Horizontal Inequality in Optimal Taxation:

Evidence from a Development Context

Magnus Eldrup
University of Copenhagen

September 2025

Abstract

This paper examines optimal income taxation in developing countries, where informal employment and tax evasion are pervasive. Traditional optimal tax models, assuming full compliance, overlook the horizontal inequality between formal and informal workers created by the tax system. By introducing untaxable agents into the Mirrlees (1971) framework, this study bridges optimal tax theory with the enforcement realities of developing countries and provides a welfarist basis for tax prescription under horizontal inequality. Applied to the Kenyan tax system using comprehensive administrative data, the analysis shows that accounting for preferences against horizontal inequality leads to optimal tax schedules with lower rates and high exemption thresholds, closely mirroring observed policy. This challenges the common recommendation of tax base broadening. The mechanism of strong horizontal inequality aversion is confirmed through a novel experimental survey that directly elicits marginal social welfare weights in Kenya. Model extensions show that neither extensive margin responses nor capacity constraints alone can explain the observed exemptions.

Funding from the Nordic Tax Research Council is gratefully acknowledged. This paper has benefited from comments from Niels Johannesen, Anders Jensen, Emmanuel Saez, Claus Thustrup Kreiner, David Atkin, Joana Naritomi and Henrik Kleven as well as from particpants at the Nordic Public Policy Symposium (2024), EDGE Jamboree (2024), UCPH IID (2025), UCPH PhD Seminar (2024) and DGPE (2024). The experiment in this paper was pre-registered at https://www.socialscienceregistry.org/trials/15719

1 Introduction

The first-best solution to the issue of income tax evasion will always be to eliminate it. On this point, there is very little disagreement. Yet, tax evasion has been a problem for as long as taxation has been a project, and it has no end in sight. The persistence can hardly be ascribed to a lack of academic interest (Allingham & Sandmo, 1972; Keen & Slemrod, 2017; Okunogbe & Pouliquen, 2022) nor to a lack of public consternation. Both surveys (Cummings et al., 2001) and experimental data (Bellemare et al., 2008; Charness & Rabin, 2002) consistently find a strong aversion to tax evasion across a broad range of countries.

Nowhere is the intractability of evasion more apparent than in developing countries. Here, large shares of workers are self-employed or employed informally with no paper trail nor third party reporting available (Besley & Persson, 2013; La Porta & Shleifer, 2014). Consequently, these workers effectively remit no taxes (Slemrod & Yitzhaki, 2002). This implies that any income tax also implicitly constitutes a transfer from formal to informal workers thus creating an additional dimension of redistribution and corresponding horizontal inequality. Such transfers violate the assumptions underpinning workhorse optimal income tax models and requires a change to how we approach our models of income taxation. Slemrod & Yitzhaki (2002) highlighted this issue over 20 years ago, yet it remains largely unresolved.

This paper investigates optimal income taxation in this second-best world. Specifically, it explores how a social planner, faced with intractable evasion, can set optimal non-linear taxes trading off both vertical and horizontal inequality concerns. That is, rather than asking *what drives evasion*, this paper examines *how we might tax optimally in the presence of evasion.* The model is applied to the Kenyan tax system using an administrative dataset on the full universe of taxpayers and empircally estimated social welfare weights. The simulations show that accounting for untaxable workers yields very different structures of optimal tax systems.

This perspective complements the large literature linking tax design to administrative capacity (Bergeron et al., 2021; Besley, 2020; Jensen, 2022). That work shows that the structure of personal income taxes is tightly connected to what governments are able to enforce. I take these capacity constraints as given and show that, by themselves, they cannot explain the prevalence of high exemption thresholds. Specifically, horizontal inequality preferences provide a necessary rationale for imposing high exemptions.

At first glance, one might think high exemptions follow automatically from informality and low incomes. Yet in 2021, 1.3 million *registered* Kenyan workers earned below the exemption threshold.

Taxing them would have increased PIT revenues by 13 billion Ksh. annually, with an additional 44 billion Ksh. from inframarginal effects above the threshold. This would constituete a 13.5 pct. increase in PIT revenue. As such, exemptions cannot be justified by capacity constraints alone.

Applications of the Mirrlees optimal tax model to development contexts are often rightly viewed with skepticism (Slemrod & Gillitzer, 2014). Many factors critical in developing countries, such as compliance costs and participation effects in the formal sector, are not easily incorporated into a Mirrleesian framework. I address these concerns by expanding the model to feature both participation effects and administrative costs. However, I also argue that despite its limitations the Mirrlees optimal tax model provides an important, theoretically consistent, framework for navigating the equity-efficiency trade-offs of taxation. Without such a foundation, no model, regardless of its complexity, will be able to produce useful tax policy prescriptions.

In this paper, I expand the Saez (2001a) formulation of the Mirrlees (1971) optimal income tax model. This model is a pillar of modern public finance owing to its elegant approach to the efficiency-equity trade-off of taxation in the presence of skill heterogeneity. I augment the model with a second dimension of heterogeneity encompassing the possibility of agents being either 'taxable' or 'untaxable', which in this application corresponds to being formal or informal. This mapping between being 'untaxable' and 'informal' is supported by the fact that only a tiny fraction of tax-liable informal workers remit any taxes. I then generalize the social welfare function (SWF) to navigate the trade-off facing the planner between horizontal equality, vertical equality and efficiency building on previous work of Auerbach & Hassett (2002).

The proposed SWF has several advantages. The function disentangles horizontal and vertical equality preferences (Stiglitz, 1982), it is derived directly from individual utilities and is as such welfarist, and it still simplifies to the Saez (2001a) optimal tax formula under common assumptions. Additionally, the function spans two intuitively appealing edge cases for horizontal inequality aversion, thus conferring reasonable bounds for optimal taxes. The edge cases are *horizontal utilitarianism* where utility is valued equally and *horizontal Rawlsianism* where any additional utility owing to tax evasion is assigned a social weight of zero.

There is a notable paucity of papers applying optimal tax theory to developing countries. This is likely because optimal tax formulas often produce stylized results that bear little resemblance to the observed tax schedules in these countries (Gordon & Li, 2009). For example, the classic result from optimal income taxation of never excluding low earners from the tax base is strongly at odds with the actual tax systems in developing countries where the exemption cutoff is routinely placed at the 8th and 9th deciles of the income distribution (Jensen, 2022). Some interpret this as evidence

that tax systems in the developing world are highly suboptimal (Fall, 2022; OECD, 2012), while others recognize that many of the pre-conditions for applying simple optimal tax formulas, such as costless taxation and horizontal equality, are violated, thus undermining their relevance.

I apply the new model to the Kenyan personal income tax (PIT) to examine whether accounting for the presence of an untaxable sector leads to more reasonable model prescriptions, closer to the observed tax system of Kenya. To achieve this, I use a combination of administrative tax records and household survey data to construct full income and skill distributions for both the taxable (formal) and the untaxable (informal) sectors.

Additionally, I verify the mechanism of horizontal inequality aversion directly by eliciting the social preferences of Kenyans using a novel dynamic survey design. Similarly to Capozza & Srinivasan (2024), I employ pairwise comparisons to deduce the relative marginal social welfare weights placed on individuals both across different incomes and across tax-compliance status. Aggregating these weights under a 'median-voter' assumption yields clear results. Whilst preferences on vertical inequality are varied and broadly moderate, respondents exhibit a very strong aversion to horizontal inequalities arising due to evasion. This also affirms that, conditional on earnings, taxable workers are perceived as horizontally disadvantaged[1].

In the simulations, a striking result emerges. For reasonable elasticities and empirically determined welfare weights, the tax schedule for Kenya prescribed by this model, moves much closer to the actual Kenyan tax system. Specifically, while the simple Mirrlees model favors high taxes across the full tax base, the adjusted formulas prescribe limited tax bases and lower taxes in the presence of horizontal inequality aversion. I find that given the estimated welfare weights it is optimal to set an exemption threshold of at 200,000 Ksh. which is close to the current threshold of 288,000 Ksh. The model thus contests the commonly held orthodoxy, that base broadening is always a good policy (IMF, 2024; Tanzi & Zee, 2001; Zolt & Bird, 2005). In doing so, the model also provides a normative underpinning for the striking empirical fact recently uncovered by Jensen (2022), that the placement of the income tax exemption threshold generally moves down the income distribution in comovement with increases in the share of formally employed workers. This fact cannot be rationalized within the simple optimal income tax framework but fits the prescriptions of this model.

[1] Historically the informal sector been seen as disadvantaged subsistence sector (Gollin, 2008; Harris & Todaro, 1970; Lewis, 1954). However, recent work gives us reason to question whether informal workers are better or worse off than what can be gleaned from their consumption relative to formal workers with the same net income (Blattman & Dercon, 2018; Maloney, 2004).

To rigorously test the robustness of the results, I numerically solve the optimal non-linear problem using high-order polynomial approximation (Heathcote & Tsujiyama, 2021a). This allows for the inclusion of both participation effects, i.e., movements to informality in response to average tax rates as well as of marginal administrative costs of taxation. In the absence of horizontal inequality aversion, neither effect can rationalize any large exemption. Barring any omitted confounders, this indicates that horizontal inequality is either an important policy objective or that developing countries set highly suboptimal tax policies owing to their large exemptions (Jensen, 2022).

In summary, this paper attempts to jointly solve a theoretical and an empirical issue. Theoretically, it is the first to propose a tractable way to incorporate the missing dimension of horizontal inequality into optimal tax models. Empirically this improves the relevance of Mirrleesian tax models in a development context as the central prescriptions of the model no longer are irreconcilable with observed policy choices and as it lends us a formalized framework that can be used to study exemption policy.

1.1 Literature Review

The analysis of this paper intersects with three distinct literatures. The model builds on previous optimal income taxation models. My discussion of horizontal inequality borrows from an extensive body of work on this concept. And the focused application of the model to developing countries draws on the broader development public finance literature.

This paper builds on the foundational optimal income tax papers (Mirrlees, 1971; Saez, 2001a). Specifically, the paper relates to the literature on taxation under multi-dimensional heterogeneity (Kleven et al., 2009; Rothschild & Scheuer, 2013), however, unlike these papers, I assume that the second dimension of heterogeneity is publicly observable. Previously, Choné & Laroque (2010) have described how negative taxes can be optimal in a Mirrleesian framework under multidimensional heterogeneity. A result similar to that in this paper. However, their result assumes no tagging and arises due to social welfare weights *increasing* with income, a conjecture not aligned with the empirical results in this paper.

By implicitly conditioning on observables, the model interacts with the burgeoning literature on 'tagging' (Akerlof, 1978; Cremer et al., 2010; Weinzierl, 2012). The 'Generalized Social Welfare Weights' approach of Saez & Stantcheva (2016) is the most prominent approach for broadening the range of admissible variables in the SWF of the planner. In their paper, they also discuss taxation in the presence of horizontal inequality. However, they stop short of formalizing the vertical-horizontal equity trade-off.

There are strong conceptual links between the tax evasion literature and the theoretical literature on horizontal inequality as a principle of justice (Duclos et al., 2003; Slemrod, 2019). This is evidenced by discussions of whether it might be possible to integrate horizontal inequality or evasion directly into the SWF of the government (Slemrod & Yitzhaki, 2002, p. 1447). Such an attempt has previously been made by Cowell (1990); his augmentations are however a deviation from strict welfarism as they disregard 'unobserved' income entirely[2]. To identify a formulation that is compatible with traditional principles of welfarism and taxation, I consult the theoretical literature on horizontal inequality and adhere to the central principles of cardinality and disentangling preference for horizontal and vertical equality (Auerbach & Hassett, 2002; Kaplow, 2000; Stiglitz, 1982). I discuss this further in section 2.2.

Optimal non-linear taxation in a Roy-style model of self-selection has been previously discussed by Rothschild and Scheuer (2013). This paper differs crucially in that the assumption of uniform taxation across sectors is discarded. The implementation of the endogenous type extension using parametric tax functions draws on the work of Heathcote and Tsujiyama (2021a).

Navigating the presence of informality, both among firms and workers (Ulyssea, 2020), is an ever-present concern in development public finance. Gordon & Li (2009) argue that informality is central to explaining the unique tax systems of developing countries. In a similar spirit Besley & Persson (2014) examine determinants of limited taxation and García Peñalosa & Turnovsky (2005) examine how informality might shape the relative weight put on labor taxes. Recently, Jensen (2022) has shown empirically that the structure of the labor force (i.e., the share of formally employed workers) correlates strongly with the size of the PIT tax base.

Tax evasion and enforcement policies are necessary components of tax policy in developing contexts (Keen & Slemrod, 2017; Slemrod, 2019; Slemrod & Yitzhaki, 2002). A lot of work within this literature builds on the 'optimal evasion model' of Allingham & Sandmo (1972). However, in this paper, I only consider evasion on the extensive margin and thus abstract away from this literature.

Recently, evidence has been mounting that taxation operates as a self-reinforcing social contract (Besley, 2020; Kresch et al., 2023; Weigel, 2020). That is, evasion might have dynamic effects on morale and compliance that exceeds its immediate fiscal impact. This suggests a strong economic

[2] Cowell (1990, p. 147) briefly mentions the idea of something akin to 'horizontal Rawlsianism' but rejects it summarily due to the informational constraint on identifying evaders. This is contradicted by a central point in this paper, which is that for optimal tax policy, we do not need to know which individuals evade, merely the income distributions of evaders and non-evaders.

rationale for social preferences against evasion. Additionally, as discussed in Pomeranz & Vila-Belda (2019), high levels of evasion implies that for a given revenue target governments are forced to place a large and undue burden on compliant taxpayers.

2 Generalized Optimal Tax Model

In this section, I introduce a model of taxation under intractable evasion. The model is a simple expansion of the Mirrlees (1971) model to a heterogeneous agent framework where one type of agent has access to a costless evasion technology. I augment the conventional Mirrlees (1971) SWF to account for horizontal equality concerns. I discuss central criticisms from the horizontal inequality literature and find the optimal tax schedule given the model. Finally, I solve edge-cases of the model to find stylized facts regarding optimal tax policy in the presence of evasion.

2.1 Model Setup:

Types and Utility. Consider an economy populated by two types of agents $j \in \{1,2\}$. The types differ only in a single respect, their tax evasion technology. Type $j = 1$ agents have no technology for evasion and are subject to costless taxation. These agents are henceforth referred to as *taxable agents*. The type $j = 2$ agents have a costless evasion technology with no detection risk and are thus *untaxable agents*. In addition to their type, *j*, agents also possess a latent skill, *n*, which determines their income earning potential. Both the type and skill of an agent are initially fixed.

All agents have identical utility functions $u(c, z)$. Here, c, denotes consumption and, z, gross labor income. Agents get utility from consumption and disutility from exerting effort to earn income, z such that the ratio between income and skill determines the disutility of effort. I assume quasi linear utility to avoid income effects and simplify the analysis.

Skill Distributions. For both types of agents, the skill distribution covers $n \in [0; \infty)$ where zero indicates no potential earnings under any effort. The pdf of skills is $f(n, 1)$ for taxable agents and $f(n, 2)$ for untaxable agents. The share of taxable agents with a given skill-level is given by $\rho(n)$, and the average share of taxable agents with a skill above n is $\rho^+(n)$ such that the economy-wide skill-distribution can be characterized as:

$$h(n) = \rho^+(0)f(n, 1) + \big(1 - \rho^+(0)\big)f(n, 2)$$

The corresponding CDF is $H(n)$.

Tax Instruments. The economy is controlled by a benevolent social planner. The social planner has full information on whether the agent is type 1 or 2. The planner does not know the skill of

any individual and can only tax, based on income and type. When setting taxes, the planner knows the skill and type distributions as well as the utility function with certainty. The planner can set a fully non-linear tax schedule, $T'(z)$ and distribute a universal lumpsum $-T(0)$ which is lower than or equal to its revenues. It is not possible for the government to target the lumpsum to a single type of worker and marginal taxes are bounded such that $T'(z) \in [0;1]$[3]. The planner knows that only taxable agents will remit taxes. The problem of the social planner is thus very close to the basic optimal tax problem (Mirrlees, 1971; Saez, 2001b).

Social Welfare Function. The social planner attempts to maximize a social welfare function. Consider the most general formulation of such a function:

$$SWF = \int_0^\infty Q\big(u(n,1), u(n,2)\big)h(n)dn \tag{1}$$

Here, $Q(\cdot)$, is the social preference function. This is similar to the well-known preference function from the optimal income tax literature (Diamond & Saez, 2011; Mirrlees, 1971; Saez, 2001b) except from the fact that the social preference function, $Q(\cdot)$, now discerns between the utility of taxable and untaxable workers $u(n,2)$[4].

2.2 Defining Horizontal Inequality

This paper employs the standard definition of horizontal inequality as 'the unequal treatment of equals' (Duclos et al., 2003; Musgrave, 1990). This leaves us the difficult task of defining "equals". The invocation of the Mirrlees model in this paper fortunately lends a short-cut. Since all agents, regardless of type, have a skill level, n, which represents underlying potential earnings, it is natural to define "equals" as agents with the same level of skill, n. Consequently, the assumption of identical utility functions implies that the difference in ex-post utility is a natural measure of horizontal inequality across agents with identical skill. This definition of horizontal inequality, which is tied directly to the utility of otherwise identical agents has the notable benefit of being consistent with a welfarist SWF[5] and applicable under endogenous labor supply. The same cannot be said for the various rank-based approaches which lose their welfarist basis by their auxiliary measure (Aronson et al., 1994; King, 1983).

[3] This restriction is discussed further in section 3.1 and loosened in appendix 9.1.
[4] Notice that the planner is maximizing over the unobservable n here, since the allocation adheres to the Spence-Mirrlees condition, there is a unique mapping for the chosen tax system from $n \to z$ and the change in notation is benign. See A9.3 for a full proof.
[5] I use the term 'welfarist SWF' in the sense that it is individualistic and Paretian (Kaplow & Shavell, 2001)

Kaplow (1989, 2000) has argued that there is no basis for invoking horizontal inequality as a distinct social objective. I argue that much of this criticism pertains to specific methods for measuring horizontal inequality proposed in the past, not to this model. Additionally, evidence from the survey presented in section 4 indicates strong preferences against horizontal inequality. However, since this model is a generalization, it includes as an edge-case the possibility of a social planner having no aversion to horizontal inequality in line with the arguments of Kaplow.

At its core, the problem of the social planner is trading of the gains of equality with the cost of redistribution. Stiglitz (1982) argues, that in order to accept horizontal inequality as a valid normative criteria, it will be necessary "to trade off horizontal equity with other desiderata of a good tax system" (Stiglitz, 1982, p. 28). The SWF presented in the subsequent section navigates this by allowing for completely independent preferences for horizontal and vertical equality. The trade-off with efficiency is captured by the Mirrleesian foundation of the model. This is similar in spirit to the approach suggested by Auerbach & Hassett (2002)[6].

The generalized social welfare weights (GSWW) approach suggested by Saez & Stantcheva (2016) has proven useful as it allows for independent weights on horizontally inequitable outcomes without the need for a globally specified social welfare function. However, as shown by Sher (2024), deviations by this approach from structural utilitarianism leads to inconsistent prescriptions. Additionally, the approach exemplified by Saez & Stantcheva (2016) only captures the case where horizontal equity is the overriding social objective and as such neglects the trade-off with vertical equality. By integrating horizontal inequality into the global social welfare function and solving it in an optimal tax context this paper resolves these issues.

Given the definitions presented, a reasonable concern is whether the assumption that taxable workers are 'horizontally disadvantaged' compared to untaxable workers of the same skill is defensible. This concern is well exemplified by asking why taxable workers are not switching to the untaxable sector to increase their utility? There are two ways to counter this concern. It could be the case that the labor market is segmented, and workers are tied to a sector by some immutable characteristic such as their place of birth. Alternatively, we can consider a case where taxable workers indeed would receive lower utility in the untaxable sector but where the social planner only cares about fairness and material well-being and thus only allows 'earnings', 'skill' and

[6] Notably Auerbach & Hassett (2002) also propose a nested structure to incorporate dual motives in a Paretian framework. However, their approach relies on classifying equals using a kernel and a functional form assumption which results in a larger number of arbitrary policy parameters. Further, their SWF is not applicable in an optimal tax context as it assumes no behavioral effects.

'consumption' as permissible variables in the SWW[7]. Fortunately, the direct elicitation of the marginal SWWs of Kenyans in section 4 circumvents this issue altogether and allows us to remain agnostic on this issue[8].

2.3 The Generalized Social Welfare Function

Building on the discussions in section 2.2 I disentangle vertical equality preferences, $G(\cdot)$, from horizontal equality preferences, $H(\cdot)$, using a nested structure.

$$SWF = \int_0^\infty H\left(G\big(u(n,1)\big), G\big(u(n,2)\big), \rho(n)\right) h(n)dn \tag{2}$$

Here $G(\cdot)$ follows the optimal income tax literature such that it has a positive first derivative and is weakly concave in utility. The horizontal preference function, $H(\cdot)$, thus navigates the *social value adjusted* utilities, $G(u)$, under the possibility of horizontal inequality. I propose the following functional form for $H(\cdot)$.

$$SWF = \int_0^\infty \left(\sum_{j=1}^{2} G(u(n,j))^\phi \cdot \rho(n,j)\right)^{\frac{1}{\phi}} h(n)dn \tag{3}$$

Here $H(\cdot)$ is the generalized power mean formula[9]. In this paper we have two types, however this formulation is easily adapted to any number of types, even continous, as is shown in appendix 9.1. The parameter ϕ captures the degree of *horizontal inequality* aversion. For $\phi = 1$, this collapses to a weighted average of the utility of the two types. As this implies a simple maximization of social value adjusted utility, I denote this *horizontal utilitarianism*. For $\phi \to -\infty$ the outer function becomes a minimum function, assigning the lowest utility to all workers of a given skill. I denote this *horizontal Rawlsianism*. As such, we can reasonably impose bounds on this such that $\phi \in (-\infty; 1]$. For a more flexible functional form, ϕ can also be endogenized to vary across skill.

Accommodating the criticisms of Stiglitz (1982), this SWF navigates the trade-off between efficiency, vertical equality, and horizontal equality as follows. If $G(\cdot)$ is linear and $\phi = 1$ then we have a standard utilitarian planner. If $G(\cdot)$ is linear and $\phi \to -\infty$ then the planner places no value on additional utility attained through tax evasion but remains strictly utilitarian across skills. For

[7] This idea is developed further in Saez & Stantcheva (2016)
[8] That is, concerns regarding non-pecuniary aspects of working in either sector as well as of the degree to which the choice of taxability is indeed voluntary, will be captured by these weights. However, there is of course no way to establish that the positively determined marginal SWWs are also the ethically correct ones.
[9] This approach is related to the Atkinson (1970) approach to inequality measurement.

$G(u) = \min(u)$ the planner is a typical Rawlsian and only assigns weight to the agent with the lowest utility.

Finding Optimal Rates

In this section, I develop the intuition for the optimal tax rates under horizontal inequality. The full proof can be found in appendix A9.3 and an intuitive perturbation-based approach in A9.2.

Using the SWF proposed, we can identify the marginal social welfare weights (SWW) assigned to any individual by the planner. This corresponds to the money-metric social value of one additional unit of consumption for an agent of type j and skill n:

$$q(n,j) = \frac{\partial SWF}{\partial c_i} \cdot \frac{1}{h(n) \cdot \rho(n,j)} = \frac{Q_{G|j}(\cdot) \cdot G_u(u(n),j) \cdot u_c(n,j)}{\lambda \cdot \rho(n,j)} \tag{4}$$

I can collapse this expression using a weighted average $\bar{q}(n) \equiv \rho(n)q(n,1) + \big(1-\rho(n)\big)q(n,2)$. The assumption of no income effects implies that the average marginal SWW across all agents has to be unity. I thus find the marginal value of public funds (MVPF) as:

$$\lambda = \int \bar{q}(n) \cdot h(n) dn \tag{5}$$

This allows for the determination of any $q(n,j)$ and I can thus find the average marginal SWW for taxable agents above a given income $q^+(n,1)$. As the optimal allocation fulfills the Spence-Mirrless condition I can recast such that $\tilde{q}^+(z) = q^+(n(z,1),1)$. Using the optimal control or the perturbation method (Diamond & Saez, 2011; Kleven, 2021; Saez & Stantcheva, 2016) we arrive at the optimal non-linear tax schedule:

$$T'(z) = \frac{1-\tilde{q}^+(z)}{1-\tilde{q}^+(z)+\alpha(z,1)\cdot e(z,1)} \tag{7}$$

This is similar to well-known sufficient statistics[10] optimal income tax formulas (Diamond & Saez, 2011; Kleven, 2021; Saez, 2001b) with two minor differences. First, the Pareto parameter α only describes the gross income distribution of taxable workers. Second, the average marginal SWWs only cover taxable workers. This implies that even for Paretian preferences the value of $q^+(z,1)$ can now exceed 1. For $q^+(z,1) \leq 1$ we observe normal behavior and $T'(z) \in [0,1]$. For $q^+(z,1) \in \big(1; 1+\alpha(z,1)\cdot e(z,1)\big)$ it is optimal for the government to set $T'(z) < 0$ and subsidize the earning

[10] As pointed out in Kleven (2021), this is not actually a *true* sufficient statistics formula as it presupposes knowledge of all the parameters at the optimum which requires a more structured approach.

of taxable workers. For $q^+(z,1) > 1 + \alpha(z,1) \cdot e(z,1)$ the formula instead identifies the local minimum of the SWF. In section 3.1 I discuss how this is of limited empirical relevance.

Horizontal Utilitarianism and Horizontal Rawlsianism

To illustrate the impact of the expanded model, I compare three special cases: a 'naïve' planner who disregards untaxable workers when setting rates (similar to current models), a 'horizontally utilitarian' planner who does not pursue horizontal inequality reduction as an intrinsic goal, and a 'horizontally Rawlsian' planner who is highly averse to horizontal inequality.

Both the naïve and the horizontally utilitarian planner will assign marginal SWWs as follows:

$$q_k(n,1) = \frac{1}{\lambda_k} \cdot G_u\big(u(n,1)\big) \tag{8}$$

Where $k \in \{0,1\}$. Notice here that the only difference in SWWs arises through the different marginal value of public funds, λ_k, between the two cases. This gives the condition for equal weights as:

$$\begin{gathered} \lambda_{horizontalUtilitarian} = \lambda_{naive} \quad <=> \\ \int G_c\big(u(n,1)\big) f(n,1) dn = \int G_c\big(u(n,2)\big) f(n,2) dn \end{gathered} \tag{9}$$

As intractable evasion only affect tax rates through $\tilde{q}^+(z)$ this implies:

> Proposition 1: The social planner can disregard untaxable agents when setting taxes only in the knife-edge case where the average marginal social utility of consumption, $G_c(u(c))$, is identical across horizontally inequitable groups.

The intuitive implication here is that when we extend our analysis to cover untaxable agents then the recipients of the lumpsum grant changes which in turn affects the marginal value of public funds. If the untaxable workers have lower socially adjusted utilities, then a utilitarian planner will have a higher marginal value of public funds and correspondingly increase taxes.

For the horizontally Rawlsian planner we see that when $\phi \to -\infty$, the outer horizontal preference function converges to $H(G_{|j=1}, G_{|j=2}) = \min(G_{|j=1}, G_{|j=2})$. This implies that $Q_{G|j=2} = 0$ and $Q_{G|j=2} = 1$ if $u(n,2) > u(n,1)$. As such, the marginal SWWs are given by:

$$q_2(n,1) = \frac{1}{\lambda_2 \cdot \rho(n)} \cdot G_u\big(u(n,1)\big) \quad and \quad q_2(n,2) = 0 \tag{10}$$

This also gives us the MVPF as $\lambda_2 = \int G_u\big(u(n,1)\big)h(n)dn$. As by definition $\rho(n) \in [0;1]$, I can conclude the following general relation:

> Proposition 2: For a given marginal value of public funds, optimal taxes are lower for a horizontally Rawlsian than a horizontally utilitarian planner.

This, all else equal, pulls in the direction of taxes decreasing with horizontal inequality aversion. However, the relation is ambiguous. In some cases, the horizontally Rawlsian planner may value funds more as she can more effectively alleviate two dimensions of inequality simultaneously.

These two propositions give some basic intuition on the impact of intractable evasion on optimal tax rates. Proposition 1 shows that even if we do not accept horizontal inequality aversion as a valid social objective then intractable evasion still has pronounced effects on optimal rates. Proposition 2 helps illuminate the two channels through which horizontal inequality aversion impacts optimal rates.

Employment Structure and the Tax Exemption

In this section I investigate an important stylized fact uncovered in Jensen (2022). He finds that over the course development, as employment shares increase, the tax exemption threshold moves down the income distribution.

The presence of untaxable workers in the expanded model breaks the classic result of non-negative optimal taxes. As discussed in 3.1, negative income taxes are broadly infeasible in a development context. To account for this, I expand the unconstrained problem with a Kuhn-Tucker inequality condition in A9.4. This yields a simple condition for the constrained tax, namely $T_c'(z) = \max[0, T'(z)]$.

Interestingly, this allows us to derive some stylized insights on optimal exemption policy. We can derive the optimal exemption as $T_u'(\bar{z}) = 0$ which is equivalent to identifying $q^+(\bar{z}, 1) = 1$. Rather than identify $\bar{z}$ directly I investigate the conditions under which the Jensen-rule holds i.e., where an increase in the employment shares yields a corresponding decrease in the tax exemption. For simplicity I take a local approach and identify when a local increase in the share of taxable workers at the exemption decreases social weight on taxable workers $\frac{\partial q_1(n)}{\partial \rho(n)} < 0$. The full derivation and interpretation is presented in appendix A9.5.

Proposition 3: Given a local increase in the share of taxable workers at or above the exemption threshold, the optimal size of the exemption decreases if and only if $\phi < 0$.

Thus, horizontal inequality can provide a normative underpinning for the Jensen-rule if governments are sufficiently averse to horizontal inequality averse such that $\phi < 0$. The cutoff, $\phi = 0$, corresponds to a horizontal preference function being a weighted geometric mean. This lends the proposition a corollary: as long as a given relative decrease in $G\big(u(n,1)\big)$ (i.e., the social utility of the horizontally disadvantaged worker) with a symmetric increase in the social utility of advantaged workers, $G\big(u(n,1)\big)$, *decreases* social welfare, then the Jensen-rule holds[11].

In section 4, I find evidence that the horizontal inequality aversion of Kenyans is far stronger than this. Interestingly, if governments have no intrinsic preference for horizontal equality, i.e., $\phi = 1$, then the model predicts no direct correspondence between the share of taxable workers and the exemption threshold.

3 Empirical Application: Personal Income Taxation in Kenya

3.1 Setting: Personal Income Taxes in LICs

Personal income tax systems in developing countries differ markedly from those in developed countries in several ways. On an aggregate level, a lower share of taxes is collected through PIT in LICs, with higher shares coming from, VAT and trade taxes (Besley & Persson, 2014). Additionally, developing countries have very small income tax bases owing, in part, to comparatively large tax exemptions. In a large cross-country sample Jensen (2022, p. 225) finds that countries with per capita incomes below 8,000 USD (2011 PPP) on average set exemption thresholds that exclude around 90 pct. of workers from the tax base. He argues that this fact owes to the high-prevalence of self-employment in these countries (Jensen, 2022).

Building on this insight, this paper defines 'untaxable workers' as the combination of self-employed workers and salaried workers not subject to tax withholding. This allows me to shed light on plausible mechanisms behind the stylized facts uncovered in Jensen (2022).

The desirability of narrow income tax bases in developing economies is contested. For developed as well as developing economies, the prevailing wisdom is that it is optimal to keep tax exemptions

[11]More intuitively, consider a vertically utilitarian government. If a symmetric relative increase in the income of the horizontally advantaged with a corresponding decrease for the disadvantaged, lowers social welfare despite increasing total earnings, then the Jensen-rule holds.

low and tax bases large (Fall, 2022; OECD, 2012). The intuition underpinning this is that narrow tax bases puts too much pressure on raising revenues from a limited number of taxpayers thus reducing efficiency. In this paper I argue that these prescriptions are generally informed by an overly simplistic view of optimal taxation in developing economies.

Negative income taxes are not generally seen in LICs as they are administratively costly and difficult to implement. Furthermore they are much more vulnerable to cheating than regular taxes.

3.2 The Kenyan Tax System

The personal income tax system in Kenya derives a large majority (98 pct.) of its tax revenues from withholding taxes (PAYE) (Kanina et al., 2025). This validates the assumption that informal (untaxable) workers do not remit taxes. These revenues predominantly stem from public employees (39 pct.) and employees of large private firms (35 pct.)[12].

In 2021, all Kenyan taxpayers were entitled to a tax relief of 28,800 Ksh/year, a sizeable expansion from the previous year. Given that the first bracket of the Kenyan tax system taxes the first 288,000 Ksh earned at a flat rate of 10%, this relief system is nearly equivalent to having a zero-rated bracket for earnings below 288,000 Ksh.[13]. As such, the Kenyan system of 2021 is essentially a two-bracket system with an exemption at 288,000 Ksh, earnings exceeding 288,000 Ksh taxed at 25 pct. and earnings exceeding 388,000 Ksh taxed at 30 pct. In addition to PIT, Kenyan workers also pay various social contributions to health and social insurance funds. For the purposes of this paper, I do not consider these expenses as taxes. I examine counter-factual tax schedules for Kenya in 2021. I choose this year for analysis as this is the only year for which survey data is available.

The administrative dataset on Kenyan personal income taxes comprises eight years from 2015 until 2022. I use a cross-checked dataset of matched employee and employer returns created by Kanina et al. (2025). The authors manage to match 66 pct. of employee records to corresponding employers. For the remaining observations only employee *or* employer data was available. Although the data is in panel format, I only make use of a single cross-section to match the static nature of the model.

I use data on 'taxable income' which is defined as the sum of 'employment income' and 'other income' net of deductions. In 2021, 55 pct. of registered Kenyan workers earned taxable income above the official exemption threshold. However, 67 pct. of registered workers remitted some

[12] These shares are averages from 2017-22 as reported in Garces et al. (Forthcoming)
[13]As the relief is applied on a monthly level, it is technically possible to incur a positive tax liability without earning above 288,000 Ksh/year. The data illustrates that this is only an issue for a small minority of taxpayers.

positive taxes. The tax payments of the workers with incomes below the exemption only comprises 0.66 pct. of total PIT revenue. In the informal sector 89 pct. of all workers earn below the exemption threshold. Given this, I infer that 9.7 pct. of the current Kenyan labor force were evading on the extensive margin in 2021[14]. The availability of administrative data simplifies analysis considerably as it removes the need to account for the fact that high earning formal workers are likely to be undersampled in household surveys.

3.3 Finding Income Distributions

A benefit of the sufficient statistics formulation of the optimal tax problem is that the model can be solved with only a few data points, namely:

- the distribution of gross incomes for taxable workers, $\tilde{f}(z,1)$
- the distribution of gross incomes for untaxable workers, $\tilde{f}(z,2)$
- the share of taxable workers in the labor force, $\rho^{+}(0)$
- the elasticity of taxable income, $e(z,1)$
- the marginal social welfare weights on taxable workers, $q(z,1)$

I find the distribution of taxable workers using the full universe of taxpayers from administrative data provided by the Kenyan Revenue Authority. I define z as gross taxable income. To identify the income distribution of untaxable workers I use the 2021 Kenya Continuous Household Survey (KNBS, 2021a). This survey has data both on whether a worker has a written contract and whether she is subject to withholding taxes (PAYE). As it is apparent that only very few non-PAYE workers pay any taxes, I define informal (untaxable) workers as any worker between the ages 18-65 who works with gross earnings above 1,500 Ksh/month who does not experience withholding taxes. Whilst the KCHS has almost complete coverage of the incomes of salaried informal workers, many self-employed workers do not report their earnings. To remediate this, I follow Kanina et al. (2025) and impute gross income for the workers with no reported income using data on the size of their household and their total consumption. I report a comparison of income distributions with and without imputations in appendix A9.6 .

The administrative records provided identify the exact number of taxable workers in 2021. I identify 2,702,080 individual taxpayers. To identify the share of taxable agents I use the Kenyan National Bureau of Statistics estimate for the total size of the Kenyan labor force in 2021 of

[14] The calculation is as follows: 11 pct. informal workers with positive tax liability multiplied with 85.6 pct share of informal workers in the total labor force.

19,106,156 workers (KNBS, 2021b) which implies that taxable agents comprise $\rho^{+}(0) = 14.1\ pct.$ of the labor force.

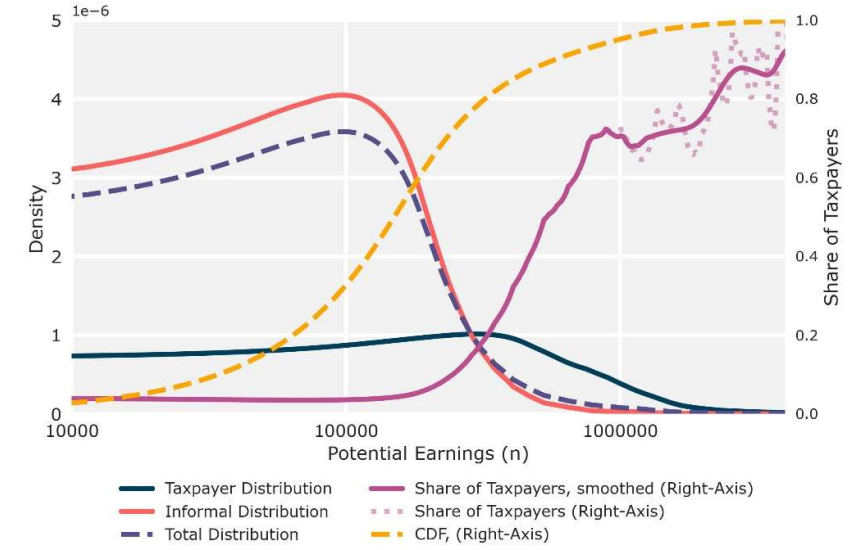
I employ the Saez (2001) calibration method to move from observed earnings to our measure of skill, potential earnings. For this I use the elasticity of taxable income for taxable workers in Kenya in 2021. In a recent study, Kanina et al. (2025) use tax reforms from 2020-21 to identify the ETI of private-sector workers with incomes in the range 375,000 – 975,000 Ksh (2022-prices) as $e = 0.51$. Including public-sector workers leads to a somewhat lower elasticity of 0.36. I use 0.51 as the main elasticity in the calculation as the lower elasticity in the public sector may be related to government wage policy as well.

Assuming of iso-elasticity, allows for the identification of the skill parameter using the actual Kenyan tax schedule as:

$$n_i = \begin{cases} \dfrac{z_i}{\left(1 - T'(z_i)\right)^{e}} & if\ j = 1 \\ z_i & if\ j = 2 \end{cases} \tag{11}$$

This implies the following skill distributions:

Figure 1: **Skill Distributions, Kenya 2021**



Note: Taxpayer and Informal Distributions are smoothed using an Epanechnikov kernel with bandwidths of 200,000 and 75,000. Gaussian smoothing is applied to the "Share of Taxpayers" for potential earnings above 900,000. The first axis is log-transformed.
Source: KCHS 2021, KRA, and own calculations

This plot shows that a large majority of low-skilled workers are informal (untaxable). Only for skills above $n = 540{,}000$ does the prevalence of taxable workers exceed that of untaxable ones.

4 Direct Elicitation of Marginal SWWs

Standard measures of redistributive preferences do not easily feed into optimal-tax models because they lack a welfare-weight interpretation[15] and until recent work by Dwenger et al. (2024) very little research on preferences regarding horizontal inequality had been available. To bridge this gap, I directly elicit marginal social welfare weights (MSWWs) via a controlled redistribution experiment capturing both vertical and horizontal inequality[16]. My approach draws on a major step forward by Capozza & Srinivasan (2024) who use series of pairwise comparisons to identify the relative marginal social welfare weights (MSWWs) assigned to agents of different incomes. This survey uses their central innovation of pairwise comparisons.

4.1 Methodology

Experimental Task

I elicit marginal social welfare weights using a continuous 'leaky-bucket' allocation task (Okun, 1975). Respondents allocate a share $s \in [0,1]$ of a hypothetical Ksh 1,000 between two recipients R1 and R2. In the experiment R2 always has lower consumption than R1. Allocating more to R2 shrinks the total endowment. Under standard envelope arguments, the slope of the frontier at the chosen allocation reveals the ratio of marginal social welfare weights:

$$\frac{q(z_1, j_1)}{q(z_2, j_2)} = -\frac{\partial c_2\big(s_1(c_1)\big)}{\partial c_1} \tag{12}$$

The endowment shrinks at rate such that the choice set, $s \in [0,1]$, covers the full set of weakly redistributive preferences from $s = 0$ where the participant assigns equal weight to both recipients, to $s = 1$ where the participant only values the consumption of the disadvantaged recipient. Allowing s to vary continuously enhances measurement precision and mitigates intransitivity issues noted in discrete designs (Capozza & Srinivasan, 2024).

Survey Design

Each respondent completes three rounds of distribution tasks, and in each round, they make two allocations. In the first allocation, they divide the leaky-bucket endowment between two tax-compliant individuals whose incomes differ ($z_1 > z_2$); each income pair is drawn at random from the empirical income distribution so that responses cover the full earnings spectrum. In the second

[15] Survey measures of fairness typically report ordinal attitudes, not the cardinal weights needed in welfare functions.(Almås et al., 2020; Bellemare et al., 2008; Charness & Rabin, 2002)[15].

[16] The experiment was pre-registered at AEA https://www.socialscienceregistry.org/trials/15719

allocation of the same round, they choose between two recipients with identical incomes ($z_1 = z_2$) but different tax-compliance statuses. This structure ensures that we observe both vertical inequality tradeoffs (across incomes) and horizontal inequality tradeoffs (across compliance) for each respondent.

The *gain-gain* framing inherent in distributing a shrinking lump-sum between two recipients is useful as it removes any endowment effects as neither recipient a priori has a bigger claim to the bundle. However, to test whether framing effects significantly impact redistributive tastes, we expose 20 pct. of respondents to an alternative *loss-gain* where respondents are able to take up to 1000 Ksh from R1 and transfer it to R2 at a cost. This framing is arguably more similar to how taxes work but also faces the risk of endowment effect biases. (Capozza & Srinivasan, 2024; Charité et al., 2022).

Aggregating preferences across respondents requires difficult methodological choices. While previous approaches (Capozza & Srinivasan, 2024) make a parametric assumption to fit a functional form capturing progressivity, this paper directly identifies non-linear social MSWWs[17].

The parametric approach needs the income of one of the recipients ("common recipient") to remain constant across all questions. Unfortunately, Capozza & Srinivasan (2024) find that the income of the 'common recipient' directly impacts the estimated progressivity of the MSWWs.

In this paper, the social preferences are not restricted to being parametric and there is thus no common recipient. Rather, for every combination of discrete incomes, the relative social welfare weights $\frac{q_1}{q_2}$ are determined as the median of the relative welfare weights assigned to the two types. In appendix *A9.10*, I formalize this aggregation approach and illustrate how it is a consistent estimator of true median MSWWs, if the underlying social preferences are transitive. This consistency result notably only holds for the non-linear approach and not for the parametric one.

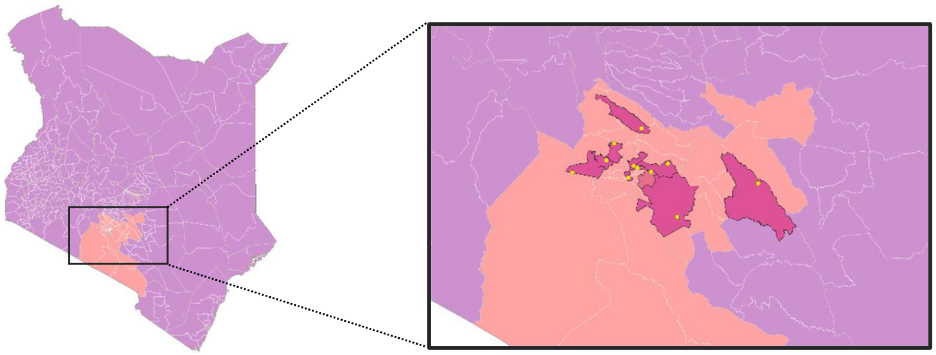
Sampling:

Data was collected in June 2025 via in-person tablet interviews in 11 subcounties across Nairobi, Kiambu, Machakos and Kajiado, covering urban, peri-urban, and rural contexts. Subcounties were selected with probability proportional to population size, and within each a random start location was drawn. Enumerators followed a randomized "walk" procedure visiting every second

[17] These papers employ a common 'Common Recipient' in their surveys meaning that each respondent across choices only sees variation in the income of one of the agents in the redistribution problem. This allows them to back out a complete ranking across these income. However, to aggregate these further to a social ranking OLS is used to represent the preferences of each individual as a single parameter value. It is not clear that the use of OLS is unproblematic as it implicitly assigns a specific cardinal interpretation to the MSWWs of an individual.

household along a spun-pin direction and interviewed one eligible adult per household. They were trained to explain the efficiency–equity trade-off underlying the task without influencing responses. The sample comprises 654 completed interviews. The sample is reweighted to match national averages. Pre- and post-weight balance is reported in A9.7.

Figure 2: **Sampling Framework**

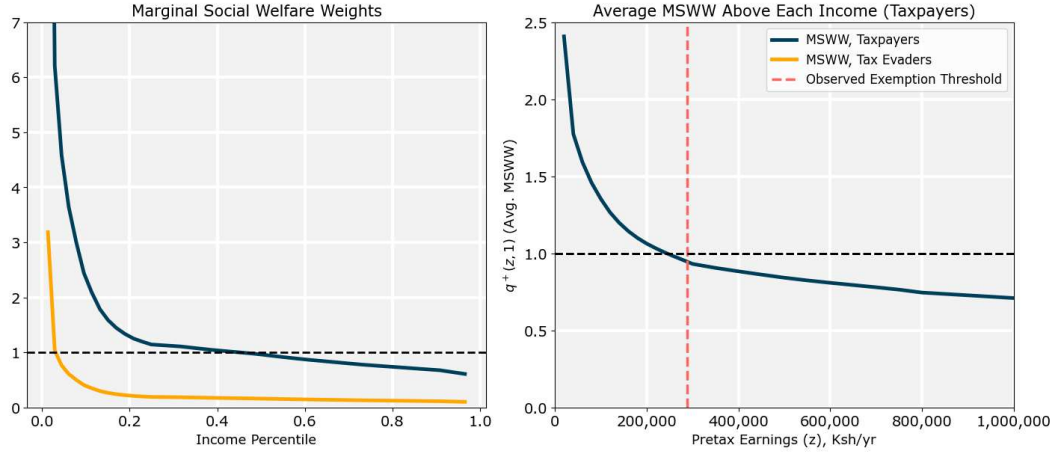


Source: Light pink denotes Nairobi, Kiambu, Machakos and Kajiado county. Dark pink denotes the sub-counties randomly selected. Yellow dots are randomly selected starting points.

Results

Using the bottom-up aggregation method as outlined in A9.10 yields the following schedules.

Figure 3: **Empirical Estimates of Marginal Social Welfare Weights, Kenya 2025**



Source: Survey data, Kenya 2025 and author's calculations.

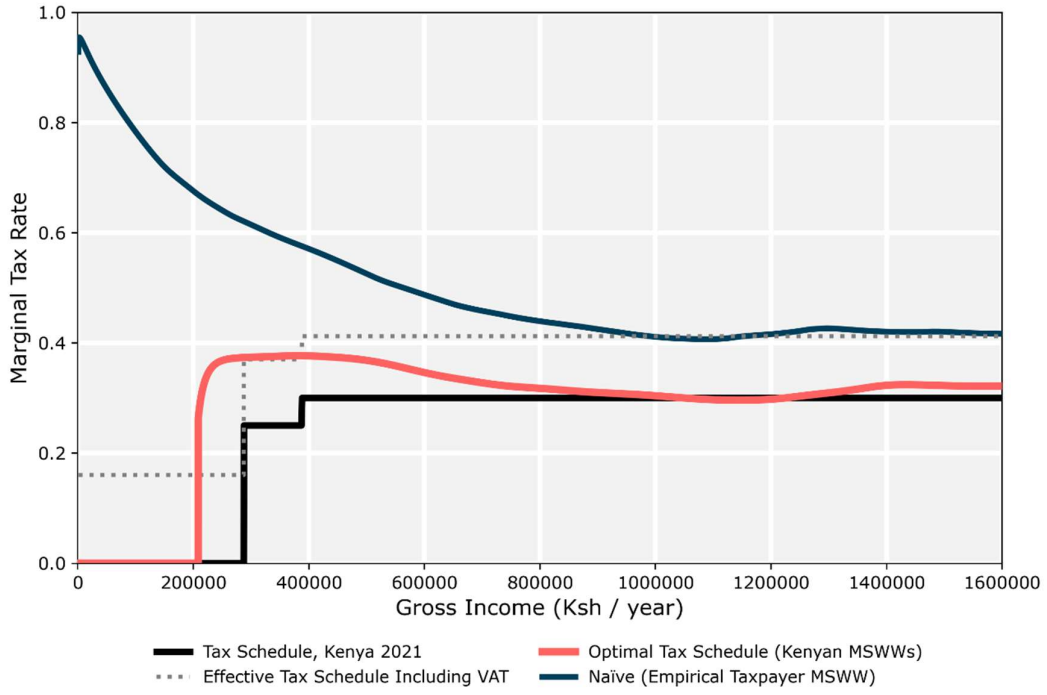
Preferences across incomes are moderately progressive, we see that a taxpayer at the 10th percentile of the distribution is given approximately three times the social weight of consumption as someone at the 40th percentile. In contrast, we observe a very strong aversion to horizontal inequality with the average ratio between social weights assigned to taxable and untaxable workers of identical incomes of 6.

Combining these schedules allows us to examine the implications for the optimal placement of the exemption threshold. The survey implies that $\tilde{q}^{+}(\bar{z}) = 1$ for $z = 200{,}000\ Ksh/yr$. This threshold is slightly lower than the observed threshold of 288,000 Ksh/yr. Using *loss-gain* framing yields almost identical results indicating no or small framing effects.

5 Simulation Results: Optimal Taxes in Kenya

In this section, I present optimal tax simulations for Kenya accounting for the presence of untaxable workers and horizontal inequality aversion. I compare the schedule implied by the empirical MSWWs with a schedule naïvely disregarding informal workers[18]. I plot actual marginal tax rates for Kenya and the marginal effective tax rate in the edge-case where everyone pays full VAT on all consumption. The true METR is likely closer to the VAT exclusive schedule especially for lower incomes (Bachas et al., 2020). In appendix A9.8 I also present simulations using standard parametric utility and social welfare functions.

Figure 4: **Tax Prescriptions, Kenya 2021**



Note: For incomes above 500,000 Ksh, I apply Gaussian smoothing to the series. "Kenyan MSWWs" indicates that the series uses MSWWs for taxpayers scaled by a MVPF accounting for untaxable workers as well, "Empircal Taxpaye MSWW" indicates MVPF only based on the MSWWs of taxable workers.
Source: Simulation results based on data from Kenyan Continuous Household Survey and Kenya Revenue Authority as well as collected survey data.

In figure 4, we observe that disregarding the presence of the informal sector leads to very

[18] In appendix A9.8 I also plot the schedules for standard parametric functional forms.

misguided tax prescriptions. I compare two tax prescriptions: *(i)* a 'naïve' prescription assuming that the planner disregards untaxable workers and uses the empirical MSWWs for taxable workers in Kenya to set rates and *(ii)* an optimal prescription where the government correctly accounts for the presence of an informal sector.

Comparing the two prescriptions, it is apparent that the naïve planner fails to impose any tax exemption, instead levying marginal tax rates in excess of 70 pct. on earnings below 200,000 Ksh/year. This is in stark contrast to the 'optimal' prescription where an exemption covering the first 200,000 Ksh/year is prescribed. The marginal tax rates of the two schedules converge towards earnings of 500,000 Ksh/year after which the 'naïve schedule' remains approximately 10 percentage points higher than the optimal schedule. Thus a 'naïve' planner sets taxes that are too high especially for lower incomes.

The exemption occurs because we observe empirically that the average MSWWs on taxpayers is above 1 ($\tilde{q}^{+}(0) = 2.4$). This indicates that the planner aims to distribute money towards taxable workers rather than away from them. An unconstrained planner would set ex-post optimal tax rates that equalize the average marginal social welfare weights between taxable and untaxable workers to maximize social welfare[19]. However, as the planner is constrained by the need to set non-negative taxes, the ex-post the social weight on taxable workers remains larger.

Unpacking the shape of the horizontally Rawlsian schedule, observe the following. The exempt region for incomes below 200,000 Ksh/year arises because $\tilde{q}^{+}(z) > 1 \; if \; z < 200{,}000 \; Ksh$. This implies that the government would prefer to set negative tax rates for these incomes. As this is not possible, an exemption is employed instead. There is no discontinuous jump in the tax rate at the exemption. Rather, as the average SWW on taxable workers is decreasing rapidly in this region, marginal tax rates increase rapidly as well. This leads to a very small gradient for $\frac{\partial z}{\partial n}$ at the threshold which means that taxes rise very rapidly as a function of gross income, z.[20]

Accounting for untaxable workers, it is optimal to exempt taxable workers with earnings under 200,000 Ksh from paying any personal income taxes. The parametric approach yields an even higher prescription of 260,000 Ksh. These are very close to the actual exemption of 288,000 Ksh. This result is in complete contradiction to the prescriptions of the conventional intensive Mirrlees tax models where it is never optimal to set tax rates at exactly zero (Diamond & Saez, 2011;

[19] To understand this, consider that when z→0 the behavioral cost of taxation embodied by $\epsilon\alpha$ also approaches zero. Thus, it is optimal to adapt taxes at the bottom until the welfare weights are equalized.

[20] Since the marginal tax function remains continuous following the imposition the non-negativity constraint this does not in itself lead to bunching at the threshold. However, the Saez (2001) simplification of local-linearity of the tax system does not hold at the exemption point which can lead to slight second-order bunching.

Mankiw et al., 2009). As such we see that solving the theoretical issue of horizontal inequality also provides us with a solution to the empirical issue of a lack of exemptions in optimal income tax models. Intuitively, the objective of the government to reduce horizontal inequality in this case proves sufficiently strong for it to be worthwhile to forego additional tax revenues in order to impose horizontal equality on low-skill agents through a sizeable exemption.

6 Expanded Model with Extensive Margin and Admin Cost

Whilst horizontal inequality aversion presents *a* possible normative underpinning for the sizeable Kenyan tax exemption, two common alternative hypotheses deserve scrutiny. First, the fixed type assumption can be interrogated. If the tax system induces shifts to informality, it is possible that this could rationalize the large exemption. Second, high marginal tax administration costs could motivate a large exemption to maintain collection efficiency. In this section I expand the model to account for these possibilities and find that neither can rationalize an exemption. This analysis is done for the parametric formulation of the problem.

To account for these effects, I move beyond the Mirrleesian solution to the optimal tax problem to instead solve the problem with numerical optimization techniques using a high-order tax polynomial following the approach of Heathcote and Tsujiyama (2021a). In appendix A9.9, I show that this technique faithfully recreates the non-linear results for the standard formulation of the problem.

6.1 Extensive Margin Effects

To accommodate the possibility of workers switching sectors in response to the tax policy, I introduce the function $\tilde{u}$ which captures the net utility gain of working in the taxable sector for a worker with skill level n compared to the untaxable sector. I include a Roy-style sectoral preference term, $\omega(n)$, in this function to capture the non-pecuniary aspects of sectoral choice.

$$\tilde{u}(n) = u(z(n,1), c(n,1), n) - u(z = n, c = n, n) + \omega(n) \tag{12}$$

Workers will choose to work in the formal sector when $0 < \tilde{u}(n,c)$. I define the distribution of $\omega(n)$ to fulfill two criteria. I force an isoelastic participation elasticity to the taxable sector. Further I impose that the distribution of $\omega(n)$ must replicate the observed shares of taxable workers for

the ATRs in Kenya in 2021[21]. To maintain comparability with previous results, the preference parameter ω does not enter the social welfare function[22].

Little is known about how large the participation elasticity is in a developing context. One study by McKay et al. (2024), looking at four East African countries, finds statistically insignificant participation elasticities close to zero. Studies of middle-income countries find statistically significant but modest participation elasticities (Haanwinckel & Soares, 2021; Ulyssea, 2020). When simulating the extensive margin model, I employ both a conservatively large elasticity of 0.5 as well as a more realistic elasticity of 0.1.

6.2 Marginal Administrative Costs

Maintaining records on, enforcing, and processing a large number of taxpayers will always come at some budgetary cost to the government. Theoretically this could constitute a rationale for the government to exclude low-earning taxpayers from the tax base (Moore, 2023). To provide a simple upper-bound estimate of the effect administrative costs on optimal taxation, I assume constant marginal administrative costs. I additionally assume that this cost applies to untaxable workers with earnings above the exemption as well. Denoting the marginal cost as γ the government budget constraint becomes.

$$T_0 = \rho^+(0)\left(\int_0^\infty \left[T'(z)\left(1-\tilde{F}(z,1)\right)dz\right] - \left(1-H(\bar{z})\right)\gamma\right) - E \qquad (13)$$

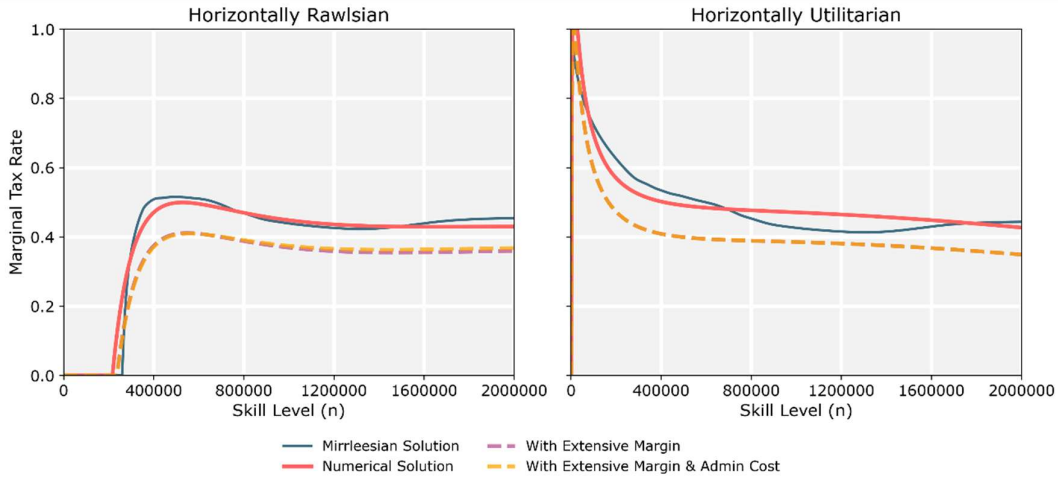
Where T_0 is the lumpsum, $1-H(\bar{z})$ is, the share of taxable workers with earnings above the exemption $\bar{z}$ and E is the exogenous budget requirement of the government. To set a credible upper bound, I assume a marginal administrative cost for each additional taxpayer of $\gamma =$ 5,000 *Ksh per year* corresponding to approximately 38 USD. This upper bound is *very* conservative seeing as Kenya operates a digital Pay-As-You-Earn system requiring limited individualized processing[23].

[21] I define the sectoral elasticity as $\zeta = \frac{dP[Formal]}{d(1-ATR)} \cdot \frac{ATR}{P[Formal]}$. For the simulations, since it is only the share of taxable workers, $\rho(n)$, which is of interest, it is not necessary to identify the specific functional form of $\omega(n)$.

[22] Including the Roy parameter would imply that the choice of tax schedule would additionally consider how sectoral satisfaction varies with income. This parameter is neither readily quantifiable nor necessarily relevant in an SWF.

[23] To exemplify this, consider the following. The total cost of collection for the Kenyan Revenue Authority in 2021 was 31 billion Ksh. If the marginal cost of including workers in the tax base was indeed 38 USD, then removing the exemption would have to imply a 19 pct. rise in the *total* cost of collections across *all* tax bases.

Figure 5: **Optimal Taxation with Extensive Margin and Admin Costs**



Note: incomes above 500,000 Ksh, I apply Gaussian smoothing to the series.
Source: Simulation results based on data from Kenyan Continuous Household Survey and Kenya Revenue Authority.

The result of the simulation is presented above in figure 5. The essential result here is that while the presence of an extensive margin does depress optimal tax rates, it cannot, by itself, rationalize the imposition of a tax exemption. This is even true under a sizeable elasticity of 0.5. The reason for this is that the effect of the margin is analogous to a higher intensive elasticity.

Further, we see that the presence of a marginal administrative cost has next to no impact on optimal rates. This result rebukes what would seem a commonsense argument, namely, that at the very least, if a formal worker with a given income produces 4,000 Ksh in tax receipts but a 5,000 Ksh admin cost then she should be covered by the exemption. The reason why this is not the case is that barring the invocation of a notched tax system, the cost of excluding this worker would not be the 4,000 Ksh she pays in taxes but the 4,000 Ksh that *everyone* with earnings above 4,000 Ksh would be exempt from paying taxes on.

7 Conclusion

This paper investigates why developing economies often set high tax exemption thresholds, frequently above the 80th income percentile (Jensen, 2022), a practice that diverges sharply from the predictions of conventional optimal tax models, which advocate for no exemptions (Diamond & Saez, 2011; Mankiw et al., 2009). This discrepancy raises the question of whether such policies reflect a failure of the models to capture reality or a pursuit of suboptimal policies by policymakers in developing countries. I argue for the former. Specifically I posit that the assumption of full compliance in the basic Mirrlees (1971) framework is particularly problematic in developing economies with high informality rates (Gordon & Li, 2009). Beyond efficiency concerns, taxing

formal workers while leaving informal workers untaxed introduces significant horizontal inequalities that are not captured by standard model.

I account for this by extending the standard optimal tax model to a two-type framework allowing for the presence of untaxable workers. I propose a new welfarist approach to the optimal tax problem with preferences on horizontal inequality. I show that if the social planner is sufficiently averse to horizontal inequality, then large exemption thresholds, close to the actual threshold observed in Kenya, become optimal.

To seriously consider exemptions as a policy tool to counter horizontal inequality, I examine several prominent alternative explanations for the observed exemptions. Using numerical approaches, I show that neither extensive margin responses nor marginal administrative costs of taxation can rationalize the imposition of an exemption within an optimal tax framework. I do not consider a margin to unemployment (Saez, 2002) as most very low income economies have little unemployment, with the informal sector and home production acting as last-resort employment opportunities instead (Adhvaryu & Nyshadham, 2017; Baah-Boateng, 2015).

To further bolster the argument that horizontal inequality provides a normative basis for tax exemption, I administer a dynamic survey to directly elicit the preferences of Kenyans on both vertical and horizontal inequality. Using pairwise comparisons, marginal social welfare weights across both income and tax compliance status can be identified. Aggregating preferences using a median-voter objective shows that while preferences on vertical inequality is varied and moderately progressive, preferences against horizontal inequality are strong enough to rationalize an exemption of 200,000 Ksh/year.

This paper doesn't argue that horizontal inequality is directly guiding policymakers in their decisions, nor that horizontal inequality should be an overriding objective. Rather, it rebuts the argument that tax exemptions are necessarily inefficient policy tools. By providing a consistent welfarist normative framework for analyzing exemption policy, future models should be able to more faithfully capture the reality facing policymakers in developing contexts.

8 References

Adhvaryu, A., & Nyshadham, A. (2017). Health, Enterprise, and Labor Complementarity in the Household. *Journal of Development Economics*, *126*, 91–111. https://doi.org/10.1016/j.jdeveco.2016.11.007

Akerlof, G. A. (1978). The Economics of "Tagging" as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning. *The American Economic Review*, *68*(1), 8–19.

Allingham, M. G., & Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics*, *1*(3–4), 323–338.

Almås, I., Cappelen, A. W., & Tungodden, B. (2020). Cutthroat Capitalism versus Cuddly Socialism: Are Americans More Meritocratic and Efficiency-Seeking than Scandinavians? *Journal of Political Economy*, *128*(5), 1753–1788. https://doi.org/10.1086/705551

Aronson, J. R., Johnson, P., & Lambert, P. J. (1994). Redistributive Effect and Unequal Income Tax Treatment. *The Economic Journal*, *104*(423), 262. https://doi.org/10.2307/2234747

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*, *2*(3), 244–263. https://doi.org/10.1016/0022-0531(70)90039-6

Auerbach, A. J., & Hassett, K. A. (2002). A New Measure of Horizontal Equity. *American Economic Review*, *92*(4), 1116–1125. https://doi.org/10.1257/00028280260344650

Baah-Boateng, W. (2015). Unemployment in Africa: How appropriate is the global definition and measurement for policy purpose. *International Journal of Manpower*, *36*(5), 650–667. https://doi.org/10.1108/IJM-02-2014-0047

Bachas, P., Gadenne, L., & Jensen, A. (2020). *Informality, Consumption Taxes and Redistribution* (No. w27429; p. w27429). National Bureau of Economic Research. https://doi.org/10.3386/w27429

Bellemare, C., Kröger, S., & Van Soest, A. (2008). Measuring inequity aversion in a heterogeneous population using experimental decisions and subjective probabilities. *Econometrica*, *76*(4), 815–839.

Bergeron, A., Tourek, G., & Weigel, J. (2021). *The State Capacity Ceiling on Tax Rates: Evidence from Randomized Tax Abatements in the DRC*.

Besley, T. (2020). State Capacity, Reciprocity, and the Social Contract. *Econometrica*, *88*(4), 1307–1335. https://doi.org/10.3982/ECTA16863

Besley, T., & Persson, T. (2013). Chapter 2—Taxation and Development. In A. J. Auerbach, R. Chetty, M. Feldstein, & E. Saez (Eds.), *Handbook of Public Economics* (Vol. 5, pp. 51–110). Elsevier. https://doi.org/10.1016/B978-0-444-53759-1.00002-9

Besley, T., & Persson, T. (2014). Why Do Developing Countries Tax So Little? *Journal of Economic Perspectives*, *28*(4), 99–120. https://doi.org/10.1257/jep.28.4.99

Blattman, C., & Dercon, S. (2018). The Impacts of Industrial and Entrepreneurial Work on Income and Health: Experimental Evidence from Ethiopia. *American Economic Journal: Applied Economics*, *10*(3), 1–38. https://doi.org/10.1257/app.20170173

Capozza, F., & Srinivasan, K. (2024). *Who Should Get Money? Estimating Welfare Weights in the US*.

Charité, J., Fisman, R., Kuziemko, I., & Zhang, K. (2022). Reference points and redistributive preferences: Experimental evidence. *Journal of Public Economics*, *216*, 104761. https://doi.org/10.1016/j.jpubeco.2022.104761

Charness, G., & Rabin, M. (2002). Understanding Social Preferences with Simple Tests*. *The Quarterly Journal of Economics*, *117*(3), 817–869. https://doi.org/10.1162/003355302760193904

Choné, P., & Laroque, G. (2010). Negative Marginal Tax Rates and Heterogeneity. *American Economic Review*, *100*(5), 2532–2547. https://doi.org/10.1257/aer.100.5.2532

Cowell, F. A. (1990). Cheating the government: The economics of evasion. *MIT Press Books*, *1*.

Cremer, H., Gahvari, F., & Lozachmeur, J.-M. (2010). Tagging and Income Taxation: Theory and an Application. *American Economic Journal: Economic Policy*, *2*(1), 31–50.

Cummings, R. G., Martinez-Vazquez, J., & McKee, M. (2001). Cross Cultural Comparisions of Tax Compliance Behavior. *International Center for Public Policy Working Paper Series, at AYSPS, GSU*, Article paper0103. https://ideas.repec.org//p/ays/ispwps/paper0103.html

Diamond, P., & Saez, E. (2011). The Case for a Progressive Tax: From Basic Research to Policy Recommendation. *Journal of Economic Perspectives*, *25*(4), 165–190. https://doi.org/10.1257/jep.25.4.165

Duclos, J.-Y., Jalbert, V., & Araar, A. (2003). Classical horizontal inequity and reranking: An integrating approach. In *Fiscal policy, inequality and welfare* (pp. 65–100). Emerald Group Publishing Limited.

Dwenger, N., Sjursen, I. H., & Vietz, J. (2024). What Is Fair? Experimental Evidence on Fair Equality vs Fair Inequality. *CESifo Working Papers*. https://www.cesifo.org/en/publications/2024/working-paper/what-fair-experimental-evidence-fair-equality-vs-fair-inequality

Fall, F. (2022). Strengthening the Tax System to Reduce Inequalities in South Africa. *OECD Working Papers*.

García Peñalosa, C., & Turnovsky, S. J. (2005). Second-best optimal taxation of capital and labor in a developing economy. *Journal of Public Economics*, *89*(5), 1045–1074. https://doi.org/10.1016/j.jpubeco.2004.06.002

Gollin, D. (2008). Nobody's business but my own: Self-employment and small enterprise in economic development. *Journal of Monetary Economics*, *55*(2), 219–233. https://doi.org/10.1016/j.jmoneco.2007.11.003

Gordon, R., & Li, W. (2009). Tax structures in developing countries: Many puzzles and a possible explanation. *Journal of Public Economics*, *93*(7), 855–866. https://doi.org/10.1016/j.jpubeco.2009.04.001

Haanwinckel, D., & Soares, R. R. (2021). Workforce Composition, Productivity, and Labour Regulations in a Compensating Differentials Theory of Informality. *The Review of Economic Studies*, *88*(6), 2970–3010. https://doi.org/10.1093/restud/rdab017

Harris, J. R., & Todaro, M. P. (1970). Migration, unemployment and development: A two-sector analysis. *American Economic Review*, *60*(1), 126.

Heathcote, J., & Tsujiyama, H. (2021a). Optimal Income Taxation: Mirrlees Meets Ramsey. *Journal of Political Economy*, *129*(11), 3141–3184. https://doi.org/10.1086/715851

Heathcote, J., & Tsujiyama, H. (2021b). *Practical Optimal Income Taxation*. https://doi.org/10.21034/sr.626

IMF. (2024). *IMF Reaches Staff-Level Agreement with Kenya on Seventh Reviews of the Extended Fund Facility and Extended Credit Facility Arrangements and the Second Review Under the Resilience and Sustainability Facility*. IMF. https://www.imf.org/en/News/Articles/2024/06/10/pr-24213-kenya-agreement-on-7th-reviews-of-eff-and-ecf-arrangements-and-2nd-review-under-rsf

Jensen, A. (2022). Employment Structure and the Rise of the Modern Tax System. *American Economic Review*, *112*(1), 213–234. https://doi.org/10.1257/aer.20191528

Kanina, J., Mugure, J., Nato, J., Urzainqui, D. G., & Fisker, P. (2025). *Assessing the Impact of Personal Income Tax Reform in Kenya with Administrative Data: Behavioral Responses and Distributional Implications*. https://publication.aercafricalibrary.org/handle/123456789/3969

Kaplow, L. (1989). HORIZONTAL EQUITY: MEASURES IN SEARCH OF A PRINCIPLE. *National Tax Journal*, *42*(2), 139–154. https://doi.org/10.1086/NTJ41788784

Kaplow, L. (2000). *Horizontal equity: New measures, unclear principles*.

Kaplow, L., & Shavell, S. (2001). Any Non-welfarist Method of Policy Assessment Violates the Pareto Principle. *Journal of Political Economy*, *109*(2), 281–286. https://doi.org/10.1086/319553

Keen, M., & Slemrod, J. (2017). Optimal tax administration. *Journal of Public Economics*, *152*, 133–142. https://doi.org/10.1016/j.jpubeco.2017.04.006

King, M. A. (1983). An Index of Inequality: With Applications to Horizontal Equity and Social Mobility. *Econometrica*, *51*(1), 99–115. https://doi.org/10.2307/1912250

Kleven, H. J. (2021). Sufficient statistics revisited. *Annual Review of Economics*, *13*, 515–538.

Kleven, H. J., Kreiner, C. T., & Saez, E. (2009). The optimal income taxation of couples. *Econometrica*, *77*(2), 537–560.

KNBS. (2021a). *Kenya—Kenya Continuous Household Survey (KCHS)—2021* [Dataset].

KNBS. (2021b). *Quarterly Labour Force Report 2021 Quarter 1*. https://www.knbs.or.ke/reports/quarterly-labour-force-report-2021-quarter-1/

Kresch, E. P., Walker, M., Best, M. C., Gerard, F., & Naritomi, J. (2023). Sanitation and property tax compliance: Analyzing the social contract in Brazil. *Journal of Development Economics*, *160*, 102954. https://doi.org/10.1016/j.jdeveco.2022.102954

La Porta, R., & Shleifer, A. (2014). Informality and Development. *Journal of Economic Perspectives*, *28*(3), 109–126. https://doi.org/10.1257/jep.28.3.109

Lewis, W. A. (1954). Economic Development with Unlimited Supplies of Labour. *The Manchester School*, *22*(2), 139–191. https://doi.org/10.1111/j.1467-9957.1954.tb00021.x

Maloney, W. F. (2004). Informality Revisited. *World Development*, *32*(7), 1159–1178. https://doi.org/10.1016/j.worlddev.2004.01.008

Mankiw, N. G., Weinzierl, M., & Yagan, D. (2009). Optimal Taxation in Theory and Practice. *Journal of Economic Perspectives*, *23*(4), 147–174. https://doi.org/10.1257/jep.23.4.147

McKay, A., Pirttilä, J., & Schimanski, C. (2024). The Tax Elasticity of Formal Work in Sub-Saharan African Countries. *The Journal of Development Studies*, *60*(2), 217–244. https://doi.org/10.1080/00220388.2023.2279477

Mirrlees, J. A. (1971). An Exploration in the Theory of Optimum Income Taxation. *Review of Economic Studies*.

Moore, M. (2023). Tax obsessions: Taxpayer registration and the "informal sector" in sub-Saharan Africa. *Development Policy Review*, *41*(1), e12649. https://doi.org/10.1111/dpr.12649

Musgrave, R. A. (1990). Horizontal equity, once more. *National Tax Journal*, *43*(2), 113–122. https://doi.org/10.1086/NTJ41788830

OECD. (2012). *What are the Best Policy Instruments for Fiscal Consolidation* (OECD Economics Department Policy Notes).

Okun, A. M. (1975). *Equality and Efficiency: The Big Tradeoff*. Brookings Institution Press. https://www.jstor.org/stable/10.7864/j.ctt13wztjk

Okunogbe, O., & Pouliquen, V. (2022). Technology, taxation, and corruption: Evidence from the introduction of electronic tax filing. *American Economic Journal: Economic Policy*, *14*(1), 341–372.

Pomeranz, D., & Vila-Belda, J. (2019). Taking State-Capacity Research to the Field: Insights from Collaborations with Tax Authorities. *Annual Review of Economics*, *11*(Volume 11, 2019), 755–781. https://doi.org/10.1146/annurev-economics-080218-030312

Ranchhod, V., & Finn, A. (2015). *Estimating the effects of South Africa's youth employment tax incentive–an update*.

Rothschild, C., & Scheuer, F. (2013). Redistributive Taxation in the Roy Model*. *The Quarterly Journal of Economics*, *128*(2), 623–668. https://doi.org/10.1093/qje/qjs076

Saez, E. (2001a). Using elasticities to derive optimal income tax rates. *The Review of Economic Studies*, *68*(1), 205–229.

Saez, E. (2001b). Using elasticities to derive optimal income tax rates. *The Review of Economic Studies*, *68*(1), 205–229.

Saez, E., & Stantcheva, S. (2016). Generalized social marginal welfare weights for optimal tax theory. *American Economic Review*, *106*(01), 24–45.

Sher, I. (2024). Generalized Social Marginal Welfare Weights Imply Inconsistent Comparisons of Tax Policies. *American Economic Review*, *114*(11), 3551–3577. https://doi.org/10.1257/aer.20211025

Slemrod, J. (2019). Tax Compliance and Enforcement. *Journal of Economic Literature*, *57*(4), 904–954. https://doi.org/10.1257/jel.20181437

Slemrod, J., & Gillitzer, C. (2014). Insights from a Tax-systems Perspective*. *CESifo Economic Studies*, *60*(1), 1–31. https://doi.org/10.1093/cesifo/ift015

Slemrod, J., & Yitzhaki, S. (2002). Tax Avoidance, Evasion, and Administration*. In A. J. Auerbach & M. Feldstein (Eds.), *Handbook of Public Economics* (Vol. 3, pp. 1423–1470). Elsevier. https://doi.org/10.1016/S1573-4420(02)80026-X

Stiglitz, J. E. (1982). *Utilitarianism and horizontal equity: The case for random taxation*.

Tanzi, V., & Zee, H. (2001). Tax Policy for Developing Countries. *IMF Economic Issues*, *27*.

Ulyssea, G. (2020). Informality: Causes and Consequences for Development. *Annual Review of Economics*, *12*(1), 525–546. https://doi.org/10.1146/annurev-economics-082119-121914

Weigel, J. L. (2020). The Participation Dividend of Taxation: How Citizens in Congo Engage More with the State When it Tries to Tax Them*. *The Quarterly Journal of Economics*, *135*(4), 1849–1903. https://doi.org/10.1093/qje/qjaa019

Weinzierl, M. C. (2012). *Why do we redistribute so much but tag so little? The principle of equal sacrifice and optimal taxation*. National Bureau of Economic Research.

Zolt, E. M., & Bird, R. M. (2005). *Redistribution Via Taxation: The Limited Role of the Personal Income Tax in Developing Countries* (SSRN Scholarly Paper No. 804704). https://papers.ssrn.com/abstract=804704

9 Appendices

9.1 Rewriting to continuous horizontal inequality

In order to rewrite the SWF as a continuous function we allow for the type to be a continuous measure such that $j \in [\underline{j}, \bar{j}]$. Here we can think of $j = \underline{j}$ as full compliance and $j = \bar{j}$ as no compliance. Now $f(n, j)$ is a two-dimensional joint pdf where it is still true that $h(n) = \int \rho(n, j) f(n, j)\, dj$. The SWF now becomes:

$$SWF = \int_0^{\infty} \int_{\underline{j}}^{\bar{j}} \left(G\big(u(n,j)\big)^{\phi} \cdot \rho(n,j)\right)^{\frac{1}{\phi}} dj\, h(n)\; dn$$

In the continuous case, we can conceptualize *j* as anything as long as we measure $\rho(n, j)$ according. For example, we can set $j \in [0,1]$ and let it capture "percent of owed taxes evaded". This would allow the model to distinguish tax evasion on the intensive margin as well.

9.2 Finding Optimal Rates Using Perturbation Approach:

A tax is locally optimal if for all values of z, a small change in marginal tax rates, $dT'(z)$ has no effect on aggregate social welfare.

Mechanical effect of tax change:

$$dM = dT'(z) \cdot [1 - H(z)]\rho^{+}(z)$$

Welfare effect of tax change:

$$dW = -q^{+}(z) \cdot dM$$

The behavioral effect is driven by the taxable workers with incomes z:

$$dB = -dT'(z) \cdot g(z)\rho(z) \cdot e \cdot z \cdot \frac{T'(z)}{1 - T'(z)}$$

Set condition for local optimality and rewrite, $dM + dW + dB = 0$

$$\big(1 - q^{+}(z)\big) dT'(z) \cdot [1 - H(z)]\rho^{+}(z) = dT'(z) \cdot g(z)\rho(z) \cdot e \cdot z \cdot \frac{T'(z)}{1 - T'(z)} \quad <=>$$

$$\frac{T'(z)}{1 - T'(z)} = \frac{\big(1 - q^{+}(z)\big)[1 - H(z)]\rho^{+}(z)}{g(z)\rho(z) \cdot e \cdot z} \quad <=>$$

$$T'(z) = \frac{1 - q^{+}(z)}{1 - q^{+}(z) + \alpha_{|j=1} \cdot e}$$

Where I use the definition of the pareto-parameter for taxable workers $\alpha_{|j=1}(z) = \frac{g(z)\rho(z)\cdot e\cdot z}{[1-H(z)]\rho^{+}(z)}$.

9.3 Proof of Optimal Tax Rates with Untaxable Workers:

In this derivation, we analyze the optimal taxation problem in an economy with two types of workers. The government aims to design an optimal tax system that maximizes social welfare while considering the behavioral responses of workers and budget constraints.

Setting Up the Hamiltonian:

I start by formulating the Hamiltonian, H, for the government's optimization problem. The costate variables θ_1 and θ_2 correspond to the IC constraints as in the standard formulation of the problem.

$$\begin{aligned} H \; = \Big(& Q\big(u(n,1), u(n,2), \rho(n)\big) \\ & + \lambda \Big[\rho(n)(n \cdot l(n,1) - [u(n,1) + v\big(l(n,1)\big)]) \\ & + \big(1 \; - \rho(n)\big) \Big(n \cdot \bar{l}(n,2) - \Big[u(n,2) + v\left(\bar{l}(n,2)\right) \Big] \Big) \Big] \Big) \, h(n) \\ & + \theta_1(n) \left[\frac{l(n,1) v'\big(\; l(n,1)\big)}{n} \right] + \theta_2(n) \left[\frac{l(n,2) v'\big(l(n,2)\big)}{n} \right] \end{aligned}$$

Where:

- $Q(u(n,1), u(n,2), \rho(n))$: Social welfare function.
- λ: Lagrange multiplier on government budget.
- $v\big(l(n,j)\big)$: Labor disutility for type (n,j).
- $\bar{l}(n,2)$: Labor supply untaxable workers.
- $h(n)$: Density function at ability n.
- $\theta_1(n), \theta_2(n)$: Costate variables associated with the incentive compatibility (IC) constraints for Type I and Type II workers, respectively.

First-Order Conditions:

We derive the FOCs by taking the derivatives of H with respect to the control and state variables.

1. $\frac{\partial H}{\partial l(n,1)} = 0$:

$$\lambda\big[\rho(n)\left(n - v'\big(l(n,1)\big)\right)\big] h(n) + \theta_1(n) \left(\frac{v'(l(n,1) + l(n,1) v''\big(l(n,1)\big)}{n} \right) = 0 \qquad (A3.1)$$

2. $\frac{\partial H}{\partial u(n,1)} = 0$:

$$\left[\frac{\partial Q}{\partial u(n,1)} - \lambda\rho(n)\right] h(n) + \frac{d\theta_1(n)}{dn} = 0 \qquad (A3.2)$$

3. $\frac{\partial H}{\partial u(n,2)} = 0$:

$$\left[\frac{\partial Q}{\partial u(n,2)} - \lambda\big(1-\rho(n)\big)\right] h(n) + \frac{d\theta_2(n)}{dn} = 0 \qquad (A3.3)$$

Solving for Costate variables:

I integrate FOCs 2 and 3 in order to find the co-states $\theta_1(n)$ and $\theta_2(n)$. I employ the transversality conditions:

$$\lim_{n\to\infty} \theta_k(n) = 0 \;\; and \;\; \theta_k(0) = 0 \qquad for\ k \in \{1,2\}$$

Finding $\theta_1(n)$ using *(A3.2)*:

$$-\frac{d\theta_1(n)}{dn} = \left[\frac{\partial Q}{\partial u(n,1)} - \lambda\rho(n)\right] h(n)$$

$$\theta_1(n) = \int_n^\infty \left[\frac{\partial Q}{\partial u(n,1)} - \lambda\rho(n)\right] h(m)dm \qquad (A3.4)$$

And similarly, for $\theta_2(n)$:

$$\theta_2(n) = \int_n^\infty \left[\frac{\partial Q}{\partial u(n,2)} - \lambda\big(1-\rho(n)\big)\right] h(m)dm \qquad (A3.5)$$

We can add the costates to get:

$$\theta_1(n) + \theta_2(n) = \int_n^\infty \left[\frac{\partial Q}{\partial u(n,1)} + \frac{\partial Q}{\partial u(n,2)} - \lambda\right] h(m)dm \qquad (A3.6)$$

Using that $\theta_1(0) + \theta_2(0) = 0$:

$$0 = \int_n^\infty \left[\frac{\partial Q}{\partial u(n,1)} + \frac{\partial Q}{\partial u(n,2)} - \lambda\right] h(m)dm$$

$$\lambda = \int_n^\infty \left[\frac{\partial Q}{\partial u(n,1)} + \frac{\partial Q}{\partial u(n,2)}\right] h(m)dm \qquad (A3.7)$$

Deriving the Optimal Tax Rate:

For type I workers the utility function is:

$$u(n,1) = c(n) - v\big(l(n,1)\big) \qquad (A3.8)$$

Subject to the budget constraint:

$$c(n) = z(n) - T\big(z(n)\big) \tag{A3.9}$$

Where $z(n) = n \cdot l(n,1)$ is gross income and $T\big(z(n)\big)$ is the tax function.

We review the FOC for utility maximization:

$$\frac{\partial u(n,1)}{\partial l(n,1)} = 0$$
$$\Big(1 - T'\big(z(n)\big)\Big)\, n - v'\big(l(n,1)\big) = 0$$

Which yields the FOC

$$v'\big(l(n,1)\big) = \Big(1 - T'\big(z(n)\big)\Big)\, n \tag{A3.10}$$

I now use the FOC from eq. *(A3.1)* and insert eq *(A3.10)*:

$$\lambda[\rho(n)nT'\big(z(n)\big)]h(n) + \theta_1(n)\left(\frac{\Big(1 - T'\big(z(n)\big)\Big)\, n + l(n,1)v''\big(l(n,1)\big)}{n}\right) = 0 \tag{A3.11}$$

Elasticity of Taxable Income:

I define the elasticity of taxable income e:

$$e = \frac{\Big(1 - T'\big(z(n)\big)\Big)\, n}{l(n,1)v''\big(l(n,1)\big)} \tag{A3.12}$$

And rewrite:

$$\frac{l(n,1)v''\big(l(n,1)\big)}{n} = \frac{1 - T'\big(z(n)\big)}{e}$$

I now use this definition to simplify eq. *(A3.11)*

$$\lambda[\rho(n)nT'\big(z(n)\big)]h(n) + \theta_1(n)\left(\Big(1 - T'\big(z(n)\big)\Big) + \frac{1 - T'\big(z(n)\big)}{e}\right) = 0$$

$$\lambda[\rho(n)nT'\big(z(n)\big)]h(n) = -\theta_1(n)\Big(1 - T'\big(z(n)\big)\Big)\left(1 + \frac{1}{e}\right) \tag{A3.13}$$

Defining Social Marginal Welfare Weights:

The marginal SWW for type I workers can be defined as:

$$q_1(n) = \frac{\frac{\partial Q}{\partial u(n,1)}}{\lambda\rho(n)} = \frac{\frac{\partial Q}{\partial c(n,1)}}{\lambda\rho(n)} \tag{A3.14}$$

Which allows us to rewrite:

$$\frac{\partial Q}{\partial u(n,1)} = q_1(n) \cdot \lambda \cdot \rho(n) \tag{A3.15}$$

Using equations *(A3.4)* and *(A3.15)*

$$-\theta_1(n) = \int_n^\infty [\lambda\rho(m) - q_1(m) \cdot \lambda \cdot \rho(m)]h(m)dm$$

$$-\theta_1(n) = \lambda \int_n^\infty \rho(m)[1 - q_1(m)]h(m)dm$$

Using that $f(m,1)\rho^+(0) = \rho(n)h(n)$ and that $\rho^+(n) = \int_n^\infty \rho(m)h(m)dm$

$$-\theta_1(n) = \lambda\rho^+(0) \int_n^\infty [1 - q_1(m)]f(m,1)dm \tag{A3.16}$$

Insert into eq. *(A3.13)*

$$\lambda[\rho(n)nT'(z(n))]h(n) = \lambda\rho^+(0)\left[\int_n^\infty [1 - q_1(m)]f(m,1)dm\right]\left(1 - T'(z(n))\right)\left(1 + \frac{1}{e}\right)$$

$$[\rho(n)nT'(z(n))]h(n) = \rho^+(0)\left[\int_n^\infty [1 - q_1(m)]f(m,1)dm\right]\left(1 - T'(z(n))\right)\left(1 + \frac{1}{e}\right) \tag{A3.17}$$

<u>Solving for Marginal Tax Rates:</u>

Rewrite eq. *(A3.17)*

$$\frac{T'(z(n))}{1 - T'(z(n))} = \rho^+(0)\frac{1 - \int_n^\infty q_1(m)f(m,1)dm}{\rho(n)nh(n)}\left(1 + \frac{1}{e}\right)$$

$$\frac{T'(z(n))}{1 - T'(z(n))} = \frac{1 - \int_n^\infty q_1(m)f(m,1)dm}{nf(n,1)}\left(1 + \frac{1}{e}\right) \tag{A3.18}$$

To relate the expression to observable income distribution, we proceed further. We can move from skills to income by using the Saez (2001) concept of a virtual density where we assume a locally linear tax schedule. Moving from skills is benign as our formulation above fulfills the Spence-Mirrlees single crossing condition ensuring unique mapping from $n \to z$:

$$z(n,1) = n \cdot l\left(n, \left(1 - T'(z(n))\right)\right) \tag{A3.19}$$

Take derivative wrt. n using the assumption that T' is locally linear:

$$\frac{dz(n,1)}{dn} = l(n,1) + n \cdot \frac{dl(n,1)}{dn}$$

Use that $\frac{dl(n,1)}{dn} = \frac{1-T'(z(n))}{v'\ (l(n,1)}$

$$\frac{dz(n,1)}{dn} = l(n,1)\left(1 + \frac{\big(1-T'(z)\big)\cdot n}{l(n,1)v''(l(n,1)}\right)$$

Use definition: $e = \frac{\big(1-T'(z(n))\big)n}{l(n,1)v''(l(n,1))}$

$$\frac{dz(n,1)}{dn} = l(n,1)\cdot(1+e) \tag{A3.20}$$

We now introduce the densities over income as $g(z)$ for the PDF and $G(z)$ for the CDF. The densities are related by:

$$g(z,1)\cdot\frac{dz(n,1)}{dn} = f(n,1) \tag{A3.21}$$

Insert *(A3.20)*:

$$g(z,1)z(n,1)(1+e) = n\cdot f(n,1) \tag{A3.22}$$

Substitute *(A3.22)* into *(A3.18)*:

$$\frac{T'\big(z(n)\big)}{1-T'\big(z(n)\big)} = \frac{1-\int_n^\infty q_1(m)f(m,1)dm}{g(n,1)z(n,1)(1+e)}\Big(\frac{1+e}{e}\Big)$$

$$= \frac{1-\int_n^\infty q_1(m)f(m,1)dm}{g(n,1)z(n,1)}\Big(\frac{1}{e}\Big)$$

$$= \frac{1-G(z,1)}{g(z,1)z(n,1)}\Big(\frac{1}{e}\Big)\frac{1-\int_n^\infty q_1(m)f(m,1)dm}{1-G(z,1)}$$

$$= \frac{1}{\alpha(z,1)}\Big(\frac{1}{e}\Big)\,1-q^+(z,1)$$

$$= \frac{1-q^+(z,1)}{e\cdot\alpha(z,1)}$$

We can isolate to find the final formula:

$$T'(z) = \frac{1-q^+(z,1)}{1-q^+(z,1)+e\cdot\alpha(z,1)} \tag{A3.23}$$

Where we employ the definitions

- Avg. marginal SWW above z is $q^+(z) = \frac{\int_n^\infty q_1(m)f(m,1)dm}{1-G(z,1)}$

- Local Pareto parameter at z: $\alpha(z) = \frac{z \cdot g(z)}{1-G(z)}$
- Cumulative distribution of income: $G(z)$, where $G(z(n,1)) = F(n,1)$

9.4 Kuhn-Tucker Conditions for Non-negative Taxes

In this section I outline the impacts of including a non-negativity constraint on taxes in the optimal control derivation outlined in section A9.3. The derivations are almost unchanged, so emphasis is reserved for the novel parts.

In order to restrict the planner to non-negative taxes we begin by augmenting the Hamiltonian with an additional costate variable:

$$\begin{aligned}\tilde{H} \;= \Big(& Q\big(u(n,1), u(n,2), \rho(n)\big) \\ &+ \lambda \Big[\rho(n)(n \cdot l(n,1) - [u(n,1) + v\big(l(n,1)\big)]) \\ &+ \big(1 \;- \rho(n)\big)\left(n \cdot \bar{l}(n,2) - \left[u(n,2) + v\left(\bar{l}(n,2)\right)\right]\right)\Big]\Big) h(n) \\ &+ \theta_1(n)\left[\frac{l(n,1)v'\big(\; l(n,1)\big)}{n}\right] + \theta_2(n)\left[\frac{l(n,2)v'\big(l(n,2)\big)}{n}\right] \\ &+ \omega(n)\left(\bar{l}(n,2) - l(n,1)\right)\end{aligned}$$

Here I use $\bar{l}(n,2) \geq l(n,1)$ as an easy proxy for non-negative taxes. As type II workers faces effective tax rates of 0 pct. and as preferences are homogenous a violation of this condition could only occur through negative taxes. We thus also have:

$$\omega(n) \geq 0 \qquad and \qquad \omega(n)\left(\bar{l}(n,2) - l(n,1)\right) = 0$$

The new costate only affects one FOC, eq. *(A3.1)*

$$\lambda\big[\rho(n)\left(n - v'\big(l(n,1)\big)\right)\big]h(n) + \theta_1(n)\left(\frac{v'(l(n,1) + l(n,1)v''\big(l(n,1)\big)}{n}\right) - \boldsymbol{\omega(n)} = 0 \;\; (A4.2)$$

$$\lambda[\rho(n)nT'\big(z(n)\big)]h(n) + \theta_1(n)\left(\frac{\left(1 - T'\big(z(n)\big)\right)n + l(n,1)v''\big(l(n,1)\big)}{n}\right) - \omega(n) = 0 (A4.2)$$

Simplify using elasticity

$$\lambda[\rho(n)nT'\big(z(n)\big)]h(n) + \theta_1(n)\left(\left(1 - T'\big(z(n)\big)\right) + \frac{1 - T'\big(z(n)\big)}{e}\right) - \omega(n) = 0$$

$$\lambda[\rho(n)nT'(z(n))]h(n) = -\theta_1(n)\left(1-T'(z(n))\right)\left(1+\frac{1}{e}\right)+\omega(n)$$

$$\lambda[\rho(n)nT'(z(n))]h(n) = \lambda\rho^{+}(0)\left[\int_n^{\infty}[1-q_1(m)]f(m,1)dm\right]\left(1-T'(z(n))\right)\left(1+\frac{1}{e}\right)+\omega(n)$$

$$[\rho(n)nT'(z(n))]h(n) = \rho^{+}(0)\left[\int_n^{\infty}[1-q_1(m)]f(m,1)dm\right]\left(1-T'(z(n))\right)\left(\frac{1+e}{e}\right)+\frac{\omega(n)}{\lambda}$$

$$[\rho(n)nT'(z(n))]h(n) = \rho^{+}(0)\left[\int_n^{\infty}[1-q_1(m)]f(m,1)dm\right]\left(1-T'(z(n))\right)\left(1+\frac{\omega(n)(e)}{\lambda\cdot(1+e)}\right)\left(\frac{1+e}{e}\right) \quad (A4.3)$$

Rewrite to find the MTR:

$$\frac{T'(z(n))}{1-T'(z(n))} = \rho^{+}(0)\frac{1-\int_n^{\infty}q_1(m)f(m,1)dm+\frac{\omega(n)(e)}{\lambda\cdot(1+e)}}{\rho(n)nh(n)}\left(1+\frac{1}{e}\right)$$

Rename $\tilde{\omega}(n) = \frac{\omega(n)e}{\lambda\cdot(1+e)}$

$$\frac{T'(z(n))}{1-T'(z(n))} = \rho^{+}(0)\frac{1-\int_n^{\infty}q_1(m)f(m,1)dm+\tilde{\omega}(n)}{\rho(n)nh(n)}\left(1+\frac{1}{e}\right) \quad (A4.4)$$

$$\frac{T'(z(n))}{1-T'(z(n))} = \rho^{+}(0)\frac{1-q^{+}(n)+\tilde{\omega}(n)}{\rho(n)nh(n)}\left(1+\frac{1}{e}\right) \quad (A4.5)$$

What we find is that when $q^{+}(n) \leq 1$ then $\omega(n) = 0 \Rightarrow \ \tilde{\omega}(n) = 0$ and we arrive at our original formula. When $q^{+}(n) > 1$ then $\tilde{\omega}(n)$ increases until $\bar{l}(n,2) = l(n,1)$ which corresponds to $\tilde{\omega}(n) = -\left(1-q^{+}(n)\right)$. As such, it is clear that the non-negativity constraint is identical to simply clamping the unrestricted taxes with the condition $T_c'(z) = \max[0, T_u'(z)]$.

9.5 Change in Exemption Threshold in Response to Increased Formality

To identify the optimal response to a small increase in formality I consider the marginal SWW on the highest skilled worker whose income falls below the exemption:

$$q_1(\bar{n}) = \frac{\frac{\partial Q}{\partial c(\bar{n},1)}}{\lambda\rho(\bar{n})} \quad (A5.1)$$

Assuming that there is an exemption in place we know that q_1 is decreasing in z and that $q_1(\bar{n}) > 1$. For a given marginal value of public funds, the necessary condition for the Jensen-rule is:

$$\frac{\partial q_1(\bar{n})}{\partial\rho(\bar{n})} < 0 \quad (A5.2)$$

Using *(A5.1)*:

$$\frac{1}{\lambda}\left(\begin{array}{l}(-\rho(\bar{n})^{-2})\dfrac{\partial Q}{\partial c(\bar{n},1)}+\dfrac{\dfrac{\partial^2 Q}{\partial c(\bar{n},1)\partial\rho(\bar{n})}}{\rho(\bar{n})}\end{array}\right)<0$$

$$\frac{1}{\lambda}\frac{\partial^2 Q}{\partial c(\bar{n},1)\partial\rho(\bar{n})}<\frac{\dfrac{\partial Q}{\partial c(\bar{n},1)}}{\lambda\rho(\bar{n})}$$

$$\frac{1}{\lambda}\frac{\partial^2 Q}{\partial c(\bar{n},1)\partial\rho(\bar{n})}<q_1(\bar{n}) \qquad (A5.3)$$

Finding $\frac{\partial^2 Q}{\partial c(\bar{n},1)\partial\rho(\bar{n})}$ for the general case:

$$\frac{\partial^2 Q}{\partial c(\bar{n},1)\partial\rho(\bar{n})}=\frac{\partial Q}{\partial c(\bar{n},1)}\left[\frac{1}{\rho(\bar{n})}+\left(\frac{1}{\phi}-1\right)\left(\frac{G(u_1)}{Q}\right)^{\phi}\right] \qquad (A5.4)$$

Insert into *(A5.3)*.

$$\frac{1}{\lambda}\frac{\partial Q}{\partial c(\bar{n},1)}\left[\frac{1}{\rho(\bar{n})}+\left(\frac{1}{\phi}-1\right)\left(\frac{G(u_1)}{Q}\right)^{\phi}\right]<q_1(\bar{n})$$

$$\frac{1}{\lambda}\frac{\partial Q}{\partial c(\bar{n},1)}\left[\left(\frac{1}{\phi}-1\right)\left(\frac{G(u_1)}{Q}\right)^{\phi}\right]<q_1(\bar{n})-q_1(\bar{n})$$

Using that $\lambda, \frac{\partial Q}{\partial c(n,1)}, \frac{G(u_1)}{Q}$ are all positive the condition becomes:

$$\frac{1}{\phi}-1<0$$

$$\frac{1}{\phi}<1 \qquad (A5.5)$$

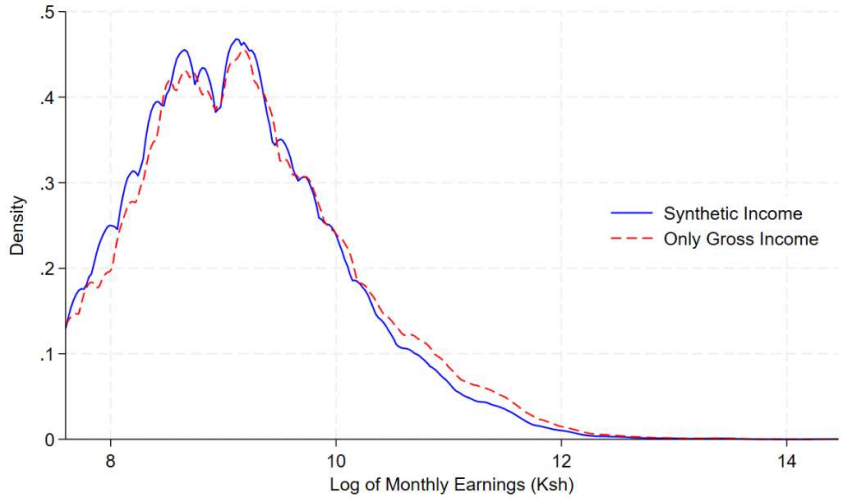
Since we have $\phi \in (-\infty;1]$ we can characterize the result as follows:

$$\begin{aligned}&\frac{\partial q_1(\bar{n})}{\partial\rho(\bar{n})}<0 \qquad \textit{for } \phi<0\\ &\frac{\partial q_1(\bar{n})}{\partial\rho(\bar{n})}>0 \quad \textit{for } \phi\in(0;1)\\ &\frac{\partial q_1(\bar{n})}{\partial\rho(\bar{n})}=0 \qquad \textit{for } \phi=1\end{aligned} \qquad (A5.6)$$

This shows that if the planner has no intrinsic preference for horizontal equality, then the exemption is not responsive to local changes in the share of taxable workers.

9.6 Distribution of Synthetic Incomes

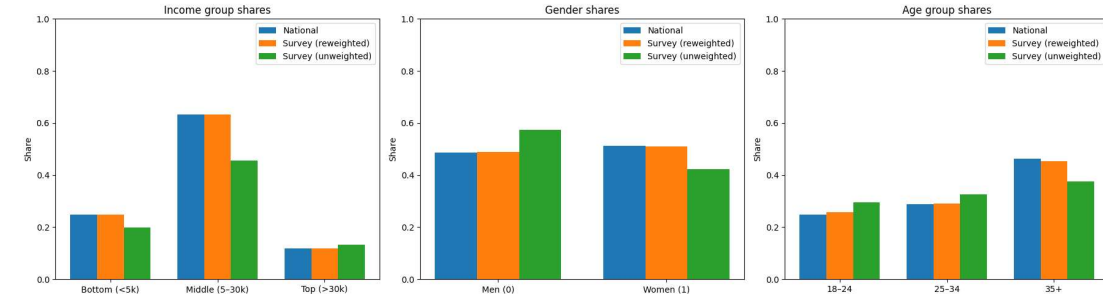
Figure A1: **Synthetic and Observed Income Distribution**



Source: Kenya Continuous Household Survey 2021

9.7 Sample Characteristics of Survey

Figure A2: **Survey Means and National Averages in Kenya: Age, Gender & Income**

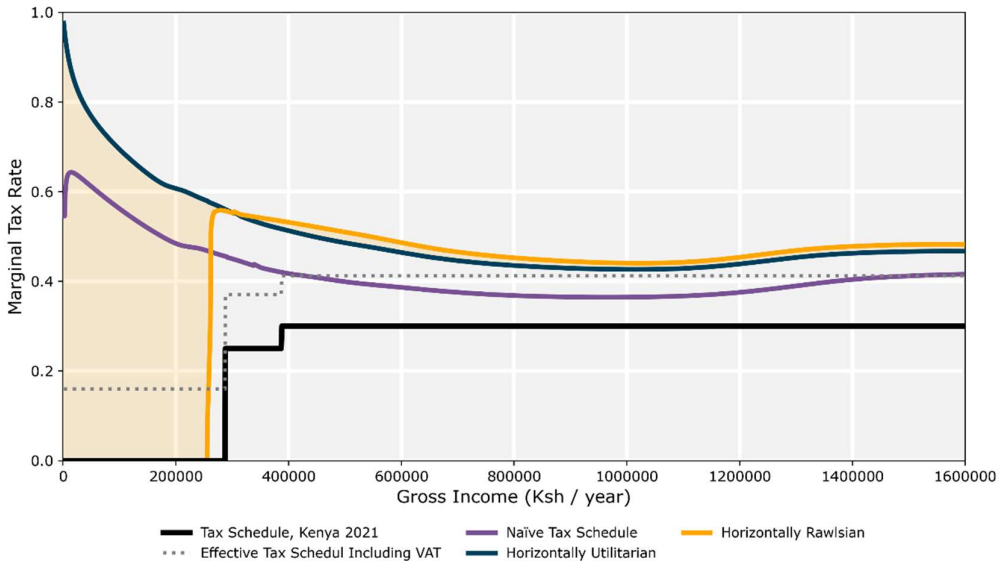


Source: Kenya National Bureau of Statistics, Census 2019 & Survey Data

9.8 Parametric Optimal Tax Simulations

All simulations are performed using the vertical preference function $G(u) = \sqrt{u}$. As the assumption of no income effects implies no diminishing marginal utility of consumption, this transformation is intended to reassert diminishing social marginal returns to consumption.

Figure A3: **Optimal Parametric Tax Prescriptions, Kenya 2021**



Note: The area highlighted in yellow captures the range of tax prescriptions over horizontal inequality preferences ranging from horizontal Rawlsianism to horizontal utilitarianism. For incomes above 500,000 Ksh, I apply Gaussian smoothing to the series. *Source:* Simulation results based on data from KCHS and KRA.

9.9 Numerical Solution to Optimal Tax Problem

To account for the limitations of solving the optimal tax problem as a multidimensional screening problem, this paper also invokes a second structural approach, by solving the problem numerically. This proves a worthwhile endeavor as it allows us to add a layer of complexity, not easily facilitated by the mechanism design approach. However, to ensure the consistency of this structural approach, I first examine how closely the numerical approach can approximate the true non-linear tax function under horizontal inequality, uncovered in this paper. This approach draws inspiration from Heathcote and Tsujiyama (2021b).

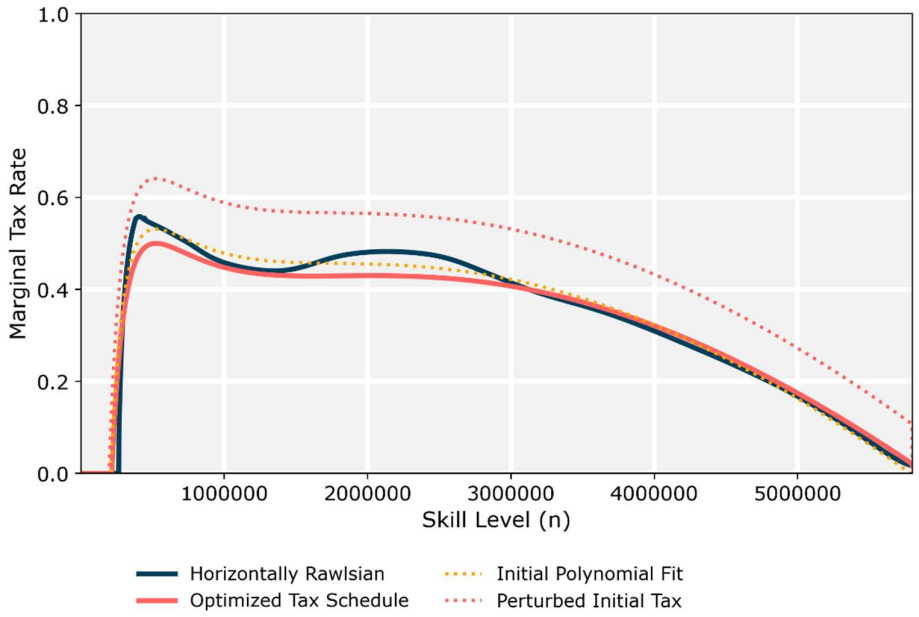
The feasibility of solving the optimal tax problem using a flexible parametric class follows from the Stone-Weierstrass theorem which states that any continuous function on a closed interval can be approximated arbitrarily closely by polynomial functions. To account for the fact that the optimal marginal tax schedule has a sharp kink at the point of the exemption threshold, I restrict the polynomial approximation such that:

$$\tilde{T}(z) = \begin{cases} 0 & z < \bar{\kappa} \\ \sum_{i=0}^{4} \kappa_i z^i & else \end{cases}$$

This leaves six parameters to solve for $\{\bar{\kappa}, \kappa_0, \kappa_1, \kappa_2, \kappa_3, \kappa_4\}$. I identify the parameters using a Nelder-Mead optimization algorithm with the social welfare function as the objective function.

I begin by simple fitting the Mirrlees tax function as closely as possible with the polynomial using simple least square and setting $\bar{k}$ equal to the exemption found in the optimization. Given this initial fit, I set wide bounds for the optimization. I then create an initializing parameter vector which is equal to the parameters from the initial fit perturbed by a random noise term. This allows me to ensure that optimizing the welfare function does move to the Mirrleesian solution whilst restricting search to a more reasonable parameter space.

Figure A4: **Ramsey Approximation of Mirrlees Optimal Tax**



Source: Simulation results based on data from KCHS and KRA.

The corresponding fit captures the overall shape of both the horizontal Rawlsian and utilitarian schedules quite well. As is evidenced by the fact that the fit of the pure polynomial approximation is similar to the numerical optimization, it is clear that the discrepancies between the numerical and Mirrleesian solutions arise due to the inherent limitations on non-linearity arising from invoking a parametric approach.

9.10 Aggregating Social Preferences

Consider a population where all voters have social preferences that can be characterized by a schedule of MSWWs, $q_i(z, j)$. The median of MSWWs of the full population is defined as:

$$q^*(z, j) \equiv \text{median}\big(q_i(z = z, j = j)\big)$$

When setting policy, a social planner wishes to respect median preferences and thus use q^* as the social objective. However, the planner cannot observe q^* nor q_i. Instead, the planner elicits votes from a representative sample of voters, each with ordinal preferences.

9.10.1 Voting Framework

When eliciting votes, the planner is constrained by the following conditions:

- **Finite Votes:** As voting is costly, each voter i is endowed with k votes $\{r_{i,1}, r_{i,2}, \dots, r_{i,k}\}$
- **Ordinal Preferences:** Voters can only express relative valuations, e.g., $q_i(z_a) > q_i(z_b)$ or $q_i(z_a)$ is twice $q_i(z_b)$ where $q_i(\cdot)$ represents social weight.
- **Single-Valuation Votes:** Each vote can only capture one comparison (e.g., $r_{i,j} = \frac{q_i(z_a)}{q_i(z_b)}$).

For tractability, the planner focuses on a finite set of incomes $\{z_1, z_2, \dots, z_N\}$.

9.10.2 Sequential Voting

The weight function q_i is inherently defined only up to a scaling constant, since all comparisons are relative. Without loss of generality, we normalize the scale of q_i by fixing the lowest income's weight as a reference. We set: $q_i(z_1) = q^*(z_1) = 1$.

We now begin a procedure for recursive aggregation of weight ratios.

Vote #1: $q^*(z_2)/q^*(z_1)$:

The planner begins by identifying the median ratio that the population assigns between agents with the two lowest incomes. A subset of the population $S_1 \in \mathcal{H}$ is asked to vote on their preferred ratio $r_{i,j} = \frac{q_i(z_2)}{q_i(z_1)}$. This gives us the voting set $R_{1,2} = \{r_{i,j}\}$. We define the median ratio:

$$\hat{r}_{1,2} = \text{median}(R_{1,2})$$

This allows us to identify $q^*(z_2) = q^*(z_1) \cdot \hat{r}_{1,2} = \hat{r}_{1,2}$.

Vote #2: $q^*(z_3)/q^*(z_1)$ and $q^*(z_3)/q^*(z_2)$:

To identify $q^*(z_3)$ the planner begins by taking two new subsamples of the population $S_2^a, S_2^b \in \mathcal{H}$. From S_2^a she elicits $R_{1,3}$ as the votes $r_{i,j} = \frac{q_i(z_3)}{q_i(z_1)}$ while from S_2^b she elicits $R_{2,3} = \left\{\frac{q_i(z_3)}{q_i(z_2)}\right\}$. The planner transforms $R_{2,3}$ using the results from Vote #1 such that $\overline{R}_{2,3} = R_{2,3} \cdot q^*(z_2)$. This allows the planner to interpret all votes as comparing workers with z_1 and z_3 as follows:

$$\hat{r}_{1,3} = \text{median}(\{R_{1,3}, R_{2,3} \cdot q^*(z_2)\})$$

We see $q^*(z_3) = q^*(z_1)\hat{r}_{1,3} = \hat{r}_{1,3}$.

Subsequent Votes:

The subsequent votes will be given as:

$$q^*(z_l) = \text{median}(\{R_{1,l}, R_{2,l} \cdot q^*(z_2), \ldots, R_{l-1,l} \cdot q^*(z_{l-1})\})$$

When ratios have been found for all n discrete income levels, the planner will have found a full characterization of social preferences, q^*.

Noise, Consistency and Order of Voting:

Since the planner relies on samples of the population, S, estimates will be subject to statistical error such that the estimated ratios $\tilde{q}^*(z),\ \tilde{r}$ will deviate from the true values, $q^*(z)$ and $\hat{r}$. Consider the first ratio:

$$\tilde{r}_{1,2} = \hat{r}_{1,2} + \epsilon_{1,2}$$

However, since the sample median is a robust estimator for the population median, we know that $\tilde{r}_{1,2} \xrightarrow{p} \hat{r}_{1,2}$. This in turn implies that $\tilde{r}_{1,3} = \text{median}(\{\tilde{r}_{1,2}, \tilde{r}_{2,3}\tilde{r}_{1,2}\}) \xrightarrow{p} \hat{r}_{1,3}$ and so forth. In other words, this bottom-up median aggregator is a consistent estimator of population preferences q^*.

For finite samples we have to account for the issue that errors will propagate upwards in the estimates through the transitivity assumption. Consider:

$$\begin{aligned}\tilde{r}_{1,3} &= \text{median}(\{\hat{r}_{1,2} + \epsilon_{1,2}, (\hat{r}_{1,2} + \epsilon_{1,2})(\hat{r}_{2,3} + \epsilon_{2,3})\}) \\ &= \text{median}(\{\hat{r}_{1,2} + \epsilon_{1,2}\ , \hat{r}_{1,2}\hat{r}_{2,3} + \hat{r}_{1,2}\epsilon_{2,3} + \hat{r}_{2,3}\epsilon_{1,2} + \epsilon_{1,2}\epsilon_{2,3}\})\end{aligned}$$

Despite the introduction of these additional interaction terms, notice the following:

$$E\left[\hat{r}_{1,2}\hat{r}_{2,3} + \hat{r}_{1,2}\epsilon_{2,3} + \hat{r}_{2,3}\epsilon_{1,2} + \epsilon_{1,2}\epsilon_{2,3}\right] = \hat{r}_{1,2}\hat{r}_{2,3}$$

In words, as long as the individual ratios are unbiased estimates, i.e., $E[\epsilon_{1,2}] = 0$ and as long as errors are uncorrelated across votes, which is true as long as sampling is random, $E[\epsilon_{1,2}\epsilon_{2,3}] = 0$ then we see that the bottom-up median aggregator is also an unbiased estimator of the true median preferences, q^*.

Thus far we have used a bottom-up voting approach. However, the order of voting does not affect the properties of the estimator. As such, it is equally valid to do a top-down approach which will also be unbiased and consistent. While this holds in expectation it is not the case that the order of

aggregation is irrelevant for a given sample. Intuitively, under a bottom-up approach $\epsilon_{1,2}$ will propagate across all estimated ratios and thus have a much larger impact on, $\tilde{q}$ than $\epsilon_{N-1,N}$ which only affects $\tilde{q}(z_N)$. The opposite is true under a top-down aggregation.

Simulation:

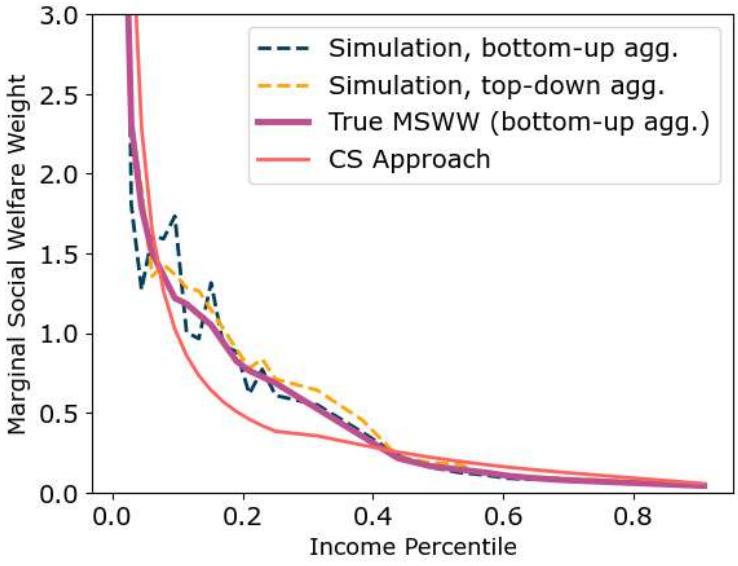
To illustrate the properties of this estimator, I perform a simulation. First, a schedule for $\tilde{q}_0$ is estimated which we impose as the true underlying preferences $\tilde{q}_0 = q^*$. Then we draw a sample of votes (N=1000) noisily distributed around the true median:

$$q_i(z_j) = (1 - \epsilon_{i,j})q^*(z_j)$$

Where $\epsilon \sim U(-1,1)$.

Without resampling, we compare the estimated preferences for three estimators. The bottom-up median aggregator, the top-down median aggregator and the OLS approach suggested by Capozza & Srinivasan (2024):

Figure A4: **Simulating Estimators for Median Preferences**



Source: Simulations.

From this the following is clear. The bottom-up and top-down approaches yield very similar results indicating that the order of voting does not affect estimates. From the bottom-up approach it is clear that the accumulation of errors *increases* rather than decreases efficiency as income increases as is apparent from the fact that the simulated MSWW converges to the true MSWW quickly as incomes increase. Lastly, when comparing to the CS approach we see that while the functional

form assumption implied in the CS approach has the notable benefit of ensuring a smooth MSWW function, this restriction leads to a consistent underestimation of MSWWs for workers with incomes between the 10^{th} and 40^{th} percentiles.