.72\%$	$t_p = 29.1\%$	$\delta = 1.042$

Table 1 presents the optimal tax rates and additional tax instruments for the optimal threshold tax, minimum tax, and generalized minimum tax. Despite the fundamental differences among the dual-regime tax systems and their distinct implications for firm behavior, the simulation results reveal comparable optimal rates for the output tax (between 7.7% and 8.2%) and the profit tax (around 29.1%) across all three systems. Therefore, when comparing the different dual-regime systems, the impact on firms' optimal output and cost overreporting decisions, as well as the resulting social costs under each system, is primarily driven by the behavioral incentives inherent in each system, rather than the tax rate levels.

Figure 7a illustrates the output distortion for firms with a given productivity level under each of the three dual-regime tax systems. The blue dots represent the threshold tax, the red dots represent the minimum tax, and the green dots represent the generalized minimum tax system. The positive output tax imposed in each of the three systems impedes firm entry into the formal sector. Firms with very low productivity cannot break even in the formal sector and are forced out of the market. Firms with sufficiently high productivity enter the market and produce under the output tax and their output decisions are distorted downward by the tax.

Figure 7: Output and Cost Overreporting under Different Dual-Regime Systems



(a) Output vs. Productivity

(b) Cost Overreporting vs. Productivity

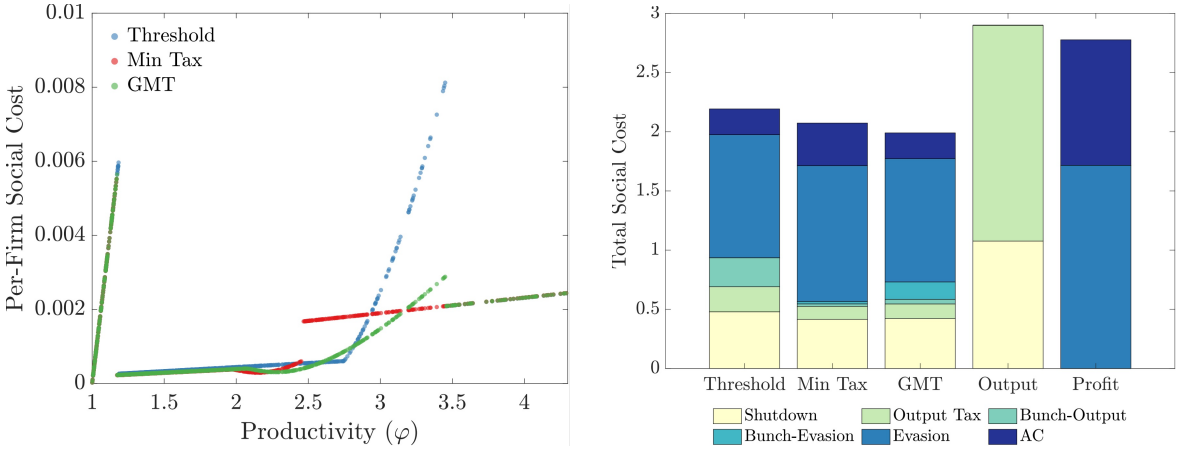
For firms bunching under the threshold tax (blue dots), output distortions become more pronounced as higher-productivity firms choose to reduce their output to the threshold level, y^*. In contrast, firms bunching under the minimum tax (red dots) and the generalized minimum tax (green dots) continuously increase their output as productivity levels rise. Under the generalized minimum tax, firms that bunch produce even above the output level that would be optimal under a (non-distortive) profit tax, resulting in overproduction.

Figure 7b shows firms' optimal cost overreporting behavior under the three dual-regime tax systems. Our modeling approach, which assumes separable overreporting costs, allows firms subject to the profit tax to overreport a fixed portion of their costs, regardless of their productivity level. This still generates the empirical finding that larger firms overreport larger amounts in absolute terms. Under the threshold tax (blue dots), firms subject to the output tax and firms that bunch at the output threshold, y^*, can by construction not overreport cost; only once firms are subject to the profit tax cost overreporting becomes attractive. This contrasts with the behavior under the minimum tax systems, where firms that bunch at the profit rate cutoff also overreport cost. By increasing both output and cost overreporting, firms maintain their profit rate at the cutoff level, making them indifferent between the output and profit tax.

As illustrated in Figure 7b, under the minimum tax (red dots), bunching firms continuously increase their cost overreporting until they reach the optimal overreporting portion that applies under the profit tax. In contrast, under the generalized minimum tax

(green dots), a discontinuity in cost overreporting occurs. Firms with higher productivity are able to bunch at the profit level cutoff as well because the profit tax liability may exceed their output tax liability by a positive amount under the generalized minimum tax. These high-productivity firms, which bunch, increase both output and cost overreporting beyond the optimal level under the profit tax to maintain a profit rate in line with the cutoff $\delta\, t_o/t_p$. This explains the discontinuity observed in the simulation results for the generalized minimum tax.

Figure 8: Social Cost under Different Dual-Regime Systems



(a) Social Cost per Firm

(b) Cumulative Social Cost

Figure 8a presents the per-firm social cost incurred under each of the three dual-regime tax systems, while Figure 8b illustrates the total social costs, along with the breakdown of the composition of these costs under each system. We distinguish between six different types of social costs. First, the positive output tax rate forces low-productivity firms to shut down and exit the market because their productivity level is insufficient to break even in the formal sector. As shown in Figure 8a, the social costs of firm shutdowns increase with firm productivity because the forgone output of exiting firms becomes larger. Second, firms operating under the output tax respond to the non-deductibility of costs by reducing their optimal output level. This distortion from the output tax becomes more severe for firms with higher productivity levels. Third, firms that bunch also distort their output. This type of social cost is especially significant under the threshold tax (blue dots in Figure 8a in the productivity range from 2.75 to 3.45), where relatively higher-productivity firms bunch at the output threshold y^*. In contrast, under both minimum tax systems, bunching firms increase their output, leading to the downward-sloping part of the u-shape pattern depicted by the red and green dots in Figure 8a. Fourth, bunching

firms under both minimum tax systems also increase their cost overreporting, resulting in socially wasteful cost of overreporting. This type of social cost offsets the positive effects of increased output of bunching firms, explaining the upward-sloping part of the u-shape in the per-firm social cost. Finally, for firms operating under the profit tax, two additional types of wasteful social costs emerge: the cost of overreporting and the administrative cost per firm. While by assumption the administrative costs per firm remain constant, the cost of overreporting increases for higher-productivity firms, explaining the rising social costs per firm under the profit tax, as illustrated in Figure 8a.

When comparing the total social costs across different tax systems, we find that each of the three dual-regime tax systems results in substantially lower social costs compared to a pure output or pure profit tax system. This outcome arises due to the ability of the dual-regime business tax systems to allocate firms optimally between the two taxes, thereby reducing the social costs associated with firm exit and output distortions as well as cost overreporting and administrative cost, compared to a system relying solely on either an output or a profit tax.

Among the three dual-regime systems, our baseline simulations using estimates from recent empirical literature indicate that the threshold system generates the highest social costs.[24] This outcome is primarily driven by the bunching behavior of relatively high-productivity firms at the output threshold, which leads to a substantial loss of output from these firms. Comparing the social cost of the output distortion of bunched firms, we find that it is almost double under the threshold tax vs. both minimum tax systems. As a result, any form of minimum tax system is preferable to the threshold tax system. Furthermore, the generalized version of the minimum tax—which has not yet been implemented in any country—has the potential to further reduce the social costs of taxation in countries that apply a minimum tax version of a dual-regime system, including those countries listed in Table 3 of Appendix A.

The possibility that a firm's profit tax liability can exceed the output tax liability by a certain amount under the generalized minimum tax before the firm is subject to the profit tax implies that more firms will bunch at the profit rate cutoff and be indifferent between the output and profit tax. This allows the government to reduce the number of firms subject to the profit tax under the generalized minimum tax, which reduces wasteful cost of overreporting and administrative cost compared to the standard minimum tax. This

[24]The robustness of this result is further examined in the Comparative Static Analysis section, where we find that only under extreme parameter values the threshold system might outperform the generalized minimum tax system.

reduction in social costs is substantial enough that even the larger wasteful cost of cost overreporting by the increased number of bunching firms under the generalized minimum tax does not affect the ranking. The generalized minimum tax still has the lowest total social cost among the three dual-regime tax systems.

7 Comparative Static Analysis

In this section, we assess the robustness of our simulation results with respect to the calibration of the key parameters, administrative cost, a, the cost evasion parameter, β, and the revenue requirement, G. To ensure that the results of our comparative static analysis are not driven by extreme productivity distributions, we conduct a Monte Carlo–type simulation, where our outcomes represent the average over 100 independent productivity draws.

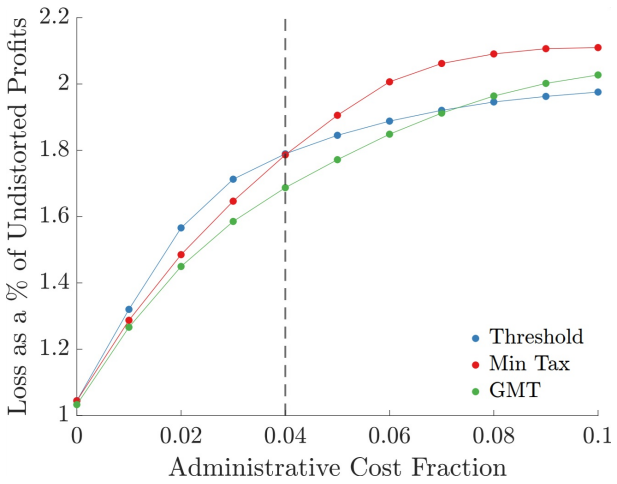
7.1 Administrative Cost, a

In the following simulations, we examine the impact of varying the level of administrative cost, defined relative to the revenue requirement G under the assumption that all firms are subject to the profit tax. The parameter a is adjusted from 0 to 0.1. Empirical evidence suggests that a reasonable estimate for this parameter is approximately 0.04 (Lobel, Scot, and Zúniga 2024), which serves as the baseline value in our baseline simulations.[25]

Figure 9 presents the change in social cost across the three dual-regime tax systems as the share of administrative costs relative to the revenue requirement increases from 0 to 0.1. The dashed vertical line at 0.04 indicates the administrative cost fraction used in our baseline simulations, based on estimates from the empirical literature (Lobel, Scot, and Zúniga, 2024). At relatively low levels of administrative cost, social costs are highest under the threshold system. However, as administrative cost continues to rise, the rate of increase in social costs under the threshold system flattens relative to both minimum tax systems. While social costs remain consistently lower under the generalized minimum tax relative to the standard minimum tax, at sufficiently high administrative cost values (approximately

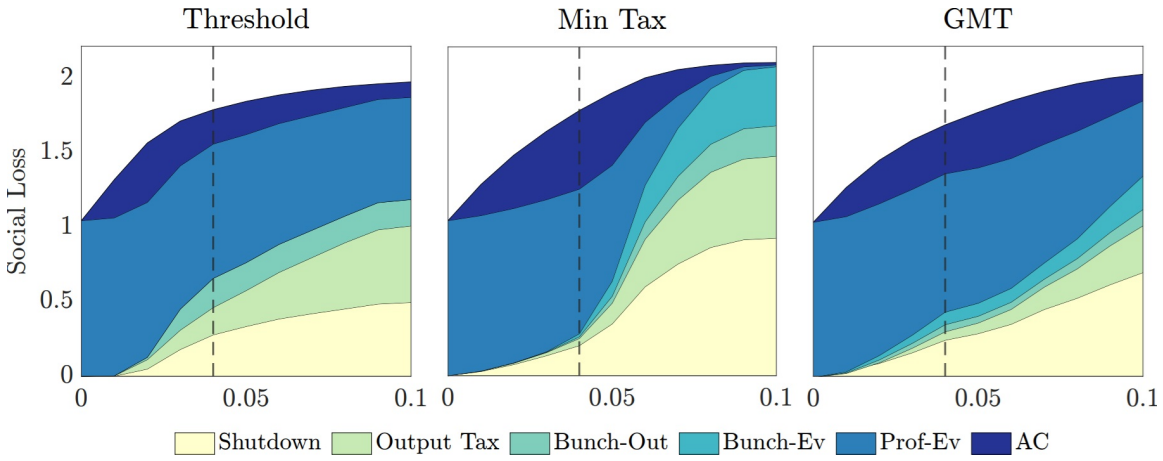
[25]Since the comparative statics are averaged over 100 draws, the equilibrium values for the baseline parameterization in the figures here do not exactly match those in previous figures, which examine an isolated draw.

Figure 9: Comparison of Social Cost as Administrative Cost Increases



0.075), the threshold system begins to outperform the generalized minimum tax in terms of lower overall social costs.

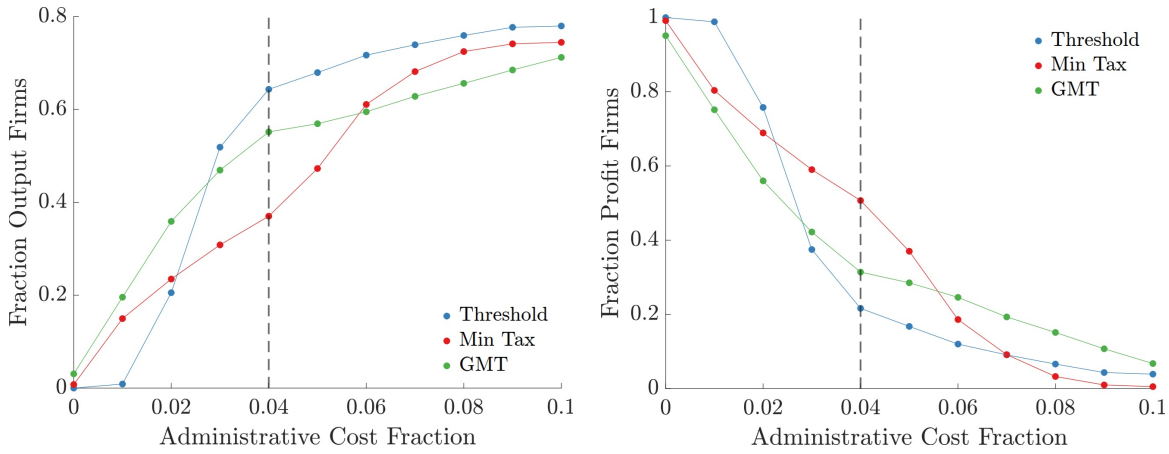
Figure 10: Composition of Social Cost as Administrative Cost Increase



To further clarify the relative performance of the three dual-regime tax systems, we analyze the composition of overall social costs as administrative costs increase (see Figure 10). When administrative cost is zero or very low, it is optimal under each dual-regime tax system to have a substantial proportion of firms operating under the profit tax. This results in relatively low overall social costs, which, at low administrative cost levels, primarily consist of cost overreporting (Prof-Ev) and administrative cost (AC). The presence of significant cost overreporting under both minimum tax systems makes it suboptimal to assign all firms to the profit tax, even in the absence of administrative cost, as depicted

in Figure 11b.[26]

Figure 11: Fraction of Firms Operating under each Tax



(a) Firms under the Output Tax

(b) Firms under the Profit Tax

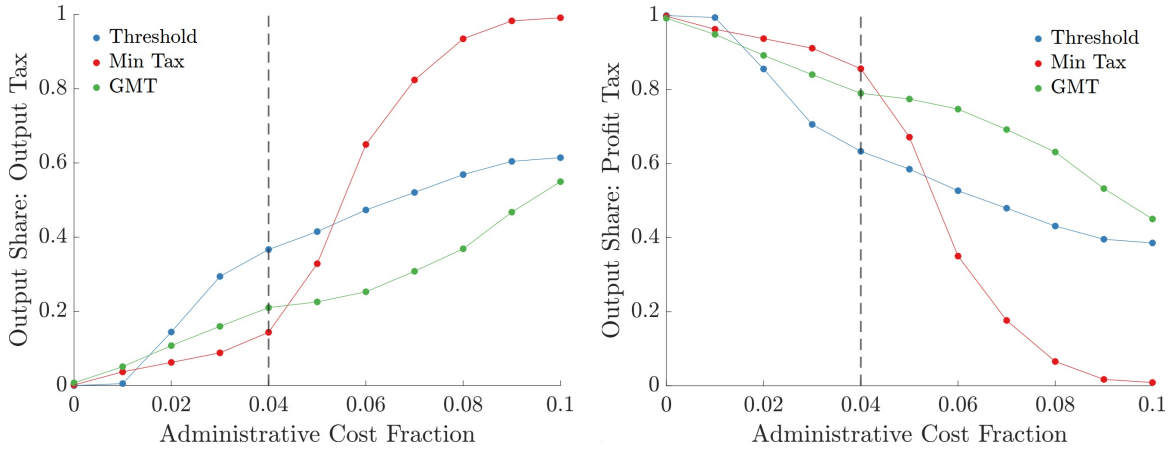
As administrative cost increases, it is optimal that fewer firms operate under the profit tax, while correspondingly a greater proportion of firms are under the output tax. Figure 11 illustrates the fraction of firms operating under the output tax (Figure 11a) and profit tax (Figure 11b), respectively, as administrative cost rises, while Figure 12 displays the output produced by firms subject to the output (Figure 12a) and the profit tax (Figure 12b) as administrative cost changes.

When administrative cost is zero or very low, the proportion of firms subject to the output tax is negligible, with nearly all firms operating under the profit tax. Accordingly, output is produced by firms almost exclusively subject to the profit tax. However, as administrative cost becomes substantial, reaching 10% of the revenue requirement, the share of firms subject to the output tax increases to approximately 75% under the threshold tax and slightly below 70% under both minimum tax systems. Under the profit tax, fewer than 10% of firms remain in the case of the generalized minimum tax, with even lower proportions observed for the threshold tax and the minimum tax (The remaining firms necessary to account for 100% exit the market, as discussed in the subsequent sections).

In terms of the output generated by firms subject to the output and profit taxes, respectively, Figure 12 shows that under very high administrative cost, slightly more than

[26]With zero administrative cost and the majority of firms operating under the profit tax, the revenue requirement can be satisfied with an optimal output tax rate of close to zero and a profit tax rate of approximately 29%, as shown in Figures 13a and 13a, respectively.

Figure 12: Output of Firms Operating under each Tax



(a) Output of Firms under the Output Tax

(b) Output of Firms under the Profit Tax

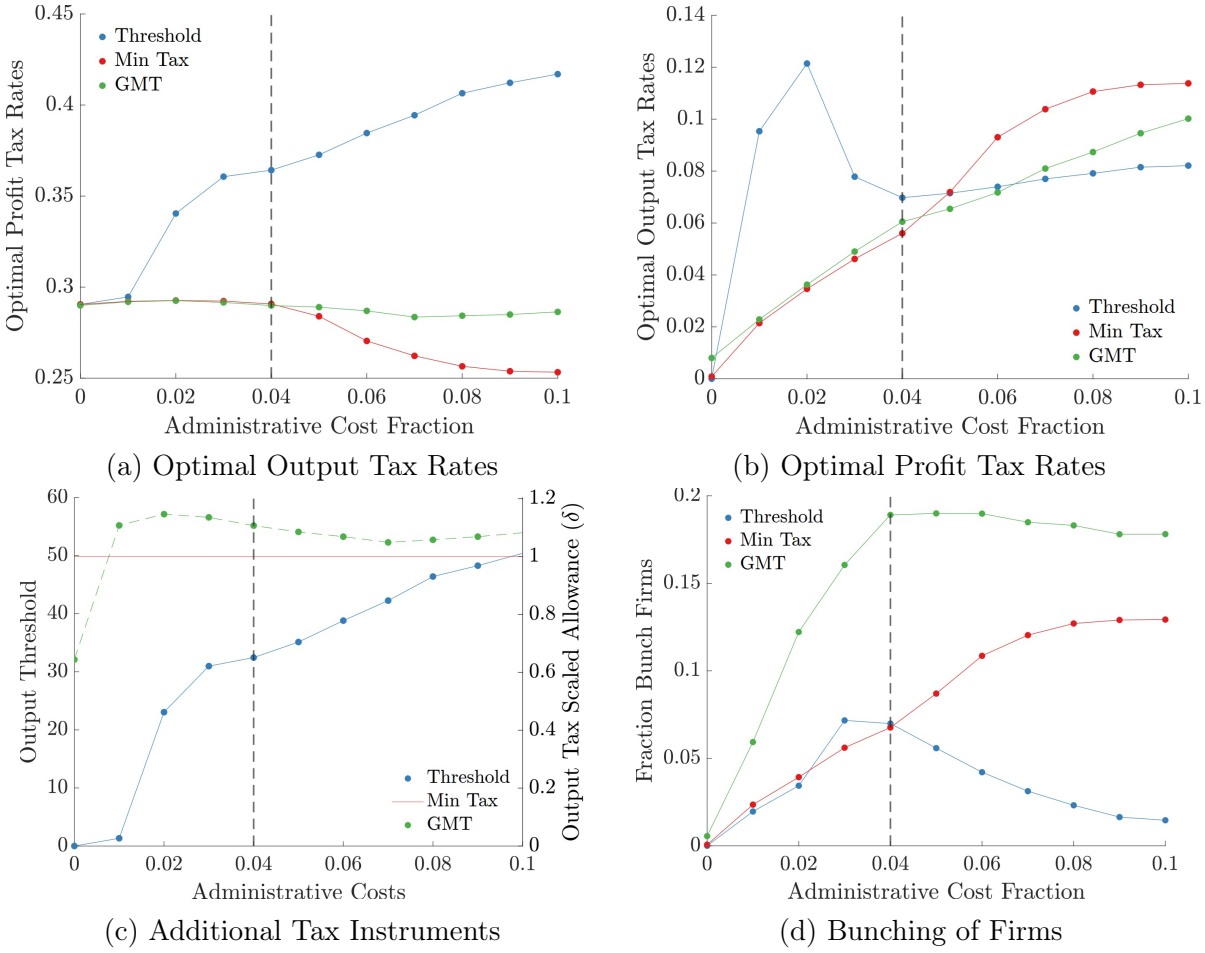
half of total output is produced by firms subject to the output tax, with the remainder produced by firms facing the profit tax. Under the minimum tax system, however, the distribution of output becomes highly skewed, with nearly all output produced by firms subject to the output tax.

To control the transition of firms from the profit tax to the output tax, the optimal output threshold (y^*), which determines the point above which firms become subject to the profit tax, is gradually adjusted upwards as administrative cost increase (see Figure 13c). To ensure the exogenous revenue requirement is satisfied, both the optimal output tax rate and the profit tax rate must increase, as demonstrated in Figures 13a and 13b, respectively.

Under the minimum tax system, no additional tax instruments are available beyond the two tax rates to control the number of firms operating under each tax. To increase the number of firms subject to the output tax, the ratio of tax rates, $\frac{t_o}{t_p}$, must rise. This requirement explains the sharp increase in the optimal output tax rate under the minimum tax as administrative cost increase. Furthermore, as illustrated in Figure 13b, a reduction in the profit tax rate becomes necessary at very high levels of administrative cost to ensure a sufficient increase in the tax rate ratio $\frac{t_o}{t_p}$, thereby guaranteeing that the optimal number of firms are subject to the output tax.

In contrast, the additional tax instrument, δ, available under the generalized minimum tax, grants the government greater control over the distribution of firms subject to the output and the profit tax. When set at a value above 1, this parameter allows a firm's profit tax liability to exceed its output tax obligation before being subject to the profit

Figure 13: Optimal Tax Rates, Tax Instruments, and Bunching of Firms



(a) Optimal Output Tax Rates

(b) Optimal Profit Tax Rates

(c) Additional Tax Instruments

(d) Bunching of Firms

tax. The additional instrument δ thereby acts as a multiplier, ensuring that even moderate adjustments in tax rates are sufficient to raise the cutoff ratio $\frac{\delta t_o}{t_p}$, which determines the proportion of firms operating under each tax.

A comparison of the optimal output and profit tax rates under the minimum and generalized minimum tax systems, as depicted in Figures 13a and 13b, reveals that when administrative cost is substantial (greater than 0.04), the generalized minimum tax requires less pronounced adjustments in tax rates to achieve the optimal allocation of firms across the output and the profit tax.

The availability and type of additional tax instruments within each of the three dual-regime tax systems have significant implications for the extent of bunching that emerges under each system, as illustrated in Figure 13d. While an increase in the optimal output threshold induces firms with higher productivity levels to bunch, the fraction

of firms bunching under the threshold system peaks at administrative cost between 3% and 4%. It then steadily declines as administrative cost continues to rise. This lower degree of bunching under the threshold system explains why overall social costs in this system increase only moderately compared to both minimum tax systems. In contrast, the increasing proportion of firms that bunch under the minimum tax system amplifies output distortions and especially evasion cost, thereby raising overall social costs above those observed in the other two dual-regime tax systems.

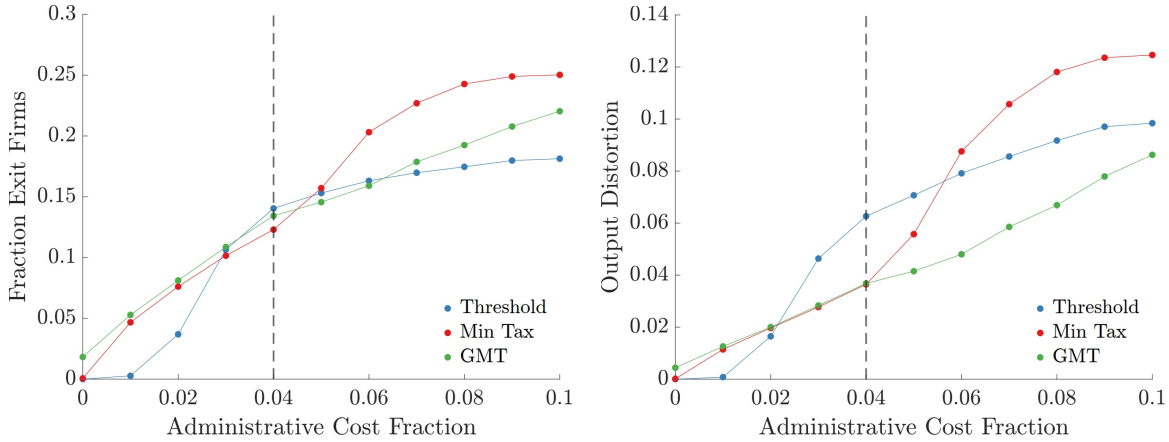
In the case of the generalized minimum tax, bunching behavior exhibits the distinct characteristics discussed in Section 5. Not only is the proportion of firms bunching under this system substantially higher than in the other two systems, but this bunching also results in overproduction, introducing an additional margin that further contributes to overall social costs.

The higher optimal output tax rates under each of the three dual-regime tax systems, as shown in Figure 13a, lead to an increase in social costs associated with the exit of less productive firms from the taxed sector (shutdown). Additionally, these higher output tax rates exacerbate tax-induced output distortions for firms that remain in the taxed sector and operate under the output tax. The change in the composition of social costs as administrative cost increases is illustrated in Figure 10. Under both minimum tax systems, these two components of social costs rise substantially with higher administrative cost, accounting for approximately or slightly less than 50% of total social costs in the cases of the minimum tax and the generalized minimum tax, respectively. This explains why both minimum tax systems perform worse, that is, result in higher social costs compared to the threshold system when administrative cost is substantial.

This conclusion is further supported by Figure 14a, which presents the fraction of firms forced to exit the taxed sector due to insufficient productivity levels that prevent them from generating positive after-tax profits and breaking even. When administrative cost reach approximately 4% of the revenue requirement, around 15% of firms exit the formal sector under each of the three dual-regime tax systems. As administrative cost continues to rise, the exit rate increases to 25% under the minimum tax and to 20% under the generalized minimum tax.

Similarly, as administrative cost increases, the higher optimal output tax rates amplify the tax-induced output distortion for firms subject to the output tax (Figure 14b). For relatively high administrative cost, this distortion is most pronounced under the minimum tax system, which also features the highest optimal output tax rate. While the generalized

Figure 14: Firm Exit and Output Distortion



(a) Firms Exiting the Y-Sector (b) Tax-Induced Output Distortion

minimum tax exhibits the second-highest output tax rate, fewer firms operate under the output tax within this system. Consequently, the aggregate tax-induced output distortion under the generalized minimum tax is lower than under the threshold system, which, despite having a slightly lower optimal output tax rate, includes a greater number of firms operating under this tax (see Figure 11a).

7.2 Cost Evasion Parameter, β

We next examine how the facility of cost overreporting affects overall social costs under each of the three dual-regime tax systems. In the baseline simulation, we set the parameter β, which determines the cost of overreporting, to 8.3. This value results in cost overreporting of approximately 11% of reported profits under the profit tax, and is indicated by a dashed vertical line in the figures.

The range of β values considered in the comparative static analysis corresponds to cost overreporting between roughly 9% and 15%, consistent with recent estimates from Lobel, Scot, and Zúniga (2024), Alejos (2018), and Best et al. (2015). Further details on the calibration are provided in Appendix C. As shown in Figure 15, our simulation results indicate that, regardless of the extent of cost overreporting, the generalized minimum tax consistently results in lower overall social costs compared to both the threshold tax and the standard minimum tax systems. The latter two yield very similar levels of social cost.

When cost overreporting is particularly severe and entails significant social costs, it is optimal to reduce the number of firms operating under the profit tax while increasing

Figure 15: Comparison of Social Cost as Overreporting Cost Increases

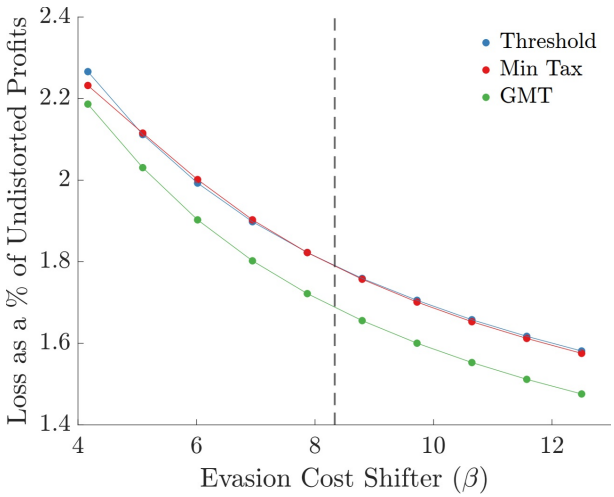
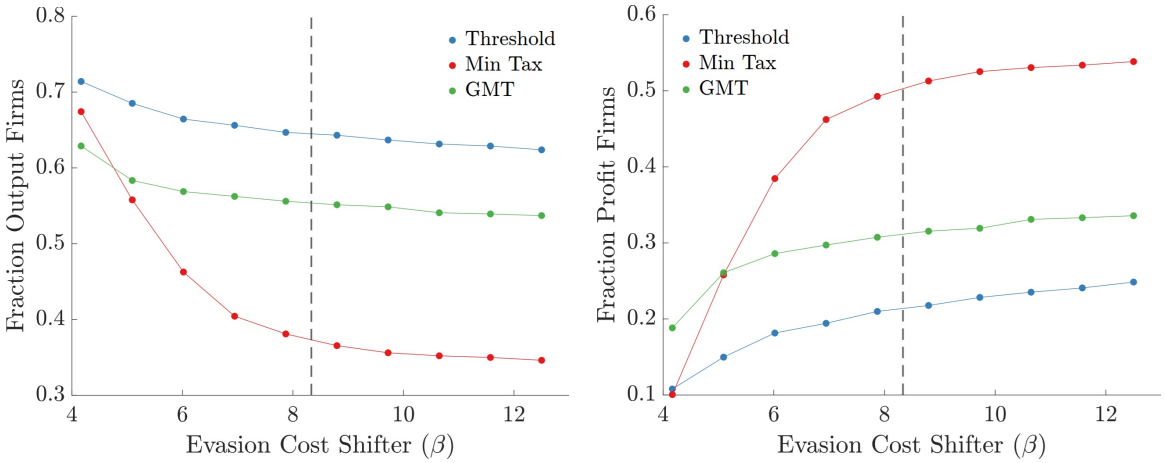


Figure 16: Fraction of Firms Operating under each Tax



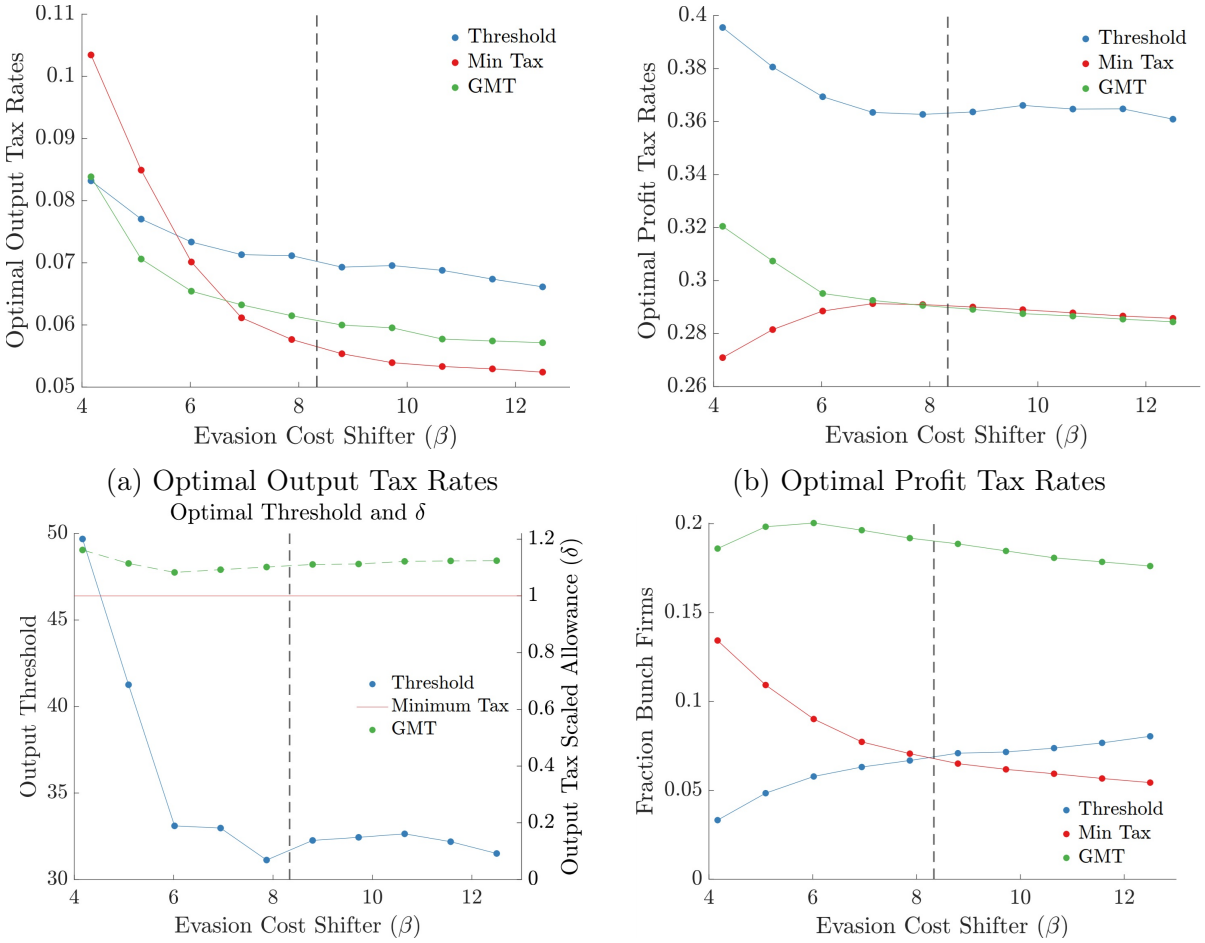
(a) Firms under the Output Tax

(b) Firms under the Profit Tax

the proportion of firms subject to the output tax. Figures 16a and 16b illustrate that under the threshold and minimum tax systems, the optimal fractions of firms operating under the output and profit taxes are approximately 70% and 10%, respectively. Under the generalized minimum tax, the distribution of firms between the two taxes is more balanced, with slightly above 60% of firms subject to the output tax and roughly 20% operating under the profit tax.

Controlling the optimal distribution of firms across the output and profit taxes requires the government to optimally set tax rates and available tax instruments, given the revenue requirement. These are illustrated in Figures 17a and 17b.

Figure 17: Optimal Tax Rates, Tax Instruments, and Bunching of Firms



(a) Optimal Output Tax Rates

(b) Optimal Profit Tax Rates

(c) Additional Tax Instruments

(d) Bunching of Firms

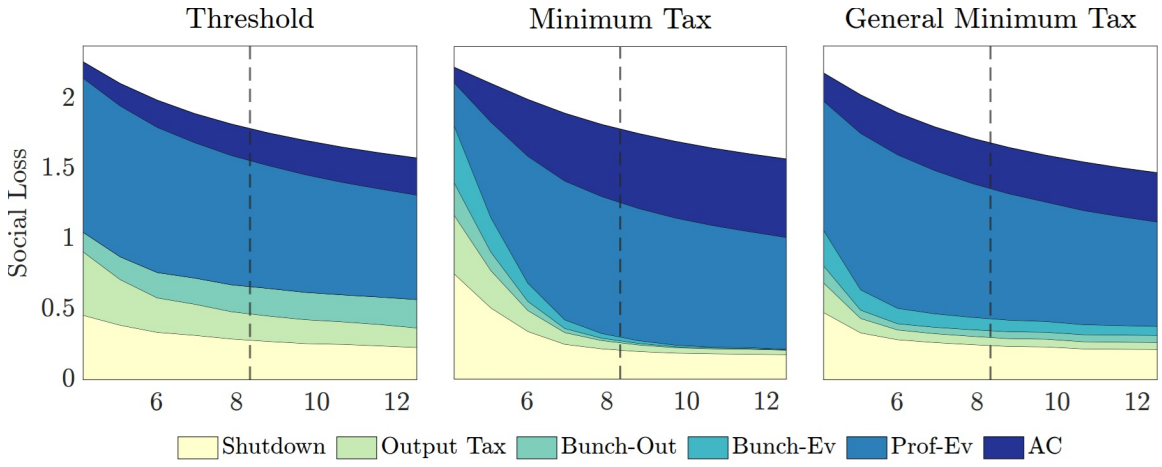
Figure 17c illustrates, that, in the case of the threshold system, the output threshold y^* is set at a relatively high level when cost overreporting is severe to prevent an excessive number of firms from operating under the profit tax. As the cost of overreporting increases and the severity of overreporting diminishes, the optimal output threshold declines, generating a greater number of firms subject to the profit tax (see Figure 16b). Furthermore, as cost overreporting decreases, both the optimal output tax rate and the profit tax rate decline. When tax evasion through cost overreporting becomes less severe, the required revenue can be collected with lower tax rates.

This pattern applies to all three dual-regime tax systems. However, in the case of the minimum tax, the two tax rates serve as the sole instruments to regulate the fraction of firms under each tax. Consequently, when cost overreporting is severe, a relatively high output tax rate and a low profit tax rate are necessary under the minimum tax system to

ensure a sufficiently high cutoff ratio $\frac{t_o}{t_p}$ and the optimal portion of firms under the output tax. As the cost of overreporting increases, thereby mitigating the extent of overreporting, the optimal output tax rate declines from over 10% to approximately 5%.

Under the generalized minimum tax, the government possesses the additional tax instrument δ, which is set greater than 1 (see Figure 17c). This ensures that a firm's profit tax liability may exceed its output tax liability by the fraction $1 - \delta$ while still being subject to the output tax. This additional tax instrument allows the government to achieve the same optimal cutoff ratio, $\frac{\delta t_o}{t_p}$, with less pronounced adjustments in the output and profit tax rates.

Figure 18: Comparison of Social Cost as Overreporting Cost Increases

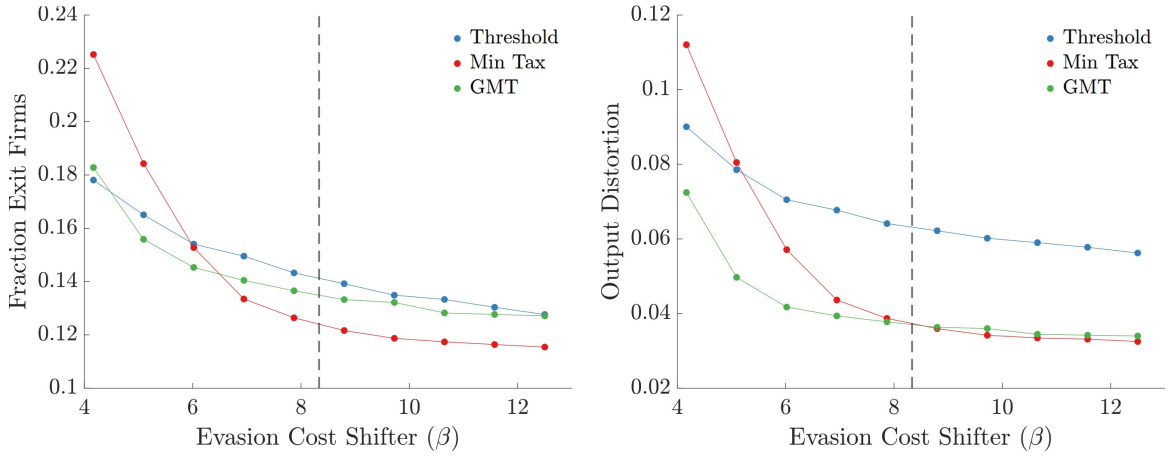


The absence of an additional tax instrument under the minimum tax necessitates a relatively high optimal output tax rate in this system when cost overreporting is severe. This leads to increased social costs, particularly in the form of firm exits from the taxed Y-sector and tax-induced output distortions under the minimum tax, as illustrated in Figure 18, which presents the composition of overall social costs for each dual-regime tax system.

Under the generalized minimum tax, social costs of firm exits from the taxed Y-sector and tax-induced output distortions are substantially lower due to the lower optimal output tax rate when cost overreporting is severe. As cost overreporting diminishes, the distribution of firms across the two tax regimes becomes more balanced under the generalized minimum tax compared to the minimum tax. This results in lower administrative cost under the former, thereby reducing overall social costs. Compared to the threshold tax, the generalized minimum tax yields lower overall social costs due to

reduced tax-induced output distortions, as demonstrated in Figure 19b.

Figure 19: Firm Exit and Output Distortion



(a) Firms Exiting the Y-Sector

(b) Tax-Induced Output Distortion

As overreporting costs increase, the optimal adjustments in tax rates and instruments lead to a growing fraction of firms that bunch under the threshold system (see Figure 17d). Although the fraction of firms bunching under the generalized minimum tax is significantly larger, it is important to note that bunching under both minimum tax systems is less detrimental. This is because it results in considerably smaller output distortions compared to the threshold system, as illustrated in Figures 19b and 18.

7.3 Revenue Requirement, G

Finally, we examine the performance of the different dual-regime tax systems under varying government revenue requirements. Specifically, we vary the revenue requirement from business taxation between 5% and 15% of undistorted firm output.

As illustrated in Figure 20, our findings indicate that for relatively low revenue requirements, the threshold tax system generates the lowest overall social costs. However, for higher revenue requirements, the generalized minimum tax system performs best. Our baseline specification assumes a revenue requirement of 10%, indicated by a dashed vertical line in the figures.

As the revenue requirement increases, the government must generate additional funds through higher tax rates. The simulations show that it is optimal under each of the three dual-regime tax systems to increase both the optimal profit tax rate (see Figure 21b) and the number of firms operating under the profit tax (see Figure 22b).

Figure 20: Comparison of Social Cost as the Revenue Requirement Increases

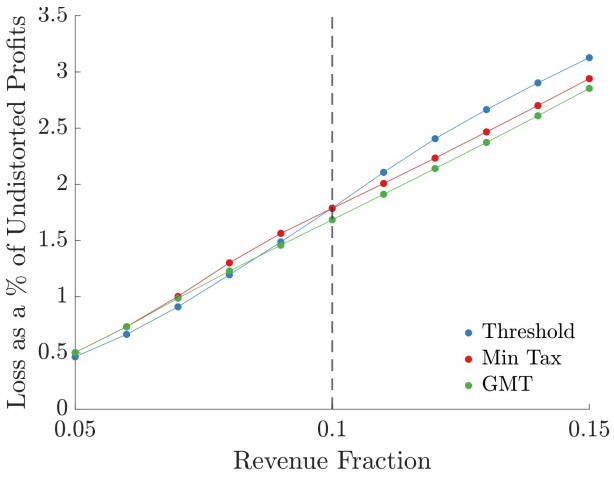
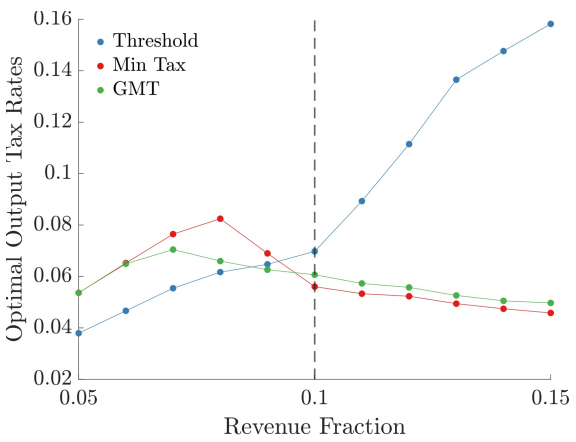
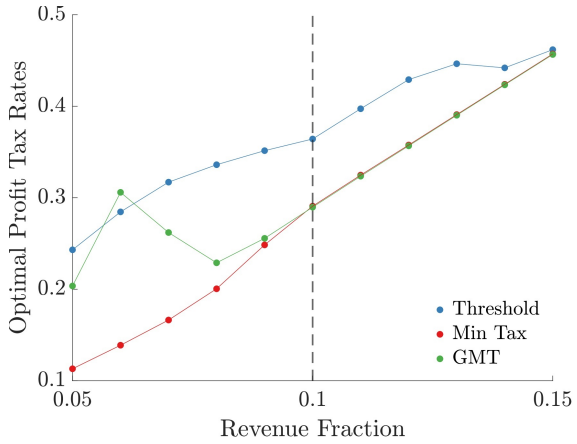


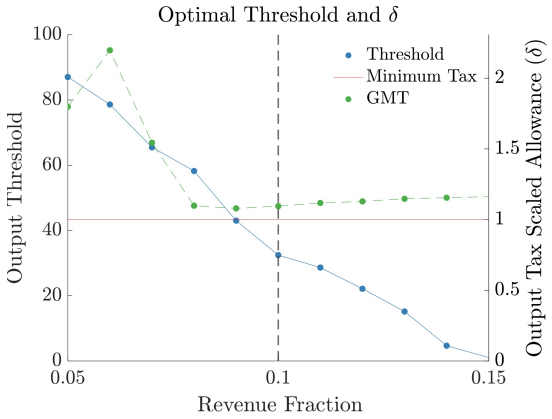
Figure 21: Optimal Tax Rates, Tax Instruments, and Bunching of Firms



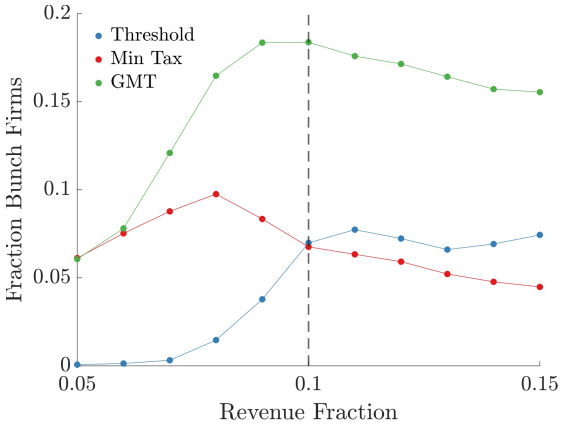
(a) Optimal Output Tax Rate



(b) Optimal Profit Tax Rate



(c) Additional Tax Instruments

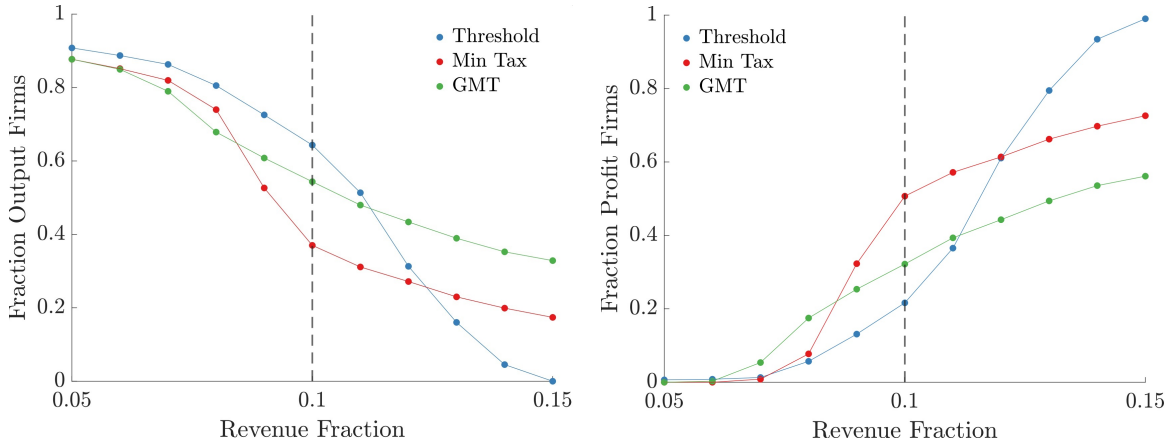


(d) Bunching of Firms

For relatively low revenue requirements, the optimal profit tax rate is approximately 10% under the minimum tax, 20% under the generalized minimum tax, and 25% under the threshold tax. As the revenue requirement rises, the optimal profit tax rate increases under each of the three dual-regime systems, reaching a maximum level of approximately 45% when the revenue requirement amounts to 15% of undistorted firm output.

To increase the number of firms subject to the profit tax, the initially high output threshold is gradually reduced under the threshold tax system. In both minimum tax systems, the increase in the profit tax rate lowers the cutoff ratio $\frac{t_o}{t_p}$, leading to a growing number of firms being subject to the profit tax. Additionally, as the revenue requirement continues to rise, the optimal output tax rate declines, further reducing the cutoff tax rate ratio.

Figure 22: Fraction of Firms Operating under Each Tax



(a) Firms under the Output Tax

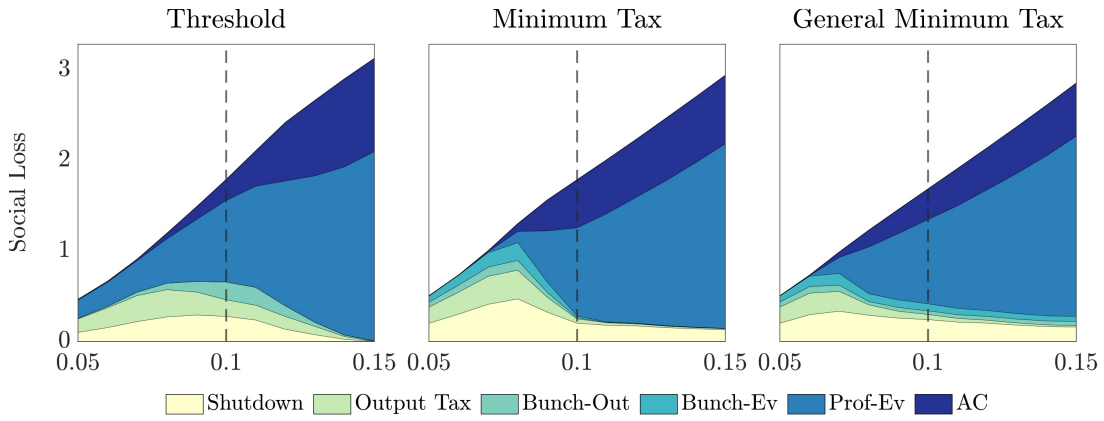
(b) Firms under the Profit Tax

In the case of the generalized minimum tax, the government also chooses the tax instrument δ, which is set close to or even above 2 when the revenue requirement is very low. This ensures that most firms remain under the output tax. However, as the revenue requirement increases, the optimal δ decreases to a value slightly above 1, shifting more firms from the output tax to the profit tax.

The optimal choice of tax rates and tax instruments under each of the dual-regime tax systems, along with the resulting distribution of firms operating under each tax, provides insight into the composition and patterns of overall social costs illustrated in Figures 20 and 23.

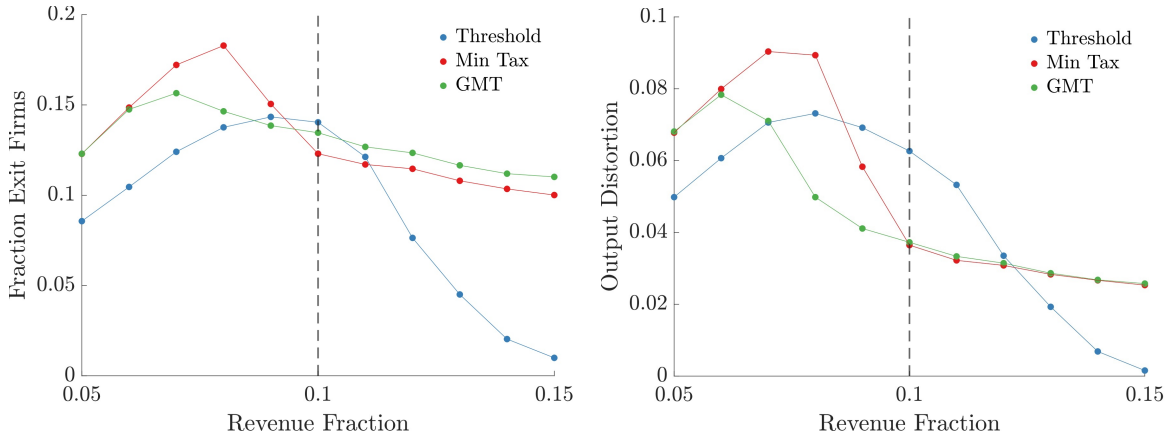
Specifically, the higher optimal output tax rates under both minimum tax systems lead to a greater proportion of firms exiting the taxed Y-sector and result in higher tax-induced

Figure 23: Composition of Social Cost as Administrative Cost Increase



output distortions of firms operating under the output tax (see Figures 24a and 24b). Furthermore, for low revenue requirements, the fraction of firms bunching under both minimum tax systems is substantially larger than under the threshold tax system (see Figure 21d). This leads to additional social costs, including increased output distortions and evasion costs for firms that bunch. These costs raise the overall social costs under both minimum tax systems above those observed under the threshold system.

Figure 24: Firm Exit and Output Distortion



(a) Firms Exiting the Y-Sector

(b) Tax-Induced Output Distortion

However, as the revenue requirement increases and it becomes optimal that a greater number of firms operate under the profit tax, both minimum tax systems exhibit an advantage over the threshold system. To generate the required revenue of 15% of firms' undistorted output, nearly all firms must operate under the profit tax in the case of the threshold system, which leads to very high evasion cost and administrative cost (Figure

23). In contrast, under both minimum tax systems, a non-negligible proportion of firms continues to operate under the output tax. This reduces administrative cost relative to the costs associated with firm exits and tax-induced output distortions, resulting in lower overall social costs under both the minimum and generalized minimum tax compared to the threshold tax.

8 Conclusion

Dual-regime business tax systems, under which smaller firms face an output tax and larger firms face a tax on profits, are pervasive around the world. Which tax a firm faces is explicit under threshold systems, where a firm's output determines which tax they face, or implicit, as under a minimum tax system where each firm owes the higher of the two tax liabilities, given their chosen output level.

Surprisingly given the pervasiveness of such business tax systems, there is almost no formal analysis of even the basic optimal design questions. In this paper, we begin such an analysis by investigating the optimal parameters of a dual-regime system and what aspects of the environment influence which kind of dual-regime business tax systems is best. We also introduce and analyze a new kind of minimum tax, which we label a generalized minimum tax, under which the law specifies the maximum percentage amount by which the profit tax liability can exceed the output tax liability without triggering the requirement that the firm is subject the profit tax.

We characterize the tax system that maximizes the welfare of a representative consumer who receives all after-tax firm profits, so there are no distributional considerations, and thus the objective is to raise the required revenue with minimal efficiency costs. Efficiency costs arise from a fixed per-firm administrative cost for each firm subject to the profit tax, from firms distorting their output to reduce their tax liability, from firms exiting the taxed sector and choosing to operate in the informal economy, and from firms subject to a profit tax incurring costs to facilitate cost overreporting. Critically, the distortions from output adjustments are not only marginal, because under either a threshold system or a minimum tax system some firms will adjust their output non-incrementally—resulting in "bunching" of output—in order to ensure they are (just) subject to the more attractive tax alternative. There is more inefficient bunching under a threshold system because all bunched firms reduce their output level to the same amount regardless of their productivity, unlike

under a minimum tax system, where output increases with productivity even for firms that bunch.

Often the stated motivation of a dual-regime business tax systems is to reduce the administrative and compliance costs incurred by small establishments, motivated by the belief (and credible evidence) that for income taxes these costs relative to measures of size are larger for small firms. In addition, turnover taxes are presumed to be less susceptible to evasion, as costs are relatively difficult for most tax administrations to monitor; thus, such a dual-regime system arguably makes efficient use of administrative resources. On the other hand, a tax on output inefficiently distorts production, while a profit tax, or at least a pure profit tax with full cost deductibility and no evasion, is non-distortionary.

We begin with a general theoretical analysis of the optimal tax problem. One immediate result is that an optimal dual-regime tax systems is better than levying only an output tax or only a profit tax. How much better and which alternative is better depends on the parameters of the model. The threshold system has a natural, mechanical advantage over the minimum tax because it features three policy parameters (the two tax rates plus the threshold), while the standard minimum tax has just two—the two tax rates; the generalized minimum tax makes the regime choice among alternative three-parameter systems.

The existence of an output threshold causes firms to bunch just below the threshold, making the threshold tax system the least favorable in terms of output distortion. Under the minimum tax system, firms bunch at the profit rate cutoff instead of an output threshold, leading to a continuous increase in the output of firms near this cutoff. The optimal generalized minimum tax system allows the profit tax liability to exceed output tax liability, and firms that bunch overproduce, raising their outputs above the level they would produce without taxation in order to reduce profits enough to qualify for the output tax rather than the profit tax.

We then turn to a numerical calibration of the model, choosing functional form and parameters that reflect the empirical literature and produce an equilibrium that shares aspects of a typical developing country. The simulation exercise reveals that, despite their differences, the threshold tax, the minimum tax, and the generalized minimum tax lead to very similar tax rate structures for both the output tax and profits tax: in our simulations, the optimal output tax is roughly between 7.7 and 8.2 percent, and the optimal profit tax rate is approximately 29.1 percent. However, the differences in the design of the threshold tax system and minimum tax systems imply that each system generates very different

output distortions and welfare effects. In the optimal versions of all the systems, a small percentage of firms is subject to the profit tax while remitting a substantial percentage of tax revenue–a pattern that is common with dual-regime business tax systems in practice.

Our baseline simulations, using parameter values from recent empirical literature, indicate that the generalized minimum tax results in the lowest social costs, followed by the standard minimum tax and the threshold tax system. The better performance of both minimum tax systems relative to the threshold system is primarily due to the lower social costs associated with bunching under these regimes. This finding holds significant policy implications, as several countries currently employ either a dual-regime threshold or minimum tax systems, yet none has adopted a generalized minimum tax. Our results suggest that implementing a generalized minimum tax could substantially enhance welfare of these countries by refining their dual-regime tax system.

Assessing the robustness of our findings, we conduct a comparative statics analysis. We find that under scenarios characterized by very high administrative costs or very low revenue requirements, the threshold tax system may yield marginally lower social costs than both minimum tax systems. This result is driven by the greater flexibility of the threshold system to "simply" adjust the output threshold y^* when targeting a lower number of firms to operate under the profit tax. In contrast, under both minimum tax regimes where firms bunch at the cutoff profit rate, reducing the number of firms subject to the profit tax requires more substantial changes in both the output and profit tax rates. These sharper adjustments in both tax rates increase the social costs associated with firm exit, output distortions, and cost overreporting.

Our analysis suggests several directions for fruitful future research. One such direction is to introduce compliance costs. Like administrative cost, the costs borne directly by firms are social costs but, unlike administrative cost, firms take them into account when making decisions. Firms close to the margin of one tax base versus the other might incur especially high compliance costs, as they need to calculate both tax liabilities to determine which is best for them. Another direction is to analyze other versions of the dual-regime tax systems; for example, some systems use assets as one of the bases. Third, the model might elaborate on the role of the informal sector. In our model, it uses only labor and produces with constant returns to scale. A large recent literature on the informal sector could be drawn on to enrich the analysis of this sector and its role in the optimal dual-regime business tax systems. Finally, other policy instruments could be investigated. One is the fraction of costs that are deductible, as investigated by Best et al. (2015). Less than

full deductibility constrains cost overreporting, but also limits the efficiency advantages of a (quasi-)profits tax. Another is the intensity of tax enforcement, which is costly but which can moderate the extent of socially costly and revenue-sapping cost overreporting.

References

[1] Alejos, L., 2018, "Firms' (Mis)-Reporting under a Minimum Tax: Evidence from Guatemalan Corporate Tax Returns," Working Paper, Available at http://www-personal.umich.edu/~lalejos/pdf/FirmsMisreporting.pdf.

[2] Almunia, M., and D. Lopez-Rodriguez, 2018, "Under the Radar: The Effects of Monitoring Firms on Tax Compliance," *American Economic Journal: Economic Policy*, 10 (1), 1–38.

[3] Bachas, P., and M. Soto, 2021, "Corporate Taxation under Weak Enforcement." *American Economic Journal: Economic Policy* 13 (4), 36–71.

[4] Besley, T., and T. Persson, 2013, "Taxation and Development," Handbook of Public Economics, Vol. 5, edited by A.J. Auerbach, R. Chetty, M. Feldstein, and E. Saez, 51–110.

[5] Best, M.C., A. Brockmeyer, H.J. Kleven, J. Spinnewijn, and M. Waseem, 2015, "Production versus Revenue Efficiency with Limited Tax Capacity: Theory and Evidence," *Journal of Political Economy*, 123 (6), 1311–1355.

[6] Devereux, M., L. Liu, and S. Loretz, 2014, "The Elasticity of Corporate Taxable Income: New Evidence from UK Tax Records," *American Economic Journal: Economic Policy*, 6 (2), 19–53.

[7] Dharmapala, D., J. Slemrod, and J.D. Wilson, 2011, "Tax Policy and the Missing Middle: Optimal Tax Remittance with Firm-Level Administrative Costs," *Journal of Public Economics*, 95 (9-10), 1036–1047.

[8] Gordon, R., and W. Li, 2009, "Tax Structures in Developing Countries: Many Puzzles and a Possible Explanation," *Journal of Public Economics*, 93 (7-8), 855–866.

[9] Grossman, G.M., and E. Helpman, 1994, "Protection for Sale," *American Economic Review*, 84 (4), 833–850.

[10] Helpman, E., Melitz, M., and Y. Rubinstein, 2008, "Estimating Trade Flows: Trading Partners and Trading Volumes," *The Quarterly Journal of Economics*, 123 (2), 441–487.

[11] Kanbur, R., and M. Keen, 2014, "Thresholds, Informality, and Partitions of Compliance," *International Tax and Public Finance*, 21 (4), 536–559.

[12] Keen, M., and J. Mintz, 2004, "The Optimal Threshold for a Value-Added Tax," *Journal of Public Economics*, 88 (3-4), 559–576.

[13] Lobel, F., T. Scot, and P. Zúniga, 2024, "Corporate Taxation and Evasion Responses: Evidence from a Minimum Tax in Honduras," *American Economic Journal*, 16 (1), 482–517.

[14] Melitz, M. J., 2003, "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71 (6), 1695–1725.

[15] Mosberger, P., 2016, "Accounting versus Real Production Responses among Firms to Tax Incentives: Bunching Evidence from Hungary," *MNB Working Papers*, No. 3.

[16] Tybout, J., 2014, "The Missing Middle, Revisited," *Journal of Economic Perspectives*, 28 (4), 235–36.

[17] Wei, F., and J.-F. Wen, 2019, "The Optimal Turnover Threshold and Tax Rate for SMEs," *IMF Working Paper*, WP/19/98.

[18] Wei, F., and J.-F. Wen, 2023, "Designing a Presumptive Income Tax Based on Turnover in Countries with Large Informal Sectors," *IMF Working Paper*, WP/23/267.

A Countries with Dual-Regime Business Tax System

Table 2 presents a non-exhaustive list of countries that, as of May 2024, apply a threshold to distinguish firms subject to a turnover tax from those subject to a profit tax. Additionally, the table reports each country's corporate tax rate (CTR), the applicable turnover tax rate, and the threshold amount in both national currency and US dollars.

Table 2: Countries with Threshold Tax Regimes

Country	CTR	Threshold Tax	Threshold	(in USD)
Africa				
Cabo Verde	22.0%	4.0% of gross sales	CVE 10,000,000	(98,'000)
Congo (DRC)[1]	30.0%	2.0% of turnover	CDF 80,000,000	(29,000)
Egypt[2]	22.5%	1.0% of turnover	EGP 10,000,000	(210,000)
Kenya	30.0%	3.0% of turnover	KES 25,000,000	(190,000)
Madagascar	20.0%	3.5% of turnover	MGA 400,000,000	(90,000)
Rwanda	30.0%	3.0% of turnover	RWF 20,000,000	(15,600)
South Africa[3]	27.0%	3.0% of turnover	ZAR 1,000,000	(54,000)
Uganda[4]	30.0%	1.5% of gross turnover	UGX 150,000,000	(40,500)
Zambia[5]	30.0%	4.0% of turnover	ZMW 800,000	(31,000)
Americas				
Ecuador	25.0%	2.0% of gross income	USD 300,000	(300,000)
Peru	29.5%	1.5% of gross income	PEN 525,000	(141,750)
Asia				
Indonesia	22.0%	0.5% of turnover	IDR 4,800,000,000	(312,000)
Uzbekistan	15.0%	4.0% of turnover	UZS 1,000,000,000	(81,000)
Europe				
Armenia[6]	18.0%	3.5% of turnover	AMD 115,000,000	(298,500)
Azerbaijan	20.0%	2.0% of gross revenue	AZN 200,000	(118,000)
France[7]	25.0%	2.2% of turnover	EUR 78,000	(85,000)
Italy[8]	24.0%	15% of turnover	EUR 85,000	(92,000)
North Macedonia[9]	10.0%	1.0% of gross income	MKD 6,000,000	(105,000)
Romania[10]	16.0%	3.0% of revenue	EUR 500,000	(540,000)

[1]Reduced rate of 1% for supply of goods; 2% applies to supply of services. [2]Lump sum payment for micro firms and progressive rates from 0.5% to 1% for small firms with turnover between USD 21,000 to 210,000. [3]Lower rates of 1% and 2% apply to turnover below ZAR 500,000 and 750,000. [4]Progressive rates increasing from 1% to 3% with turnover up to the threshold. [5]Persons working in the gig economy must pay 4% of the gross income on income between ZMW 12,000 and 800,000 per annum. [6]3.5% of turnover for production activities, 5% of turnover for trading activities. [7]Threshold and rate apply to self-employed doing professional activities; a higher threshold of EUR 189,000 and lower rate of 1.7% of turnover apply for commercial, trade or service activities. [8]Flat tax scheme applies to self-employed only. [9]Exemption from corporate income taxation for firms with gross income below MKD 3,000,000 (USD 52,500). [10]1% of revenue if revenues do not exceed EUR 60,000.

Source: PWC Worldwide Tax Summaries Online (accessed May 2024). https://taxsummaries.pwc.com/

In most of the African countries with a threshold tax in place, the threshold is set at around USD 50,000, which the exception of Egypt and Kenya. The latter two countries

have much higher thresholds of around USD 200,000. Firms with an annual turnover below the threshold are subject to a tax liability of between 1 and 4 percent of turnover. Firms with turnover exceeding the threshold are subject to the standard corporate tax regime. Several European countries as well have a threshold tax system in place. In most European countries the threshold is set at around USD 100,000 (except for Armenia and Romania) and turnover tax rates are between 1 and 3 percent.

Table 3: Countries with Minimum Tax Regimes

Country	Corporate Tax Rate	Minimum Tax
Africa		
Cameroon[1]	33.0%	5.5% of turnover
Chad	35.0%	1.5% of turnover
Congo (DRC)	40.0%	1.0% of turnover
Equatorial Guinea	35.0%	1.5% of turnover
Gabon[2]	30.0%	1.0% of annual revenue
Ivory Cost[3]	25.0%	0.5% of turnover
Mauritania	25.0%	2.0% of gross income
Senegal[4]	30.0%	0.5% of turnover
Americas		
Bolivia	25.0%	3.0% of sales revenue
Dominican Republic	27.0%	1.0% of turnover
Guyana[5]	40.0%	2.0% of turnover
Honduras	25.0%	1.0% of gross income
Nicaragua[6]	30.0%	3.0% of turnover
Asia		
India[7]	30.0%	15% of adj. book profits
Pakistan	29.0%	1.25% of turnover
South Korea[8]	24.0%	17% of taxable income
Europe		
Austria	23.0%	fixed minimum amount
Hungary	9.0%	2.0% of total revenues
Romania[9]	26.0%	1.0% of net turnover

[1]Firms with annual turnover below XAF (=0.0016 USD) 3 bn. are subject to a lower corporate tax rate of 27.5% or 2.2% of annual turnover. [2]Minimum tax cannot be lower than a flat minimum tax of XAF (=0.0016 USD) 1 m. [3]Minimum tax is subject to a maximum of XOF (=0.0017 USD) 35 m. [4]Tax on turnover is capped at XOF (=0.0017 USD) 5 m.[5]Any excess of the minimum tax can be carried forward to offset subsequent corporate income tax liability.[6]Lower minimum tax of 1.0% (2.0%) for income up to C$ (=0.027 USD) 60 m. (between 60 and 160 m). [7]Tax rates exclusive of any surcharges. [8]Taxable income before certain deductions. [9]Applies only to firms with turnover exceeding EUR 50 m.

Source: PWC Worldwide Tax Summaries Online (accessed May 2024). https://taxsummaries.pwc.com/

Table 3 presents a non-exhaustive list of countries that, as of May 2024, apply a minimum tax system. The table includes each country's corporate tax rate as well as the applicable minimum tax.

Under the minimum tax system, not a threshold determines whether a firm pays the output tax or the profit tax, but each firm must pay the higher of the two taxes, given their chosen output level. As it becomes apparent from Table 3, the structure of the minimum tax systems is quite similar around the world (except for India and South Korea). Countries levy tax rates of between 1 and 2 percent on turnover and corporate tax rates range between 25 and 35 percent. Only in India and South Korea much higher tax rates of around 15 percent are applied, however, not on gross turnover but on some adjusted base measure.

B Proof of Propositions

Proposition 1: *If administrative costs and overreporting are sufficiently low, then a dual-regime threshold system is optimal.*

Proof. We know that a profit tax is optimal if there are no administrative and overreporting costs. Thus, we can demonstrate the optimality of a dual regime threshold system by introducing a small amount of administrative costs and showing that output taxation for firms producing at or below a threshold y^* is desirable. We start by setting $t_o = 0$ and placing the threshold y^* at the lowest productivity of firms subject to the profit tax: $y^* = y_p(\varphi_l(0))$. Since type-φ_l firms earn zero profits and therefore face no profit tax liabilities, this policy does not affect welfare. Next, we keep $t_o = 0$ and increase y^* above $y_p(\varphi_l)$. Firms with a productivity φ where $y_p(\varphi) = y^*$ now find they can reduce their tax liabilities to zero by switching from the profit tax to the output tax. Consequently, firms start bunching at y^*. However, the extent of bunching can be controlled by keeping y^* close to $y_p(\varphi_l)$. Thus, we can set y^* at a level where a positive measure of firms bunch, but only those with productivities where profit tax payments are less than the administrative cost a: $t_p T_p(\varphi) < a$. This policy increases total after-tax profits while generating a government budget surplus. Finally, we can generalize the welfare-improving policy to include output tax rates below some small upper bound $\epsilon > 0$. The output tax reduces the extent of bunching, but a positive amount of welfare-improving bunching will still occur for sufficiently low values of t_o. Additionally, the output tax creates behavioral effects for firms that previously faced a zero output tax, resulting in a rise in φ_l, and a reduction in the outputs of non-bunched firms paying the output tax. However, the envelope theorem implies that these behavioral changes have a zero first-order effect on after-tax profits, and they also have a zero first-order effect on government revenue when

evaluated at $t_o = 0$. Thus, we conclude that welfare can be increased by moving from a profit tax system to a tax system where some firms pay a positive output tax. □

Proposition 2: *Starting from an output tax system with budget-balancing tax rate t_o^s, design a threshold system for a division of firms between the two taxes, where firms with productivities less (greater) than φ^* pay the output tax (profit tax) and $\varphi^* > \varphi_l(t_o^s)$. If administrative costs and overreporting are positive but sufficiently small, budget-balancing tax rates, t_o^d and t_p^d, and threshold y^* can be chosen so that $\varphi^* > \varphi_l(t_o^d)$, resulting in a dual-regime threshold system that provides higher welfare than the output tax system.*

Proof. Given the assumption of arbitrarily small administrative costs and overreporting, we may set them equal to zero. Given t_o^s, first choose y^* and the profit tax t_p so that there is no bunching of firms at y^*, that is $y(t_o^s, varphi^*) = y^*$, implying $\varphi^{**} = \varphi^*$. A type-φ^* firm's tax payments must be higher under the profit tax than under the output tax because this firm faces the same after-tax profits under the output tax and the profit tax. However, its before-tax profits are higher under the profit tax because its output is not distorted. Under assumption (7), all firms with productivities greater than φ^* also make higher tax payments than they would under the output tax system. It follows that the government budget is in surplus. To bring the government budget back into balance, reduce the taxes to t_o^d and t_p^d, with these rates chosen to leave unchanged the productivity of the indifferent firm, φ^*, once y^* is increased to prevent bunching from occurring, i. e., $y(t_o^d, \varphi^*) = y^*$. With all firms now facing higher before-tax profits but total tax payments unchanged, total after-tax profits must be higher under the threshold system. Finally, note that the reduction in the output tax from t_o^s to t_o^d brings new firms into the formal sector, reducing the lowest productivity from $\varphi_l(t_o^s)$ to $\varphi_l(t_o^d)$. Thus, φ^* continues to lie above the lowest productivity, $\varphi_l(t_o^d)$, in which case the threshold system is indeed a dual-regime tax system. □

C Parameterized Model Algebra

This appendix provides the algebra for the parameterized version of the model used in the numerical exercises. For these exercises, we use a quadratic cost function $c(\varphi) = 0.5\frac{1}{\varphi}y^2 - f$. Note that we solve a discretized version of the model where we draw random firm productivities and so any model aggregates are computed numerically. Hence, we only require the algebra for the firm-level choices and thresholds here.

Threshold System:
Under a profit tax, the firm maximizes setting $y = \varphi$ and earns a profit of $(1-t_\pi)(0.5\varphi - f)$. Under an output tax, the firm maximizes setting $y = (1-t_o)\varphi$ and earns a profit of: $(1-t_o)^2 0.5\varphi - f$. A firm that bunches produces $y = y^*$ with profits $(1-t_o)y^* - 0.5\frac{1}{\varphi}y^{*2} - f$. The upper end of the bunching threshold, φ^*, is defined by: $(1-t_\pi)(0.5\varphi - f) = (1-t_o)y^* - 0.5\frac{1}{\varphi}y^{*2} - f$. The lower end is $\varphi^{**} = \frac{y^*}{1-t_o}$.

Minimum Tax System:
Much of this algebra is the same as under the threshold system. Under a profit tax, the firm maximizes setting $y = \varphi$ and earns a profit of $(1-t_\pi)(0.5\varphi - f)$. Under an output tax, the firm maximizes setting $y = (1-t_o)\varphi$ and earns a profit of: $(1-t_o)^2 0.5\varphi - f$.

The difference is in the firms that bunch and their output levels when bunching. The output level of a bunching firm will be such that tax payments are equalized under the two systems: $t_y y = t_\pi \left(y - 0.5\frac{1}{\varphi}y^2 - f\right)$. We can re-write this as: $0.5\frac{1}{\varphi}y^2 - \frac{t_\pi - t_y}{t_\pi}y + f = 0$. This can be solved using the quadratic formula: $y_b = \varphi\left[\frac{t_\pi - t_y}{t_\pi} + \left(\left(\frac{t_\pi - t_y}{t_\pi}\right)^2 - 2\frac{1}{\varphi}f\right)^{0.5}\right]$.[27]

Next, we can find the bunching region, φ^* and φ^{**}. The type φ^* is producing its profit maximizing output while producing at the bunching level, i.e $y_b = \varphi$. Using the y_b expression above, we can obtain the threshold: $\varphi^* = 2f\left\{\left(\frac{t_\pi - t_y}{t_\pi}\right)^2 - \left(\frac{t_y}{t_\pi}\right)^2\right\}^{-1}$. The lower end of the threshold, φ^{**}, is choosing the same output it would under the output tax, i.e. $y_b = (1-t_y)\varphi$. This gives us $\varphi^{**} = 2f\left\{\left(\frac{t_\pi - t_y}{t_\pi}\right)^2 - \left[(1-t_y) - \frac{t_\pi - t_y}{t_\pi}\right]^2\right\}^{-1}$.

Calibration of Profit Underreporting:
To account for cost overreporting, we use the profit underreporting function $e(\gamma) = \beta\gamma^\varsigma$, and the first order condition for the optimal fraction of pre-tax profits reported is: $\beta\varsigma\gamma^{\varsigma-1} = t_p$.

In the baseline, $\beta = 8.3$ and $\varsigma = 3$, which yields a baseline 10% overreporting rate under a 25% profit tax while ensuring a plausible fraction of firms under the profit and output tax for all three systems. Since our baseline optimal tax rates are closer to 29%, the implies ends up implying an overreporting rate of about 11%. The range of β values used for the comparative statics is 50% of this value to 150% of this value. This range

[27] While in principle there could be two roots, only the larger root will be relevant. For firms in the bunching thresholds, it can be shown through simple algebra that the roots will always be real.

spans cost overreporting rates of 9-15%, which is consistent with Best et al. (2015), Alejos (2018) and Lobel, Scot, and Zúniga (2024).