INTERNATIONAL MONETARY FUND

Designing a Presumptive Income Tax Based on Turnover in Countries with Large Informal Sectors

Feng Wei and Jean-François Wen

WP/23/267

2023
DEC

WORKING PAPER

 WP/23/267

IMF Working Paper

Fiscal Affairs Department

Designing a Presumptive Income Tax Based on Turnover in Countries with Large Informal Sectors

Prepared by Feng Wei and Jean-François Wen

Authorized for distribution by Mario Mansour
December 2023

ABSTRACT

Turnover (sales) is frequently used in developing countries as a presumptive income tax base, to economize on the costs of tax administration and taxpayer compliance. We construct a simple model where a size threshold separates firms paying turnover tax from those paying profit tax (regular income tax), and where firms have the option of producing in the untaxed, informal sector. The optimal turnover tax rate trades off two policy concerns: reducing informality and avoiding strategic reductions in sales by firms seeking to remain below the threshold for the profit tax. We provide analytical results and calibrate the model to compute the optimal policy using realistic parameter values. The optimal turnover tax rate for countries with large informal sectors is found to be around 2.5% across most scenarios, while the threshold separating the turnover tax regime from profit tax lies for the most part between $65,000 and $95,000. Introducing an optimally designed turnover tax reduces the rate of informality of businesses by about 12 percentage points in the calibrated model.

RECOMMENDED CITATION: Wei, Feng, and Jean-François Wen. 2023. Designing a Presumptive Income Tax Based on Turnover in Countries with Large Informal Sectors. IMF Working Paper 23/267. Washington, DC: International Monetary Fund.

JEL Classification Numbers:	H21, H25, H32, O17
Keywords:	Turnover Tax; Corporate Income Tax; Compliance Costs; Informality
Author’s E-Mail Address:	jwen@imf.org; fwei@sdu.edu.cn

* The authors would like to thank Thomas Benninger, Brendan Crowley, Anna Ivanova, Alexander Klemm, Li Liu, Simon Naitram, Rafael Machado Parente, Hyunmin Park, Alexandra Solovyeva, Anthony Ramarozatovo for their helpful comments on a draft of the paper.

WORKING PAPERS

Designing a Presumptive Income Tax Based on Turnover in Countries with Large Informal Sectors

Prepared by Feng Wei and Jean-François Wen

Contents

I. INTRODUCTION

Turnover taxes are prevalent in developing countries, as a simple form of presumptive taxation of business income.[1] Taxes levied on turnover (sales), instead of a profit tax – that is, corporate income tax (CIT) for legal entities, personal income tax (PIT) for unincorporated individual entrepreneurs – can reduce the relatively high compliance costs of micro and small enterprises (MSEs), which might otherwise discourage entrepreneurs from formalizing their activities and paying taxes. Moreover, for the tax administration, sales are easier to ascertain than profits, which require deductions to be verified. Although tax collection from turnover tax regimes is typically below 0.25 percent of GDP, their main objective is to promote greater inclusiveness within the formal sector of the economy. In countries where voluntary compliance with the income tax laws is strained by a lack of taxpayer knowledge and capacity, a turnover tax can be a useful policy to reduce informality and encourage tax compliance, when it is combined with taxpayer education and tax collection enforcement. More than 35 developing, emerging market, and transition economies currently use turnover as a presumptive income tax base for qualifying entrepreneurs and enterprises (see Table 1 below). Even among developed countries a turnover-based presumptive income tax is available for individual entrepreneurs or businesses (e.g., Austria, France, Italy, Portugal). However, the tax literature offers little guidance on choosing the tax rate and threshold.

We construct a simple model where a size threshold separates firms paying standard business income tax from firms below the threshold who pay a proportional tax on their turnover. We make two extreme assumptions to allow us to study, in the simplest possible way, the optimal design of a turnover tax regime. Following Kanbur and Keen (2014), we suppose that firms can costlessly adjust their sales downward from an exogenous maximum potential amount, for example to qualify for the presumptive income tax regime. Furthermore, firms can escape taxation altogether by becoming 'ghosts' in the shadow economy.[2] The production inefficiencies associated with the presumptive regime and with informal sector production are also modeled very simply, as exogenous additions to the marginal cost of production. The assumptions highlight how the option to produce in the untaxed, informal sector constrains the turnover tax rate and how the turnover tax interacts with the standard income tax regime to create 'notches' in the production levels of firms.

The advantage of such a streamlined model is that it is easy to calibrate and generates practical tax policy guidelines and explicit solutions for the optimal policy. We show how the

[1] Presumptive income taxation refers to the use of indicators (e.g., sales, assets, floor space, number of employees, etc.) to assess and tax the 'presumed profit' of the enterprise or individual. This paper focuses on a specific design of presumptive taxation. For a general review of presumptive taxes, see chapter 12 in Thuronyi (1996) and chapter 2.5 in World Bank Group (2007).

[2] Waseem (2018) shows that the number of enterprises operating in the formal sector responds to tax pressures. When Pakistan increased the tax on unincorporated partnerships to a level comparable to the standard corporate tax, within three years of the tax increase, the number of partnerships in Pakistan had declined to 36% of the baseline level. Very few of these companies became incorporated. It is plausible that a portion of them became informal.

Table 1. International practices on turnover thresholds and tax rates (2021)

Country	Threshold (USD)	Turnover tax rate (%)	Corporate tax rate (%)
Africa			
Algeria	58,000	5 (goods), 12 (other)	26
Benin	95,000	2	30
Congo (DRC)	40,000	1 (goods), 2 (services)	30
Congo (Brazzaville)	174,000	5	28
Egypt	637,000	0.5 – 1 (progressive rates)	22.5
Ethiopia	25,000	2	30
Guinea	16,000	5	25
Kenya	442,000	1	30
Liberia	21,000	4	25
Madagascar	51,000	5	20
Mauritania	83,000	3	25
Niger	190,000	3 (trade), 5 (other)	30
Rwanda	20,000	3	30
Senegal	86,000	2 (mfg.), 5 (services)	30
Seychelles	68,000	1.5	25
Tanzania	42,000	3 – 5.3	30
Uganda	41,000	1.5	30
Zambia	48,000	4	30
Asia			
Indonesia	336,000	1	22
Mongolia	17,000	1	25
Tonga	44,000	2	25
Central and South America			
Brazil	21,000	3 – 5 (progressive rates)	34
Ecuador	300,000	2	25
Guatemala	19,000	5	25
Peru	132,000	1.5	29.5
Eastern Europe			
Armenia	240,000	5 (trading), 3.5 (production)	18
Azerbaijan	117,000	2	20
Belarus	196,000	6	18
Georgia	162,000	1	15
Kyrgyzstan	354,000	4 (mfg., comm., ag. proc.), 6 (other)	10
North Macedonia	110,000	1	10
Romania	880,000	1 (3 if no employees)	16
Russia	32,000	4 (B2C sales), 6 (B2B sales)	20
Ukraine	278,000	5 (for 3 VAT payer)	18
Uzbekistan	92,000	4	15

Sources: Tax codes and IBFD library for 2021. *Notes*: US dollar exchange rate on January 1, 2021. Corporate income tax rates are provided for the sake of comparison. Personal income tax rates are also pertinent but are omitted due to space constraints.

optimal turnover tax rate and threshold vary in response to the profit tax rate, the cost of informal sector production, and the costs of tax compliance. The optimal policy results in bunching behavior or in an absence of bunching, with the former case occurring only at relatively high profit tax rates. While the solution for the optimal threshold in the model is relatively simple and intuitive, the solution for the optimal tax rate is complicated, reflecting the balancing act between two behavioral margins. That is, if the tax rate is too high then more firms will choose to produce in the informal, untaxed, sector; but if it is too low, some firms will migrate away from the standard regime by bunching just below the threshold. Bunching can arise from discrete changes in the tax liability at an earnings cutoff (Kleven and Waseem, 2013) or due to an increase in taxpayer compliance costs upon entering the standard regime (Keen and Mintz, 2004). Simulations of a calibrated version of the model (set largely on evidence from countries in Sub-Saharan Africa) generate realistic shares of informal enterprises in the economy and provide numerical values for the optimal design of the turnover tax. Introducing an optimally designed turnover tax reduces the rate of informality of businesses by about 12 percentage points in the model calibrated for countries with large informal sectors. The optimal turnover tax rate is found to be around 2.5% across most scenarios, while the threshold separating the turnover tax regime from profit tax lies for the most part between $65,000 and $95,000. We show that the effective tax rate on profit implied by the optimal turnover tax rate may be lower than the effective or statutory CIT rate at the threshold. The finding contravenes a frequently heard policy recommendation, that the effective tax rate in the presumptive regime should equal or exceed that in the regular regime. The conventional wisdom sometimes has force (e.g., when the CIT rate is moderate), but more generally the turnover tax rate must address not only the margin at the threshold but also the margin between informality and registering for the presumptive tax regime in the first place. From this perspective, the optimal presumptive tax regime does not aim to replicate the effective tax burden of the regular regime, but instead forms part of an optimal tax strategy to mitigate compliance costs while dealing with the informal sector.

While our analysis is related to the literature on the the effects of the VAT threshold on the behavior of firms, all of these studies take the VAT tax rate as exogenous.[3] Endogenizing the turnover tax rate in our model is important because it is one of the policy levers that can be used to balance the dual concerns over bunching and escaping into informality. The major contribution of this study is to optimize not only the threshold but the tax rate, thereby improving our understanding of how to design presumptive tax regimes for economies characterized by high levels of informality. We abstract from the VAT in our analysis of the turnover tax. Kanbur and Keen (2014) show that it is generally not optimal to set the same threshold for multiple taxes, because the size of the notch leads to greater bunching than if thresholds are set separately. Moreover, some countries have VAT thresholds of zero or do not have a VAT. These observations justify analyzing the turnover tax threshold and tax rate independently of the VAT. However, if the turnover tax's threshold is pegged to the VAT registration

[3]The seminal paper on the optimal VAT threshold is Keen and Mintz (2004). Kanbur and Keen (2014) reexamine the optimal VAT threshold when firms can escape taxation by becoming informal or can reduce their tax burdens by misreporting sales. de Mel et al. (2010) analyze the incentive for VAT evasion when there are informality chains. Liu et al. (2021) examine bunching below the VAT threshold and voluntary VAT registration by firms in the presence of tax cascading.

threshold, as is sometimes recommended in policy advising, then the model can be used to find the optimal turnover tax rate for the given threshold.

Other related studies include Dharmapala et al. (2011) and Best et al. (2015). The first of these shows that, in the presence of per-firm administrative costs, it can be optimal to exclude small firms from taxation by setting a size threshold, even if this causes somewhat larger firms to bunch below the threshold. In contrast to their study, we examine the application of a presumptive tax on firms below the threshold. The second paper analyzes a tax system whereby firms are taxed on profits, provided the tax liability is greater than an alternative minimum tax levied on turnover. The purpose of the minimum tax is to reduce the opportunity for evading the corporate income tax. The turnover tax in this case does not economize on compliance costs, since every firm must calculate and report its liabilities under the regular regime. Thus, our paper addresses a different policy issue.

Section II presents the model. Section III describes the optimal tax policy and provides comparative statics analysis. Section IV calculates the optimal turnover threshold and turnover tax rate for alternative calibrations. Section V concludes. Proofs of lemmas and propositions are contained in an appendix.

II. A Simple Model of Firms' Behavior

A. Sales and profits

A population of firms is endowed with differing levels of *maximum* potential sales, $Z \in (0, Z^M)$, where Z is exogenous and distributed according to a twice differentiable distribution function $H(Z)$ with strictly positive density $h(Z)$. Denote the derivative of the density function by $h'(Z)$. The upper-support of the distribution, Z^M, may be finite or infinite. Firms can choose to adjust (costlessly) their sales to any level below their maximum potential, for instance, to reduce their tax burden. The output price is normalized to one, so output is the same as sales. The marginal cost of production is assumed to be a constant, $C < 1$. Hence, a firm producing sales of Z has a pre-tax profit of $(1-C)Z$.

The tax system is as follows. Firms with actual sales equal to or exceeding a sales threshold $\bar{Z}$ face a tax on profits at the rate t^c (henceforth, the 'regular regime', with the superscript intimating the CIT). Firms selling below the threshold are taxed on their sales, with no deduction for costs, at the rate t (henceforth, the 'presumptive regime'). Firms face fixed compliance costs of Γ and Γ' in the regular and presumptive regimes, respectively. The government faces corresponding fixed administrative costs of A and A'. Since sales are easier to record and audit than profits, let $\Gamma > \Gamma' > 0$ and $A > A' > 0$. The profit tax rate is taken to be exogenous, as it is determined by the wider prevailing fiscal framework of the country. The turnover tax rate t and the threshold $\bar{Z}$ are the policy choices in the model. A commonly encountered policy decision in developing countries is how to adjust the turnover tax regime in response to changes in the profit tax rate.

We assume that there is an extra marginal cost α, on top of C, incurred by firms in the presumptive tax regime. The extra cost of production can represent higher borrowing costs due to a lack of verifiable information on profit (since the tax base is sales), or production inefficiencies caused by the non-deductibility of costs under presumptive taxation.[4] Finally, as in Kanbur and Keen (2014), we assume that firms can escape taxation altogether by 'disappearing' into the informal sector. In this case the marginal cost is increased by an amount λ, on top of C, where λ is a parameter representing the inefficiencies associated with producing in the informal sector. For example, informal enterprises face inconveniences in having to produce in ways that avoid detection by tax officials.[5] We suppose that $\lambda > \alpha > 0$ to reflect the high cost of informal production. We also assume that $1 - C > \lambda > t^c(1-C)$, where the first inequality means that firms can earn positive profits in the informal sector, while the second precludes the informal sector from dominating the regular regime for all firms. Lastly, the fixed compliance costs, Γ and Γ', are presumed not to be so large as to preclude the possibility that even large producers (i.e., for Z in a neighborhood of Z^M) are unable to earn strictly greater net profit in the regular regime (given t^c) and in the presumptive regime (at $t = 0$), than in the informal sector.

We can summarize the structure of the economy by specifying the after-tax profit functions of four types of firms: those producing at their maximum sales level in the regular regime (*regulars*, earning π^R); those who adjust their sales downward to just below the threshold (*adjusters*, earning π^A);[6] those producing at their maximum sales in the presumptive regime (*presumptives*, earning π^P); and those producing at their maximum sales but escaping taxation by remaining informal (*informals*, earning π^I). Each firm chooses how to behave to maximize its after-tax profits. The net profit functions are:

$$\pi^R(Z) = (1-t^c)(1-C)Z - \Gamma \tag{1}$$
$$\pi^A(\bar{Z}) = (1-t-C-\alpha)\bar{Z} - \Gamma' \tag{2}$$
$$\pi^P(Z) = (1-t-C-\alpha)Z - \Gamma' \tag{3}$$
$$\pi^I(Z) = (1-C-\lambda)Z \tag{4}$$

[4] In a standard model of a firm's input choices, a pure profit tax is non-distortionary, while a tax on turnover would comparatively reduce the firm's optimal inputs and output. Therefore, for any turnover tax rate that collects the same tax revenues as the profit tax, the turnover tax would result in a relatively lower after-tax profit (ignoring compliance costs). In our model, α mimics this effect. That is, for any t^c, setting $t = t^c(1-C)$, the resulting after-tax profit is lower in the turnover tax regime than in the regular profit tax regime.

[5] Informal firms have less scope for marketing or they locate in obscure locations to avoid attracting the attention of the law (Bruhn and McKenzie, 2014). Advances in electronic technology have facilitated coordination between government agencies, such that a taxpayer identification number (TIN) can be required to access government services (e.g., obtaining passports or driver's licenses, registering cars and property, using public schools or hospitals, or subscribing to public utility services). These initiatives can serve to increase the costs of operating outside the tax system (Bird and Zolt, 2008).

[6] Their sales are below but arbitrarily close to the threshold, and their profit is below but arbitrarily close to π^A.

For simplicity, the fixed compliance cost is being treated as non-deductible in the net profit equation for the regular regime. This is not unrealistic in the case of small enterprises, where most of the time spent for tax compliance activities is attributable to the owners (Smulders et al., 2012). In any case, with t^c fixed, making the compliance cost deductible is a simple matter of rescaling the size of Γ by $(1-t^c)$. Similarly, if there are variable compliance costs that are proportional to turnover, they can be easily incorporated in the model's calibration by appropriately redefining the parameters C, α, and λ.[7]

B. Partitioning the distribution of firms

1. Informal and formal sectors

Recall that firms have the option of becoming informal. Firms with potential sales below the threshold prefer informality to the presumptive regime if and only if

$$(1-C-\lambda)Z > (1-t-C-\alpha)Z - \Gamma' \tag{5}$$

which defines a cutoff sales level,

$$Z^{IP} = \frac{\Gamma'}{\lambda - t - \alpha} \tag{6}$$

All firms with $Z < Z^{IP}$ prefer informality over presumptive taxation. Note that the proportion of informals increases with the turnover tax rate and with the cost of tax compliance. The denominator of (6) must be positive or else all firms prefer informality to the presumptive tax regime.[8]

Firms with potential sales exceeding the threshold also could choose to join the informal sector rather than face the regular regime. For completeness, use the profit functions to define a cutoff Z^{IR} such that all $Z < Z^{IR}$ choose the informal sector over the regular regime; and similarly, Z^{AI} is such that all firms with Z in the interval $\bar{Z} < Z < Z^{AI}$ prefer adjusting their sales down to (just below) the threshold $\bar{Z}$ over producing their maximum sales Z in the informal sector.[9]

2. Separating and bunching

Firms with maximum sales of at least $\bar{Z}$ decide between producing at their maximum and being subjected to the regular regime, or reducing their output to just below the threshold and

[7] For example, let v denote a per-unit, tax deductible, variable compliance cost in the regular regime. Then, in π^R, replace C with $\hat{C} = C + v$; in π^A and π^P, change C to $\hat{C}$ while replacing α with $\hat{\alpha} = \alpha - v$; and in π^I use $\hat{C}$ in the place of C and substitute $\hat{\lambda} = \lambda - v$ for λ.

[8] That is, $\lambda - t - \alpha < 0$ would imply $1 - C - \lambda > 1 - t - C - \alpha$.

[9] The explicit solutions for Z^{IR} and Z^{AI} will not be needed for presenting the analysis.

paying the presumptive tax. Given a threshold $\bar{Z}$ and tax rates t^c and t, adjusting dominates the regular regime whenever

$$(1-t-C-\alpha)\bar{Z}-\Gamma' > (1-t^c)(1-C)Z-\Gamma \tag{7}$$

which defines a cut-off sales level

$$\hat{Z} = \frac{(1-t-C-\alpha)\bar{Z}+\Gamma-\Gamma'}{(1-t^c)(1-C)} \tag{8}$$

All firms with sales below $\hat{Z}$ prefer adjusting over being regulars. Since firms can only adjust their sales downward from their potential level, bunching cannot occur when $\hat{Z} < \bar{Z}$. In that case, i.e., when bunching is absent, all firms produce at their maximum sales and the threshold separates them into two tax groups (abstracting from the option of informality for the moment). Firms with sales below the threshold earn π^P while those at or above the threshold earn π^R. We refer to this situation as a 'separating' equilibrium and illustrate it with Figure 1.

In the contrary case of $\hat{Z} > \bar{Z}$, a bunching equilibrium occurs, which is illustrated with Figure 2. Firms in the segment $[\bar{Z}, \hat{Z})$ have $\pi^A(\bar{Z}) > \pi^R(Z)$ and will adjust their production to (just below) the threshold, earning the same profit, (arbitrarily close to) $\pi^A(\bar{Z})$. When bunching occurs, sales levels between $\bar{Z}$ and $\hat{Z}$ will not be observed.[10] When $\hat{Z} > \bar{Z}$ the measure of firms bunching is given by $H(\hat{Z}) - H(\bar{Z})$. In the case of a uniform distribution for $H(Z)$ on the interval $(0, Z^M)$, the proportion of bunchers is given by

$$\frac{\hat{Z}-\bar{Z}}{Z^M} = \frac{[t^c(1-C)-(t+\alpha)]\bar{Z}}{(1-t^c)(1-C)Z^M} + \frac{\Gamma-\Gamma'}{(1-t^c)(1-C)Z^M} \tag{9}$$

Thus, bunching is driven by the differences in tax rates and compliance costs between the regular and presumptive regimes. Other things being the same, bunching is expected at relatively high rates of profit tax. In a practical setting, a separating equilibrium is identified by an absence of bunching in the empirical distribution of turnover size.

[10] Figures 1 and 2 are drawn under the assumption that the slope of π^R exceeds the slope of π^P. The relative slopes are assured if $t^c(1-C) < \alpha$, but it is necessary to assume this for the analysis. The only implication of π^P being steeper than π^R is that $\hat{Z}$ must exceed $\bar{Z}$, so then only bunching equilibria exist.

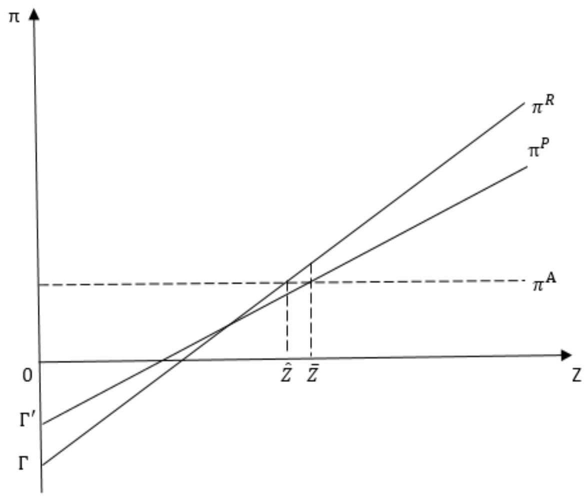

Figure 1. 'Separating' Case

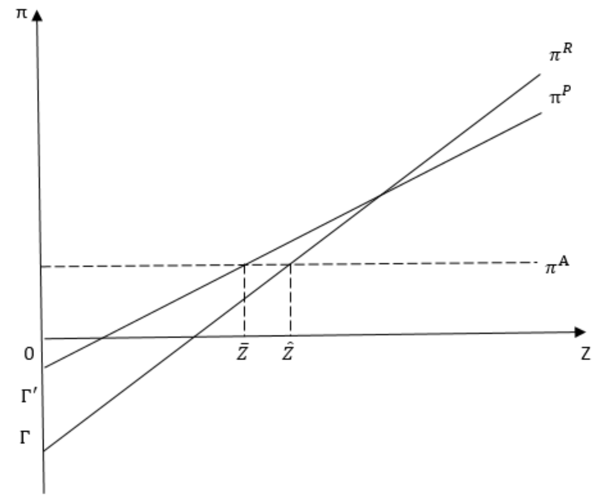

Figure 2. 'Bunching' Case

3. Partitions

Policy choices generate various configurations of the cutoffs, which in turn partition the space $(0, Z^M)$ according to the choices made by the firms. An exposition of all possibilities would be tedious. Instead, we focus on the partitions in which the smallest firms are in the informal sector, with intermediate firms in the presumptive regime, and the largest firms are in the regular regime.[11]

Lemma 1. *Let the turnover tax policy be such that* $\hat{Z} < \bar{Z}$ *with* $0 < Z^{IP} < \bar{Z}$ *and* $Z^{IR} < \bar{Z}$. *Then all firms produce their maximum output (separating case) and*

1. $\forall Z \in (0, Z^{IP})$*, firms locate in the informal sector*
2. $\forall Z \in [Z^{IP}, \bar{Z})$*, firms locate in the presumptive regime*
3. $\forall Z \in [\bar{Z}, Z^M)$*, firms locate in the regular regime*

Lemma 2. *Let the turnover tax policy be such that* $0 < Z^{IP} < \bar{Z} < \hat{Z} \leq Z^{AI}$. *Then some firms will restrict their output (bunching case) and*

[11]Most configurations of the cutoffs cannot arise without generating a contradiction in the preference orderings. For example, it cannot be the case that $Z^{AI} < Z^{IR} \leq \hat{Z}$. To see this, let $j \succ k$ represent "j is preferred to k," where j and k correspond to the actions defined by the cutoff Z^{jk}. By the definition of the various cutoffs, $\forall Z \in (Z^{AI}, \hat{Z})$, $I \succ A \succ R$, while $\forall Z \in (Z^{IR}, \hat{Z}]$, $R \succ I$, which is a contradiction. A valid alternative partition is $\hat{Z} < Z^{AI} \leq Z^{IR}$. In this case, firms with sales between Z^{AI} and Z^{IR} will prefer informality over either tax regime, even though larger firms ($Z > Z^{IR}$) prefer the regular regime and smaller firms ($Z < Z^{AI}$) prefer the presumptive regime. Since informal production is inefficient, a change in the turnover tax rate or threshold can be expected to generate an increase in welfare by inducing the informal firms into one of the tax regimes, provided the fixed administrative cost is not too large. Similarly, if $\bar{Z} < Z^{IP}$, then the presumptive regime would be irrelevant since all firms with sales below the threshold would opt for informality.

1. *$\forall Z \in (0, Z^{IP})$, firms locate in the informal sector*

2. *$\forall Z \in [Z^{IP}, \bar{Z})$, firms locate in the presumptive regime*

3. *$\forall Z \in [\bar{Z}, \hat{Z})$, firms bunch just below the threshold*

4. *$\forall Z \in [\hat{Z}, Z^{M})$, firms locate in the regular regime*

Note that it is unnecessary to specify the position of Z^{IR} in the partition of Lemma 2. Since $\hat{Z} \leq Z^{AI}$, it cannot be the case that $\hat{Z} < Z^{IR}$ because this would generate a contradiction in the preference ordering. That is, from the definitions of $\hat{Z}$ and Z^{AI}, firms with potential sales between these two cutoffs prefer being regulars to informals. But those with potential sales between $\hat{Z}$ and Z^{IR} prefer being informals to regulars, which cannot be. Thus, $Z^{IR} \in [0, \hat{Z}]$ but its precise location in the interval is irrelevant for the equilibrium.

III. Social Welfare Analysis

Social welfare is the sum of private net incomes (Π) and net tax revenue (G), with the latter multiplied by a factor $\delta > 1$, representing the marginal social value of tax revenues. The choice variables in the social welfare function are the turnover threshold $\bar{Z}$ and the tax rate t. Hence,

$$SW(\bar{Z}, t) = \Pi(\bar{Z}, t) + \delta G(\bar{Z}, t) \tag{10}$$

The welfare function will consist of a series of integrals with limits of integration determined by the relevant partition of sales in $(0, Z^{M})$. For analytical purposes, we construct the welfare function in accordance with the partitions in lemma 1 or lemma 2.[12]

A. Welfare in the separating equilibrium: $\bar{Z} \geq \hat{Z}$

We first construct the social welfare function for a separating equilibrium described by the partition in Lemma 1. The aggregate profit function is

$$\Pi(\bar{Z}, t) = \int_{0}^{Z^{IP}} \pi^{I} h(Z) dZ + \int_{Z^{IP}}^{\bar{Z}} \pi^{P} h(Z) dZ + \int_{\bar{Z}}^{Z^{M}} \pi^{R} h(Z) dZ \tag{11}$$

and the total net tax revenue is

$$G(\bar{Z}, t) = \int_{Z^{IP}}^{\bar{Z}} (tZ - A') h(Z) dZ + \int_{\bar{Z}}^{Z^{M}} [t^{c}(1 - C)Z - A] h(Z) dZ \tag{12}$$

Since the informal sector is an untaxed sector, there are no tax revenues collected from there. Firms with sales below the threshold (including bunchers) pay tax based on their sales, while

[12] In the numerical simulations of the model, we consider all admissible partitions.

firms with sales at or above the threshold pay tax based on their profit. The welfare function in a separating configuration is then given by

$$\begin{aligned} SW(\bar{Z},t) = & \int_0^{Z^{IP}} \pi^I h(Z)dZ + \int_{Z^{IP}}^{\bar{Z}} \pi^P h(Z)dZ + \int_{\bar{Z}}^{Z^M} \pi^R h(Z)dZ \\ & + \delta\{\int_{Z^{IP}}^{\bar{Z}} (tZ - A')h(Z)dZ + \int_{\bar{Z}}^{Z^M} [t^c(1-C)Z - A]h(Z)dZ\} \end{aligned} \tag{13}$$

The first-order condition for the threshold $\bar{Z}$ can be written as

$$\{[(1-t-C-\alpha)+\delta t]\bar{Z} - \Gamma' - \delta A'\}h(\bar{Z}) = \{[(1-t^c)(1-C)+\delta t^c(1-C)]\bar{Z} - \Gamma - \delta A\}h(\bar{Z}) \tag{14}$$

A small increase in the threshold causes the marginal firm to be moved from the regular regime to the presumptive regime. Since there is no bunching of firms in the separating case, the identity of the marginal firm is unique. Thus the optimum occurs when the net welfare gain from the new arrival in the presumptive regime, given by the left-hand side of (14), balances with the net welfare loss from the firm exiting the regular regime, given by the right-hand side. Rearranging the terms generates an expression for $\bar{Z}$.

Proposition 1. *For given tax rates t and t^c, the optimal threshold in a separating equilibrium (i.e., when $\bar{Z} > \hat{Z}$) is given by*

$$\bar{Z} = \frac{(\Gamma + \delta A) - (\Gamma' + \delta A')}{(\delta - 1)[t^c(1-C) - t] + \alpha} \tag{15}$$

Proposition 1 provides a convenient formula for the optimal threshold at fixed (but not necessarily optimal) tax rates, t and t^c, when the market equilibrium is characterized by an absence of bunching. In a practical setting, adjustments to the turnover threshold are often recommended for the existing tax rates. The formula (15) is akin to the 'benchmark' expression for the optimal VAT in Kanbur and Keen (2014) and Keen and Mintz (2004), which is interpreted there as an optimality condition when compliance is perfect and there are no behavioral responses (and the VAT tax rate is fixed). However, in the case of (15), behavioral responses of firms (by adjusting their sales downward) are not precluded. Instead, the formula arises as an equilibrium outcome of the model under policies, such as low to moderate profit tax rates, that result in a separating equilibrium.[13]

From the first-order condition of welfare with respect to the tax rate t in a separating equilibrium, we obtain the following optimality condition.

Proposition 2. *For a given threshold $\bar{Z}$ and profit tax rate t^c, the optimal turnover tax rate in a separating equilibrium is given by*

$$(\delta - 1)\int_{Z^{IP}}^{\bar{Z}} Zh(Z)dZ = \delta(tZ^{IP} - A')h(Z^{IP})\frac{dZ^{IP}}{dt} \tag{16}$$

[13]The proposition assumes the denominator of (15) is positive. If it is non-positive, then the regular regime is dominated by the presumptive regime and $\bar{Z}$ goes to infinity.

The left-hand side of (16) is the marginal social benefit of the increased tax revenues collected from firms in the presumptive regime, as a result of raising the tax rate. On the right-hand side is the marginal social cost of the lost tax revenues, net of administrative costs, resulting from firms now choosing to vanish from the formal sector into the informal sector ($dZ^{IP}/dt > 0$). The latter amount is weighted by the density of firms at the margin of indifference between informality and operating in the presumptive regime. The optimal tax rate requires $tZ^{IP} - A'$ to be positive, because the left-hand side of (16) is positive. Then, provided $h'(Z^{IP})$ is non-negative (as with a uniform distribution), if an interior solution for t exists it must be unique, because the left-hand side of (16) is strictly decreasing in t while the right-hand side is strictly increasing in t. An interior solution for t means it is non-negative and generates a partition that is consistent with a separating equilibrium defined by Lemma 1.

In the case of $H(Z)$ uniformly distributed on $(0, Z^M)$, the first-order condition for t given by (16) results in a cubic equation. The cubic can be written in general form as $at^3 + bt^2 + ct + d = 0$, where the coefficients a, b, c, and d are functions of the parameters of the model and the threshold $\bar{Z}$. An examination of the discriminant of the cubic in the numerical simulations reported in Table 3 in section IV, in the case where $H(Z)$ is the uniform distribution and a separating equilibrium is observed, confirms the existence of a unique real root. This observation enables us to write an explicit solution for the optimal turnover tax rate in terms of the threshold.

Proposition 3. *If $H(Z)$ is uniformly distributed, then the welfare-maximizing turnover tax rate in a separating equilibrium, for a given threshold $\bar{Z}$, is determined by a cubic equation. If its discriminant is positive, there exists a unique real root given by*

$$t = \lambda - \alpha + S + T \tag{17}$$

where $S = \sqrt[3]{R + \sqrt{Q^3 + R^2}}$, $T = \sqrt[3]{R - \sqrt{Q^3 + R^2}}$, $Q = \frac{3ac - b^2}{9a^2}$, $R = \frac{9abc - 27a^2d - 2b^3}{54a^3}$, and the coefficients a, b, c, and d are defined by

- $a = -M_0$
- $b = 3M_0(\lambda - \alpha)$
- $c = -3(\lambda - \alpha)^2 M_0 - M_1 - M_2$
- $d = M_0(\lambda - \alpha)^3 + M_1(\lambda - \alpha) + M_3$
- $M_0 = \frac{(\delta - 1)\bar{Z}}{2}$
- $M_1 = \Gamma'^2 + \delta A'\Gamma' - \frac{\delta - 1}{2}\Gamma'^2$
- $M_2 = (\delta - 1)\Gamma'^2$
- $M_3 = (\alpha - C)\Gamma'^2$

The discriminant of the cubic equation is given by $\Delta = Q^3 + R^2$.

The formula provides guidance for selecting the turnover tax rate in a commonly encountered policy situation, where the threshold is pegged a priori to the VAT threshold for reasons of administrative simplicity. The simultaneous solutions for (17) and (15) can be readily computed with spreadsheets.[14]

[14] Equation (15) could be combined with (17) to write the optimal threshold in closed-form, but the expression would be unintelligible.

Lemma 3. *If $H(Z)$ is uniformly distributed, then the optimal turnover tax rate approaches the difference between the marginal cost of production in the informal sector and the presumptive regime, as the cost of compliance in the presumptive tax regime approaches zero ($\lim_{\Gamma' \to 0} t^* \to \lambda - \alpha$).*

The lemma shows clearly how the existence of an informal sector constrains the government's choice of the turnover tax rate. We now turn to comparative statics analysis for further insights.

1. Comparative statics for a separating equilibrium

We assume an interior optimum for $\bar{Z}$ and t exists and consider the effects of small changes in the parameter values on the optimal policy. The parameters of interest in our comparative statics analysis are λ, Γ' and Γ, t^c, and α on the optimal values of $\bar{Z}$ and t. The effects of δ are ambiguous. We first provide the *partial* effects of parameter changes on the optimal tax rate t, holding $\bar{Z}$ fixed, and the optimal threshold $\bar{Z}$, holding t fixed. These partial effects show the direct impacts of the parameter changes on each policy variable while abstracting from feedback between the policy variables. The results can be useful for understanding the direction of optimal policy reform when only one policy variable is being considered for reform, such as a change in the desired threshold when the tax rate is not up for discussion, or vice versa. Then we present the *full* comparative statics, which can guide an overall reform in the presumptive tax regime.

Lemma 4. *In the case of a separating equilibrium, the following are the partial effects of parameter changes on the optimal threshold, holding the turnover tax rate fixed, if the derivative of the density function $h'(Z)$ is non-negative at the optimal threshold.*

1. *The optimal threshold is unaffected by the marginal cost of informal production.*
2. *The optimal threshold is decreasing in the profit tax rate.*
3. *The optimal threshold is decreasing (increasing) in the fixed cost of complying with the presumptive (regular) regime.*
4. *The optimal threshold is decreasing in the marginal cost of producing in the presumptive regime.*

Note that the uniform distribution satisfies the requirement in the proposition, since $h'(Z) = 0$ for all Z. It is a sufficient condition, but not a necessary one, for the comparative statics analysis. Lemma 4 implies that any change in the optimal threshold as a result of changes in the marginal cost of informality would only arise indirectly via changes in the turnover tax rate.

Lemma 5. *In the case of a separating equilibrium, the following are the partial effects of parameter changes on the optimal turnover tax rate, holding the threshold fixed, if the derivative of the density function $h'(Z)$ is non-negative at the point of indifference between informality and the presumptive regime (Z^{IP}).*

1. *The optimal turnover tax rate is increasing in the marginal cost of informal production.*
2. *The optimal turnover tax rate is unaffected by the profit tax rate.*
3. *The optimal turnover tax rate is decreasing in the fixed cost of complying with the presumptive regime, but independent of the fixed cost of the regular regime.*
4. *The optimal turnover tax rate is decreasing in the marginal cost of producing in the presumptive regime.*

Again, a uniform distribution for $H(Z)$ is a sufficient condition to determine the signs of the derivatives. Lemma 5 highlights how, holding the threshold constant, the optimal turnover tax rate increases as informality becomes less relatively attractive (due to higher λ or lower Γ'); and a change in the profit tax rate has no effect on the optimal turnover tax rate because concern over bunching does not come into play in a separating equilibrium.

The full comparative statics analysis is given next, assuming the first- and second-order conditions for a maximum are satisfied.

Proposition 4. *In the case of a separating equilibrium, the following are the full effects of parameter changes on the optimal threshold, if the derivative of the density function $h'(Z)$ is non-negative at the optimal threshold and at Z^{IP}.*

1. *The optimal threshold and tax rate are both increasing in the marginal cost of informal production.*
2. *The optimal threshold and tax rate are both decreasing in the profit tax rate.*
3. *The optimal threshold and tax rate are both decreasing (increasing) in the fixed cost of complying with the presumptive (regular) regime.*
4. *The optimal threshold and tax rate are both decreasing in the marginal cost of producing in the presumptive regime.*

The results are explained as follows. An increase in the marginal cost of informal production makes the informal sector less attractive, allowing the government to increase the tax rate in the presumptive regime. Although the cost of informal production has no direct effect on the optimal threshold, the higher tax rate itself induces an optimal raising of the threshold to increase the number of firms in the presumptive regime. An implication of Proposition 4 is that countries with rampant informal sector activity, which can be interpreted in the model as economies with a low λ, should set a relatively small threshold *and* a low turnover tax rate, provided that the equilibrium continues to be characterized by an absence of bunching.

Countries with high profit tax rates also should set lower levels of t and $\bar{Z}$. An increase in the profit tax rate makes the regular regime more lucrative for the government, inducing a lowering of the threshold. While the higher profit tax rate has no direct effect on the optimal turnover tax rate, the lower threshold implies that the presumptive regime now contains fewer firms, which in turn reduces the revenue sacrifice from marginally lowering t to attract more firms from the informal sector. A higher marginal cost of producing in the presumptive regime would induce the government to lower the turnover tax rate to attract informal firms by reducing the size of the tax-disincentive to formalizing. The optimal threshold is thereby reduced to reflect the weakened revenue-generating capacity of the presumptive regime.

An increase in the fixed compliance cost of the presumptive regime renders the informal sector relatively more attractive. This induces a lower turnover tax rate to counter the movement of firms out of the presumptive tax regime, which in turn favors lowering the threshold. A higher compliance cost or administrative cost in the regular regime induces an expansion of the threshold to reduce the number of firms paying profit tax, which in turn favors raising the turnover tax rate because the proportion of firms in that regime has increased. As for the fixed administrative cost of the presumptive regime, the comparative statics results are ambiguous because there are conflicting forces. On the one hand, higher A' diminishes the net tax benefit of attracting firms away from the informal sector, thereby militating in favor of a higher turnover tax rate. On the other hand, higher A' means there is a cost saving from decreasing the threshold, which itself induces a lower turnover tax rate. The effect of a change in the marginal social value of public funds is also ambiguous. This can already be seen from (15), where the sign of the direct effect of changing δ on the optimal threshold for a given t depends on the levels of the parameters and the policy variables.

The various effects are weighted by the mass of firms affected by the changes. It is for this reason that the condition on $h'(Z^{IP})$ is used to identify the net effects, although it is not a necessary condition for the results. In a separating equilibrium, optimal policy is determined solely by the margin between the presumptive regime and the informal sector, because firms are not induced to bunch below the threshold. In contrast, in a bunching equilibrium, the government's turnover tax policy will need to trade off the margins between the presumptive regime and informality and between the presumptive and regular regimes.

B. Welfare in the bunching equilibrium: $\bar{Z} < \hat{Z}$

We now turn to constructing the welfare function in the case of the bunching equilibrium partition described in Lemma 2.

Total profit is then,

$$\begin{aligned}\Pi(\bar{Z},t) = &\int_0^{Z^{IP}} \pi^I h(Z)dZ + \int_{Z^{IP}}^{\bar{Z}} \pi^P h(Z)dZ \\ &+ \int_{\bar{Z}}^{\hat{Z}} \pi^A h(Z)dZ + \int_{\hat{Z}}^{Z^M} \pi^R h(Z)dZ\end{aligned} \tag{18}$$

while the total net tax revenue is

$$\begin{aligned}G(\bar{Z},t) = &\int_{Z^{IP}}^{\bar{Z}} [tZ - A']h(Z)dZ + \int_{\bar{Z}}^{\hat{Z}} [t\bar{Z} - A']h(Z)dZ \\ &+ \int_{\hat{Z}}^{Z^M} [t^c(1-C)Z - A]h(Z)dZ\end{aligned} \tag{19}$$

Social welfare in the bunching partition is then given by

$$\begin{aligned}SW(\bar{Z},t) = &\int_0^{Z^{IP}} \pi^I h(Z)dZ + \int_{Z^{IP}}^{\bar{Z}} \pi^P h(Z)dZ + \int_{\bar{Z}}^{\hat{Z}} \pi^A h(Z)dZ \\ &+ \int_{\hat{Z}}^{\infty} \pi^R h(Z)dZ + \delta\{\int_{Z^{IP}}^{\bar{Z}} (tZ - A')h(Z)dZ \\ &+ \int_{\bar{Z}}^{\hat{Z}} (t\bar{Z} - A')h(Z)dZ + \int_{\hat{Z}}^{Z^M} [t^c(1-C)Z - A]h(Z)dZ\}\end{aligned} \tag{20}$$

The first-order condition with respect to $\bar{Z}$ can be rearranged to obtain the following optimality condition.

Proposition 5. *For given tax rates t and t^c, the optimal threshold in a bunching equilibrium is given by*

$$\begin{aligned}&[(1-t-C-\alpha)+\delta t][H(\hat{Z})-H(\bar{Z})] + \delta(t\bar{Z} - A')h(\hat{Z})\frac{d\hat{Z}}{d\bar{Z}} \\ =&\delta[t^c(1-C)\hat{Z} - A]h(\hat{Z})\frac{d\hat{Z}}{d\bar{Z}}\end{aligned} \tag{21}$$

The left-hand side of (21) is the marginal benefit from a small increase in the threshold, while the right-hand side is the marginal cost. As the threshold increases, the already existing adjusters would increase their sales and bunch just below the new threshold, generating marginal gains in profit and tax revenue in the presumptive regime. These are represented by the first term in (21). In addition, some firms that used to stay unconstrained in the regular regime now become bunchers ($d\hat{Z}/d\bar{Z} > 0$) – their move creates a net increase in tax revenue from the presumptive regime, the social value of which is the second term on the left-hand side. On the right-hand side of (21) is the social value of the net tax revenue loss from the firms leaving the regular regime.

An expression for the optimal threshold under bunching can be derived, for given tax rates t and t^c, in the case that $H(Z)$ is the uniform distribution. Then (21) can be simplified to give

$$\bar{Z}^* = \frac{(1-t-C-\alpha)[(\Gamma+\delta A)-(\Gamma'+\delta A')]+[\delta t-\delta(t^c/(1-t^c))(1-t-C-\alpha)](\Gamma-\Gamma')}{[(1-t-C-\alpha)+\delta t][(t+\alpha)-t^c(1-C)]+\delta(t^c/(1-t^c))(1-t-C-\alpha)^2-\delta t(1-t-C-\alpha)} \tag{22}$$

This optimal threshold in the case of bunching provides an analogue to the expression in (15) for the case of a separating equilibrium (though in the latter expression for the optimal threshold for a given turnover tax rate, it was not assumed that $H(Z)$ is uniform).

Turning to the optimal turnover tax rate in a bunching equilibrium, the first-order condition for optimizing welfare with respect to t can be written as follows.

Proposition 6. *For a given threshold $\bar{Z}$ and profit tax rate t^c, the optimal turnover tax rate in a bunching equilibrium is given by*

$$\begin{aligned}&(\delta-1)\int_{Z^{IP}}^{\bar{Z}} Zh(Z)dZ+(\delta-1)\int_{\bar{Z}}^{\hat{Z}} \bar{Z}h(Z)dZ-\delta[t^c(1-C)\hat{Z}-A]h(\hat{Z})\frac{d\hat{Z}}{dt}\\ &=-\delta(t\bar{Z}-A')h(\hat{Z})\frac{d\hat{Z}}{dt}+\delta(tZ^{IP}-A')h(Z^{IP})\frac{dZ^{IP}}{dt}\end{aligned} \tag{23}$$

In addition to the terms present in the first-order condition for t already seen previously in (16) for the case of a separating equilibrium, the expression (23) contains three new terms. The two new terms on the left-hand side (i.e., the second and third ones) represent the social value of the gain in revenues collected from existing bunchers as a result of the increased tax rate and the new tax revenues collected from the former bunchers who have switched to become regulars, due to the tax rate increase, as $d\hat{Z}/dt < 0$. On the right-hand side, the new term (i.e., the first one) corresponds to the social value of the revenue loss in the presumptive regime, due to fewer bunchers. The first-order condition for the turnover tax rate is clearly more complex when there is bunching behavior, than in its absence.

In the case where $H(Z)$ is the uniform distribution, (23) can be written, for a given $\bar{Z}$, as a quartic equation in t. If the discriminant of the quartic equation is negative, as is the case in the simulations (see section IV), then there are two distinct real roots (and two imaginary roots) (Irving, 2013, Theorem 6.5). Since the real roots are solutions to the first-order condition for t, one root is welfare-maximizing and the other is welfare-minimizing. The following proposition provides an expression for the optimal turnover tax rate based on root that maximizes social welfare.

Proposition 7. *If $H(Z)$ is uniformly distributed, then the welfare-maximizing turnover tax rate in a bunching equilibrium, for a given threshold $\bar{Z}$, is determined by a quartic equation. If its discriminant is negative, the optimal turnover tax rate is given by*

$$t=(1/4)(\lambda-\alpha)^3+(3/4)(\lambda-\alpha)+J-\frac{1}{2}\sqrt{-4J^2-2j-\frac{k}{J}} \tag{24}$$

where

- $j = \frac{8ac-3b^2}{8a^2}$; $k = \frac{b^3-4abc+8a^2d}{8a^3}$
- $J = \frac{1}{2}\sqrt{-\frac{2}{3}j + \frac{1}{3a}(K + \frac{\Delta_0}{K})}$; $K = \sqrt[3]{\frac{\Delta_1 + \sqrt{\Delta_1^2 - 4\Delta_0^3}}{2}}$
- $\Delta_0 = c^2 - 3bd + 12ae$; $\Delta_1 = 2c^3 - 9bcd + 27b^2e + 27ad^2 - 72ace$
- $a = -Q$; $b = (\lambda-\alpha)^3Q + 3(\lambda-\alpha)Q - M$; $c = -3(\lambda-\alpha)^2Q + 3(\lambda-\alpha)M + N$
- $d = (\lambda-\alpha)^3Q - 3(\lambda-\alpha)^2M + S$; $e = (\lambda-\alpha)^3M + V$
- $S = -\frac{(\delta+1)}{2}\Gamma'^2 - \delta\Gamma'A' - \delta\Gamma'^2$; $V = [\delta\Gamma'A' - \frac{\delta-1}{2}\Gamma'^2](\lambda-\alpha) + [\omega(1-\lambda-C) - (1-C-\alpha)]\Gamma'^2$
- $M = -\frac{\delta-1}{2}\bar{Z}^2 - \frac{\delta(A-A')\bar{Z}}{(1-t^c)(1-C)} + \frac{(\delta-1)\bar{Z}[(1-C-\alpha)\bar{Z}+\Gamma-\Gamma']}{(1-t^c)(1-C)} + \frac{\delta t^c(1-C)\bar{Z}[(1-C-\alpha)\bar{Z}+\Gamma-\Gamma']}{[(1-t^c)(1-C)]^2}$
- $N = -\Gamma'$; $Q = -\frac{(2\delta-1)\bar{Z}^2}{(1-t^c)(1-C)} - \frac{\delta t^c(1-C)\bar{Z}^2}{[(1-t^c)(1-C)]^2}$

The discriminant of the quartic equation is given by $\Delta' = -({\Delta_1}^2 - 4{\Delta_0}^3)/27$.

Further, with a uniform distribution for potential sales, if the first-order condition for $\bar{Z}$ holds, then there can exist only one interior solution for t. An interior solution for t means it is non-negative and generates an equilibrium partition that is consistent with Lemma 2.[15] Moreover, given the solution for t, (22) ensures a unique value of $\bar{Z}^*$. The result is stated as a lemma.

Lemma 6. *Suppose potential sales follow the uniform distribution. If an interior solution for the policy variables exists (i.e., the solutions to the first-order conditions for t and $\bar{Z}$ are non-negative and satisfy the partition for a bunching equilibrium defined by Lemma 2), then it is the unique interior solution.*

1. Comparative statics for a bunching equilibrium

The comparative statics for the bunching case are more complicated than for the separating case and some effects of parameter changes on the optimal turnover threshold or tax rate cannot be signed a priori. We assume that the first- and second-order conditions for a welfare maximum are satisfied. We first consider the direct or partial effects of small parameter changes before considering the full comparative statics.

Lemma 7. *The following are the partial effects of parameter changes on the optimal threshold, holding the turnover tax rate fixed, in a bunching equilibrium, if $h'(Z)$ is non-negative at $\bar{Z}$, Z^{IP} and $\hat{Z}$.*

[15] A interior solution for t requires $Z^{IP} > 0$, i.e., $\lambda - t - \alpha > 0$. Consequently, the 'second' real root for the quartic equation in t, discussed above in the context of Proposition 7, violates the condition $\lambda - t - \alpha > 0$ and minimizes welfare.

1. *The optimal threshold is independent of the marginal cost of informal production.*

2. *The optimal threshold is decreasing in the profit tax rate, if $t^c(1-C-\alpha) > \lambda - \alpha$.*

The partial comparative statics effects of the fixed compliance costs and the additional marginal cost of production in the presumptive regime are ambiguous.

Lemma 8. *The following are the partial effects of parameter changes on the optimal turnover tax rate, holding the threshold fixed, in a bunching equilibrium, if $h'(Z)$ is non-negative at $\bar{Z}$, Z^{IP} and $\hat{Z}$.*

1. *The optimal turnover tax rate is increasing in the marginal cost of informal production.*

2. *The optimal turnover tax rate is increasing in the profit tax rate, if $t^c(1-C-\alpha) > \lambda - \alpha$.*

3. *The optimal turnover tax rate is decreasing in the marginal cost of producing in the presumptive regime.*

The effects of the fixed compliance costs are ambiguous.

The full comparative statics effects of the parameters have mostly ambiguous signs. The next proposition provides sufficient conditions to determine the effect of changes in the marginal cost of informal production on the optimal policy.

Proposition 8. *In a bunching equilibrium, if $h'(Z)$ is non-negative at $\bar{Z}$, Z^{IP} and $\hat{Z}$, and if $t^c(1-C-\alpha) > \lambda - \alpha$, then the optimal turnover tax rate and threshold are both increasing in the marginal cost of informal production.*

The immediate effect of an increase in the marginal cost of informal sector production is to induce a higher turnover tax rate. This, in turn, encourages a higher threshold. Although the higher threshold itself impacts the proportion of firms bunching, generating further adjustments, the overall effect of a higher λ can be determined under the stated conditions; the last condition is used to ensure that the marginal shift in the cutoff for bunching generates a net gain in tax revenues.

In the case of an increase in the profit tax rate, the revenue-generating power of the regular regime is improved, which has a mechanical effect – i.e., ignoring behavioral changes – of reducing the optimal threshold, which has a knock-on effect of lowering the optimal turnover tax rate because of the reduced revenue-generating power of the presumptive regime. At the same time, however, a higher profit tax rate encourages more firms to bunch, which creates an incentive for the government to curb bunching by raising the turnover tax rate. The net effects of these changes on bunching behavior will depend on the levels of the threshold and turnover tax rate and on the mass of firms at each margin of behavioral response. In general, it is not possible to sign the final effects of the change in t^c, nor changes in the fixed compliance costs of each regime, even if the size distribution of firms is assumed to be uniform.

The optimal policy responses to parameter changes in a bunching equilibrium are clearly more complicated than in the case of a separating equilibrium. In section IV, the impacts of changing the parameters of the model are explored further with the aid of numerical simulations.

IV. Numerical solutions for the optimal policies

Numerical simulations are used to identify the welfare maximizing design of the turnover tax regime. Optimal policy is determined by direct search optimization based on random numbers. We take 4000 draws from a given frequency distribution and determine the optimal behavior of each firm (e.g., informality, unconstrained in the presumptive regime, bunching, or unconstrained in the regular regime), taking into account the sales constraint imposed by the threshold. All possible types of partitions of the cutoffs are considered. We calculate total welfare by aggregating the individual contributions of each firm. Both the 'separating' and 'bunching' cases emerge as equilibria, depending on the value of the (exogenous) profit tax rate.

A. Calibration

The parameter values are based on a synthesis of survey results for different countries that have turnover tax regimes. Thus the results do not represent any particular country. However, most of the evidence used for the calibration is based on countries in Sub-Saharan Africa.

1. Sales

In most countries, the presumptive regime is restricted to sole proprietorships, partnerships, and close corporations, while excluding publicly traded corporations. Thus, the relevant sizes of companies for the purposes of calibrating the model are micro, small, and medium-sized enterprises. LaPorta and Shleifer (2008) provide statistics on average sales from different World Bank surveys in low-income countries. The Micro survey targets areas of a country where there is a high concentration of businesses with fewer than five employees, but randomly selects all establishments in the area. In this survey, which includes firms both unregistered and registered with the central government, the average value of sales is about $51,000 in 2006. The Enterprise survey drops firms with fewer than five employees and includes many firms with more than 100 employees. The average size of firms in the Enterprise survey for the same countries and year as the Micro survey is about $1.1 million. The overall weighted-average sales across the two World Bank surveys is about $815,000. We calibrate a lognormal distribution to approximately this mean.

2. Compliance costs

Taxpayer compliance costs vary widely across countries due to differences in tax designs and administrative practices, and in the salaries used to value in-house costs of compliance. Sapiei et al. (2014) estimates the corporate income tax compliance cost as between 0.05% and 15% of taxable turnover in developing and transition economies. We base our estimates of the cost of tax compliance on three surveys that each permit separate calculations for the profit tax and the turnover tax.

Smulders et al. (2012) estimated average taxpayer costs for the small business sector in South Africa from a survey in 2011. Enterprises with turnovers below ZAR 1 million (about $102,445 in 2022)[16] could opt for paying a turnover tax (with a progressive marginal rate structure) in lieu of profit tax and VAT. The average in-house cost of complying with the profit tax was ZAR 15,822 ($1,621 in 2022), while for the turnover tax it was ZAR 14,030 ($1,473 in 2022).[17] The average external cost of compliance for all taxes combined was estimated at ZAR 9,882 ($1,012 in 2022), but smaller businesses (turnover of less than ZAR 1 million) tended not to incur these expenses. According to Coolidge and Ilic (2009), most of the outsourced tax compliance-related activities in South Africa are associated with the income tax returns. Hence, the total compliance costs for profit tax is obtained by adding the full external cost to the in-house cost.[18]

World Bank Group (2017) estimated average company tax compliance costs in Kyrgyzstan from survey results for the years 2012, 2014, and 2016. Of the sampled firms, about 60 percent had turnovers below KGS 1 million ($18,223 in 2022), and 80 percent of these paid the turnover tax. The average company cost of tax compliance (in-house plus external) reported over the three surveys was KGS 18,716 ($307 in 2022) for the general regime without VAT, and KGS 13,343 ($220 in 2022) for the turnover tax regime. The assessments represent lower bounds against the actual costs of tax accounting, because the in-house time-costs were evaluated by survey respondents using the formal salary of their employees, whereas the actual level of labor costs exceed the formally declared salary.

CIAT et al. (2015) provides estimates of tax compliance costs for small and medium enterprises in Brazil based on a 2012 survey. It distinguishes between taxpayers of corporate income tax and payers of the simplified system, Simples Nacional, which is an optional tax regime for micro and small enterprises. It combines several regular taxes and applies progressive marginal rates to gross revenues rather than business income. The study finds average total compliance costs to be BRL 19,770 ($5,908 in 2022) for payers of the tax on actual income and BRL 7,563 ($2,260 in 2022) for payers of Simples. However, the methodology does not disentangle tax costs from total costs charged by accounting firms for their services.

[16] The USD figure applies a cumulative inflation factor from the survey year to 2022 and converts from local currency units to US dollars at the exchange rate on January 1, 2022. A similar adjustment is applied to the other local currency figures stated below.

[17] The reported averages are after trimming the 5% extremes of the survey responses.

[18] Because South Africa's turnover tax has a progressive rate structure, the cost of tax complying with it may be relatively high, compared to the more typical proportional turnover tax.

The authors note that the federation of Brazilian accounting firms attributes only one-quarter of the total accounting costs to strictly tax compliance activities. This suggests an average total compliance cost of $1,477 for the profit tax and $565 for the turnover tax in Brazil.

Taxpayer compliance costs are likely to scale with the size of the business. Keen and Mintz (2004) assume that variable compliance costs for VAT are 0.1 percent of turnover. Similarly, based on information in the surveys on the distribution of firm sizes, we deduct variable costs equal to 0.1 percent of the average turnover of firms in each tax regime to derive the estimate of the fixed cost of compliance. Based on our assessment of the available evidence, we set the fixed compliance cost to $1,400 for the profit tax and $300 for the turnover tax. In the simulations, we assume that variable compliance costs are 0.1 percent of each firm's turnover.

3. Administration costs

The cost of tax administration is based on a study on Slovakia (Nemec et al. 2015), which calculates the cost for the presumptive form of PIT at 2.25 percent of the presumptive PIT revenues.[19] and the cost of administering CIT at 1.65 percent of CIT revenues, while noting that Slovakia's cost is considerably higher than in the Czech Republic where it is 0.67 percent of CIT revenues. We calibrate the fixed cost of tax administration to achieve, in the (base case) simulations, approximately the latter figure for profit tax revenues and the presumptive PIT figure for turnover tax revenues.

4. Cost of informal production

The additional marginal cost associated with producing informally, λ, is set at 0.075 to fix the relative size of the informal sector at realistic values. Based on surveys of enterprises conducted by the World Bank's Enterprise Analysis Unit in nine cities across four African countries, Aga et al. (2020) report within-country average rates of informal businesses ranging from 66 percent to 90 percent, and a cross-country average of 77.25. In our base case, about 67 percent of firms are in the informal, untaxed, sector.[20] The additional marginal cost associated with being in the presumptive regime, α, is set to 0.025, while C is 0.85, implying a profit margin of 15 percent for firms in the regular regime.[21]

[19] Entrepreneurs in Slovakia can deduct 60% of their revenues as a presumptive cost deduction, making the presumptive PIT essentially a tax on turnover.

[20] Under the classification of informality used in the surveys reported by Aga et al. (2020), which is based on the non-registration of businesses (i.e., informal from a legal standpoint), some firms will be considered informal even if they pay taxes (i.e., are not fiscally informal). In our simulation results, a portion of firms paying the presumptive tax may be regarded as informal in the legalistic sense of Aga et al. (2020), in addition to the purely informal, untaxed, segment of firms.

[21] The figure for α is consistent with evidence in Benjamin and Mbaye (2012) on total factor productivity (TFP) differences between formal firms and what they refer to as 'large informal' firms. The two categories are similar, except that the latter firms under-report their sales to qualify for the presumptive tax regime. They find that the formal firms have higher total factor productivity (TFP) than the large informal firms but that the difference is small. In our model, TFP corresponds to the inverse of the unit cost of production.

Table 2 summarizes the base case calibration of the model.

Table 2. Baseline parameter values

Description	Notation	Value
Profit margin in regular regime	$(1-C)$	15%
Profit margin in presumptive regime	$(1-C-\alpha)$	12.50%
Profit margin in informal sector	$(1-C-\lambda)$	7.50%
Average potential sales	$E(Z)$	$815,000
Fixed compliance cost in regular regime	Γ	$1,400
Fixed compliance cost in presumptive regime	Γ'	$300
Fixed administrative cost (per firm) in regular regime	A	$900
Fixed administrative cost (per firm) in presumptive regime	A'	$25
Mean of lognormal distribution	μ	8
Standard deviation of lognormal distribution	σ	3.35

B. Results

Table 3 shows the optimal turnover tax rate and threshold for different values of the profit tax rate, at the base case calibration. It also provides sensitivity analysis using various alternative assumptions. The table identifies the type of equilibrium – separating or bunching – associated with the policy optimum.

In the base case, the optimal threshold is near $80,000 and the turnover tax rate is about 2.5 percent when the profit tax rate is 25 percent. Bunching equilibria are observed at profit tax rates of 25 percent or higher. When the base case equilibria exhibit bunching, the optimal turnover tax rate increases with the profit tax rate. Intuitively, a higher profit tax rate creates a greater incentive for bunching, which is dampened but not eliminated by raising the turnover tax rate. In contrast, when there is a separating equilibrium, the optimal turnover tax rate decreases with the profit tax rate as a consequence of the reduced optimal threshold. The optimal thresholds are decreasing with the profit tax rate throughout the range of profit tax rates in the base case.

Case 1 repeats the base case except with the fixed compliance cost being deductible for profit tax.[22] An interesting observation, illustrated for the base case and Case 1 in Table 3, is that, at

[22] In the base case and in Cases 2 to 8, the fixed compliance cost is treated as non-deductible for profit tax. Making it deductible would somewhat obscure the effect of changing the profit tax rate, because the profit tax rate has direct effects on bunching but it also alters after-tax fixed costs of compliance. Furthermore, the South Africa survey of micro and small enterprises shows that owners do much of the in-house accounting. If owners are unincorporated or are incorporated but do not receive wages for their time spent on tax accounting, then the compliance costs are non-deductible.

the higher end of profit tax rates, the effective tax rate (defined as the ratio of taxes to before-tax profit) in the regular regime tends to exceed the effective tax rate in the presumptive regime when evaluated at the turnover threshold. (The effective tax rate in the regular regime equals the statutory tax rate t^c if Γ is tax deductible, and exceeds it if Γ is non-deductible.) To discourage bunching, or, in other words, to encourage firms to remain in the regular regime, it is often recommended to make the effective tax rate equal or higher in the presumptive regime, compared to the regular regime. However, this omits consideration of tax compliance costs, the reduction of which is the very purpose of the presumptive regime. Furthermore, the concern over possible bunching ignores the other relevant margin, which is the incentive to be informal when the turnover tax rate is high.

The base case in Table 3 is supplemented by Table 4, which provides additional information on the various cutoffs, the proportion of businesses operating informally, and the percentage point reduction in informality relative to a simulation with the same parameter values but which excludes the turnover tax regime (equivalent to forcing the threshold to be zero). About two-thirds of the businesses operate in the informal sector in equilibrium. In the absence of a presumptive tax regime, about 78 percent of firms are in the informal sector (these are firms with sales below $Z^{IP} = \$11,583$). Thus, the introduction of the presumptive regime reduces the amount of informality by about 12 percentage points. The proportion of firms bunching just under the threshold is below 1 percent, which is realistic.[23]

Cases 2 to 8 consider modifications of parameters values. If the fixed cost of complying with the regular regime decreases (Case 2), then the optimal threshold and tax rate fall relative to the base case. If the cost of producing informally rises (Case 3), the optimal turnover tax rate and threshold both increase. The higher tax rate exploits the diminished attractiveness of the informal sector to generate greater revenues, and the higher threshold, in turn, reflects the improved revenue-productivity of the presumptive regime. Assigning a higher marginal cost to producing in the presumptive regime (Case 4) induces a lower turnover tax rate to compensate and coax firms out of the informal sector. The lower tax rate is accompanied by a lower threshold because of the diminished revenue-generating capacity of the turnover tax. Case 5 replaces the lognormal distribution for potential sales with a uniform distribution with the same mean. Compared to the base case, the optimal turnover tax rate is about 3/4 of a percentage point higher and the threshold is about \$10,000 higher. The comparatively lower tax rate and threshold under the lognormal distribution reflects its positive skew which means there are more firms with low sales, for whom the informal sector is attractive. In Case 6, the lognormal distribution is again used but with a mean that is higher than in the base case.[24] The effect of the higher mean on the turnover tax rate is minimal, while the threshold slightly

[23] Bruhn and Loeprick (2016) found that about 3 percent of firms filing under the new turnover tax regime for small businesses in Georgia reported sales just below the threshold a year after it was introduced, and close to none reported sales just above the threshold. However, bunching had already been observed at the threshold prior to the introduction of the new small business regime, attributable to the effects of the VAT threshold. By comparing the distribution of sales before and after the introduction of the small business tax regime, the authors estimated that about 0.04 percent of the 30,000 previously existing firms had reacted strategically to the new regime by declaring revenues just below the threshold.

[24] A higher mean of the logarithm of sales raises both the mean and variance of the level of sales. The expected value of potential sales in Case 6 is double that of the base case.

increases at the lower values of the profit tax rate but slightly decreases at the higher profit tax rates, where bunching behavior is observed.[25] In Case 7, the fixed cost of administering the presumptive regime is quadrupled, resulting in a lower threshold but a higher turnover tax rate. Finally, Case 8 assumes a higher marginal value of public funds. This increases the optimal threshold and tax rate in the separating equilibria; in the bunching equilibria, the optimal threshold falls but the effect on the turnover tax rate is unsystematic.

The sensitivity analysis provided by Cases 2 to 8 span a wide spectrum of plausible parameter values of the model for countries with large informal sectors. Overall, the optimal turnover tax rate is found to be around 2.5% across most scenarios (with a low of 2.1 percent and a high of 3.6 percent), while the threshold separating the turnover tax regime from profit tax lies for the most part between \$65,000 and \$95,000 (with a low of \$53,000 and a high of \$112,000). The results show that the model provides realistic optimal policy advice: the optimal tax rates and thresholds in the base case and the sensitivity analysis are within the support of the distribution of observed policy choices (see Table 1). However, turnover tax regimes in many countries differ substantially from these values, suggesting opportunities for beneficial tax reforms.

[25] The effect of economic development on the optimal threshold arises only indirectly. On the one hand, the time spent by taxpayers to comply with the tax laws may be relatively greater in poor countries, due to lower rates of education. On the other hand, the financial or opportunity cost of each hour of compliance activities will tend to be smaller, given lower wages (productivity) in poorer countries. If the total money value of time spent on compliance and administrative activities is approximately constant as per capita GDP varies, then lower per capita GDP would translate to a higher relative cost of compliance and administrative activities and hence a higher turnover threshold relative to GDP. In the model, lower values for GDP per capita correspond to smaller values of the mean of the distribution of potential sales.

Table 3. Simulation Results

Profit Tax Rate	15%	20%	25%	30%	35%
Base case					
Optimal turnover threshold (thousand $)	95.8	85.4	78.9	70.1	66.1
Optimal turnover tax rate	2.81%	2.61%	2.41%	2.61%	2.87%
Effective tax rate in regular regime	16.62%	21.48%	28.35%	34.61%	40.75%
Effective tax rate in presumptive regime	23.06%	22.45%	19.88%	21.62%	23.83%
Type of equilibrium	Separating	Separating	Bunching	Bunching	Bunching
Case 1. Deductible compliance cost					
Optimal turnover threshold (thousand $)	82.7	72.2	63.1	54	46.9
Optimal turnover tax rate	2.41%	2.31%	2.31%	2.38%	2.49%
Effective tax rate in regular regime	15.00%	20.00%	25.00%	30.00%	35.00%
Effective tax rate in presumptive regime	19.86%	19.11%	19.21%	19.93%	20.99%
Type of equilibrium	Separating	Separating	Separating	Bunching	Bunching
Case 2. Lower cost of compliance in regular regime					
Optimal turnover threshold (thousand $)	84.9	75.3	70.3	57.4	54.8
Optimal turnover tax rate	2.61%	2.31%	2.31%	2.33%	2.73%
Type of equilibrium	Separating	Separating	Bunching	Bunching	Bunching
Case 3. Higher cost of informality					
Optimal turnover threshold (thousand $)	101.6	87.3	79.7	72	67.8
Optimal turnover tax rate	3.24%	2.81%	2.81%	2.88%	3.12%
Type of equilibrium	Separating	Separating	Bunching	Bunching	Bunching
Case 4. Higher extra marginal cost in the presumptive tax regime					
Optimal turnover threshold (thousand $)	73.8	69	61.8	57.7	53.4
Optimal turnover tax rate	2.23%	2.23%	2.10%	2.30%	2.35%
Type of equilibrium	Separating	Separating	Separating	Bunching	Bunching
Case 5. Uniform distribution					
Optimal turnover threshold (thousand $)	106.9	95.9	85.7	78.3	70.3
Optimal turnover tax rate	3.60%	3.47%	3.41%	3.34%	3.47%
Type of equilibrium	Separating	Separating	Separating	Bunching	Bunching
Case 6. Lognormal distribution with larger mean					
Optimal turnover threshold (thousand $)	96.4	88.1	78.2	68.7	65.8
Optimal turnover tax rate	2.87%	2.87%	2.59%	2.87%	2.87%
Type of equilibrium	Separating	Separating	Separating	Bunching	Bunching
Case 7. Higher cost of administrative cost in the presumptive tax regime					
Optimal turnover threshold (thousand $)	91.8	83.7	75.9	68.8	64
Optimal turnover tax rate	2.81%	2.81%	2.73%	2.74%	2.87%
Type of equilibrium	Separating	Separating	Separating	Bunching	Bunching
Case 8. Higher marginal value of public funds					
Optimal turnover threshold (thousand $)	112.2	96.5	80.8	69.3	53.2
Optimal turnover tax rate	3.00%	3.00%	2.81%	2.81%	2.81%
Type of equilibrium	Separating	Separating	Separating	Bunching	Bunching

Notes: In each alternative case, the parameters are identical to the base case, except for the parameter indicated: Case 1: Γ is deductible under profit tax; Case 2: Γ = 1200; Case 3: λ = 0.07; Case 4: $\alpha = 0.03$; Case 5: H(Z) uniformly distributed on (0, 1,600,000); Case 6: $\mu = 8.68$; Case 7: $A' = 100$; Case 8: $\delta = 1.5$.

Table 4. Detailed results for base case

Profit tax rate	**15%**	**20%**	**25%**	**30%**	**35%**
			With presumptive tax regime		
Optimal turnover tax rate	2.81%	2.61%	2.41%	2.61%	2.87%
$\bar{Z}$	95800	85400	78900	70100	66100
$\hat{Z}$	81435.45	79550.50	80542.31	76503.71	76568.51
Z^{IP}	13698.63	12552.30	11583.01	12552.30	14084.51
Z^{AI}	119773.60	108614.13	102146.80	88438.53	80872.40
Z^{IR}	26666.67	31111.11	37333.33	46666.67	62222.22
Total private income	4.1309E+08	3.8901E+08	3.6484E+08	3.4056E+08	3.1626E+08
Total tax revenue	7.2742E+07	9.6855E+07	1.2103E+08	1.4530E+08	1.6956E+08
Total welfare	5.0766E+08	5.1492E+08	5.2218E+08	5.2944E+08	5.3670E+08
Comparison between $\bar{Z}$ and $\hat{Z}$	$\bar{Z} > \hat{Z}$	$\bar{Z} > \hat{Z}$	$\bar{Z} < \hat{Z}$	$\bar{Z} < \hat{Z}$	$\bar{Z} < \hat{Z}$
Proportion of informal firms	68.28%	59.28%	66.55%	67.45%	68.50%
Percentage point reduction in informality	7.13%	17.66%	11.75%	12.31%	13.23%
			Without presumptive tax regime		
Sales cutoff for informality	34583.26	31186.73	37436.34	46676.87	62357.35
Proportion of informal firms	75.40%	76.93%	78.30%	79.76%	81.73%
Net tax revenue from regular regime	7.2131E+07	9.6472E+07	1.2078E+08	1.4503E+08	1.6917E+08

V. CONCLUSIONS

The emphasis of this paper is on how the informal sector both motivates and constrains the design of turnover-based presumptive income taxes. It contributes to a better comprehension of the interaction of presumptive and regular tax systems in the context of development, which is key to bringing about reforms that would permit an increase in tax revenue and a reduction in the prevalence of informal sector activities. We analyze the optimal design of a proportional tax on turnover applied to firms with sales below a threshold. Firms in the model make strategic choices for tax purposes, regarding their sales level and whether to produce formally or to evade taxes altogether by disappearing into the informal sector.

The analysis of the model generates formulas for the optimal threshold and the tax rate (when the size distribution of firms in uniformly distributed), while numerical simulations are provided to further guide policy choices. The formulas are implementable with spreadsheets, making them useful for policy analysts.[26] The results suggest an optimal threshold in the range of $65,000 to $95,000 and a turnover tax rate of around 2.5 percent for countries with large informal sectors. The optimal turnover tax induces about 12 percentage points of the enterprises to migrate from operating in the informal sector to registering for the turnover tax regime, compared to the situation where the turnover tax system is absent. We show that the effective tax rate on profit implied by the optimal turnover tax rate may be lower than the corresponding rate in the regular regime. This reflects the twin concerns driving the design of the presumptive tax regime: the risk of causing bunching below the threshold and the risk of encouraging firms to produce in the informal sector.

Setting the right turnover threshold and tax rate depends critically on the cost of tax compliance activities and the relative cost of producing informally. Both of these considerations are influenced by actions of the tax administration. Steps can be taken to reduce the cost of tax filing and tax payments to encourage firms to become fiscally formal, on the one hand, and campaigns to identify informal enterprises can increase the chance of detecting evasion, on the other hand.

Several caveats apply. First, tax evasion only occurs in the model by firms producing informally – that is, completely outside of the view of the tax authorities. In reality, some formal firms may under-report their sales in order to remain below the size threshold separating the standard income tax regime and the presumptive tax regime. Second, we abstract from a co-existing VAT threshold. Although it can be argued that economies of scope in taxpayer compliance exist when the VAT registration threshold coincides with the turnover tax threshold, Kanbur and Keen (2014) have shown that there are game-theoretic reasons for setting the two thresholds far apart. That is, setting one threshold above the other might induce firms to profitably expand their sales until they are just below the higher threshold; hence they have now crossed the lower threshold and pay more tax on that tax instrument. If the thresholds were identical, the same firms might choose to produce just below both thresholds to avoid

[26]The determination of whether the equilibrium is a separating one or a bunching one can be done using equation (9).

the higher tax burdens associated with crossing the common threshold for both tax instruments. Third, the simplifying assumption that the marginal cost penalty of informal production (λ) is fixed, rather than a rising function of output, likely underestimates an enterprise's marginal cost of medium and large scale informal production. Other things being the same, a rising marginal cost of operating informally can be expected to reduce the number and size of firms in the informal sector, which would translate into a higher optimal turnover tax rate. Fourth, the model assumes that the presumptive tax is obligatory for firms with sales below the threshold. In practice, countries decide whether small taxpayers can voluntarily opt for the regular regime. Finally, the model's general economic setting features several limitations. It assumes that registered and unregistered firms produce the same good and hence potential general equilibrium effects across sectors are abstracted from. It assumes further that the maximum supply of each firm is exogenous and therefore unaffected by the turnover tax rate. Future work could integrate these various considerations into the analysis.

Turnover taxes are a particularly simple form of small business taxation. More complicated alternatives include applying profit tax with simplified accounting rules (e.g., cash accounting), or allowing a presumptive cost deduction as a proportion of turnover. The latter case is essentially a turnover tax but it enables the application of the statutory profit tax rate. For near-subsistence level sellers, a fixed fee system is likely more suitable than a tax on turnover. A well-designed turnover tax should have a single, or very few, tax rates to avoid arbitrage by the reclassifications of activities and must be accompanied by anti-splitting rules. A proportional turnover tax rate will, however, result in horizontal tax inequity unless all firms have identical profit margins. Despite the drawbacks and risks associated with turnover tax regimes, they have proven to be an enduring component of tax policy in developing and transition economies. Several countries have abandoned turnover taxes for MSEs only to return to them a few years later (Engelschalk and Loeprick, 2016). Finally, it should be stressed that reducing the cost of tax compliance through presumptive taxation may help encourage formalization, but it cannot be successful without accompanying improvements in tax inspections and taxpayer education.

REFERENCES

Aga, A., Jolevski, M., and Muzi, S. (2020). Insights from enterprise surveys: Including the informal economy in policy responses to covid-19. *World Bank Blog*.

Benjamin, N. C. and Mbaye, A. A. (2012). The informal sector, productivity, and enforcement in west africa: A firm-level analysis. *Review of Development Economics*, 16(4):664–680.

Best, M. C., Brockmeyer, A., Kleven, H. J., Spinnewijn, J., and Waseem, M. (2015). Production versus revenue efficiency with limited tax capacity: theory and evidence from pakistan. *Journal of Political Economy*, 123(6):1311–1355.

Bird, R. M. and Zolt, E. M. (2008). Technology and taxation in developing countries: From hand to mouse. *SSRN Electronic Journal*.

Bruhn, M. and Loeprick, J. (2016). Small business tax policy and informality: evidence from georgia. *International Tax and Public Finance*, 23:834–853.

Bruhn, M. and McKenzie, D. (2014). Entry regulation and the formalization of microenterprises in developing countries. *The World Bank Research Observer*, 29(2):186–201.

CIAT, Receita Federal, and SEBRAE (2015). Measuring tax transaction costs in small and medium enterprises. *New York: United Nations*.

Coolidge, J. and Ilic, D. (2009). Tax compliance perceptions and formalization of small businesses in south Africa. *World Bank Policy Research Working Paper*, (4992).

de Mel, S., McKenzie, D., and Woodruff, C. (2010). Who are the microenterprise owners? evidence from sri lanka on tokman v. de soto. *International Differences in Entrepreneurship*, pages 63–88.

Dharmapala, D., Slemrod, J., and Wilson, J. D. (2011). Tax policy and the missing middle: Optimal tax remittance with firm-level administrative costs. *Journal of Public Economics*, 95(9-10):1036–1047.

Engelschalk, M. and Loeprick, J. (2016). The taxation of micro and small businesses in transition economies: Country experience of the introduction of special tax regimes. *Journal of Tax Administration*, 2(1):145–197.

Irving, R. (2020). *Beyond the quadratic formula*, volume 62. American Mathematical Soc.

Kanbur, R. and Keen, M. (2014). Thresholds, informality, and partitions of compliance. *International Tax and Public Finance*, 21(4):536–559.

Keen, M. and Mintz, J. (2004). The optimal threshold for a value-added tax. *Journal of Public Economics*, 88(3-4):559–576.

Kleven, H. J. and Waseem, M. (2013). Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from pakistan. *The Quarterly Journal of Economics*, 128(2):669–723.

LaPorta, R. and Shleifer, A. (2008). The unofficial economy and economic development. *Brookings Papers on Economic Activity*, 2008:275–352.

Liu, L., Lockwood, B., Almunia, M., and Tam, E. H. (2021). Vat notches, voluntary registration, and bunching: Theory and uk evidence. *Review of Economics and Statistics*, 103(1):151–164.

Nelson, D. (2008). *The Penguin dictionary of mathematics*. Penguin UK.

Nemec, J., Pompura, L., and Sagat, V. (2015). Administrative costs of taxation in slovakia. *European Financial and Accounting Journal*, 10(2):51–61.

Sapiei, N. S., Abdullah, M., and Sulaiman, N. A. (2014). Regressivity of the corporate taxpayers' compliance costs. *Procedia - Social and Behavioral Sciences*, 164:26–31.

Smulders, S., Sitglingh, M., Franzen, R., and Fletcher, L. (2012). Tax compliance costs for the small business sector in south africa: Establishing a baseline. *eJTR*, 10:184.

Thuronyi, V. T. (1996). *Tax Law Design and Drafting*. Washington DC: International Monetary Fund.

Waseem, M. (2018). Taxes, informality and income shifting: Evidence from a recent pakistani tax reform. *Journal of Public Economics*, 157:41–77.

World Bank Group (2007). *Designing a Tax System for Micro and Small Businesses: Guide for Practitioners*. Washington DC: International Finance Corporation.

World Bank Group (2017). *Assessment of Tax Compliance Costs for Businesses in the Kyrgyz Republic*. World Bank.

APPENDIX

Proof of Lemma 1

The proof follows immediately from the definitions of Z^{IP}, $\hat{Z}$, $\bar{Z}$, and Z^{IR}.

Proof of Lemma 2

The proof follows immediately from the definitions of Z^{IP}, $\hat{Z}$, $\bar{Z}$, and Z^{AI}.

Proof of Proposition 1

The derivative of the welfare function (for a separating equilibrium) (13) with respect to the threshold is given by

$$\begin{aligned}\frac{dSW}{d\bar{Z}} &= \pi^P(\bar{Z})h(\bar{Z}) - \pi^R(\bar{Z})h(\bar{Z}) \\ &+ \delta\{(t\bar{Z} - A')h(\bar{Z}) - [t^c(1-C)\bar{Z} - A]h(\bar{Z})\} \\ &= \{[(1-t-C-\alpha) + \delta t]\bar{Z} - \Gamma' - \delta A'\}h(\bar{Z}) \\ &- \{[(1-t^c)(1-C) + \delta t^c(1-C)]\bar{Z} - \Gamma - \delta A\}h(\bar{Z})\end{aligned} \tag{25}$$

Setting the last expression to zero, the first-order condition can be rearranged to generate the result for $\bar{Z}$.

Proof of Proposition 2

The derivative of the social welfare function (for a separating equilibrium) (13) with respect to the tax rate is

$$\begin{aligned}\frac{dSW}{dt} &= \pi^I(Z^{IP})h(Z^{IP})\frac{dZ^{IP}}{dt} + \int_{Z^{IP}}^{\bar{Z}}(-Z)h(Z)dZ - \pi^P(Z^{IP})h(Z^{IP})\frac{dZ^{IP}}{dt} \\ &+ \delta[\int_{Z^{IP}}^{\bar{Z}} Zh(Z)dZ - (tZ^{IP} - A')h(Z^{IP})\frac{dZ^{IP}}{dt}] \\ &= (\delta - 1)\int_{Z^{IP}}^{\bar{Z}} Zh(Z)dZ - \delta(tZ^{IP} - A')h(Z^{IP})\frac{dZ^{IP}}{dt}\end{aligned} \tag{26}$$

where the second equality uses $\pi^I(Z^{IP}) = \pi^P(Z^{IP})$ by the definition of Z^{IP}. Setting the derivative to zero, the first-order condition can be rearranged to yield the result.

Proof of Proposition 3

If $H(Z)$ is the uniform distribution, then after replacing Z^{IP} by (6) in (16) and rearranging terms, the equation becomes

$$(\delta-1)\bar{Z}^2 = \frac{[(\delta+1)\Gamma'^2+2\delta\Gamma'A']t-[2\delta\Gamma'A'-(\delta-1)\Gamma'^2](\lambda-\alpha)}{(\lambda-t-\alpha)^3} \tag{27}$$

and Equation (27) can be written as a cubic equation $at^3+bt^2+ct+d=0$ with the coefficients given in the proposition. The expression $\Delta = Q^3+R^2$ is the discriminant of the equation, where $Q=\frac{3ac-b^2}{9a^2}$ and $R=\frac{9abc-27a^2d-2b^3}{54a^3}$. If it is positive, then there is a unique real root for the cubic equation (Nelson, 2008), given by the expression in the proposition (Cardano's formula).

Proof of Lemma 3

As Γ' goes to 0 the parameters of the cubic function for t simplify to:

- $a=-(\delta-1)\bar{Z}^2$
- $b=3(\delta-1)\bar{Z}^2(\lambda-\alpha)$
- $c=-3(\delta-1)(\lambda-\alpha)^3\bar{Z}^2$
- $d=(\delta-1)(\lambda-\alpha)^3\bar{Z}^3$

Using these in the expressions for Q, R, it can be seen that S and T approach 0, hence $t=\lambda-\alpha$.

Proof of Lemma 4

Let $dSW/d\bar{Z}=0$ denote the first-order condition for maximizing the social welfare function (for a separating equilibrium) with respect to the threshold. Partiallu differentiating of the first-order condition and rearranging terms yields,

1. $\frac{d\bar{Z}}{d\lambda}|_t = -\frac{\partial^2 SW}{\partial\bar{Z}\partial\lambda}\Big/\frac{\partial^2 SW}{\partial\bar{Z}^2}=0$
2. $\frac{d\bar{Z}}{dt^c}|_t = -\frac{\partial^2 SW}{\partial\bar{Z}\partial t^c}\Big/\frac{\partial^2 SW}{\partial\bar{Z}^2}<0$
3. $\frac{d\bar{Z}}{d\Gamma'}|_t = -\frac{\partial^2 SW}{\partial\bar{Z}\partial\Gamma'}\Big/\frac{\partial^2 SW}{\partial\bar{Z}^2}<0$; $\frac{d\bar{Z}}{d\Gamma}|_t = -\frac{\partial^2 SW}{\partial\bar{Z}\partial\Gamma}\Big/\frac{\partial^2 SW}{\partial\bar{Z}^2}>0$

4. $\frac{d\bar{Z}}{d\alpha}|_t = -\frac{\partial^2 SW}{\partial \bar{Z}\partial \alpha} \Big/ \frac{\partial^2 SW}{\partial \bar{Z}^2} < 0$

The second-order partial derivatives of the welfare function are obtained by partially differentiating (25) with respect to $\bar{Z}$, t, λ, Γ', Γ, t^c and α.

- $\partial^2 SW/\partial \bar{Z}^2 = -\{(\delta-1)[t^c(1-C)-t]+\alpha\}h(\bar{Z}) + \{(\pi^P(\bar{Z})-\pi^R(\bar{Z}))+\delta(t\bar{Z}-A')-\delta[t^c(1-C)\bar{Z}-A]\}h'(\bar{Z})$. This is assumed to be negative as part of the second-order conditions for optimality.
- $\partial^2 SW/\partial \bar{Z}\partial t = (\delta-1)\bar{Z}h(\bar{Z}) > 0$
- $\partial^2 SW/\partial \bar{Z}\partial \lambda = 0$
- $\partial^2 SW/\partial \bar{Z}\partial \Gamma' = -h(\bar{Z}) < 0$
- $\partial^2 SW/\partial \bar{Z}\partial \Gamma = h(\bar{Z}) > 0$
- $\partial^2 SW/\partial \bar{Z}\partial t^c = -(\delta-1)(1-C)\bar{Z}h(\bar{Z}) < 0$
- $\partial^2 SW/\partial \bar{Z}\partial \alpha = -\bar{Z}h(\bar{Z}) < 0$

Proof of Lemma 5

Let $dSW/dt = 0$ denote the first-order condition for maximizing the social welfare function (for a separating equilibrium) with respect to the turnover tax rate. Partially differentiating of the first-order condition and rearranging terms yields,

1. $\frac{dt}{d\lambda}|_{\bar{Z}} = -\frac{\partial^2 SW}{\partial t\partial \lambda} \Big/ \frac{\partial^2 SW}{\partial t^2} > 0$
2. $\frac{dt}{dt^c}|_{\bar{Z}} = -\frac{\partial^2 SW}{\partial t\partial t^c} \Big/ \frac{\partial^2 SW}{\partial t^2} = 0$
3. $\frac{dt}{d\Gamma'}|_{\bar{Z}} = -\frac{\partial^2 SW}{\partial t\partial \Gamma'} \Big/ \frac{\partial^2 SW}{\partial t^2} < 0$; $\frac{dt}{d\Gamma}|_{\bar{Z}} = -\frac{\partial^2 SW}{\partial t\partial \Gamma} \Big/ \frac{\partial^2 SW}{\partial t^2} = 0$
4. $\frac{dt}{d\alpha}|_{\bar{Z}} = -\frac{\partial^2 SW}{\partial t\partial \alpha} \Big/ \frac{\partial^2 SW}{\partial t^2} < 0$

The second-order partial derivatives of the welfare function (for a separating equilibrium) are obtained by partially differentiating (26) with respect to $\bar{Z}$, t, λ, Γ', Γ, t^c and α.

- $\partial^2 SW/\partial t\partial \bar{Z} = (\delta-1)\bar{Z}h(\bar{Z}) > 0$

- $d^2SW/dt^2 < 0$:

$$d^2SW/dt^2 = (\delta-1)(-Z^{IP})h(Z^{IP})\frac{dZ^{IP}}{dt} - \delta(Z^{IP}+t\frac{dZ^{IP}}{dt})h(Z^{IP})\frac{dZ^{IP}}{dt} \\ -\delta(tZ^{IP}-A')[h'(Z^{IP})(\frac{dZ^{IP}}{dt})^2 + h(Z^{IP})\frac{d^2Z^{IP}}{dt^2}] \tag{28}$$

The first row is negative since $\frac{dZ^{IP}}{dt} > 0$. The term $tZ^{IP} - A'$ is positive by the first-order condition for t (setting (26) to zero). Hence, the second row is negative if $h'(Z^{IP}) \geq 0$.

- $\partial^2SW/\partial t\partial\lambda > 0$:

$$\partial^2SW/\partial t\partial\lambda = (\delta-1)(-Z^{IP})h(Z^{IP})\frac{dZ^{IP}}{d\lambda} - \delta th(Z^{IP})\frac{dZ^{IP}}{d\lambda}\frac{dZ^{IP}}{dt} \\ -\delta(tZ^{IP}-A')[h'(Z^{IP})\frac{dZ^{IP}}{d\lambda}\frac{dZ^{IP}}{dt} + h(Z^{IP})\frac{d^2Z^{IP}}{dtd\lambda}] \tag{29}$$

The first row is positive since $\frac{dZ^{IP}}{d\lambda} < 0$. The term $tZ^{IP} - A'$ is positive by the first-order condition for t. Hence, the second row is positive if $h'(Z^{IP}) > 0$ since $\frac{d^2Z^{IP}}{dtd\lambda} < 0$.

- $\partial^2SW/\partial t\partial\Gamma' < 0$

$$\partial^2SW/\partial t\partial\Gamma' = (\delta-1)(-Z^{IP})h(Z^{IP})\frac{dZ^{IP}}{d\Gamma'} - \delta th(Z^{IP})\frac{dZ^{IP}}{d\Gamma'}\frac{dZ^{IP}}{dt} \\ -\delta(tZ^{IP}-A')[h'(Z^{IP})\frac{dZ^{IP}}{d\Gamma'}\frac{dZ^{IP}}{dt} + h(Z^{IP})\frac{d^2Z^{IP}}{dtd\Gamma'}] \tag{30}$$

The first row is negative since $\frac{dZ^{IP}}{d\Gamma'} > 0$. The term $tZ^{IP} - A'$ is positive by the first-order condition for t. Hence, the second row is negative if $h'(Z^{IP}) > 0$ since $\frac{d^2Z^{IP}}{dtd\Gamma'} > 0$.

- $\partial^2SW/\partial t\partial\Gamma = 0$
- $\partial^2SW/\partial t\partial t^c = 0$
- $\partial^2SW/\partial t\partial\alpha < 0$

$$\partial^2SW/\partial t\partial\alpha = (\delta-1)(-Z^{IP})h(Z^{IP})\frac{dZ^{IP}}{d\alpha} - \delta th(Z^{IP})\frac{dZ^{IP}}{d\alpha}\frac{dZ^{IP}}{dt} \\ -\delta(tZ^{IP}-A')[h'(Z^{IP})\frac{dZ^{IP}}{d\alpha}\frac{dZ^{IP}}{dt} + h(Z^{IP})\frac{d^2Z^{IP}}{dtd\alpha}] \tag{31}$$

The first row is negative since $\frac{dZ^{IP}}{d\alpha} > 0$. The term $tZ^{IP} - A'$ is positive by the first-order condition for t. Hence, the second row is negative if $h'(Z^{IP}) > 0$ since $\frac{d^2Z^{IP}}{dtd\alpha} > 0$.

Proof of Proposition 4

Write the system of totally differentiated first-order conditions in matrix form:

$$\begin{bmatrix} \frac{\partial^2 SW}{\partial \bar{Z}^2} & \frac{\partial^2 SW}{\partial \bar{Z}\partial t} \\ \frac{\partial^2 SW}{\partial t \partial \bar{Z}} & \frac{\partial^2 SW}{\partial t^2} \end{bmatrix} \begin{bmatrix} d\bar{Z} \\ dt \end{bmatrix} = \begin{bmatrix} -\frac{\partial^2 SW}{\partial \bar{Z}\partial \lambda} & -\frac{\partial^2 SW}{\partial \bar{Z}\partial \Gamma'} & -\frac{\partial^2 SW}{\partial \bar{Z}\partial \Gamma} & -\frac{\partial^2 SW}{\partial \bar{Z}\partial t^c} & -\frac{\partial^2 SW}{\partial \bar{Z}\partial \alpha} \\ -\frac{\partial^2 SW}{\partial t\partial \lambda} & -\frac{\partial^2 SW}{\partial t\partial \Gamma'} & -\frac{\partial^2 SW}{\partial t\partial \Gamma} & -\frac{\partial^2 SW}{\partial t\partial t^c} & -\frac{\partial^2 SW}{\partial t\partial \alpha} \end{bmatrix} \begin{bmatrix} d\lambda \\ d\Gamma' \\ d\Gamma \\ dt^c \\ d\alpha \end{bmatrix} \tag{32}$$

Let H denote the Hessian of second derivatives of welfare and $|H|$ its determinant. It is assumed that H is negative definite at the solutions to the first-order conditions to satisfy the second-order sufficiency conditions for a welfare maximum. H is negative definite when $|H| > 0$ and $\frac{\partial^2 SW}{\partial \bar{Z}^2} < 0$ and $\frac{\partial^2 SW}{\partial t^2} < 0$. The latter two inequalities are satisfied automatically with the condition stated in the proposition, that $h'(Z)$ is non-negative at the optimum. Hence, the additional assumption for an interior welfare maximum is that $|H| > 0$, where

$$|H| = \begin{vmatrix} \frac{\partial^2 SW}{\partial \bar{Z}^2} & \frac{\partial^2 SW}{\partial \bar{Z}\partial t} \\ \frac{\partial^2 SW}{\partial t \partial \bar{Z}} & \frac{\partial^2 SW}{\partial t^2} \end{vmatrix} = (\frac{\partial^2 SW}{\partial \bar{Z}^2}\frac{\partial^2 SW}{\partial t^2}) - (\frac{\partial^2 SW}{\partial \bar{Z}\partial t}\frac{\partial^2 SW}{\partial t\partial \bar{Z}}) \tag{33}$$

Applying the Cramer's rule yields the comparative statics results.

1. $\frac{d\bar{Z}}{d\lambda} = \frac{-(\frac{\partial^2 SW}{\partial \bar{Z}\partial \lambda}\frac{\partial^2 SW}{\partial t^2}) + (\frac{\partial^2 SW}{\partial \bar{Z}\partial t}\frac{\partial^2 SW}{\partial t\partial \lambda})}{|H|} > 0$ since $\frac{\partial^2 SW}{\partial \bar{Z}\partial \lambda} = 0$, $\frac{\partial^2 SW}{\partial t^2} < 0$, $\frac{\partial^2 SW}{\partial \bar{Z}\partial t} > 0$ and $\frac{\partial^2 SW}{\partial t\partial \lambda} > 0$.
2. $\frac{d\bar{Z}}{dt^c} = \frac{-(\frac{\partial^2 SW}{\partial \bar{Z}\partial t^c}\frac{\partial^2 SW}{\partial t^2}) + (\frac{\partial^2 SW}{\partial \bar{Z}\partial t}\frac{\partial^2 SW}{\partial t\partial t^c})}{|H|} < 0$ since $\frac{\partial^2 SW}{\partial \bar{Z}\partial t^c} < 0$, $\frac{\partial^2 SW}{\partial t^2} < 0$, $\frac{\partial^2 SW}{\partial \bar{Z}\partial t} > 0$ and $\frac{\partial^2 SW}{\partial t\partial t^c} = 0$.
3. $\frac{d\bar{Z}}{d\Gamma'} = \frac{-(\frac{\partial^2 SW}{\partial \bar{Z}\partial \Gamma'}\frac{\partial^2 SW}{\partial t^2}) + (\frac{\partial^2 SW}{\partial \bar{Z}\partial t}\frac{\partial^2 SW}{\partial t\partial \Gamma'})}{|H|} < 0$ since $\frac{\partial^2 SW}{\partial \bar{Z}\partial \Gamma'} < 0$, $\frac{\partial^2 SW}{\partial t^2} < 0$, $\frac{\partial^2 SW}{\partial \bar{Z}\partial t} > 0$ and $\frac{\partial^2 SW}{\partial t\partial \Gamma'} < 0$.
4. $\frac{d\bar{Z}}{d\Gamma} = \frac{-(\frac{\partial^2 SW}{\partial \bar{Z}\partial \Gamma}\frac{\partial^2 SW}{\partial t^2}) + (\frac{\partial^2 SW}{\partial \bar{Z}\partial t}\frac{\partial^2 SW}{\partial t\partial \Gamma})}{|H|} > 0$ since $\frac{\partial^2 SW}{\partial \bar{Z}\partial \Gamma} > 0$, $\frac{\partial^2 SW}{\partial t^2} < 0$, $\frac{\partial^2 SW}{\partial \bar{Z}\partial t} > 0$ and $\frac{\partial^2 SW}{\partial t\partial \Gamma} = 0$.
5. $\frac{d\bar{Z}}{d\alpha} = \frac{-(\frac{\partial^2 SW}{\partial \bar{Z}\partial \alpha}\frac{\partial^2 SW}{\partial t^2}) + (\frac{\partial^2 SW}{\partial \bar{Z}\partial t}\frac{\partial^2 SW}{\partial t\partial \alpha})}{|H|} < 0$ since $\frac{\partial^2 SW}{\partial \bar{Z}\partial \alpha} < 0$, $\frac{\partial^2 SW}{\partial t^2} < 0$, $\frac{\partial^2 SW}{\partial \bar{Z}\partial t} > 0$ and $\frac{\partial^2 SW}{\partial t\partial \alpha} < 0$.

Similarly,

1. $\frac{dt}{d\lambda} = \frac{-(\frac{\partial^2 SW}{\partial \bar{Z}^2}\frac{\partial^2 SW}{\partial t\partial \lambda}) + (\frac{\partial^2 SW}{\partial \bar{Z}\partial \lambda}\frac{\partial^2 SW}{\partial t\partial \bar{Z}})}{|H|} > 0$ since $\frac{\partial^2 SW}{\partial \bar{Z}^2} < 0$, $\frac{\partial^2 SW}{\partial t\partial \lambda} > 0$, $\frac{\partial^2 SW}{\partial \bar{Z}\partial \lambda} = 0$ and $\frac{\partial^2 SW}{\partial t\partial \bar{Z}} > 0$.
2. $\frac{dt}{d\Gamma'} = \frac{-(\frac{\partial^2 SW}{\partial \bar{Z}^2}\frac{\partial^2 SW}{\partial t\partial \Gamma'}) + (\frac{\partial^2 SW}{\partial \bar{Z}\partial \Gamma'}\frac{\partial^2 SW}{\partial t\partial \bar{Z}})}{|H|} < 0$ since $\frac{\partial^2 SW}{\partial \bar{Z}^2} < 0$, $\frac{\partial^2 SW}{\partial t\partial \Gamma'} < 0$, $\frac{\partial^2 SW}{\partial \bar{Z}\partial \Gamma'} < 0$ and $\frac{\partial^2 SW}{\partial t\partial \bar{Z}} > 0$.

3. $\frac{dt}{d\Gamma} = \frac{-(\frac{\partial^2 SW}{\partial \bar{Z}^2}\frac{\partial^2 SW}{\partial t \partial \Gamma}) + (\frac{\partial^2 SW}{\partial \bar{Z} \partial \Gamma}\frac{\partial^2 SW}{\partial t \partial \bar{Z}})}{|H|} > 0$ since $\frac{\partial^2 SW}{\partial \bar{Z}^2} < 0$, $\frac{\partial^2 SW}{\partial t \partial \Gamma} = 0$, $\frac{\partial^2 SW}{\partial \bar{Z} \partial \Gamma} > 0$ and $\frac{\partial^2 SW}{\partial t \partial \bar{Z}} > 0$.

4. $\frac{dt}{dt^c} = \frac{-(\frac{\partial^2 SW}{\partial \bar{Z}^2}\frac{\partial^2 SW}{\partial t \partial t^c}) + (\frac{\partial^2 SW}{\partial \bar{Z} \partial t^c}\frac{\partial^2 SW}{\partial t \partial \bar{Z}})}{|H|} < 0$ since $\frac{\partial^2 SW}{\partial \bar{Z}^2} < 0$, $\frac{\partial^2 SW}{\partial t \partial t^c} = 0$, $\frac{\partial^2 SW}{\partial \bar{Z} \partial t^c} < 0$ and $\frac{\partial^2 SW}{\partial t \partial \bar{Z}} > 0$.

5. $\frac{dt}{d\alpha} = \frac{-(\frac{\partial^2 SW}{\partial \bar{Z}^2}\frac{\partial^2 SW}{\partial t \partial \alpha}) + (\frac{\partial^2 SW}{\partial \bar{Z} \partial \alpha}\frac{\partial^2 SW}{\partial t \partial \bar{Z}})}{|H|} < 0$ since $\frac{\partial^2 SW}{\partial \bar{Z}^2} < 0$, $\frac{\partial^2 SW}{\partial t \partial \alpha} < 0$, $\frac{\partial^2 SW}{\partial \bar{Z} \partial \alpha} < 0$ and $\frac{\partial^2 SW}{\partial t \partial \bar{Z}} > 0$.

Proof of Proposition 5

The derivative of the social welfare function (20) (for a bunching equilibrium) with respect to the threshold yields,

$$\begin{aligned}\frac{dSW}{d\bar{Z}} = {} & \pi^P(\bar{Z})h(\bar{Z}) - \pi^A(\bar{Z}) + \int_{\bar{Z}}^{\hat{Z}} (1-t-C-\alpha)h(Z)dZ \\ & - \pi^R(\hat{Z})h(\hat{Z})\frac{d\hat{Z}}{d\bar{Z}} \\ & + \delta\{(t\bar{Z}-A')h(\bar{Z}) + (t\bar{Z}-A')h(\hat{Z})\frac{d\hat{Z}}{d\bar{Z}} - (t\bar{Z}-A')h(\bar{Z}) \\ & + \int_{\bar{Z}}^{\hat{Z}} th(Z)dZ - [t^c(1-C)\hat{Z}-A]h(\hat{Z})\frac{d\hat{Z}}{d\bar{Z}}\}\end{aligned} \tag{34}$$

Using the relationships $\pi^P(\bar{Z}) = \pi^A(\bar{Z})$ and $\pi^A(\bar{Z}) = \pi^R(\hat{Z})$, Equation (34) can be simplified and the first-order condition written as

$$\begin{aligned}\frac{dSW}{d\bar{Z}} = {} & \int_{\bar{Z}}^{\hat{Z}} (1-t-C-\alpha+\delta t)h(Z)dZ \\ & + \delta\{(t\bar{Z}-A')] - [t^c(1-C)\hat{Z}-A]\}h(\hat{Z})\frac{d\hat{Z}}{d\bar{Z}} \\ = {} & 0\end{aligned} \tag{35}$$

Rearranging terms generates the result.

Proof of Proposition 6

The derivative of the social welfare function (20) (for a bunching equilibrium) is

$$\begin{aligned}\frac{dSW}{dt} &= [\omega\pi(Z^{IP}) - \pi^P(Z^{IP})]h(Z^{IP})\frac{dZ^{IP}}{dt} \\ &+ [\pi^A(\bar{Z}) - \pi^R(\hat{Z})]h(\hat{Z})\frac{d\hat{Z}}{dt} - \int_{Z^{IP}}^{\bar{Z}}(-Z)h(Z)dZ - \int_{\bar{Z}}^{\hat{Z}}(-\bar{Z})h(Z)dZ \\ &+ \delta\{\int_{Z^{IP}}^{\bar{Z}} Zh(Z)dZ + \int_{\bar{Z}}^{\hat{Z}} \bar{Z}h(Z)dZ - (tZ^{IP} - A')h(Z^{IP})\frac{dZ^{IP}}{dt} \\ &+ (t\bar{Z} - A')h(\hat{Z})\frac{d\hat{Z}}{dt} - [t^c(1-C)\hat{Z} - A]h(\hat{Z})\frac{d\hat{Z}}{dt}\}\end{aligned} \tag{36}$$

Using the relationships $\pi^P(Z^{IP}) = \pi I(Z^{IP})$ and $\pi^A(\bar{Z}) = \pi^R(\hat{Z})$ to cancel terms, Equation (36) can be simplified and the first-order condition written as

$$\begin{aligned}\frac{dSW}{dt} &= (\delta-1)\int_{Z^{IP}}^{\bar{Z}} Zh(Z)dZ + (\delta-1)\int_{\bar{Z}}^{\hat{Z}} \bar{Z}h(Z)dZ \\ &- \delta(tZ^{IP} - A')h(Z^{IP})\frac{dZ^{IP}}{dt} + \delta\{(t\bar{Z} - A') - [t^c(1-C)\hat{Z} - A]\}h(\hat{Z})\frac{d\hat{Z}}{dt} \\ &= 0\end{aligned} \tag{37}$$

Proof of Proposition 7

If $H(Z)$ is the uniform distribution, then using the definitions of Z^{IP} and $\hat{Z}$ from (6) and (8) and expanding terms, the equation becomes

$$\begin{aligned}&\{-\frac{\delta-1}{2} + (\delta-1)\frac{(1-t-C-\alpha)}{(1-t^c)(1-C)} + \delta t^c(1-C)\frac{(1-t-C-\alpha)}{[(1-t^c)(1-C)]^2} - \delta\frac{t}{(1-t^c)(1-C)}\}\bar{Z}^2 \\ &+\{(\delta-1)\frac{\Gamma-\Gamma'}{(1-t^c)(1-C)} + \delta t^c(1-C)\frac{\Gamma-\Gamma'}{[(1-t^c)(1-C)]^2} - \delta\frac{A-A'}{(1-t^c)(1-C)}\}\bar{Z} \\ &-\delta t\frac{\Gamma'^2}{(\lambda-t-\alpha)^3} + \delta\frac{\Gamma'A'}{(\lambda-t-\alpha)^2} - \frac{\delta-1}{2}\frac{\Gamma'^2}{(\lambda-t-\alpha)^2} = 0\end{aligned} \tag{38}$$

which can be rearranged into the standard quartic form $at^4 + bt^3 + ct^2 + dt + e = 0$ with the coefficients given in the proposition. $\Delta' = -({\Delta_1}^2 - 4{\Delta_0}^3)/27$ is the discriminant of the quartic equation. If it is negative, then there are two distinct real roots (and two imaginary roots) (Irving, 2020, Theorem 6.5) satisfying the first-order condition (37). One root is welfare-maximizing and the other is welfare-minimizing. The numerical solutions enable us to identify the analytical root for t that maximizes social welfare. The solution for the optimal t, for any given threshold, is then given by the formula in the proposition.

Proof of Lemma 6

First, rewrite (21) as

$$[1-t-C-\alpha+\delta t][H(\hat{Z})-H(\bar{Z})] = \quad \delta[t^c(1-C)\hat{Z}-A-(t\bar{Z}-A')]h(\hat{Z})\frac{d\hat{Z}}{dt} \tag{39}$$

An interior solution requires $t > 0$ and $Z^{IP} > 0$. The latter condition is satisfied if and only if $\lambda - t - \alpha > 0$ or equivalently, $1-t-C-\alpha > 1-C-\lambda$. Since $1-C-\lambda > 0$ by assumption, and $\hat{Z} > \bar{Z}$ by the definition of a bunching equilibrium, we can conclude that the left-hand side of (39) is positive. Consequently, satisfaction of the first-order condition requires the right-hand side of (39) to also be positive.

Now rewrite (23) as

$$\begin{aligned}&(\delta-1)\int_{Z^{IP}}^{\bar{Z}} Zh(Z)dZ + (\delta-1)\int_{\bar{Z}}^{\hat{Z}} \bar{Z}h(Z)dZ - \delta(tZ^{IP}-A')h(Z^{IP})\frac{dZ^{IP}}{dt} \\ =&\delta[t^c(1-C)\hat{Z}-A-(t\bar{Z}-A']]h(\hat{Z})\frac{d\hat{Z}}{dt}\end{aligned} \tag{40}$$

On the right-hand side of (40), $h(\hat{Z})\frac{d\hat{Z}}{dt} < 0$ and the term multiplying it is positive, as previously noted. Hence, the right-hand side of (40) is negative. Hence, satisfaction of (40) will require the left-hand side to also be negative. But on the left-hand side of (40) both integrals are positive and, because $\frac{dZ^{IP}}{dt} > 0$, the only way the left-hand side can be negative is if $t\hat{Z} - A' > 0$. Using this fact, it is easy to show by calculating the derivatives with respect to t that the left-hand side of (40) is strictly decreasing in t, while the right-hand side is strictly increasing in t, if $h'(Z^{IP}$ is non-negative. This implies that if an interior solution exists for t when $\bar{Z}$ is at its optimal value, then it is a unique interior solution. In turn, (22) ensures that there is a unique $\bar{Z}^*$ for the optimal t.

Proof of lemma 7

Let $dSW/d\bar{Z} = 0$ denote the first-order condition for maximizing social welfare function (for a bunching equilibrium) with respect to the threshold. Partial differentiation of the equation $dSW/d\bar{Z} = 0$ and rearranging terms yields,

1. $\frac{d\bar{Z}}{d\lambda}_{|t} = -\frac{\partial^2 SW}{\partial\bar{Z}\partial\lambda}\Big/\frac{\partial^2 SW}{\partial\bar{Z}^2} = 0$
2. $\frac{d\bar{Z}}{dt^c}_{|t} = -\frac{\partial^2 SW}{\partial\bar{Z}\partial t^c}\Big/\frac{\partial^2 SW}{\partial\bar{Z}^2} < 0$

The second-order partial derivatives of the welfare function (for a bunching equilibrium) are obtained by partially differentiating (34) with respect to $\bar{Z}$, t, λ, Γ, Γ', t^c and α.

- $\frac{\partial^2 SW}{\partial \bar{Z}^2} = [(1-t-C-\alpha)+\delta t][h(\hat{Z})\frac{\partial \hat{Z}}{\partial \bar{Z}} - h(\bar{Z})] + \delta t h(\hat{Z})\frac{\partial \hat{Z}}{\partial \bar{Z}} - \delta t^c(1-C)h(\hat{Z})(\frac{\partial \hat{Z}}{\partial \bar{Z}})^2 + \delta\{(t\bar{Z}-A') - [t^c(1-C)\hat{Z}-A]\}h'(\hat{Z})(\frac{\partial \hat{Z}}{\partial \bar{Z}})^2$. This is assumed to be negative as part of the second-order conditions for optimality.

- $\frac{\partial^2 SW}{\partial \bar{Z}\partial t} > 0$:

$$\begin{aligned}\frac{\partial^2 SW}{\partial \bar{Z}\partial t} &= (\delta-1)(H(\hat{Z})-H(\bar{Z})) + [(1-t-C-\alpha)+\delta t]h(\hat{Z})\frac{\partial \hat{Z}}{\partial t} \\ &+ \delta \bar{Z} h(\hat{Z})\frac{\partial \hat{Z}}{\partial \bar{Z}} - \delta t^c(1-C)\frac{\partial \hat{Z}}{\partial t}h(\hat{Z})\frac{\partial \hat{Z}}{\partial \bar{Z}} \\ &+ \delta\{(t\bar{Z}-A') - [t^c(1-C)\hat{Z}-A]\}[h'(\hat{Z})\frac{\partial \hat{Z}}{\partial t}\frac{\partial \hat{Z}}{\partial \bar{Z}} + h(\hat{Z})\frac{\partial^2 \hat{Z}}{\partial \bar{Z}\partial t}]\end{aligned} \tag{41}$$

  The term $(1-t-C-\alpha)h(\hat{Z})\frac{d\hat{Z}}{dt}$ can be written as $-\bar{Z}h(\hat{Z})\frac{d\hat{Z}}{d\bar{Z}}$ using the definition of $\hat{Z}$ given by (8). Then

$$\begin{aligned}\partial^2 SW/\partial \bar{Z}\partial t &= (\delta-1)\bar{Z}h(\hat{Z})\frac{d\hat{Z}}{d\bar{Z}} + (\delta-1)(H(\hat{Z})-H(\bar{Z})) \\ &+ \delta t h(\hat{Z})\frac{d\hat{Z}}{dt} - \delta t^c(1-C)\frac{d\hat{Z}}{d\bar{Z}}h(\hat{Z})\frac{d\hat{Z}}{dt} \\ &+ [\delta(t\bar{Z}-A') - \delta(t^c(1-C)\hat{Z}-A)][h'(\hat{Z})\frac{d\hat{Z}}{dt} + h(\hat{Z})\frac{d^2\hat{Z}}{d\bar{Z}dt}]\end{aligned} \tag{42}$$

  The first and third rows are positive if $h'(\hat{Z}) > 0$ since $\frac{d\hat{Z}}{d\bar{Z}} > 0$, $\frac{d\hat{Z}}{dt} < 0$, $\frac{d^2\hat{Z}}{d\bar{Z}dt} < 0$ and $t^c(1-C)\hat{Z} - A > t\bar{Z} - A'$ by (21). For the second row, $\delta t h(\hat{Z})\frac{d\hat{Z}}{dt} - \delta t^c(1-C)\frac{d\hat{Z}}{d\bar{Z}}h(\hat{Z})\frac{d\hat{Z}}{dt} = \delta t h(\hat{Z})\frac{d\hat{Z}}{dt} - \delta t^c\frac{1-t-C-\alpha}{1-t^c}h(\hat{Z})\frac{d\hat{Z}}{dt} = (t - t^c\frac{1-t-C-\alpha}{1-t^c})\delta h(\hat{Z})\frac{d\hat{Z}}{dt}$. Thus, if $t(1-t^c) - t^c(1-t-C-\alpha) < 0$ or, equivalently, $t < t^c(1-C-\alpha)$, then the second row must be positive since $\frac{d\hat{Z}}{dt} < 0$. Because $t < \lambda - \alpha$, t will be smaller than $t^c(1-C-\alpha)$ as long as $t^c(1-C-\alpha) > \lambda - \alpha$, which is the assumption in the proposition.

- $\frac{\partial^2 SW}{\partial \bar{Z}\partial \lambda} = 0$

- $\frac{\partial^2 SW}{\partial \bar{Z}\partial \Gamma'} = [(1-t-C-\alpha)+\delta t]h(\hat{Z})\frac{\partial \hat{Z}}{\partial \Gamma'} + \delta(t\bar{Z}-A')h'(\hat{Z})\frac{\partial \hat{Z}}{\partial \Gamma'}\frac{\partial \hat{Z}}{\partial \bar{Z}} - \delta t^c(1-C)\frac{\partial \hat{Z}}{\partial \Gamma'}h(\hat{Z})\frac{\partial \hat{Z}}{\partial \bar{Z}} - \delta[t^c(1-C)\hat{Z}-A]h'(\hat{Z})\frac{\partial \hat{Z}}{\partial \Gamma'}\frac{\partial \hat{Z}}{\partial \bar{Z}}$, the sign is ambiguous

- $\frac{\partial^2 SW}{\partial \bar{Z}\partial \Gamma} = [(1-t-C-\alpha)+\delta t]h(\hat{Z})\frac{\partial \hat{Z}}{\partial \Gamma} + \delta(t\bar{Z}-A')h'(\hat{Z})\frac{\partial \hat{Z}}{\partial \Gamma}\frac{\partial \hat{Z}}{\partial \bar{Z}} - \delta t^c(1-C)\frac{\partial \hat{Z}}{\partial \Gamma}h(\hat{Z})\frac{\partial \hat{Z}}{\partial \bar{Z}} - \delta[t^c(1-C)\hat{Z}-A]h'(\hat{Z})\frac{\partial \hat{Z}}{\partial \Gamma'}\frac{\partial \hat{Z}}{\partial \bar{Z}}$, the sign is ambiguous

- $\frac{\partial^2 SW}{\partial \bar{Z} \partial t^c} < 0$:

$$\begin{aligned}\partial^2 SW/\partial \bar{Z}\partial t^c = & [(1-t-C-\alpha)+\delta t]h(\hat{Z})\frac{\partial \hat{Z}}{\partial t^c} \\ & -\delta[(1-C)\hat{Z}+t^c(1-C)\frac{\partial \hat{Z}}{\partial t^c}]h(\hat{Z})\frac{\partial \hat{Z}}{\partial \bar{Z}} \\ & +\delta\{(t\bar{Z}-A')-[t^c(1-C)\hat{Z}-A]\}[h'(\hat{Z})\frac{\partial \hat{Z}}{\partial t^c}\frac{\partial \hat{Z}}{\partial \bar{Z}}+h(\hat{Z})\frac{\partial^2 \hat{Z}}{\partial \bar{Z}\partial t^c}]\end{aligned} \tag{43}$$

The term in braces in the third row is negative by the first-order condition (21). To show that the sum of the terms in the first and second rows is negative if $t^c(1-C-\alpha) > \lambda - \alpha$:

$$\begin{aligned} & [(1-t-C-\alpha)+\delta t]h(\hat{Z})\frac{\partial \hat{Z}}{\partial t^c} - \delta[(1-C)\hat{Z}+t^c(1-C)\frac{\partial \hat{Z}}{\partial t^c}]h(\hat{Z})\frac{\partial \hat{Z}}{\partial \bar{Z}} \\ = & \frac{1-t-C-\alpha}{1-t^c}\hat{Z}h(\hat{Z}) + \delta\frac{t}{1-t^c}\hat{Z}h(\hat{Z}) \\ & -\delta\frac{1-t-C-\alpha}{1-t^c}\hat{Z}h(\hat{Z}) - \delta\frac{t^c(1-C)}{1-t^c}\frac{1-t-C-\alpha}{(1-t^c)(1-C)}\hat{Z}h(\hat{Z}) \\ = & -(\delta-1)\frac{1-t-C-\alpha}{1-t^c}\hat{Z}h(\hat{Z}) - \frac{\delta}{1-t^c}[\frac{t^c(1-t-C-\alpha)}{1-t^c}-t]\hat{Z}h(\hat{Z})\end{aligned} \tag{44}$$

where the first equality uses the expressions for $d\hat{Z}/d\bar{Z}$ and $d\hat{Z}/dt^c$ derived from (8). The first term in the last row of the equation above is negative, and the second term is also negative under $t^c(1-C-\alpha) > \lambda - \alpha$ because $\lambda - \alpha > t$. Then, $\partial^2 SW/\partial \bar{Z}\partial t^c < 0$.

- $\frac{\partial^2 SW}{\partial \bar{Z}\partial \alpha} = -[H(\hat{Z})-H(\bar{Z})] - \delta t^c(1-C)\frac{d\hat{Z}}{d\alpha}h(\hat{Z})\frac{d\hat{Z}}{d\bar{Z}} + \delta\{(t\bar{Z}-A')-[t^c(1-C)\hat{Z}-A]\}[h'(\hat{Z})\frac{d\hat{Z}}{d\alpha}\frac{d\hat{Z}}{d\bar{Z}}+h(\hat{Z})\frac{\partial^2\hat{Z}}{\partial\bar{Z}\partial\alpha}]$, the sign is ambiguous

Proof of lemma 8

Let $dSW/dt = 0$ denote the first-order condition for maximizing the social welfare function (for a bunching equilibrium) with respect to the turnover tax rate. Partially dfferentiating of the first-order condition and rearranging terms yields,

1. $\frac{dt}{d\lambda}|_{\bar{Z}} = -\frac{\partial^2 SW}{\partial t\partial\lambda}\Big/\frac{\partial^2 SW}{\partial t^2} > 0$
2. $\frac{dt}{dt^c}|_{\bar{Z}} = -\frac{\partial^2 SW}{\partial t\partial t^c}\Big/\frac{\partial^2 SW}{\partial t^2} > 0$
3. $\frac{dt}{d\alpha}|_{\bar{Z}} = -\frac{\partial^2 SW}{\partial t\partial\alpha}\Big/\frac{\partial^2 SW}{\partial t^2} < 0$

The second-order partial derivatives of the welfare function (for a bunching equilibrium) are obtained by partially differentiatingn (36) with respect to $\bar{Z}$, t, λ, Γ', Γ, t^c and α.

- $\frac{\partial^2 SW}{\partial t \partial \bar{Z}} > 0$ by the symmetry of the cross-partial derivatives (it was shown previously that $\frac{\partial^2 SW}{\partial \bar{Z}} \partial t > 0$.)

- $\frac{\partial^2 SW}{\partial t^2} = (\delta-1)(-Z^{IP})h(Z^{IP})\frac{\partial Z^{IP}}{\partial t} + (\delta-1)\bar{Z}h(\hat{Z})\frac{\partial \hat{Z}}{\partial t} - \delta t^c(1-C)\frac{\partial \hat{Z}}{\partial t}h(\hat{Z})\frac{\partial \hat{Z}}{\partial t} - \delta[t^c(1-C)\hat{Z}-A]h'(\hat{Z})(\frac{\partial \hat{Z}}{\partial t})^2 + \delta\bar{Z}h(\hat{Z})\frac{\partial \hat{Z}}{\partial t} + \delta(t\bar{Z}-A')h'(\hat{Z})(\frac{\partial \hat{Z}}{\partial t})^2 - \delta(Z^{IP} - t\frac{\partial Z^{IP}}{\partial t})h(Z^{IP})\frac{\partial Z^{IP}}{\partial t} - \delta(tZ^{IP}-A')h'(Z^{IP})(\frac{\partial Z^{IP}}{\partial t})^2 - \delta(tZ^{IP}-A')h(Z^{IP})\frac{\partial^2 Z^{IP}}{\partial^2 t}$. This is assumed to be negative as part of the second-order conditions for optimality.

- $\frac{\partial^2 SW}{\partial t \partial \lambda} > 0$

$$\partial^2 SW/\partial t\partial\lambda = (\delta-1)(-Z^{IP})h(Z^{IP})\frac{\partial Z^{IP}}{\partial \lambda} - \delta t\frac{\partial Z^{IP}}{\partial \lambda}h(Z^{IP})\frac{\partial Z^{IP}}{\partial t} - \delta(tZ^{IP}-A')[h'(Z^{IP})\frac{\partial Z^{IP}}{\partial \lambda}\frac{\partial Z^{IP}}{\partial t} + h(Z^{IP})\frac{\partial^2 Z^{IP}}{\partial t\partial\lambda}] \quad (45)$$

The terms in the first row are negative. The second row is also positive if $h'(Z^{IP}) \geq 0$, since $tZ^{IP} - A' > 0$ by (23).

- $\frac{\partial^2 SW}{\partial t \partial \Gamma'} = (\delta-1)(-Z^{IP})h(Z^{IP})\frac{\partial Z^{IP}}{\partial \Gamma'} + (\delta-1)\bar{Z}h(\hat{Z})\frac{\partial \hat{Z}}{\partial \Gamma'} - \delta(tZ^{IP}-A')[h'(Z^{IP})\frac{\partial Z^{IP}}{\partial \Gamma'} + h(Z^{IP})\frac{\partial^2 Z^{IP}}{\partial t\partial\Gamma'}] - \delta t^c(1-C)\frac{\partial \hat{Z}}{\partial \Gamma'}h(\hat{Z})\frac{\partial \hat{Z}}{\partial t} + \delta\{(t\bar{Z}-A') - [t^c(1-C)\hat{Z}-A]\}[h'(Z^{IP})\frac{\partial Z^{IP}}{\partial \Gamma'} + h(Z^{IP})\frac{\partial^2 Z^{IP}}{\partial t\partial\Gamma'}]$, the sign is ambiguous

- $\frac{\partial^2 SW}{\partial t \partial \Gamma} = (\delta-1)\bar{Z}h(\hat{Z})\frac{d\hat{Z}}{d\Gamma} - \delta t^c(1-C)\frac{d\hat{Z}}{d\Gamma}h(\hat{Z})\frac{d\hat{Z}}{dt} + \delta\{(t\bar{Z}-A') - [t^c(1-C)\hat{Z}-A]\}h'(\hat{Z})\frac{d\hat{Z}}{d\Gamma}$, the sign is ambiguous

- $\frac{\partial^2 SW}{\partial t \partial t^c} > 0$

$$\partial^2 SW/\partial t\partial t^c = (\delta-1)\bar{Z}h(\hat{Z})\frac{\partial \hat{Z}}{\partial t^c} - \delta[(1-C)\hat{Z} + t^c(1-C)\frac{\partial \hat{Z}}{\partial t^c}]h(\hat{Z})\frac{\partial \hat{Z}}{\partial t} - \delta[t^c(1-C)\hat{Z}-A][h'(\hat{Z})\frac{\partial \hat{Z}}{\partial t^c}\frac{\partial \hat{Z}}{\partial t} + h(\hat{Z})\frac{\partial^2 \hat{Z}}{\partial t\partial t^c}] + \delta(t\bar{Z}-A')[h'(\hat{Z})\frac{\partial \hat{Z}}{\partial t^c}\frac{\partial \hat{Z}}{\partial t} + h(\hat{Z})\frac{\partial^2 \hat{Z}}{\partial t\partial t^c}] \quad (46)$$

The terms in the first row are positive. The sum of the terms in the second and third rows is also positive since $t^c(1-C)\hat{Z} - A > t\bar{Z} - A'$ by the first-order condition for $\bar{Z}$ (21).

- $\frac{\partial^2 SW}{\partial t \partial \alpha} < 0$

$$\begin{aligned}\partial^2 SW/\partial t\partial\alpha = &(\delta-1)(-Z^{IP})h(Z^{IP})\frac{\partial Z^{IP}}{\partial\alpha}+(\delta-1)\bar{Z}h(\hat{Z})\frac{\partial\hat{Z}}{\partial\alpha}\\ &-\delta t\frac{\partial Z^{IP}}{\partial\alpha}h(Z^{IP})\frac{\partial Z^{IP}}{\partial t}-\delta t^c(1-C)\frac{\partial\hat{Z}}{\partial\alpha}h(\hat{Z})\frac{\partial\hat{Z}}{\partial t}\\ &-\delta(tZ^{IP}-A')[h'(Z^{IP})\frac{\partial Z^{IP}}{\partial\alpha}\frac{\partial Z^{IP}}{\partial t}+h(Z^{IP})\frac{\partial^2 Z^{IP}}{\partial t\partial\alpha}]\\ &+\delta\{(t\bar{Z}-A')-[t^c(1-C)\hat{Z}-A]\}h'(\hat{Z})\frac{\partial\hat{Z}}{\partial\alpha}\frac{\partial\hat{Z}}{\partial t}\end{aligned} \quad (47)$$

The first three rows are negative if $h'(Z^{IP}) > 0$. The term in braces in the forth row is negative by the first-order condition for $\bar{Z}$ (21). Hence, the forth row is negative if $h'(\hat{Z}) > 0$.

Proof of Proposition 8

Write the system of totally differentiated first-order conditions in matrix form:

$$\begin{bmatrix}\frac{\partial^2 SW}{\partial\bar{Z}^2} & \frac{\partial^2 SW}{\partial\bar{Z}\partial t}\\ \frac{\partial^2 SW}{\partial t\partial\bar{Z}} & \frac{\partial^2 SW}{\partial t^2}\end{bmatrix}\begin{bmatrix}d\bar{Z}\\ dt\end{bmatrix}=\begin{bmatrix}-\frac{\partial^2 SW}{\partial\bar{Z}\partial\lambda} & -\frac{\partial^2 SW}{\partial\bar{Z}\partial t^c}\\ -\frac{\partial^2 SW}{\partial t\partial\lambda} & -\frac{\partial^2 SW}{\partial t\partial t^c}\end{bmatrix}\begin{bmatrix}d\lambda\\ dt^c\end{bmatrix} \quad (48)$$

Let H denote the Hessian of second derivatives and $|H|$ its determinant. It is assumed that H is negative definite at the solutions to the first-order conditions, to satisfy the second-order sufficiency conditions for a welfare maximum. H is negative definite when $|H| > 0$ and $\frac{\partial^2 SW}{\partial\bar{Z}^2} < 0$ and $\frac{\partial^2 SW}{\partial t^2} < 0$. The latter two inequalities are satisfied automatically with the condition stated in the proposition, that $h'(Z)$ is non-negative at the optimum. Hence, the additional assumption for an interior welfare maximum is that $|H| > 0$, where

$$|H| = \begin{vmatrix}\frac{\partial^2 SW}{\partial\bar{Z}^2} & \frac{\partial^2 SW}{\partial\bar{Z}\partial t}\\ \frac{\partial^2 SW}{\partial t\partial\bar{Z}} & \frac{\partial^2 SW}{\partial t^2}\end{vmatrix} = (\frac{\partial^2 SW}{\partial\bar{Z}^2}\frac{\partial^2 SW}{\partial t^2})-(\frac{\partial^2 SW}{\partial\bar{Z}\partial t}\frac{\partial^2 SW}{\partial t\partial\bar{Z}}) \quad (49)$$

Applying the Cramer's rule yields the comparative statics results.

1. $\frac{d\bar{Z}}{d\lambda} = \frac{-(\frac{\partial^2 SW}{\partial\bar{Z}\partial\lambda}\frac{\partial^2 SW}{\partial t^2})+(\frac{\partial^2 SW}{\partial\bar{Z}\partial t}\frac{\partial^2 SW}{\partial t\partial\lambda})}{|H|} > 0$ since $\frac{\partial^2 SW}{\partial\bar{Z}\partial\lambda} = 0$, $\frac{\partial^2 SW}{\partial t^2} < 0$, $\frac{\partial^2 SW}{\partial\bar{Z}\partial t} > 0$ and $\frac{\partial^2 SW}{\partial t\partial\lambda} > 0$.

2. $\frac{d\bar{Z}}{dt^c} = \frac{-(\frac{\partial^2 SW}{\partial\bar{Z}\partial t^c}\frac{\partial^2 SW}{\partial t^2})+(\frac{\partial^2 SW}{\partial\bar{Z}\partial t}\frac{\partial^2 SW}{\partial t\partial t^c})}{|H|}$, the sign is ambiguous since $\frac{\partial^2 SW}{\partial\bar{Z}\partial t^c} < 0$, $\frac{\partial^2 SW}{\partial t^2} < 0$, $\frac{\partial^2 SW}{\partial\bar{Z}\partial t} > 0$ and $\frac{\partial^2 SW}{\partial t\partial t^c} > 0$.

Similarly,

1. $\frac{dt}{d\lambda} = \frac{-(\frac{\partial^2 SW}{\partial \bar{Z}^2}\frac{\partial^2 SW}{\partial t \partial \lambda}) + (\frac{\partial^2 SW}{\partial \bar{Z} \partial \lambda}\frac{\partial^2 SW}{\partial t \partial \bar{Z}})}{|H|} > 0$ since $\frac{\partial^2 SW}{\partial \bar{Z}^2} < 0$, $\frac{\partial^2 SW}{\partial t \partial \lambda} > 0$, $\frac{\partial^2 SW}{\partial \bar{Z} \partial \lambda} = 0$ and $\frac{\partial^2 SW}{\partial t \partial \bar{Z}} > 0$.

2. $\frac{dt}{dt^c} = \frac{-(\frac{\partial^2 SW}{\partial \bar{Z}^2}\frac{\partial^2 SW}{\partial t \partial t^c}) + (\frac{\partial^2 SW}{\partial \bar{Z} \partial t^c}\frac{\partial^2 SW}{\partial t \partial \bar{Z}})}{|H|}$, the sign is ambiguous since $\frac{\partial^2 SW}{\partial \bar{Z}^2} < 0$, $\frac{\partial^2 SW}{\partial t \partial t^c} > 0$, $\frac{\partial^2 SW}{\partial \bar{Z} \partial t^c} < 0$ and $\frac{\partial^2 SW}{\partial t \partial \bar{Z}} > 0$.