

CEBI WORKING PAPER SERIES

Working Paper 19/26

HORIZONTAL INEQUALITY IN OPTIMAL
TAXATION: EVIDENCE FROM A DEVELOPMENT
CONTEXT

Magnus Eldrup

ISSN 2596-447X

CEBI

Department of Economics
University of Copenhagen
www.cebi.ku.dk

HORIZONTAL INEQUALITY IN OPTIMAL TAXATION: EVIDENCE FROM A DEVELOPMENT CONTEXT*

Magnus Eldrup†

Abstract

Exemptions are a key feature of income tax systems in developing countries, yet standard optimal tax models cannot explain them. By assuming full compliance, these models ignore horizontal inequality between formal and informal workers, understating the social cost of taxing the former. Extending the optimal tax framework, I incorporate informality and horizontal inequality aversion. I conduct a survey experiment in Kenya, finding that respondents place substantially higher welfare weight on formal taxpayers than otherwise identical informal workers. Combining these empirical preferences with administrative tax data yields prescriptions of high exemption thresholds matching observed policy, a result alternative explanations cannot rationalize.

JEL codes: H21, H26, D63, O17, C93

*Funding from the Nordic Tax Research Council and the Danish Ministry of Foreign Affairs (MFA project ref 360 No.:25/05740) is gratefully acknowledged.

Thank you to Eric Odhiambo, Ali Mwateuzi, Eugene Amadi, Evangelyne Wanjiru, Levi Gatimu, Suleiman Daddy and Victor Maina for extraordinary enumeration support. This paper has benefited from comments from Niels Johannesen, Anders Jensen, Augustin Bergeron, Raj Chetty, Emmanuel Saez, Claus Thustrup Kreiner, David Atkin, Christopher Hoy, Joana Naritomi, Giulia Barletta and Henrik Kleven as well as from seminar participants at University of Copenhagen, Harvard University, Kiel-CEPR, CSAE and University College Dublin.

This experiment was registered in the American Economic Association's registry for randomized controlled trials under ID AEARCTR-0015719 (2025). The experiment received IRB approval from the University of Copenhagen.

†Department of Economics, University of Copenhagen, Øster Farimagsgade 5, 1353 Copenhagen K, Denmark. Email: mel@econ.ku.dk. +45 21732017

1 Introduction

Large exemptions are a central feature of income tax systems in developing countries. In Kenya, for example, 80 percent of workers have earnings below the exemption threshold, a level comparable to most developing countries. Such large thresholds have first-order implications for tax revenues and account for much of the variation in tax-to-GDP ratios across the developing world (Jensen, 2022). Yet standard optimal tax models prescribe no exemptions. This no-exemption result relies on full compliance. If every worker pays their taxes, the government can always do better by taxing a broad base than by leaving any portion of earnings untaxed (Mankiw et al., 2009; Sadka, 1976; Seade, 1982).

In practice, tax compliance is far from universal in developing countries, where informal employment is ubiquitous. Large shares of workers operate entirely outside the formal tax system and pay no income taxes, even when earning above the exemption threshold (La Porta & Shleifer, 2014). Consequently, income taxes also imply redistribution from formal to informal workers. This additional dimension of redistribution, absent from the standard model, creates horizontal inequality, commonly defined as the unequal treatment of equals (Musgrave, 1990). In this setting, horizontal inequality arises because workers with identical skill pay different amounts of tax depending on whether they are in the formal or informal sector. Such disparities are large in practice and a salient public concern across developing countries (Bachas et al., 2026).

The natural response to this evasion is, of course, to eliminate it altogether. Yet tax evasion has persisted as long as taxation itself, despite decades of enforcement efforts and policy interventions. Tax design in developing countries should therefore take evasion as a constraint, accounting for it in policy design rather than assuming it can be immediately eliminated.

This paper investigates the design of optimal income taxation in this second-best world. I examine how a planner should set taxes when evasion is pervasive. I extend the Mirrlees (1971) framework by introducing informal (untaxed) workers and by incorporating a Paretian social welfare function that separately accounts for vertical and horizontal inequality. I implement the framework for Kenya, combining a survey experiment that empirically identifies marginal social welfare weights with administrative data on the universe of taxpayers. The survey documents the prevalence of strong horizontal inequality aversion in Kenya. Using the estimated weights, the model prescribes optimal schedules with high exemption thresholds, deviating from the prescriptions of the standard model but aligning more closely with observed Kenyan policy.

First, I show theoretically that the presence of informal workers breaks the classic Sadka–Seade (1976; 1982) no-exemption result in a novel way¹. Intuitively, when some workers are untaxed, exemptions can emerge if formal workers carry higher average marginal social welfare weights than informal workers. In that case, taxing low-income formal workers is welfare reducing as the welfare costs exceed the revenue gains. Unable to set negative taxes, the optimal policy response is an exemption. This protects formal workers from bearing a disproportionate tax burden relative to their informal counterparts. Critically, the mechanism requires formal workers to have sufficiently high welfare weights, which depends on the strength of horizontal inequality aversion.

Second, I test whether such horizontal inequality aversion is present in practice by fielding a survey experiment in Kenya that directly elicits marginal social welfare weights. Respondents allocate resources between hypothetical recipients who vary in income level or tax compliance. By making redistribution costly, the approach identifies the relative welfare weights assigned by the respondent (Capozza & Srinivasan, 2024). Aggregating responses yields fully non-parametric income- and type-specific marginal social welfare weights. The results show moderately progressive vertical preferences, in line with previous estimates, but very pronounced broad horizontal inequality aversion. The median respondent assigns formal workers about four and a half times the marginal welfare weight of informal workers with an even higher mean ratio of seven. This preference is robust across respondent incomes and alternative framings.

Finally, I characterize optimal policy for Kenya. I derive sufficient statistic formulas in an extended model allowing for evasion responses to taxation through endogenous sorting between sectors, productivity *and* amenity costs of informality as well as costly tax administration. Combining the estimated welfare weights with administrative tax records and household survey data, I identify optimal tax schedules for Kenya. Even with no extensive margin and no administrative costs, the inclusion of informal workers and horizontal inequality aversion endogenously generates a sizeable optimal exemption. The prescribed threshold of 149,000 Ksh is below the observed Kenyan threshold of 288,000 Ksh but very close to the previous 2020 threshold of 147,580 Ksh. This alignment stands in sharp contrast to the standard Mirrleesian benchmark, which favors broad bases and high taxes covering all workers, and it directly challenges the conventional policy view that base broadening is always desirable (IMF, 2024; Tanzi & Zee, 2001; Zolt & Bird, 2005).

I use the extended model to examine other mechanisms that could drive the observed exemptions. First, I investigate the impact of extensive margin effects, i.e., movements between the formal and

¹ Choné and Laroque (2010) also show that Sadka–Seade can fail under multidimensional heterogeneity. Section 1.1 situates my mechanism relative to theirs.

informal sectors in response to average tax rates. I find that in the absence of horizontal inequality aversion, extensive margin effects depress optimal rates but cannot by themselves rationalize sizeable exemptions even under implausibly large elasticities.

Second, I examine whether administrative capacity constraints can rationalize exemptions (Bergeron et al., 2024; Besley, 2020; Jensen, 2022; Kleven et al., 2016). While capacity constraints clearly shape what governments are able to enforce, it is not clear a priori whether they provide a normative rationale for exemptions. To investigate, I include a marginal administrative cost of taxation for all formal workers earning above the exemption. I find that large marginal administrative costs cannot rationalize exemptions, even when collection costs apply to informal workers too. This follows from the high implied cost of forgoing revenue from all inframarginal workers above the threshold.

Finally, I consider the mechanical revenue argument, that high exemptions simply reflect low earnings and widespread informality. The data does not support this explanation. In 2021, 1.3 million workers registered in the PAYE-system earned less than the exemption threshold and applying the statutory schedule to the workers would mechanically increase annual PIT revenue by around 13.5 percent. This shows that the exemption cannot be explained solely as a second-order concern for constrained governments.

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, it improves the relevance of Mirrleesian tax models in a development context by providing a novel and empirically grounded route around the core no-exemption result, reconciling our workhorse model with observed policy choices and disciplining simple analyses calling for blind base broadening.

1.1 Literature Review

This paper contributes to the optimal income tax, social preferences, and development public finance literatures. It studies horizontal inequality between observationally similar workers who differ in their taxable status, elicits new survey evidence on how people trade off horizontal against vertical equity, and brings these elicited welfare weights into the design of an optimal income tax schedule in a developing-country context.

The paper contributes to the optimal tax literature by identifying sufficient statistic formulas for taxation in the presence of informality and by showing the conditions under which the seminal Sadka-Seade (1976; 1982) no-exemption result is broken. The closest precedent is Choné and

Laroque (2010), who show that the Sadka-Seade result can fail under multidimensional heterogeneity² of workers. In their framework, negative marginal rates at the bottom are optimal when low-skill workers receive below-average social weights, a theoretical possibility for which they concede there is no empirical guidance. I build on their approach by showing how limited compliance allows the average social welfare weight of all taxpayers to exceed one even when social weights are decreasing in income, in contrast to Choné and Laroque, whose result requires the poorest workers to receive below-average weights. I further provide a direct empirical foundation for this channel, eliciting the relevant welfare weights rather than assuming them.

A central contribution of this paper is showing how horizontal inequality can be flexibly incorporated in the full optimal tax framework without violating the Pareto condition. The need for such a measure has been widely discussed (Slemrod & Yitzhaki, 2002) but thus far without a solution. In section 2.2, I show how the social welfare function proposed in this paper resolves the issues facing other prominent approaches to this problem (Auerbach & Hassett, 2002; Kaplow, 2000; Saez & Stantcheva, 2016).

A small but growing literature brings horizontal-equity considerations into tax design in lower-income countries (Bachas et al., 2024). Studying property taxation in Brazil, Best et al. (2025) show that horizontal inequality reduces tax compliance. Meanwhile, the closest complement to this paper is Bachas et al. (2026), who document large horizontal inequities within the personal income tax system, between employees and the self-employed. The two papers take complementary approaches: while Bachas et al. (2026) study how horizontal inequality affects the choice between the VAT and the PIT, this paper takes the tax mix as given and studies how the structure of the PIT changes with horizontal inequality.

Informality spans a spectrum, from a voluntary escape from taxation to an involuntary consequence of low productivity, and in neither form is it *a priori* costless (La Porta & Shleifer, 2014; Maloney, 2004; Meghir et al., 2015; Perry et al., 2007; Ulyssea, 2018, 2020). This distinction matters for how a planner should treat workers it cannot tax. I build on this literature in two ways. First, I develop a tractable framework in which taxable and untaxable workers of the same skill enter the planner's problem side by side, so that informality is neither ignored nor reduced to

² While this paper incorporates multidimensional heterogeneity, its structure remains closer to foundational optimal income tax papers (Mirrlees, 1971; Saez, 2001) than to papers attempting to solve the double-deviation problem (Kleven et al., 2009; Rothschild & Scheuer, 2013). This is because, in this model, the planner observes worker sector and knows that informal workers will remit less taxes, thus effectively creating a 'tagged' system (Akerlof, 1978; Cremer et al., 2010; Saez & Stantcheva, 2016; Weinzierl, 2012) rather than a multi-dimensional screening problem.

costless evasion. Second, I provide a mechanism for the empirical relationship between informality and tax exemptions documented by Jensen (2022).

This paper intersects with the literature on tax evasion. Unlike Keen and Slemrod (2017), I do not treat enforcement investment as a policy variable, studying optimal tax design conditional on enforcement. Withholding and third-party reporting raise compliance on formal wages (Kleven et al., 2011), so under Kenya's PAYE system the scope for intensive misreporting is narrower than for self-employment income. Such misreporting is nonetheless possible (Feinmann et al., 2022), but the paper's mechanism does not depend on ruling it out: horizontal inequality arises from evasion on the extensive margin, between taxable and untaxable workers, not from marginal concealment by formal workers. The paper thus offers a framework for incorporating horizontal inequality into the welfare analysis of tax policy under evasion.

Finally, by empirically eliciting marginal social welfare weights, this paper contributes to the growing literature on measurement of social preferences (Capozza & Srinivasan, 2024; Saez & Stantcheva, 2016; Weinzierl, 2014) and to the small literature on measuring preferences on horizontal inequality (Bachas et al., 2026; Dwenger et al., 2024; Hvidberg et al., 2023). I advance these literatures in two ways. First, by eliciting redistribution tradeoffs through a connected network of pairwise comparisons, I recover the full schedule of welfare weights non-parametrically without normalizing to a single reference individual, thus avoiding the "common recipient" problem (Capozza & Srinivasan, 2024). Second, I provide the first estimates of marginal social welfare weights that separately identify aversion to vertical and horizontal inequality.

2 Generalized Optimal Tax Model

In this section, I first introduce a stylized model of optimal taxation under costless evasion. The model is a simple extension of Mirrlees (1971) to a framework where one type of agent has access to a costless evasion technology. I augment the conventional Mirrlees (1971) social welfare function to capture horizontal inequality concerns and discuss central criticisms from the horizontal inequality literature. I derive the optimal tax schedule and analyze limiting cases that connect the model to stylized facts about exemption policy. I then generalize the stylized model to account for endogenous extensive margin evasion, costs of informality, and administrative costs associated with taxation. I present closed form sufficient statistics for the generalized model and discuss implications for optimal policy.

2.1 Stylized 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. These agents are henceforth referred to as "taxable" workers. In the stylized model, type $j = 2$ agents have a costless evasion technology with no detection risk and are thus "untaxable" workers. These types can be thought of as analogous to formal and informal workers, as I will argue in section 3.2. In addition to their type, j , agents also possess a latent skill, n , which determines their income earning potential. In the generalized model in section 2.6 I let types be endogenous and allow for differential productivity across sectors for workers of identical skill. In online appendix O1 I derive optimal taxes when untaxable workers instead only evade a share of their income.

All agents have identical utility functions $u(c, z; n)$. Here, c , denotes consumption and, z , gross labor income. I define the indirect utility of an agent as $v_j(n) = \max_{z \geq 0} (u(z - T_j(z), z; n))$. 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. Specifically, I adopt the Saez (2001) iso-elastic normalization.

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 $f_1(n)$ for taxable agents and $f_2(n)$ 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:

$$h(n) = \bar{\rho}f_1(n) + (1 - \bar{\rho})f_2(n)$$

Where $\bar{\rho} = \rho^+(0)$. 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 2³. The planner does not know the skill of any individual and can only tax based on reported income. When setting taxes, the planner knows the skill and type distributions as well as the utility function. The planner can set a fully non-linear marginal 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

³ This assumption is benign as any worker that files taxes is naturally type 1 and any agent that does not is type 2.

of worker and marginal taxes are bounded such that $T'(z) \in [0; 1]^4$. The planner knows that only taxable agents will remit taxes implying that $T_1(z) = T(z)$ while $T_2(z) = T(0)$. The problem of the social planner is thus very close to the basic optimal tax problem (Mirrlees, 1971; Saez, 2001) and collapses to this when $\bar{\rho} = 1$.

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

$$SWF = \int_0^\infty Q(v_1(n), v_2(n), \rho(n)) h(n) dn \quad (1)$$

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, 2001) except that the social preference function, $Q(\cdot)$, now discerns between the utility of taxable and untaxable workers. In the next section, I discuss some of the central attributes we require from this function.

2.2 Measuring Horizontal Inequality

Horizontal inequality is commonly defined as “unequal treatment of equals”. In a Mirrleesian setting, “equals” can naturally be defined as agents with the same skill level n , since skill represents potential earnings and is invariant to tax policy. Consequently, assuming identical utility functions, differences in ex-post utility across agents with the same skill provide a natural measure of horizontal inequality.

This choice of equivalence class is of course itself a normative judgment. Any characteristic on which the planner can condition tax liability, in this case formality but one could consider gender and civil status in other settings, is liable to carry an intrinsic social preference over the category alongside the pure equal-treatment concern. The elicitation of this paper in section 4 thus captures a broader measure of horizontal inequality encompassing the full relative social preference across a category the planner otherwise treats as equal rather than the pure Musgravian ideal.

Existing approaches to modelling horizontal inequality face well-known limitations. Many earlier measures impose penalties for rank reversals or unequal treatment within income classes (Duclos et al., 2003; King, 1983), but these introduce non-utility information and violate welfarism, allowing negative marginal weights (Kaplow & Shavell, 2000). Auerbach and Hassett (2002) proposed a nested vertical–horizontal aggregator, which resolved many of these issues, but because it depends on pre-tax incomes, it fails the strong Pareto criterion. Additionally, it is not applicable

⁴ This restriction is discussed further in section 3.1.

when labor supply is endogenous. Saez and Stantcheva's (2016) local approach avoids some of these issues but their definition forces horizontal equality to be the overriding objective. Further, their local approach does not guarantee a globally consistent welfare ranking (Sher, 2024).

The formulation used in this paper addresses these concerns. It retains the nested structure of Auerbach and Hassett (2002) but is fully Paretian because it aggregates only utility-based information. Defining horizontal equality with respect to the invariant skill parameter allows the measure to operate within a Mirrleesian model with behavioral responses. Unlike the local approach of Saez and Stantcheva (2016), the framework is globally defined and yields consistent welfare comparisons for finite reforms. It nests the utilitarian benchmark when horizontal inequality carries no weight and the Rawlsian limit when it becomes the overriding concern and any flexible combination of vertical and horizontal preferences as requested by Stiglitz (1982). As such, it provides the first fully welfarist and operational measure of horizontal-inequality aversion applicable to optimal tax design.

A remaining practical concern is whether taxable workers can meaningfully be viewed as horizontally disadvantaged⁵ relative to untaxable workers of the same skill, and why they do not simply switch sectors to equalize utility. This concern matters because it is equivalent to asking whether sectoral differences represent genuine horizontal inequality or endogenous choice. The generalized model of Section 2.6 makes this precise by treating sectoral choice as endogenous. Free movement only equalizes utility for the *marginal* worker at the cutoff; inframarginal formal workers can remain strictly disadvantaged. The horizontal inequality the planner responds to is therefore consistent with free switching.

A further consideration is the welfarist scope. Even with switching fully endogenous, a planner may restrict attention to "socially relevant" utility components such as consumption, skill, and effort, while treating utility from sectoral amenities as normatively irrelevant. Under such a censored SWF, the utility associated with sectoral choice is disregarded.

Empirically, none of this hinges on a claim about whether formal or informal work is intrinsically more desirable. The direct elicitation of marginal social welfare weights in Section 4 allows the analysis to remain agnostic on that issue by capturing how respondents themselves value redistribution across sectors. Kenyan respondents expressed a clear aversion to inequality arising from differences in tax status. These responses reflect both perceived fairness across sectors,

⁵ The informal sector has historically been viewed as disadvantaged (Gollin, 2014; Harris & Todaro, 1970; Lewis, 1954), though recent evidence is mixed (Blattman & Dercon, 2018; Maloney, 2004).

perceived sectoral productivity differentials and any embedded (dis)amenity associated with working in the taxable sector.

2.3 The Generalized Social Welfare Function

I disentangle vertical equality preferences, $G(\cdot)$, from horizontal equality preferences, $H(\cdot)$, using a nested structure.

$$SWF = \int_0^\infty H \left(G(v_1(n)), G(v_2(n)); \rho(n) \right) h(n) dn \quad (2)$$

Here $G(\cdot)$ follows the optimal income tax literature, it is positive, it has a positive first derivative and is weakly concave in utility. The horizontal preference function, $H(\cdot)$, navigates the vertically transformed utilities, $G(v)$, under horizontal inequality. I propose the functional form for $H(\cdot)$.

$$SWF = \int_0^\infty \left(\sum_{j=1}^2 G(v_j(n))^\phi \cdot \rho_j(n) \right)^{\frac{1}{\phi}} h(n) dn \quad (3)$$

Here $H(\cdot)$ is the generalized power mean formula. While this paper only discerns two types, the formulation is easily adapted to any number of types. 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 vertically transformed utility, I denote this *horizontal utilitarianism*. For $\phi \rightarrow -\infty$ 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 $\phi \rightarrow -\infty$ then the planner places no value on additional utility attained through tax evasion but remains strictly utilitarian across skills. For $G(v) = \min(v)$ the planner is a typical Rawlsian and only assigns weight to the agent with the lowest utility in the population. Note that the case with linear G and quasi-linear utility is degenerate as it leaves the planner no vertical motive to redistribute, regardless of ϕ .

Marginal Social Welfare Weights

We can derive the marginal social welfare weights (MSWW) using the general formulation of the SWF, eq. (3). Let $\lambda > 0$ be the marginal value of public funds, determined by the government budget constraint.

$$q_j(n) = \frac{1}{\lambda \rho_j(n)} H_{G|j}(\cdot) G_v(v_j(n)) u_{c|j}(n) \quad (4)$$

where $H_{G|j}(\cdot) \equiv \frac{\partial H(G(v_1), G(v_2); \rho)}{\partial G(v_j)}$ is the sensitivity of the horizontal aggregator to a change in the vertically transformed utility of type j . The MSWW corresponds to the per-capita money-metric social value of one additional unit of consumption for an agent of type j and skill n :

The assumption of no income effects implies that λ must adjust to ensure unit value of a uniform transfer $\int [\rho(n)q_1(n) + (1 - \rho(n))q_2(n)]h(n)dn = 1$. When using weights over income instead of skill we use the notation $\tilde{q}_j^+(z) \equiv q_j^+(n_j(z))$. Finally, it is useful to define average weights above a given income as:

$$\tilde{q}_1^+(z) = \frac{\int_z^\infty \tilde{q}_1(y) \tilde{f}_1(y) dy}{1 - \tilde{F}_1(z)} \quad (5)$$

2.4 Conditions for Tax Exemptions

Before turning to optimal rates, this section shows, intuitively, how the presence of untaxable workers can break the Sadka-Seade (1976; 1982) result, by generating exemptions at the bottom of the income distribution. Maintaining assumptions of (i) no income effects, (ii) $T'(z) \in [0, 1]$, and (iii) undifferentiated lump-sum, we get the following proposition:

Proposition 1: If average MSWWs are higher for taxable than for untaxable workers, $\tilde{q}_2^+(0) < \tilde{q}_1^+(0)$, then the optimal schedule features an exemption $\bar{z} > 0$ for taxable workers.

The complete mechanism-design proof appears in A8. To build intuition, I derive the same result intuitively using a local perturbation argument.

Perturbation argument:

The argument proceeds by asking whether it is welfare-improving to introduce a small positive marginal tax at the bottom of the taxable income distribution. If the social value of consumption

among taxable workers is sufficiently high, even an infinitesimal tax increase is welfare-reducing, implying a zero marginal rate over a positive range.

Consider first the benchmark in which all agents are taxable, $\bar{\rho} = 1$. Let z_0 denote the lowest observed income and choose $z_1 = z_0 + \epsilon$ with $\epsilon > 0$ small. Starting from $T'(z_1) = 0$, consider an infinitesimal increase $dT'(z_1) > 0$ applied over a vanishing interval around z_1 .

The mechanical revenue raised from all agents earning above z_1 is:

$$dM = dT'(z_1) \cdot \bar{\rho} \cdot [1 - \tilde{F}_1(z_1)]$$

The associated money-metric welfare effect of the mechanical effect is:

$$dW = - \int_{z_1}^{\infty} \tilde{q}_1(z) dT'(z_1) \bar{\rho} \tilde{f}_1(z) dz = -\tilde{q}_1^+(z_1) \cdot dM$$

With quasi-linear utility and $T'(z_1) = 0$, there is no behavioral effect of the perturbation, $dB = 0$.

A tax increase will be welfare improving if and only if $dM + dW + dB \geq 0$ which yields:

$$-dW \leq dM \Rightarrow \tilde{q}_1^+(z_1) \leq 1$$

In the standard Mirrlees setting $\tilde{q}_1^+(z_0) = 1$ and, by monotonicity, $\tilde{q}_1^+(z) < 1$ for all $z > z_0$ so no exemption arises⁶. Now introduce the untaxable workers. This only changes the normalization condition of social welfare weights:

$$\bar{\rho} \tilde{q}_1^+(z_0) + (1 - \bar{\rho}) \tilde{q}_2^+(z_0) = 1$$

If untaxable workers receive lower weights, $\tilde{q}_2^+(z_0) < \tilde{q}_1^+(z_0)$, normalization implies $\tilde{q}_1^+(z_0) > 1$. By continuity, $\tilde{q}_1^+(z_1) \geq 1$ for sufficiently small $\epsilon > 0$, so raising marginal taxes on the lowest taxable earners would reduce welfare. The planner therefore sets $T'(z_1) = 0$ for a positive income range $[0, \bar{z}]$, producing an exemption. This result holds even in the presence of consumption taxes⁷.

Intuitively, when some workers with low MSWWs cannot be taxed, the planner must assign greater social value to consumption within the taxed population to satisfy the overall normalization. This raises the average social value of consumption for taxable workers above unity, making any positive

⁶ Choné & Laroque (2010) show that this conclusion can fail under multidimensional heterogeneity, when the lowest-skill workers receive below-average social weights. They generate this through heterogeneity in the opportunity cost of work. My mechanism instead preserves weights that are decreasing in income.

⁷ A uniform VAT is common to both types and hence horizontally neutral; under quasi-linearity it leaves the relative welfare weights and the taxable-income distribution unchanged. The exemption condition of Proposition 1, $\tilde{q}_1^+(0) > 1$, is therefore unaffected. I confirm this empirically in online appendix O7

marginal tax on them welfare-reducing. The exemption thus functions as the feasible counterpart to a negative transfer when $T'(z) < 0$ is infeasible.

2.5 Optimal Marginal Tax Rates

I now characterize the optimal marginal tax rates implied by the planner's problem. Let $\tilde{q}_1^+(z)$ denote the average MSWWs for taxable agents earning above z , and let $\alpha_1(z)$ and $e_1(z)$ denote, respectively, the local Pareto parameter and the elasticity of taxable income with respect to the net-of-tax rate of the taxable workers.

Proposition 2: The unconstrained optimal marginal tax rate at income z satisfies

$$T'(z) = \frac{1 - \tilde{q}_1^+(z)}{1 - \tilde{q}_1^+(z) + \alpha_1(z) \cdot e_1(z)} \quad (6)$$

The proof for this proposition can be found in appendix A8. The expression closely mirrors the standard sufficient-statistics solutions to the problem (Diamond & Saez, 2011; Kleven, 2021; Saez & Stantcheva, 2016), but is defined solely over the taxable agents as reflected in $\tilde{q}_1^+$, α_1 and e_1 .⁸

The simplicity of this FOC might seem counter-intuitive given the inclusion of a second, untaxable type of worker. Unpacking this we see that although the social welfare function aggregates utilities across all workers, untaxable workers do not respond to changes in marginal tax rates and do not generate fiscal externalities; hence they drop out of the planner's problem. Their presence affects optimal taxation only indirectly, by altering the marginal value of public funds. If the requirement that untaxable workers remit *no* taxes is relaxed, the problem still has a closed-form solution and proposition 1 still holds. I derive the formula for this case in online appendix O1.

As shown in proposition 1, the value of $\tilde{q}_1^+(z)$ can exceed 1, even though preferences are Paretian, leading to negative marginal taxes. Extending the solution to the more empirically relevant constrained problem where marginal rates are non-negative simply clamps the schedule, producing a flat segment at the bottom given as $T'_c(z) = \max[0, T'(z)]$. In appendix A9 I derive this formula by solving the constrained problem with a Kuhn–Tucker inequality condition.

⁸ For $\tilde{q}_1^+(z) \leq 1$ we observe standard behavior and $T'(z) \in [0, 1]$. For $\tilde{q}_1^+(z) \in (1; 1 + \alpha_1(z) \cdot e_1(z))$ it is optimal for the government to set $T'(z) < 0$, leading to tax exemptions in the constrained problem. For $\tilde{q}_1^+(z) > 1 + \alpha_1(z) \cdot e_1(z)$ the formula instead identifies the local minimum of the SWF. This case is of no empirical importance.

Vertical Preferences and the Placement of the Exemption

Given the FOC in eq. (6), it is clear that since the exemption is defined by the crossing point $\tilde{q}^+(z) = 1$, the vertical preferences, governing the rate at which $\tilde{q}^+(z)$ declines, matter for the placement of the exemption. Consider two edge-cases. Under Rawlsian preferences, $\tilde{q}^+(z)$ is zero at all income levels above the lowest, leading to no exemption. Conversely, if $\tilde{q}^+(z)$ is constant and larger than one, the optimal tax system exempts all workers⁹.

In other words, while the existence of an exemption depends critically on the degree of horizontal inequality aversion, it is the interplay between horizontal and vertical preferences that determines the exact placement of the exemption threshold.

2.6 Endogenous Sorting, Productivity and Administrative Costs

Until now the characterization has held sectoral status fixed to focus on the horizontal inequality mechanism. However, three additional margins merit consideration. First, in reality, sectoral affiliation is not fixed, and workers can switch to the informal sector when faced with higher tax burdens. Second, there may be a latent productivity differential between the two sectors such that an equally skilled worker putting in the same effort may earn a lower income in the informal sector. Finally, tax collection is costly. As such, the marginal administrative cost of widening the tax base may have an impact on optimal policy. To account for this I generalize the model as follows.

Workers are indexed by skill n and an idiosyncratic informal-sector opportunity θ . Formal employment delivers indirect utility $v_1(n; T)$ and informal employment $v_2(n, \theta)$, where informal output carries a productivity wedge $\kappa(n) \in [0, 1]$ and a non-pecuniary Roy-style term $\psi(n, \theta)$. A worker selects formal employment when $v_1(n; T) \geq v_2(n, \theta)$. The administrative cost per formal taxpayer above the exemption is γ , so the fiscal value of a worker who newly enters formality is $T(z) - \gamma$ above the threshold. In online appendix O2.7, I let informal workers incur an administrative cost too. Let $\tilde{\eta}_\rho(z)$ denote the semi-elasticity of the formal share with respect to formal disposable income, expressed in earnings space, and define the Roy fiscal correction

$$\Gamma^+(z; \bar{z}) = \frac{1}{1 - \tilde{F}_1(z)} \int_z^\infty [T(x) - \gamma \mathbf{1}_{\{z(x) > \bar{z}\}}] \tilde{\eta}_\rho(x) \tilde{f}_1(x) dx$$

⁹ Note that we cannot generally say that flatter MSWWs lead to higher exemptions ex-post, as this depends on both the renormalization of the weights and the exact shape of the MSWW schedule.

This captures the fiscal effect of a tax change from inframarginal workers changing their sectoral affiliation in response to changes in their average tax rate. Here $T(z) - \gamma$ is the revenue loss net of administrative costs.

Proposition 3: Under endogenous sorting and administrative costs, the optimal unconstrained marginal tax rate at income z and the optimal exemption $\bar{z}$ satisfy

$$T'(z) = \frac{1 - \tilde{q}_1^+(z) - \Gamma^+(z; \bar{z})}{1 - \tilde{q}_1^+(z) - \Gamma^+(z; \bar{z}) + \alpha_1(z) \cdot e_1(z)} \quad (7)$$

$$\gamma \tilde{f}_1(\bar{z}) = T'(\bar{z}) [1 - \tilde{F}_1(\bar{z})] [1 - \tilde{q}_1^+(\bar{z}) - \Gamma^+(\bar{z}; \bar{z})] \quad (8)$$

Here, eq. (7) characterizes optimal rates, while eq. (8) pins down the optimal exemption conditional on a continuous tax schedule. A sketch of this proof can be found in A10 while the full proof using the Jacquet & Lehmann (2021) perturbation approach is in online appendix O2. Note that this condition nests proposition 2 in that when $\tilde{\eta}_\rho = 0$ the two conditions coincide. The ex-post optimal schedule is found as a fixed point in both the endogenous right-hand-side objects and the exemption boundary $\bar{z}$.

Naturally, the new Roy correction term $\Gamma^+(z; \bar{z})$ in eq. (7) and the cost term $\gamma \tilde{f}_1(\bar{z})$ in eq. (8) now allow for the possibility that exemptions can occur, even when $\tilde{q}_1^+ < 1$. In section 6, I investigate this empirically and find that, without horizontal inequality aversion, the extended model does not produce sizeable exemptions under plausible assumptions.

The general model also illustrates that the impact of any cross-sectoral productivity differential only occurs indirectly through other observables.

Proposition 4: When preferences, formal incomes, and elasticities are observed, any productivity differential between formal and informal work has no direct impact on optimal tax rates.

Proof is in appendix O2.5. Intuitively, we see that productivity differentials affect taxes in two main ways. Either through the relative welfare of workers across sectors, which is captured in $\tilde{q}_1^+$, or through a fiscal impact from switchers. However, since informal workers pay no income taxes regardless of their income, the productivity differential has no independent impact here.

3 Institutional Setting

3.1 Setting: Personal Income Taxes in Low-Income Countries

Personal income tax systems in developing countries differ markedly from those in high-income countries. Low-income countries collect a smaller share of revenue from personal income tax (PIT) and rely more heavily on VAT and trade taxes (Besley & Persson, 2014). They also have very narrow PIT bases owing to their large tax exemptions. Countries with GDP per capita between 1,600 and 3,000 USD, the relevant range for Kenya, have a median PIT exemption covering 91 pct. of workers (Jensen, 2022). Even the smallest exemption in this group still excludes 76 pct. of workers. Against this backdrop, Kenya's exemption, which covers 80 pct. of workers, is comparatively modest¹⁰.

The desirability of narrow income tax bases in developing economies is contested. The prevailing wisdom is that it is optimal to keep tax exemptions low and tax bases large (Fall, 2022; OECD, 2012). The standard intuition is that narrow tax bases place excessive revenue pressure on a limited number of taxpayers thus reducing efficiency. This paper challenges that view by identifying a plausible mechanism under which exemptions are welfare-improving.

While it is theoretically feasible to extend the tax schedule below zero and provide negative income taxes at the bottom of the distribution, such policies are almost never implemented in low-income countries. This is because negative taxes require accurate income verification, strong administrative capacity, and reliable payment systems, features that are typically weak or missing in these settings. Negative taxes also create incentives for misreporting or income fabrication, making them more vulnerable to fraud than a conventional positive tax system.

3.2 The Kenyan Tax System

The personal income tax system in Kenya derives almost all (98 pct.) of its tax revenues from withholding taxes (PAYE) (Kanina et al., 2025). This validates the assumption that informal (untaxable) workers effectively remit no personal income taxes. Tax revenues predominantly stem from public employees (39 pct.) and employees of large private firms (35 pct.)¹¹.

Although informal workers remit next to no personal income taxes, they may in principle be subject to other income taxes. In practice, however, such taxes play a negligible role for income-based redistribution. Kenya's turnover tax (TOT), which targets small, unincorporated businesses, has historically had extremely low take-up, with fewer than 10,000 active filers per year prior to

¹⁰ This group consists of Timor Leste, Côte d'Ivoire, Tanzania (Mainland), Burkina Faso, Uganda, Mali, Cameroon, Bangladesh, Cambodia, and Kenya. The lowest exemption rate in this set is observed in Côte d'Ivoire.

¹¹ These shares are averages from 2017-22 as reported in Kanina et al. (2025)

2020 and contributing less than one percent of total tax revenue (Hoy et al., 2024). Even following recent enforcement efforts, the number of TOT filers remains small relative to the size of the informal workforce. I therefore abstract from simplified regimes in the simulations.

Informal workers also remit VAT on consumption. As such, the results in this paper characterize residual PIT conditional on a common VAT rather than the consolidated wedge across all instruments. Because effective VAT rates rise with income in these settings (Bachas, Gadenne, et al., 2024), the VAT supplies its smallest wedge precisely at the bottom, where the exemption binds, and so cannot by itself generate it. This focus on the income-tax schedule rather than the tax mix is also where the public locates the horizontal-equity margin as shown in Bachas et al. (2026). I confirm in online appendix O7 that the optimal exemption is invariant to any VAT wedge.

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 effectively equivalent to a zero-rated bracket for earnings below 288,000 Ksh¹². 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 (PwC Kenya, 2025). For the purposes of this paper, these contributions are not treated as taxes. I examine counterfactual tax schedules for Kenya in 2021, the only year for which survey data are available.

In 2021, 55 pct. of PAYE-registered (formal) Kenyan workers earned taxable income above the official exemption threshold (KRA, unpublished data, 2021). Meanwhile, in the informal sector 89 pct. of all workers earn below the exemption threshold (KNBS, 2021a). This implies that 9.4 pct. of the current Kenyan labor force were evading on the extensive margin in 2021¹³. Throughout, I classify workers as taxable if they had any PAYE-registered employment during the year. Some workers therefore appear with low annual formal income, which could reflect either part-year formal work or remuneration below the statutory minimum, a widespread practice in the region¹⁴.

4 Direct Elicitation of Marginal Social Welfare Weights

¹²As 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.

¹³ The calculation is as follows: 11 pct. informal workers with positive tax liability multiplied with 85.9 pct share of informal workers in the total labor force.

¹⁴ Bhorat et al. (2017) find that only around 50 pct. of Kenyan wage workers comply with the minimum wage.

Analysis of optimal income taxation necessarily rests on assumptions about social preferences, yet such preferences are difficult to measure. In the optimal tax framework, social preferences enter policy design through marginal social welfare weights (MSWWs), which summarize how society values marginal transfers across individuals and therefore constitute a more tangible empirical object¹⁵. Unfortunately, conventional survey measures of redistributive preferences do not feed easily into optimal tax models because they lack a welfare-weight interpretation¹⁶, and until recent work by Dwenger et al. (2024), there has been little evidence on preferences regarding horizontal inequality.

To address this gap, I directly elicit MSWWs via a controlled redistribution experiment capturing preferences over both vertical and horizontal inequality¹⁷. The approach builds on the pairwise comparison method of Capozza and Srinivasan (2024), who show that repeated allocations of a shrinking endowment can identify relative welfare weights and recover the full MSWW schedule of the respondent. By eliciting social preferences directly, I obtain an empirical test of whether horizontal inequality aversion is quantitatively strong enough to rationalize the large income tax exemptions observed in developing countries.

4.1 Methodology

Experimental Task

I elicit marginal social welfare weights using a ‘leaky-bucket’ allocation task (Okun, 1975). Each respondent divides a hypothetical Ksh 1,000 between two recipients, R1 and R2. In every choice R2 has lower consumption than R1. Let c_1 and c_2 denote the post-allocation consumption levels of R1 and R2, respectively. Allocating more to R2 reduces the total endowment, implying that redistribution is costly.

Under standard envelope arguments, the slope of the frontier at the chosen allocation, i.e., the units of consumption the respondent is willing to take from the R1 to give one more unit to R2, reveals the ratio of marginal social welfare weights:

$$\frac{q_{j_2}(z_2)}{q_{j_1}(z_1)} = -\frac{\partial c_1}{\partial c_2} \quad (9)$$

¹⁵ In this section I discuss the *positive* aspect of social preferences. There is of course no guarantee that the median preferences in Kenya are *normatively* correct.

¹⁶ 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).

¹⁷ The experiment was pre-registered at AEA <https://www.socialscisceregistry.org/trials/15719>

The task spans the full range of weakly redistributive preferences from no taste for redistribution where equal weight is assigned to both recipients, $\frac{\partial c_1}{\partial c_2} = -1$, to the strongest possible preference where the participant only values the consumption of the disadvantaged recipient, $\frac{\partial c_1}{\partial c_2} = -\infty$. The full choice-set is presented in online appendix O3. Allowing choices to vary on a fine discrete grid 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 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.

Because the horizontal comparisons are constructed using observed tax liabilities, they are identified for income levels where tax liability is positive. Below the exemption, the task does not generate variation in tax status. This motivates the structural SWF estimation, which allows consistent evaluation of policy changes across the full distribution for counterfactual changes in the exemption threshold. Anchoring the elicitation on observed liabilities is what allows the structural SWF to be fitted to reproduce the weights implied by the current tax system, which in turn permits consistent evaluation of counterfactual changes in the exemption threshold, with the sub-threshold weights an output of the fitted SWF rather than an imposed extrapolation.

The horizontal comparison is constructed at equal gross income rather than equal skill, even though the model defines horizontal inequality over skill. This is deliberate as the planner only observes income and recovering skill requires an assumed elasticity of taxable income. Eliciting at equal skill would make the elicitation conditional on a pre-specified ETI, whereas comparing at equal income recovers the policy-relevant object directly. Under the Spence-Mirrlees condition the income-skill mapping is invertible, so online appendix O6 plots the weights across skills and confirms that the income-skill mapping does not meaningfully affect the ratio of the weights.

Framing considerations matter for redistributive preference elicitation. Prior work shows that stated preferences for redistribution are sensitive to how choices are presented and what

information is made salient (Alesina et al., 2018; Kuziemko et al., 2015). The benchmark design here uses a “gain–gain” framing: respondents divide a shrinking Ksh 1,000 pot between two recipients neither of whom has any prior claim to the pot. This avoids the endowment effects and moral entitlement cues that arise when one recipient is framed as initially owning the resource.

To test whether results are sensitive to this framing, 20 percent of respondents are randomly assigned to a “loss–gain” framing in which they may take up to Ksh 1,000 from the higher-income recipient and transfer it to the lower-income recipient at a cost. This version more closely resembles the structure of a tax but risks inducing reference-dependence and loss aversion (Capozza & Srinivasan, 2024; Charité et al., 2022). The average allocation choices are nearly identical across framings; differences in vertical and horizontal preferences are statistically insignificant ($p = 0.60$ and $p = 0.85$, respectively) and economically insignificant as well.¹⁸

A second framing concern is how the distinction between taxable and untaxable workers is communicated. To avoid priming respondents with moral judgments about compliance, the survey presents only two pieces of information about each recipient: their gross income and how much they remit in income taxes. No explanations, labels, or narratives are provided. This “minimal” framing captures how respondents weight individuals who differ only in compliance status, without the survey imposing assumptions about legitimacy, evasion motives, or deservingness. Consistent with Section 2.2, the elicited ratio should be read as the full relative social preference across tax status rather than a pure Musgravian equal-treatment judgment, and it is precisely the object that enters the marginal social welfare weights. The framing is presented in appendix O3.

Vertical preferences are determined by comparing “tax-compliant” workers across the income distribution. To avoid embedding horizontal preferences in these comparisons, both recipients in the vertical arm are presented as “Pays his taxes,” even where one worker is below the threshold, and no tax amounts are stated. Stating that one worker remits more than another could invite equal-sacrifice reasoning rather than the marginal welfare weights the model requires (Weinzierl, 2014). The contrast with the horizontal arm, where tax amounts are shown, is by design: the vertical comparison varies income holding compliance fixed, while the horizontal comparison varies compliance holding income fixed.

¹⁸ Point estimates are nearly identical: the mean pooled vertical MSWW ratio changes from 3.88 to 3.97 and the mean horizontal ratio from 6.77 to 6.65. Economically, 3.88 means respondents were willing to take 3.88 Ksh from the richer recipient to give 1 Ksh to the poorer; the alternative framing raises this by only 0.09 Ksh.

All responses in this survey are unincentivized, meaning that the chosen allocation by the respondents does not affect any real-world outcomes. Capozza and Srinivasan (2024) test for whether including real monetary incentives has any effect on their estimates and find no effects.

Aggregating preferences across respondents requires difficult methodological choices. While previous approaches (Capozza & Srinivasan, 2024) impose a parametric structure on social welfare weights and recover a single progressivity parameter, this paper instead directly identifies non-linear, income-specific MSWWs. The parametric approach requires the income of one recipient (“common recipient”) to be held fixed across all questions, and Capozza and Srinivasan (2024) show that the choice of this income mechanically shifts the estimated degree of progressivity, creating an identification problem¹⁹.

In this paper, no parametric restrictions are imposed to find MSWWs and no common recipient is needed. For each pair of discrete incomes and types, the relative social welfare weight $\frac{q_1}{q_2}$ is estimated as the median of respondents’ log-ratios, which is the consistent estimator of the true population median MSWW under weak conditions on the comparison graph and co-monotonicity. This consistency guarantee does not hold for the parametric estimator. The median is the natural aggregation statistic as it corresponds to the preference a majority-rule planner would implement, and because relative valuations are unbounded above. These pairwise medians are then combined to recover the full set of estimated MSWWs, $\hat{q}_j^*(z)$, over the income-type grid, using the ‘all-pair’ identification procedure described in Appendix A11, which uses all available pairwise comparisons jointly rather than chaining ratios along a single path, so that no single comparison is multiplicatively propagated across the grid.

The empirical MSWWs, $\tilde{q}_1^*(z)$, are then used to fit the parameters of the Kenyan social welfare function, Θ , by matching the SWF-implied MSWWs to these estimated schedules via the loss function defined in Appendix A11. This delivers a structural SWF that reproduces current preferences under the existing tax system while remaining fully endogenous to policy changes. The structural fit implies that all workers below the threshold have identical MSWWs across sectors. The counterfactual ratios are pinned down by the fit above the threshold and the CRRA functional form. As a robustness check, in online appendix O8, I force horizontal inequality aversion to move towards horizontal utilitarianism below the threshold and find no impact on optimal exemptions.

¹⁹ 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 incomes. 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.

Sampling:

Data was collected in June 2025 via in-person tablet interviews in 10 subcounties across Nairobi, Kiambu, Machakos and Kajiado, covering 20 percent of the total population across 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 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. As the survey was administered in person, task comprehension was ensured through enumerator-led explanations. Additionally, all respondents completed a comprehension module illustrating the trade-offs they would face in the survey.

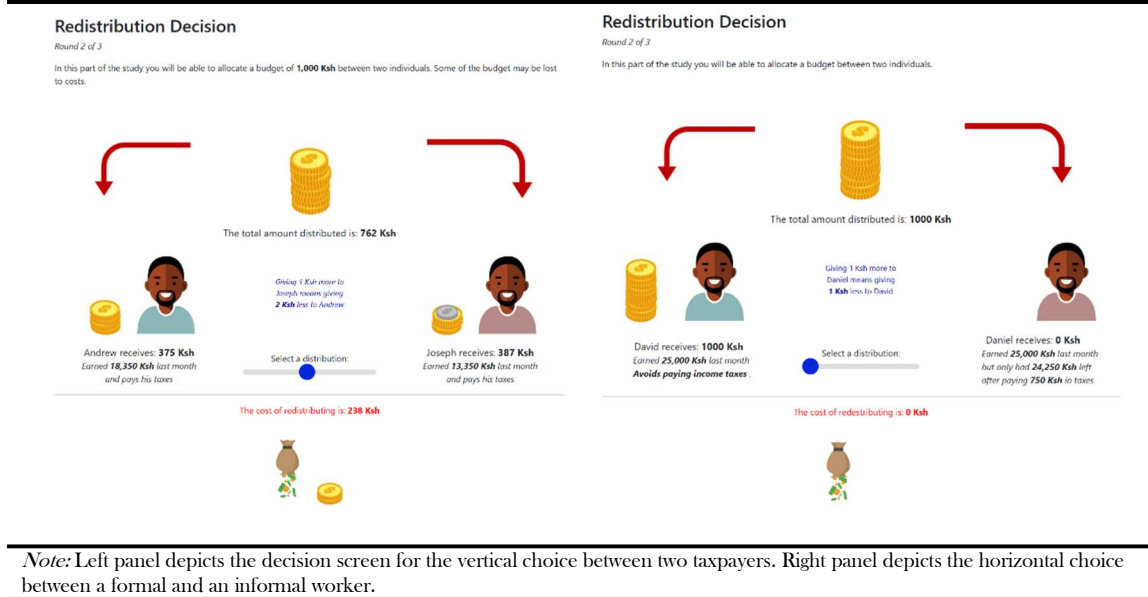
The gross sample comprises 652 completed interviews. Weights combine subcounty design weights with raking adjustments to match national age–gender–income margins. Pre- and post-weight balance along with the full survey methodology is presented in online appendix O3.

Data Quality

Standard data quality checks were applied prior to analysis. Incomplete responses and observations associated with enumerator non-compliance were excluded, and responses with implausibly short completion times were dropped to ensure adequate engagement with the task. After these filters, the cleaned sample consists of 530 respondents across both framings; the main analysis sample is the 422 gain-gain respondents, consistent with the pre-registration.

To support comprehension and data quality, all monetary amounts were displayed both numerically and using animated coin stacks that shrink and grow proportionally with allocation choices. This reduces reliance on numeracy and facilitates engagement with the allocation task.

Figure 1: Survey Decision Screens



To avoid biasing responses, each decision screen in the gain-gain framing has a random initial allocation. I leverage this to test for anchor bias and find no correlation for horizontal choices and a very slight positive correlation for vertical choices ($R^2=0.003$). Consistent with this, only 6.5 pct. of respondents select the initial allocation, 1.9 percentage points above the random benchmark, indicating limited anchoring.

As a further quality check, I examine internal consistency across choices. I define a *nested choice* as a pair of vertical allocation problems faced by the same respondent in which one income pair strictly dominates another, in the sense that the high-income recipient has lower income and the low-income recipient has higher income in one problem than in the other. Under standard assumptions, the implied MSWW ratio in the nested problem must be weakly lower than in the outer problem. Among the 512 nested choices observed, two-thirds of responses satisfy this condition. When restricting attention to economically meaningful deviations, three-quarters of responses are consistent. For comparison, in a related design, Capozza and Srinivasan (2024) report inconsistency (non-monotonicity) in 75 pct. of responses, although the two measures are not directly comparable.

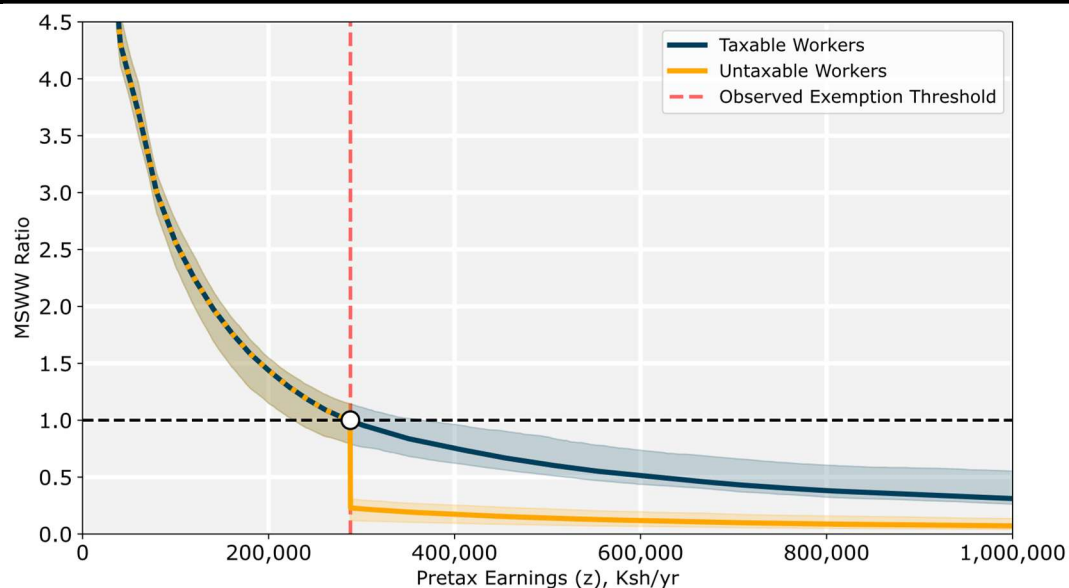
Finally, I test whether vertical and horizontal inequality aversion change across rounds and cannot reject the null of stable preferences ($p = 0.82$ and $p = 0.77$ respectively). I then pool the MSWW ratios implied by all vertical and horizontal allocation choices respectively and compare their averages. The null that preferences are identical across differences in income and differences in

compliance is strongly rejected ($p < 0.0001$). In online appendix O3, I present the full data cleaning and validation procedure.

Results

Aggregating survey responses yields schedules of welfare weight ratios which can be normalized to marginal social welfare weights $\tilde{q}_1(z)$ and $\tilde{q}_2(z)$ for taxable and untaxable workers respectively. I report histograms of pooled responses in O3. In subsequent policy simulations I account for uncertainty via bootstrap resampling as outlined in A11. In line with the pre-registration, only the 422 respondents that received the gain-gain framing are used for the main analysis. Below I plot the MSWWs relative to the worker earning just below the exemption to illustrate the relationship between horizontal and vertical preferences.

Figure 2: Empirical Estimates of MSWW Ratios, Kenya 2025



Note: Each series plots the ratio of marginal social welfare weights (MSWWs) relative to the worker just below the exemption threshold. The dotted portion of the untaxable schedule below the threshold is an imputation, as tax status does not vary in the elicitation task at those income levels. Shaded bands indicate 95% bootstrap confidence intervals from $N=1000$ draws. The dashed horizontal line marks equal MSWW to the benchmark worker. The dashed vertical line indicates the observed exemption threshold. As noted in the pre-registration, only workers who received the gain-gain framing are included in the analysis.

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

Figure 2 shows that preferences across incomes are moderately progressive. A taxpayer earning 100,000 Ksh receives about 1.8 times the weight of someone earning 200,000 Ksh. Fitting the Capozza & Srinivasan (2024) income elasticity measure of progressivity to the MSWWs of taxable workers yields $\nu = -0.54^{20}$. This is slightly less progressive, albeit close to, the progressivity estimated for the US ($\nu = -0.67$).

²⁰ This comparison measure is computed by simple OLS: $\ln \hat{q}_1^*(z) = \alpha + \nu \ln z$

In contrast to this, the results indicate a very strong aversion to horizontal inequality with the weighted (representative) median ratio between social weights assigned to taxable and untaxable workers of identical incomes of 4.4 above the threshold. This is lower than the unweighted median of 6.6. This means that the median respondent would be indifferent between giving 1 Ksh to a taxpayer or 4.4 Ksh to an informal worker of identical income. Following Section 2.2, this ratio captures the full relative social preference across tax status and not a pure equal-treatment judgment, reflecting any fairness, productivity, or amenity considerations respondents attached to compliance. This is precisely the broad object the marginal social welfare weights require. I verify that the result is not an artifact of the income–skill mapping: online appendix O6 replicates the plot across skill levels with almost identical results. In appendix O5, I recover MSWWs in levels under different normalization assumptions.

To examine heterogeneity, I test whether preferences vary systematically with gender or income. I fail to reject the null that preferences are identical in both dimensions.

5 Results

This section computes optimal nonlinear income tax schedules for Kenya.

5.1 Data and Distributions

A key advantage of the sufficient-statistics formulation of the optimal tax problem is that it can be simulated using a small set of empirically observable objects. In the present setting, these are:

- the distribution of gross incomes for taxable workers, $\tilde{f}_1(z)$
- the distribution of gross incomes for untaxable workers, $\tilde{f}_2(z)$
- the share of taxable workers in the labor force, $\bar{\rho}$
- the elasticity of taxable income, $e_1(z)$
- the marginal social welfare weights on taxable workers, $\tilde{q}_1(z)$

I construct the taxable-worker income distribution from administrative data provided by the Kenya Revenue Authority. Specifically, I use a cross-checked dataset of matched employee–employer returns (Kanina et al., 2025), which covers the universe of PAYE taxpayers in 2021. To define gross taxable income, z , I use data on 'taxable income' defined as the sum of 'employment income' and 'other income'. To match the household survey, I restrict to using data from 2021.

To identify the income distribution of untaxable workers I use the 2021 Kenya Continuous Household Survey (KNBS, 2021a). This survey records both written contracts and PAYE status.

Consistent with the institutional setting, I classify informal untaxable workers as individuals aged 18–65, salaried or self-employed, with gross monthly earnings above 1,500 Ksh who are not subject to PAYE.

While the KCHS has almost complete coverage of salaried informal incomes, some self-employed workers do not report earnings directly. To address this, I follow Kanina et al. (2025) and impute gross income for workers with missing income using a consumption-based measure. This imputation does not affect the distribution or the final results (c.f. appendix O4).

The administrative records identify the exact number of taxable workers in 2021, totaling 2,702,080 individual entries. To compute the share of taxable agents I combine this with the Kenyan National Bureau of Statistics estimate of the total labor force in 2021 of 19,106,156 workers (KNBS, 2021b). This implies that taxable agents comprise $\bar{\rho} = 14.1$ *pct.* of the labor force.

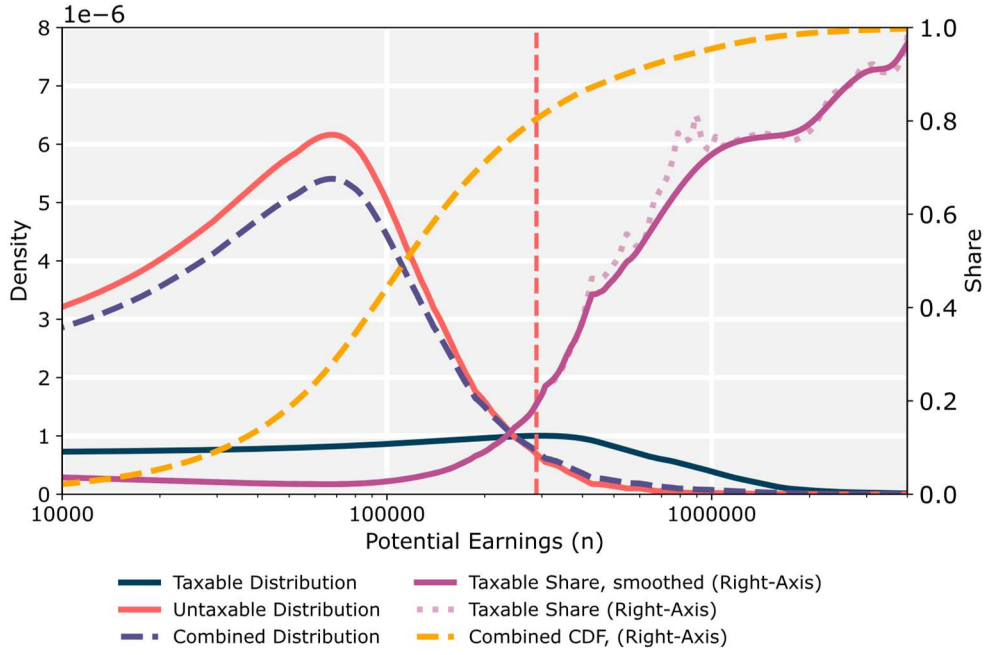
To map observed earnings into the skill parameter (potential earnings), I employ the standard Saez (2001) calibration method (see e.g., Jacquet et al., 2013; Lockwood, 2020). This requires an estimate of the elasticity of taxable income. In a recent study, Kanina et al. (2025) exploit tax reforms from 2020–21 to estimate an ETI of $e = 0.51$ for private-sector workers with incomes in the range 375,000 – 975,000 Ksh (2022-prices). Including public-sector workers leads to a somewhat lower elasticity of 0.36, potentially reflecting wage-setting constraints. I use the private-sector elasticity of 0.51 as the baseline²¹.

Inverting the earnings condition $z = n(1 - T')^e$, the implied skill parameter is given by $n_i = \frac{z_i}{(1 - T'(z_i))^e}$ for taxable workers and $n_i = z_i$ for untaxable workers where z_i is observed earnings and $T'(z_i)$ is the marginal tax rate²².

²¹ Under quasi-linearity and the non-negativity constraint, the exemption threshold solves for $\tilde{q}_1^+(z) = 1$, so for given weights and income distributions, ETI only affects rates above z , not the crossing point to first order.

²² Observations at the 288,000 Ksh kink, where T' is not single-valued, are assigned the left-limit (zero) marginal rate; the recovered skill density is then smoothed (Figure 3), limiting local distortion

Figure 3: Skill Distributions, Kenya 2021



Note: Observed incomes are mapped into skill n using the Saez (2001) transformation with marginal tax rates $T'(z)$. The log x-axis highlights the lower tail. The vertical line marks the potential earnings equivalent to the 2021 exemption threshold ($n = 288,000$). The figure shows that informality dominates at low skill levels, while taxable workers become more prevalent only above $n \approx 540,000$, with substantial overlap across the distribution. Distributions are smoothed with Epanechnikov kernels (bandwidths 200,000 and 75,000).
Source: KCHS 2021, KRA, and own calculations

Figure 3 shows a clear shift between the two resulting skill distributions: untaxable workers are concentrated at low skills, while taxable workers are concentrated at higher skill levels and only become more prevalent above $n \approx 540,000$. At the same time, there is meaningful overlap in the middle of the distribution, implying that workers with similar skills can fall on either side of the tax boundary. As such, taxability (formality) is not exclusively a high-skill trait. This overlap is what generates horizontal inequality in the current system.

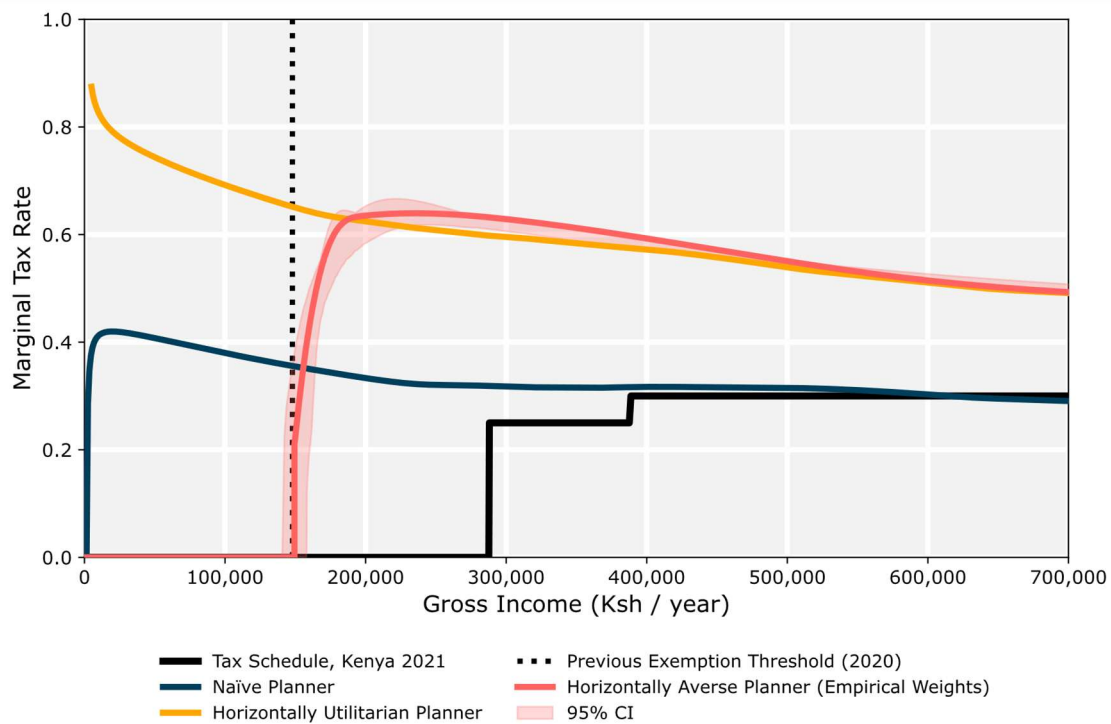
5.2 Simulation Results: Optimal Taxes in Kenya

In this section, I present optimal tax simulations for Kenya accounting for the presence of untaxable workers and horizontal inequality aversion. I maintain the assumptions of no extensive margin responses and no administrative costs. I relax these in section 6. I compare prescriptions for three cases: (i) a naïve planner who incorrectly ignores the existence of informal workers ($\bar{\rho} = 1$), (ii) a horizontally utilitarian planner who recognizes informality but assigns no intrinsic value to horizontal equality ($\phi = 1$), and (iii) a planner who correctly accounts for the informal sector *and* incorporates the empirically estimated horizontal inequality aversion. All specifications use the empirically estimated aversion to vertical inequality, $\tilde{\psi}$, but only (iii) employs the empirical horizontal inequality parameter, $\tilde{\phi}$.

To compute the optimal tax schedules, I proceed in two steps. First, the empirical MSWWs obtained in Section 4 are used to recover the structural parameters of the social welfare function, Θ^* , by matching the SWF-implied MSWWs to the estimated schedules (see Appendix A11). This delivers a welfare function that reproduces the observed MSWWs under the current tax system.

Second, given these fitted parameters, I solve for the ex-post optimal nonlinear tax schedule. Starting from the statutory 2021 schedule, I iteratively update the candidate tax function, recompute the implied income distributions and MSWWs, and repeat until the schedule, distributions, and weights jointly converge. The entire computation is repeated for each of the 1,000 bootstrap resamples of the MSWWs, generating pointwise confidence bands around the optimal schedule.

Figure 4: Optimal Tax Schedules with Informality and Horizontal Inequality



Note: All series plot ex-post optimal marginal tax rates using the empirically estimated vertical preference parameter $\tilde{\psi}$. "Naïve Planner" ignores informality and corresponds to the standard Mirrleesian benchmark. "Horizontal Utilitarian" incorporates informality but imposes, $\phi = 1$, implying no horizontal inequality aversion. "Empirical Weights, Kenya" accounts for the presence of an informal sector and for the observed aversion to horizontal inequality embodied in $\tilde{\phi}$. Pointwise 95% confidence bands are based on 1000 bootstrap resamplings of the survey-elicited MSWWs. All bootstrap series are lightly spline-smoothed over the ability distribution before computing percentiles.

Figure 4 presents the resulting optimal tax prescriptions. Neither the naïve nor the horizontally utilitarian prescriptions feature an exemption. On the contrary, in both cases, optimal marginal tax rates are high across the bottom of the distribution. Interestingly, moving from the naïve to the horizontal utilitarian case, marginal tax rates *increase*, reflecting the fact that formal workers in Kenya earn substantially higher incomes than informal workers and therefore receive lower welfare

weights in the absence of horizontal inequality aversion. Informality thus acts as a tag allowing the planner to concentrate taxes on relatively richer, formal workers.

Incorporating horizontal inequality aversion through the full empirical MSWWs fundamentally alters this prescription. These weights yield an optimal tax system with a sizable exemption threshold of approximately 149,000 Ksh. While this lies below the statutory 2021 threshold of 288,000 Ksh, it is close to the 2020 threshold of 147,580 Ksh. Although these magnitudes should be interpreted as suggestive rather than exact, the qualitative implication is clear: accounting for horizontal inequality aversion implies that the presence of a sizeable exemption in Kenya is likely welfare-*improving*.

The mechanism for this is straightforward. The empirical MSWWs imply that without an exemption, average weights for taxable workers at the bottom would exceed one, $\tilde{q}_1^*(0) > 1$, indicating that the planner aims to distribute towards taxable workers rather than away from them. A positive marginal tax on low-income formal workers would therefore be welfare-reducing. Absent a non-negativity constraint, the planner would set negative marginal tax rates at the bottom to equalize average welfare weights across taxable and untaxable workers²³. However, the non-negativity constraint forces the planner to implement a zero marginal rate over a range of incomes, generating the exemption. Horizontal inequality aversion thus provides a sufficiently strong incentive to equalize utility across formal and informal workers of the same skill, even at the cost of foregone revenue.

Taken together, these results isolate horizontal inequality aversion as the decisive force behind the exemption. Holding vertical preferences fixed and maintaining the non-negativity constraint, neither the naïve nor the horizontal utilitarian planner prescribes an exemption. Absent horizontal inequality aversion, optimal marginal rates remain positive across the bottom regardless of vertical inequality aversion. The exemption emerges only once empirically estimated horizontal inequality aversion is accounted for. Horizontal inequality aversion thus breaks the standard Sadka–Seade (1976; 1982) no-exemption result, and empirically relevant aversion is sufficient to rationalize sizable exemptions in high-informality settings.

6 Optimal Taxes with Extensive Margin and Administrative Cost

The previous section illustrates that horizontal inequality aversion provides a coherent normative explanation for the existence of sizable tax exemptions. However, two widely discussed alternative

²³ To understand this, consider that when $z \rightarrow 0$ 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.

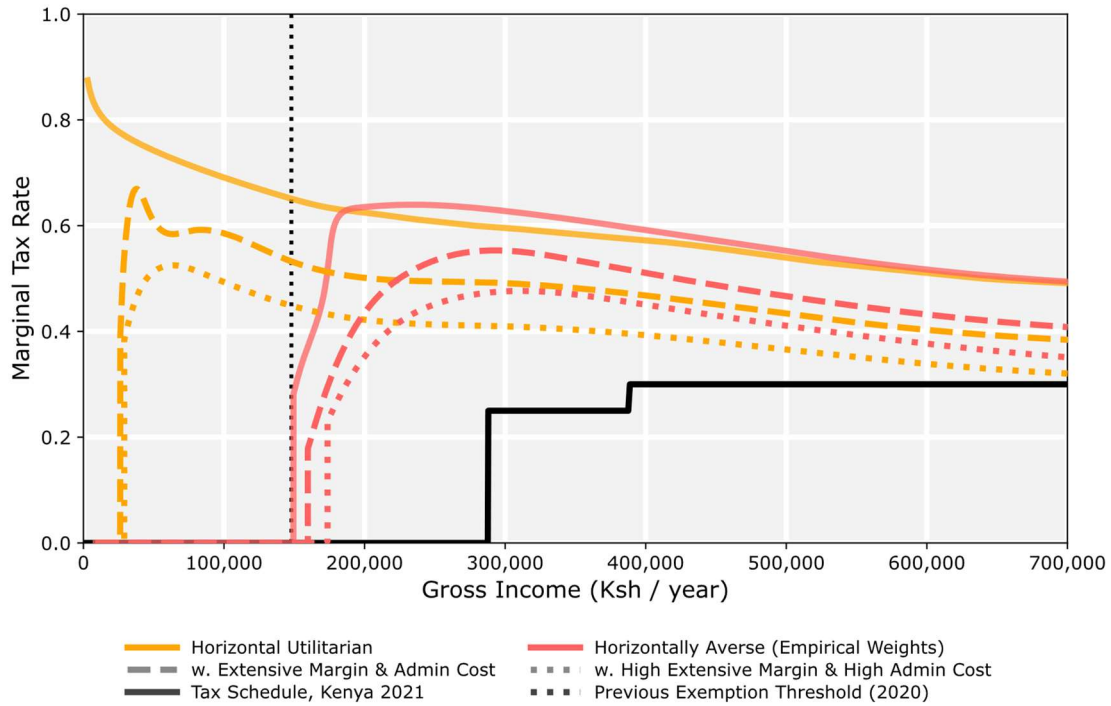
mechanisms also warrant examination. First, the fixed-type assumption may be too strong. If higher taxes induce shifts to informality, extensive-margin responses could, in principle, justify an exemption. Second, high marginal tax administration costs could motivate the imposition of an exemption to maintain collection efficiency. I test the empirical relevance of these mechanisms using the general optimal tax characterization from proposition 3.

Evidence on the extensive margin elasticity in developing contexts is very limited. One study by McKay et al. (2024), looking at four East African countries, finds statistically insignificant elasticities close to zero, while studies of middle-income countries find statistically significant but modest elasticities (Haanwinckel & Soares, 2021; Ulyssea, 2020). Despite the near-zero East African estimates, I adopt a conservatively large elasticity of $\tilde{\eta}_p = 0.5$ in the simulation, at the upper end of the middle-income range. For further robustness I also plot an elasticity of 1. In online appendix O9 I also consider a variable elasticity declining from 1 to 0.5 with the CDF of formal workers.

Administrative costs could in principle justify excluding low-earning taxpayers from the tax base (Moore, 2023). To provide a simple upper-bound estimate of the effect of administrative costs on optimal taxation, I assume a constant marginal administrative cost of $\gamma = 5,000 \text{ Ksh}$ ($\approx 38 \text{ USD}$) per taxpayer per year. I argue that this is sufficiently large to be conservative seeing as Kenya operates a digital Pay-As-You-Earn system that requires limited individualized processing²⁴. For robustness, I also consider a scenario with $\gamma = 10,000 \text{ Ksh}$. Given that informal workers in Kenya currently face essentially no enforcement, there is little reason to attribute enforcement costs to them regardless of whether they fall within the tax base. Nonetheless, for robustness, I also consider a global enforcement scenario in which an informal worker in the base incurs a cost of $\gamma_I = 1,500 \text{ Ksh}$, capturing the authority's fruitless cost of attempting to bring eligible informal earners into the net. I report this in online appendix O9.

²⁴ There is no data available to directly estimate the *marginal* cost of collection. We can, however compare this to estimates of average costs of collection across all taxbases relative to total revenues which is 2.2% for Kenya (KRA, 2023) which is similar to estimates for Rwanda (Rwanda Revenue Authority, 2024). Because this average embeds fixed system costs already sunk in the digital infrastructure, it overstates the true marginal cost of an additional taxpayer. Even so, imposing this average cost as a constant marginal cost would imply an administrative cost of including one additional worker earning the median income of a formal worker of $\approx 1000 \text{ Ksh}$, far below the assumed level. This of course does not account for compliance cost on the worker side. Assuming that filing taxes takes two hours and using median hourly wage ($\approx 225 \text{ Ksh}$) as the outside option would increase the total marginal cost by $\approx 450 \text{ Ksh}$ still far below the estimate.

Figure 5: Optimal Taxation with Extensive Margin and Administrative Costs



Note: Schedules denoted "w. Extensive Margin" feature a participation elasticity of $\tilde{\eta} = 0.5$ while schedules with "High Extensive Margin" have $\tilde{\eta}_p = 1$. Schedules with "Adm. Cost" and "High Adm. Cost" have marginal inclusion cost of an additional worker in the taxbase of $\gamma = 5,000$ Ksh and $\gamma = 10,000$ Ksh respectively. All series plot ex-post optimal marginal tax rates using the empirically estimated vertical preference parameter $\tilde{\psi}$. "Horizontal Utilitarian" incorporates informality but imposes, $\phi = 1$, implying no horizontal inequality aversion. "Empirical Weights, Kenya" accounts for the presence of an informal sector and for the observed aversion to horizontal inequality embodied in $\tilde{\phi}$.

Figure 5 shows that under horizontal utilitarianism, only very small exemptions arise, even in the presence of large administrative costs. The reason sizeable administrative costs cannot rationalize an exemption is structural. With a continuous nonlinear schedule and no notch, excluding one worker means removing the tax liability on the first unit of taxable income for every worker above that threshold. This parallels Yitzhaki's (1979) commodity-tax result: under a uniform rate, exempting one good distorts the relative prices of all others, so it is generally not optimal to do so locally. The planner therefore cannot exempt a single type without disturbing the entire tax base, and this holds even for very high marginal costs. The exemption in the main specification cannot be rationalized by administrative costs or extensive-margin responses alone.

7 Conclusion

Developing economies 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 prescribe no exemptions at all (Diamond & Saez, 2011; Mankiw et al., 2009). A common interpretation of this is that policymakers in developing countries simply

implement suboptimal policies. This paper highlights a different explanation. 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. Beyond efficiency concerns, taxing formal workers while leaving informal workers untaxed introduces significant horizontal inequalities altering the equity–efficiency trade-off in ways that can rationalize sizable exemption thresholds.

This paper extends the standard optimal tax model to a two-type framework allowing for the presence of untaxable workers. In this environment, taxing formal workers while informal workers remain outside the tax base generates horizontal inequalities that a horizontally averse planner may seek to offset. I show that once these concerns are incorporated, sizable exemption thresholds are optimal.

This paper does not 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 coherent normative structure that incorporates horizontal concerns, the approach allows us to more accurately reflect the constraints and objectives inherent in high-informality economies.

Appendix

8 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.

$$\begin{aligned} H = & (Q(v_1(n), v_2(n), \rho(n)) \\ & + \lambda[\rho(n)(n \cdot l_1(n) - [v_1(n) + b(l_1(n))]) \\ & + (1 - \rho(n))(n \cdot l_2(n) - [v_2(n) + b(l_2(n))])])h(n) \\ & + \theta_1(n) \left[\frac{l_1(n)b'(l_1(n))}{n} \right] + \theta_2(n) \left[\frac{l_2(n)b'(l_2(n))}{n} \right] \end{aligned}$$

Notation follows section 2 and l_j denotes labor supply.

First-Order Conditions:

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

1. $\frac{\partial H}{\partial l_1(n)} = 0$:

$$\lambda[\rho(n)(n - b'(l_1(n)))]h(n) + \theta_1(n) \left(\frac{b'(l_1(n)) + l_1(n)b''(l_1(n))}{n} \right) = 0 \quad (A1)$$

2. $\frac{\partial H}{\partial v_1(n)} = 0$:

$$\left[\frac{\partial Q}{\partial v_1(n)} - \lambda\rho(n) \right] h(n) + \frac{d\theta_1(n)}{dn} = 0 \quad (A2)$$

3. $\frac{\partial H}{\partial v_2(n)} = 0$:

$$\left[\frac{\partial Q}{\partial v_2(n)} - \lambda(1 - \rho(n)) \right] h(n) + \frac{d\theta_2(n)}{dn} = 0 \quad (A3)$$

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:

$$\lim_{n \rightarrow \infty} \theta_k(n) = 0 \quad \text{and} \quad \theta_k(0) = 0 \quad \text{for } k \in \{1, 2\}$$

Finding $\theta_1(n)$ using (A2):

$$\theta_1(n) = \int_n^\infty \left[\frac{\partial Q}{\partial v_1(m)} - \lambda \rho(m) \right] h(m) dm \quad (A4)$$

And symmetrically for $\theta_2(n)$. We can add the costates and use that $\theta_1(0) + \theta_2(0) = 0$:

$$\theta_1(n) + \theta_2(n) = \int_n^\infty \left[\frac{\partial Q}{\partial v_1(m)} + \frac{\partial Q}{\partial v_2(m)} - \lambda \right] h(m) dm \quad (A5)$$

$$\lambda = \int_0^\infty \left[\frac{\partial Q}{\partial v_1(m)} + \frac{\partial Q}{\partial v_2(m)} \right] h(m) dm \quad (A6)$$

Deriving the Optimal Tax Rate:

For type I workers the utility function is:

$$v_1(n) = c(n) - b(l_1(n)) \quad (A7)$$

Subject to the budget constraint:

$$c(n) = z(n) - T(z(n)) \quad (A8)$$

Where $z(n) = n \cdot l_1(n)$ is gross income and $T(z(n))$ is the tax function.

We review the FOC for utility maximization $\frac{\partial v_1(n)}{\partial l_1(n)} = 0$ which yields the FOC

$$b'(l_1(n)) = (1 - T'(z(n))) n \quad (A9)$$

Elasticity of Taxable Income:

I define the elasticity of taxable income e :

$$e = \frac{(1 - T'(z(n))) n}{l_1(n) b''(l_1(n))} \quad (A10)$$

And rewrite:

$$\frac{l_1(n) b''(l_1(n))}{n} = \frac{1 - T'(z(n))}{e} \quad (A11)$$

Inserting (A9) into the FOC (A1) and using this definition gives

$$\lambda[\rho(n)nT'(z(n))]h(n) = -\theta_1(n) \left(1 - T'(z(n))\right) \left(1 + \frac{1}{e}\right) \quad (A12)$$

Defining Social Marginal Welfare Weights:

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

$$q_1(n) = \frac{\frac{\partial Q}{\partial v_1(n)}}{\lambda\rho(n)} = \frac{\frac{\partial Q}{\partial c_1(n)}}{\lambda\rho(n)} \quad (A13)$$

Which allows us to rewrite:

$$\frac{\partial Q}{\partial v_1(n)} = q_1(n) \cdot \lambda \cdot \rho(n) \quad (A14)$$

Using equations (A4) and (A14)

$$-\theta_1(n) = \lambda \int_n^\infty \rho(m)[1 - q_1(m)]h(m)dm \quad (A15)$$

Using that $f_1(n)\bar{\rho} = \rho(n)h(n)$ and that $\rho^+(n) = \int_n^\infty \rho(m)h(m)dm$

$$-\theta_1(n) = \lambda\rho^+(0) \int_n^\infty [1 - q_1(m)]f_1(m)dm \quad (A16)$$

Insert into eq. (A12)

$$[\rho(n)nT'(z(n))]h(n) = \rho^+(0) \left[\int_n^\infty [1 - q_1(m)]f_1(m)dm \right] \left(1 - T'(z(n))\right) \left(1 + \frac{1}{e}\right) \quad (A17)$$

Solving for Marginal Tax Rates:

Rewrite eq. (A17)

$$\frac{T'(z(n))}{1 - T'(z(n))} = \frac{(1 - F_1(n))(1 - q_1^+(n))}{nf_1(n)} \left(1 + \frac{1}{e}\right) \quad (A18)$$

Where I use the definition: $q_1^+(n) = \frac{\int_n^\infty q_1(m)f_1(m)dm}{1 - F_1(n)}$

To express this over observable income, apply the standard Saez (2001) change of variables.

Under the Spence–Mirrlees condition the skill–income map $n \rightarrow z$ is invertible, and for a locally linear schedule the virtual density satisfies

$$g_1(z)z_1(n)(1 + e) = n \cdot f_1(n) \quad (A19)$$

Substituting (A19) into (A18), and using $G_1(z_1(n)) \equiv F_1(n)$ together with the local Pareto parameter $\alpha_1(z) \equiv \frac{zg_1(z)}{1-G_1(z)}$, yields the optimal marginal tax:

$$\begin{aligned} \frac{T'(z(n))}{1-T'(z(n))} &= \frac{(1-G_1(z))(1-\tilde{q}_1^+(z))}{g_1(z)z_1(1+e)} \left(\frac{1+e}{e} \right) \\ &= \frac{1}{\alpha_1(z)} \left(\frac{1}{e} \right) (1-\tilde{q}_1^+(z)) \end{aligned}$$

We can isolate to find the final formula:

$$T'(z) = \frac{1-\tilde{q}_1^+(z)}{1-\tilde{q}_1^+(z)+e \cdot \alpha_1(z)} \quad (\text{A20})$$

9 Kuhn-Tucker Conditions for Non-negative Marginal Taxes

Imposing $T'(z) \geq 0$ on the derivation of Section 8 establishes Proposition 1. The steps are unchanged except for the added constraint. I outline the changes here.

Augment the Hamiltonian of Section 8 with a multiplier $\omega(n)$ on the non-negativity constraint, written $l_2 \geq l_1(n)$, since type-2 workers face a zero effective rate, it binds only under a negative marginal rate:

$$\begin{aligned} \tilde{H} &= H + \omega(n)(l_2(n) - l_1(n)) \\ \text{st. } \omega(n) &\geq 0 \quad \text{and} \quad \omega(n)(l_2(n) - l_1(n)) = 0 \end{aligned}$$

The multiplier enters only the l_1 first-order condition (A1), adding a term $-\omega(n)$. Carrying it through the same steps as (A1) to (A18) gives

$$\frac{T'(z(n))}{1-T'(z(n))} = \frac{(1-F_1(n))(1-\tilde{q}_1^+(z)) \left(1 + \frac{1}{e}\right) + \frac{\omega(n)}{\lambda \bar{\rho}}}{nf_1(n) - \frac{\omega(n)}{\lambda \bar{\rho}}} \quad (\text{A21})$$

When $\tilde{q}_1^+(n) \leq 1$ the constraint slacks ($\omega = 0$) and (A21) collapses to the unconstrained formula. When $\tilde{q}_1^+(n) > 1$ the unconstrained rate would be negative, so $\omega(n) > 0$ rises until $l_2 = l_1(n)$. The non-negativity constraint therefore acts as a simple clamp,

$$T'_c(z) = \max[0, T'_u(z)]$$

producing a zero-rate region as outlined in Proposition 1.

10 Sketch: Proof of General Extensive Margin Model

Full derivation using the tax-perturbation method of Jacquet and Lehmann (2021) in online Appendix O2. Below follows a brief sketch- The general model adds three objects to the baseline.

Roy sectoral choice. Formal earnings and the sector cutoff,

$$z_1(n; T) = \operatorname{argmax}_z \{z - T(z) - b(\frac{z}{n})\} \quad (\text{A22})$$

give the formal share $\rho(n; T) = F_{(\theta|n)}(\bar{\theta}(n; T) | n)$, with informal productivity wedge $\kappa(n)$ and admin cost γ per formal taxpayer above $\bar{z}$.

Responses. Intensive and extensive margins summarized by two sufficient statistics,

$$z_R(n) = \frac{e(n)z_1(n)}{1 - T'(z_1)} R'(z_1(n)) \quad \text{and} \quad \rho_R(n)/\rho(n) = \tilde{\eta}_\rho(z) R(z) \quad (\text{A23})$$

Optimality. A perturbation $T_x = T - xR$ with zero first-order welfare effect gives

$$0 = \bar{\rho} \int \left\{ [\tilde{q}_1(z) - 1 + \Gamma(z; \bar{z})] R(z) + \frac{T'(z)}{1 - T'(z)} \tilde{e}(z) z R'(z) \right\} g(z) dz \quad (\text{A24})$$

Integrating by parts delivers the interior rate (7), with Roy correction $\Gamma(z; \bar{z}) \equiv [T(z) - \gamma \mathbf{1}_{\{z > \bar{z}\}}] \tilde{\eta}_\rho(z)$ and optimizing over $\bar{z}$ delivers the boundary condition (8). Since κ enters (7) only through $\{\alpha, \tilde{e}, \tilde{q}_1^+, I^*\}$, it has no direct effect on rates (Proposition 4). Imposing $T' \geq 0$ clamps the schedule to $T_c = \max\{0, T_u\}$.

11 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:

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

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 ordinal votes from a representative sample of voters.

11.1.1 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., $q_i(z_a, j_a) = \beta \cdot q_i(z_b, j_b)$ where $q_i(\cdot)$ represents social weight.
- **Single-Valuation Votes:** Each vote can only capture one comparison (e.g., $r_{a \rightarrow b} = \frac{q_i(z_a, j_a)}{q_i(z_b, j_b)}$).

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

11.1.2 Normalization and Identification

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 a reference cell. We set: $q^*(z_0, j_0) = 1$.

All other values are identified relative to this cell.

11.1.3 Estimation Strategy:

All votes happen between two grid cells, e.g., $a = (z_k, j_t)$ and $b = (z_l, j_s)$. There are N total cells. Let $\Gamma \subseteq N \times N$ be all the observed cell comparisons, i.e., $(a \rightarrow b)$ where $a \neq b$. We transform votes, $r_{a \rightarrow b}$, into log-ratios to simplify analysis while keeping median targets well-behaved. Define $\ell_{a \rightarrow b} = \log(r_{a \rightarrow b})$, this way products of ratios add $\ell_{a \rightarrow b} = \ell_{a \rightarrow c} + \ell_{c \rightarrow b}$ this allows for the estimation of one set of differences rather than chaining products, eliminating path dependence. As the log function is strictly increasing, taking medians on this scale preserves order. If ballots follow $r = r^* e^\varepsilon$ with $\text{median}(\varepsilon) = 0$, then $\text{median}(\log(r)) = \log(r^*)$ and the log-median is an unbiased target for the population median on a multiplicative scale.

For each observed pair, we define the corresponding log-median ratio as: $m_{a \rightarrow b} \equiv \text{median}\{\log(r_{a \rightarrow b})\} = \text{median}\{\ell_{a \rightarrow b}\}$ where $r_{i, b \rightarrow a} = \frac{1}{r_{i, a \rightarrow b}} \Rightarrow m_{b \rightarrow a} = -m_{a \rightarrow b}$. We allow for the inclusion of a weight, $w_{a \rightarrow b}$, to signify precision, inverse sampling probability of importance. We attempt to estimate our log-transformed true median as $v_{k,t}^* \equiv \log(q^*(z_k, j_t))$ by solving:

$$\hat{v} = \underset{v}{\operatorname{argmin}} \left[\sum_{(a \rightarrow b) \in \Gamma} w_{a \rightarrow b} [(v_a - v_b) - m_{a \rightarrow b}]^2 \right] \quad (A25)$$

Where $v_{n_0} = 0$ fixes scale at a reference cell n_0 . Identification requires that the comparison graph (N, Γ) is connected.

To recover MSWWs in levels rather than logs we simply transform $\hat{q}^*(z_k, j_t) = \exp(\hat{v}_{k,t})$. Estimating $\hat{v}$ simply requires solving for the set of $\hat{v}$ that minimizes the loss function presented above.

Separability:

If preferences for horizontal and vertical equality are separable such that we can write $q^*(z, j) = q_Z^*(z)q_J^*(j)$ and if I contains no cross-comparisons (i.e., where both z and j differ in a ballot), then the problem simplifies. Use vertical ballots pooled across types to estimate $q_Z^*(z)$ via a 1-D weighted least squares on the income grid (anchor $q_Z^*(z_{k_0}) = 1$). For horizontal ballots, pool across incomes to estimate $q_J^*(j)$ analogously (anchor $q_J^*(j_{t_0}) = 1$). Combine as $q^*(z, j) = q_Z^*(z) q_J^*(j)$.

Fitting Social Welfare Function:

As subsequent simulations will need to use endogenous weights, we use the estimates of $\hat{q}^*(z_k, j_t)$ to parameterize the generalized social welfare function:

$$SWF = \int_0^\infty \left(\sum_{j=1}^2 \rho_j(n) \left(\frac{v_j(n)^{1-\psi}}{1-\psi} \right)^{\phi(n)} \right)^{\frac{1}{\phi(n)}} h(n) dn \quad (A26)$$

Where the horizontal inequality aversion is allowed to vary across the skill distribution according to $\phi(n) = \alpha_0 + \alpha_1 \ln(n) + \alpha_2 \ln(n)^2$. The parameter vector to estimate is $\Theta = \{\alpha_0, \alpha_1, \alpha_2, \eta\}$.

Where $\psi < 1$ and $v_j > 0$. As utility is quasi-linear the MSWW implied becomes:

$$\tilde{q}(z, j) = \frac{1}{\lambda} \frac{\partial SWF(\Theta)}{\partial \tilde{v}_j(z(n))} \quad (A27)$$

Where λ is given by the normalization condition. Before fitting, the elicited weights are rescaled onto the same population-wide normalization as the structural weights. The parameter vector is chosen to minimize the loss function:

$$\begin{aligned} L_H(z_i) &= \left((\hat{q}^*(z_i, 1) - \hat{q}^*(z_i, 2)) - (\tilde{q}(z_i, 1) - \tilde{q}(z_i, 2)) \right)^2 g(z) \frac{1}{1 - G(\bar{z})} \\ L_V(z_i) &= (\hat{q}^*(z_i, 1) - \tilde{q}(z_i, 1))^2 g(z) \\ L &= \int_0^\infty \mathbf{1}_{\{z > \bar{z}\}} L_H(z) + L_V(z) dz \end{aligned}$$

The resulting parameter vector, Θ^* , thus recreates the current MSWWs given the current allocation but is determined endogenously and thus adapts to policy changes.

Bootstrapping:

To identify confidence bounds on the MSWWs and, by extension, on the optimal tax prescriptions, the paper employs a bootstrap that resamples the survey at the subcounty–participant level. Each bootstrap draw proceeds in three steps.

First, subcounties are resampled with replacement, preserving the number of primary sampling units.

Second, within each selected subcounty, participants are resampled with replacement. If a participant appears k times in the draw, all of that participant’s observations enter with multiplicity k . This produces a new resampled set of comparisons S^b that preserves the original A–B structure.

Third, for this new set of comparisons, weights are recomputed as

$$w_i^b = base_{h(i)} \cdot mult_i^b \cdot raking_i^b$$

Here, $base_{h(i)}$ is the design weight derived from subcounty population counts and sample sizes, $mult_i^b$ reflects the participant’s resampling multiplicity, and $raking_i^b$ is the post-stratification adjustment that forces the weighted distribution of income, gender, and age in the bootstrap sample to match the population margins (Online Appendix O3.1.2). Because the composition of each bootstrap sample differs, the raking step is rerun for every bootstrap draw.

The full MSWW estimation pipeline is then rerun using these weights, generating bootstrap distributions for q_1, q_2 . These are then subsequently used to estimate parameter vectors, $\tilde{\Theta}_i^*$, for the social welfare functions. Finally, the full simulation procedure is run to generate 1000 optimal tax prescriptions, $T_i(z|\tilde{\Theta}_i^*)$. The bootstrapped tax schedules are aligned on a common income support and used to construct pointwise 95% confidence bands for the optimal marginal tax rate profile.

References

- 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.
- Alesina, A., Stantcheva, S., & Teso, E. (2018). Intergenerational Mobility and Preferences for Redistribution. *American Economic Review*, 108(2), 521–554. <https://doi.org/10.1257/aer.20162015>
- 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>
- 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>
- Bachas, P., Gadenne, L., & Jensen, A. (2024). Informality, Consumption Taxes, and Redistribution. *Review of Economic Studies*, 91(5), 2604–2634. <https://doi.org/10.1093/restud/rdad095>

- Bachas, P., Hoy, C., Jensen, A., & Shaukat, M. (2026). *Horizontal Equity of Taxation: Citizen Beliefs and Policy Preferences*. Washington, DC: World Bank. <https://doi.org/10.1596/44754>
- Bachas, P., Jensen, A., & Gadenne, L. (2024). Tax Equity in Low- and Middle-Income Countries. *Journal of Economic Perspectives*, 38(1), 55–80. <https://doi.org/10.1257/jep.38.1.55>
- 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. L. (2024). The State Capacity Ceiling on Tax Rates: Evidence From Randomized Tax Abatements in the DRC. *Econometrica*, 92(4), 1163–1193. <https://doi.org/10.3982/ECTA19959>
- 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. (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>
- Best, M., Caloi, L., Gerard, F., Kresch, E. P., Naritomi, J., & Zoratto, L. (2025). *Greener on the Other Side: Inequity and Tax Compliance* (No. W34062; p. w34062). National Bureau of Economic Research. <https://doi.org/10.3386/w34062>
- Bhorat, H., Kanbur, R., & Stanwix, B. (2017). Minimum Wages in Sub-Saharan Africa: A Primer. *The World Bank Research Observer*, lkw007. <https://doi.org/10.1093/wbro/lkw007>
- 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>
- Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9, 88–97.
- 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>
- 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.
- 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*.
- Feinmann, J., Hsu Rocha, R., & Lauletta, M. (2022). *Payments Under the Table: Employer-Employee Collusion in Brazil* (SSRN Scholarly Paper No. 4270665). Social Science Research Network. <https://doi.org/10.2139/ssrn.4270665>

- Gollin, D. (2014). The Lewis model: A 60-year retrospective. *Journal of Economic Perspectives*, 28(3), 71–88.
- 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.
- Hoy, C., Scot, T., Oguso, A., Custers, A., Zalo, D., Doino, R., Karver, J., & Pillai, N. O. (2024). *Trade-Offs in the Design of Simplified Tax Regimes: Evidence from Sub-Saharan Africa* (SSRN Scholarly Paper No. 4989818). Social Science Research Network. <https://doi.org/10.2139/ssrn.4989818>
- Hvidberg, K. B., Kreiner, C. T., & Stantcheva, S. (2023). Social Positions and Fairness Views on Inequality. *Review of Economic Studies*, 90(6), 3083–3118. <https://doi.org/10.1093/restud/rdad019>
- 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>
- Jacquet, L., & Lehmann, E. (2021). Optimal Income Taxation with Composition Effects. *Journal of the European Economic Association*, 19(2), 1299–1341. <https://doi.org/10.1093/jeea/jvaa022>
- Jacquet, L., Lehmann, E., & Van der Linden, B. (2013). Optimal redistributive taxation with both extensive and intensive responses. *Journal of Economic Theory*, 148(5), 1770–1805. <https://doi.org/10.1016/j.jet.2013.07.019>
- 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.aercafricallibrary.org/handle/123456789/3969>
- Kaplow, L. (2000). *Horizontal equity: New measures, unclear principles*.
- Kaplow, L., & Shavell, S. (2000). Fairness versus welfare. *Harv. L. Rev.*, 114, 961.
- 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., Knudsen, M. B., Kreiner, C. T., Pedersen, S., & Saez, E. (2011). Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark. *Econometrica*, 79(3), 651–692. <https://doi.org/10.3982/ECTA9113>
- Kleven, H. J., Kreiner, C. T., & Saez, E. (2009). The optimal income taxation of couples. *Econometrica*, 77(2), 537–560.
- Kleven, H. J., Kreiner, C. T., & Saez, E. (2016). Why can modern governments tax so much? An agency model of firms as fiscal intermediaries. *Economica*, 83(330), 219–246.
- KNBS. (2019). *2019 Kenya population and housing census*. Kenya National Bureau of Statistics.
- 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/>
- KRA. (2023). *Annual Revenue Collection and Costs [Dataset]*.

- Kuziemko, I., Norton, M. I., Saez, E., & Stantcheva, S. (2015). How Elastic Are Preferences for Redistribution? Evidence from Randomized Survey Experiments. *American Economic Review*, 105(4), 1478–1508. <https://doi.org/10.1257/aer.20130360>
- 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>
- Lockwood, B. B. (2020). Optimal Income Taxation with Present Bias. *American Economic Journal: Economic Policy*, 12(4), 298–327. <https://doi.org/10.1257/pol.20180762>
- 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>
- Meghir, C., Narita, R., & Robin, J.-M. (2015). Wages and Informality in Developing Countries. *American Economic Review*, 105(4), 1509–1546. <https://doi.org/10.1257/aer.20121110>
- 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>
- Perry, G. E., Arias, O., Fajnzylber, P., Maloney, W. F., Mason, A., & Saavedra-Chanduvi, J. (2007). *Informality: Exit and Exclusion*. The World Bank. <https://doi.org/10.1596/978-0-8213-7092-6>
- PwC Kenya. (2025). *Kenya—Individual—Other taxes*. <https://taxsummaries.pwc.com/kenya/individual/other-taxes>
- 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>
- Rwanda Revenue Authority. (2024). *RRR Annual Report 2023/24*.
- Sadka, E. (1976). On Income Distribution, Incentive Effects and Optimal Income Taxation. *The Review of Economic Studies*, 43(2), 261–267. <https://doi.org/10.2307/2297322>
- Saez, E. (2001). 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.
- Seade, J. (1982). On the Sign of the Optimum Marginal Income Tax. *The Review of Economic Studies*, 49(4), 637–643. <https://doi.org/10.2307/2297292>
- 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., & 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](https://doi.org/10.1016/S1573-4420(02)80026-X)

- Stiglitz, J. E. (1982). Utilitarianism and horizontal equity: The case for random taxation. *Journal of Public Economics*.
- Tanzi, V., & Zee, H. (2001). Tax Policy for Developing Countries. *IMF Economic Issues*, 27.
- Ulyssea, G. (2018). Firms, informality, and development: Theory and evidence from Brazil. *American Economic Review*, 108(8), 2015–2047.
- 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>
- Weinzierl, M. (2014). The promise of positive optimal taxation: Normative diversity and a role for equal sacrifice. *Journal of Public Economics*, 118, 128–142. <https://doi.org/10.1016/j.jpubeco.2014.06.012>
- 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.
- Yitzhaki, S. (1979). A Note on Optimal Taxation and Administrative Costs. *The American Economic Review*, 69(3), 475–480.
- 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>

Online Appendix

1 Optimal Tax Rates with Employed and Self-employed Workers

In this section I make a simple perturbation derivation of optimal taxes using a delineation between "Employed" and "Self-employed" workers. I maintain the same notation where worker type follows from the subscript such that x_1 and x_2 refer to employed and self-employed respectively.

Employees always pay their full tax liability, whilst the self-employed can costlessly evade a share $1 - \sigma$ of their liability such that they only have to pay $\sigma T(z)$. Their utility function is thus:

$$u_2(c, z; n) = z - \sigma T(z) - b(z, n)$$

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:

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

Where $H(z)$ is the CDF of earnings and $\rho^+(z)$ the share of workers above income level z who are "Employed".

Notice, we can define $1 - \chi = (\rho^+(z) + \sigma(1 - \rho^+(z)))$ as the net of the marginal evasion rate.

$$dM = dT'(z) \cdot [1 - H(z)] \cdot (1 - \chi)$$

We can also simply decompose it into the effect from employees dM_1 and self-employed dM_2

$$dM = dM_1 + dM_2$$

Welfare effect of tax change depends simply on MSWW for each type:

$$dW = -(q_1^+(z) \cdot dM_1 + q_2^+(z) \cdot dM_2)$$

The behavioral effect depends on the elasticities of each type, e_j , the local density of workers, $h(z)$, and the share paid by self-employed, σ . The elasticity of employed workers, e_1 , is both their structural and reduced form elasticity while e_2 denotes the reduced-form elasticity of self-employed reported income with respect to the statutory net-of-tax rate $1 - T'(z)$. Because the self-employed

bear only the effective rate $\sigma T'(z)$, this equals their structural elasticity (defined over their own net-of-tax rate $1 - \sigma T'(z)$) times the pass-through σ .

$$dB = - \left(dT'(z)h(z)\rho(z) \cdot e_1 z \frac{T'(z)}{1 - T'(z)} + \sigma dT'(z)h(z)(1 - \rho(z))e_2 z \frac{T'(z)}{1 - T'(z)} \right)$$

$$dB = -(h(z)\rho(z)e_1 + \sigma h(z)(1 - \rho(z))e_2) \left(dT'(z) \cdot z \cdot \frac{T'(z)}{1 - T'(z)} \right)$$

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

$$(1 - q_1^+(z))dM_1 + (1 - q_2^+(z))dM_2$$

$$- (h(z)\rho(z)e_1 + \sigma h(z)(1 - \rho(z))e_2) \left(dT'(z) \cdot z \cdot \frac{T'(z)}{1 - T'(z)} \right) = 0$$

$$\frac{(1 - q_1^+(z))dM_1 + (1 - q_2^+(z))dM_2}{(h(z)\rho(z)e_1 + \sigma h(z)(1 - \rho(z))e_2) \cdot z} = dT'(z) \cdot \frac{T'(z)}{1 - T'(z)}$$

$$\frac{(1 - q_1^+(z))\rho^+(z) + (1 - q_2^+(z))\sigma(1 - \rho^+(z))}{(\rho(z)e_1 + \sigma(1 - \rho(z))e_2)} \cdot \frac{[1 - H(z)]}{h(z)z} = \frac{T'(z)}{1 - T'(z)}$$

Which leads us to our general equation for optimal taxes

$$\frac{T'(z)}{1 - T'(z)} = \frac{(1 - q_1^+(z))\rho^+(z) + (1 - q_2^+(z))\sigma(1 - \rho^+(z))}{(\rho(z)e_1 + \sigma(1 - \rho(z))e_2)\alpha(z)}$$

Exemption result: by construction we know that:

$$(1 - q_1^+(0))\bar{\rho} + (1 - q_2^+(0))(1 - \bar{\rho}) = 0$$

That is, giving one dollar to all agents has a mechanical gain of 1 and a welfare cost of 1. However, note that as long as the self-employed do not comply fully with taxes, i.e. $\sigma < 1$, then we see that:

$$(1 - q_1^+(0))\bar{\rho} + (1 - q_2^+(0))\sigma(1 - \bar{\rho}) \neq 0$$

Correspondingly this implies the potential for exemptions, i.e.,

$$(1 - q_1^+(\bar{z}))\rho^+(\bar{z}) + (1 - q_2^+(\bar{z}))\sigma(1 - \rho^+(\bar{z})) = 0, \quad \text{for } \bar{z} > 0$$

As long as $q_1^+(\bar{z})$ is larger than one and σ is less than one, it will be optimal to have some level of exemption $\bar{z} > 0$.

Sufficient statistic:

We can finish rewriting the sufficient statistic formula by defining $1 - Q(z)$ as the tax share weighted net marginal social welfare weight $(1 - q_1^+(z))\rho^+(z) + (1 - q_2^+(z))\sigma(1 - \rho^+(z))$. We can then define $\epsilon(z) = (\rho(z)e_1 + \sigma(1 - \rho(z))e_2)$ simply as the reduced form ETI of all income spanning both the employed and self-employed. We can then rewrite:

$$T'(z) = \frac{1 - Q(z)}{1 - Q(z) + \alpha \cdot \epsilon(z)}$$

2 Proof: General Extensive Margin Model

In this appendix I relax the following assumptions from the baseline model: *(i)* fixed sectoral affiliation, *(ii)* identical productivity across formal and informal sectors, and *(iii)* costless tax administration. The proof uses the tax-perturbation method of Jacquet and Lehmann (2021). I first prove the condition for the unconstrained problem, then impose a Kuhn-Tucker non-negativity condition.

2.1 Setup with Roy-style Sectoral Choice

Workers differ in two dimensions indexed by (n, θ) where n is skills and θ is a continuously distributed informal-sector opportunity. We let $h(n)$ denote the density of skill and $F_{\theta|n}(\theta|n)$ and $F_{n|\theta}(n|\theta)$ the conditional distributions of θ and n respectively. The corresponding densities are $f_{\theta|n}$ and $f_{n|\theta}$.

The indirect utility of formal employment is $v_1(n; T)$ while the indirect utility of informal work, $v_2(n, \theta)$, also depends on θ . Maintaining quasi-linearity we see that conditional on being formally employed a worker chooses taxable earnings

$$z_1(n; T) = \operatorname{argmax}_z \left\{ z - T(z) - b \left(\frac{z}{n} \right) \right\} \quad (\text{B1})$$

The corresponding indirect utility is

$$v_1(n; T) = z_1(n; T) - T(z_1) - b \left(\frac{z_1(n)}{n} \right) \quad (\text{B2})$$

For an interior formal earnings choice,

$$1 - T'(z_1(n; T)) = \frac{1}{n} b' \left(\frac{z_1(n; T)}{n} \right) \quad (\text{B3})$$

Informal employment is untaxed. To separate productivity from Roy utility heterogeneity, let $\kappa(n) \in [0, 1]$ denote the informal productivity wedge and let $\psi_\theta(n, \theta) > 0$ denote an idiosyncratic informal opportunity term.

$$v_2(n, \theta) = \max_{\ell} \{ \kappa(n)n\ell - b(\ell) + \psi(n, \theta) \} > 0 \quad (\text{B4})$$

The informal labor choice solves $\kappa(n)n = b'(\ell^*(n))$ and workers choose formal employment if and only if

$$v_1(n; T) \geq v_2(n, \theta) \quad (\text{B5})$$

Define the Roy cutoff, $\bar{\theta}$, by

$$v_2(n, \bar{\theta}(n; T)) = v_1(n; T) \quad (\text{B6})$$

Workers are formal if $\theta \leq \bar{\theta}(n; T)$. The formal share at skill n is therefore.

$$\rho(n; T) = F_{(\theta|n)}(\bar{\theta}(n; T) | n) \quad (\text{B7})$$

Let $\bar{z}$ denote the exemption threshold and let $\gamma > 0$ be the administrative cost per selected formal taxpayer above that threshold. The government budget constraint is

$$\int \rho(n; T) T(z_1(n; T)) h(n) dn - \gamma \int \rho(n; T) \mathbf{1}_{\{z_1(n; T) > \bar{z}\}} h(n) dn \geq \mathcal{R} \quad (\text{B8})$$

Where $\mathcal{R}$ denotes the revenue requirement. The first term is tax revenue from formal workers. The second term is the administrative cost of collecting from formal taxpayers above the exemption threshold. At the end of this proof, I present the case where the cost accrues to all workers in the taxbase.

2.2 Tax Perturbation

Let $R(z)$ be a smooth reform direction and define a continuous tax function

$$T_x(z) = T(z) - xR(z) \quad (\text{B9})$$

A positive R is a tax cut. In line with the assumption of undifferentiated lump-sums, we restrict to perturbations with $R(0) = 0$. Define the changes in labor supply and formality given a perturbation.

$$z_R(n) \equiv \left. \frac{d}{dx} z_{1,x}(n) \right|_{x=0} \quad \text{and} \quad \rho_R(n) \equiv \left. \frac{d}{dx} \rho(n; T_x) \right|_{x=0} \quad (\text{B10})$$

By the envelope theorem,

$$v_{1,R}(n) = R(z_1(n)) \quad (\text{B11})$$

Differentiating the formal earnings FOC to perturbation intensity x at $x = 0$ gives

$$R'(z_1(n)) - T''(z_1(n))z_R(n) = \frac{1}{n^2}b''\left(\frac{z_1(n)}{n}\right)z_R(n) \quad (\text{B12})$$

Hence the first-order response of earnings to a small movement in the direction R is

$$z_R(n) = \frac{R'(z_1(n))}{T''(z_1(n)) + \frac{1}{n^2}b''\left(\frac{z_1(n)}{n}\right)} \quad (\text{B13})$$

Define the formal earnings elasticity by

$$e(n) \equiv \frac{1 - T'(z_1)}{z_1(n)} \left[T''(z_1(n)) + \frac{1}{n^2}b''\left(\frac{z_1(n)}{n}\right) \right]^{-1} \quad (\text{B14})$$

Substituting this definition gives

$$z_R(n) = \frac{e(n)z_1(n)}{1 - T'(z_1)} R'(z_1(n)) \quad (\text{B15})$$

The perturbed Roy cutoff is the identity $v_2(n, \bar{\theta}_x(n)) = v_1(n; T_x)$. Differentiate both sides with respect to x at $x = 0$. Since informal utility is tax-invariant, the left-hand side changes only through the moving cutoff:

$$v_{2,\theta}(n, \bar{\theta}_x(n))\bar{\theta}_R(n) = v_{1,R}(n) \quad (\text{B16})$$

Using the envelope result in (B11)

$$\theta_R(n) = \frac{R(z_1(n))}{v_{2,\theta}(n, \theta(n))} \quad (\text{B17})$$

With the additive representation in (B4) this simplifies $\theta_R(n) = \frac{R(z_1(n))}{\psi_\theta(n, \theta(n))}$. Finally, differentiating the perturbed formal share, $\rho(n; T_x)$, gives

$$\rho_R(n) = \frac{f_{\theta|n}(\theta(n) | n)}{v_{2,\theta}(n, \theta(n))} R(z_1(n)) \quad (\text{B18})$$

Define the semi-elasticity of the formal share with respect to formal disposable income as

$$\eta_\rho(n) \equiv \frac{1}{\rho(n; T)} \frac{f_{\theta|n}(\theta(n) | n)}{v_{2,\theta}(n, \theta(n))} \quad (\text{B19})$$

Then the relative change in formality with respect to a given perturbation is

$$\frac{\rho_R(n)}{\rho(n; T)} = \eta_\rho(n) R(z_1(n)) \quad (\text{B20})$$

2.3 Welfare Envelope

Let M be a differentiable transformation of individual utility capturing vertical equity preferences and let Ω be a differentiable aggregator over the within-skill welfare index capturing horizontal preferences. Define

$$A(n; T) \equiv \int_{-\infty}^{\bar{\theta}(n; T)} M(v_1(n; T)) f_{\theta|n}(\theta|n) d\theta + \int_{\bar{\theta}(n; T)}^{\infty} M(v_2(n, \theta)) f_{\theta|n}(\theta|n) d\theta \quad (\text{B21})$$

Social welfare is

$$W(T) = \int_{\mathcal{N}} \Omega(A(n; T)) h(n) dn \quad (\text{B22})$$

Let $\lambda > 0$ denote the multiplier on the government budget constraint. The per-capita marginal social welfare weight on a formal worker is as noted $q_1(n) \equiv \frac{1}{\lambda} \Omega_A(A(\cdot)) M'(v_1(\cdot))$

Differentiating (B22) with regards to the intensity of the perturbation we use that $v_1(n) = v_2(n, \bar{\theta}(n))$ at the cutoff such that the welfare effect simplifies to

$$\frac{1}{\lambda} \frac{dW(T_x)}{dx} \Big|_{x=0} = \int_{\mathcal{N}} \rho(n) q_1(n) R(z_1(n)) h(n) dn \quad (\text{B23})$$

2.4 Sufficient Statistics

Assume $n \mapsto z_F(n; T)$ is differentiable and strictly monotone. Let $\hat{n}(z)$ denote its inverse. The normalized formal earnings density is $g(z)$ and the total formal share is $\bar{\rho}$ as defined in section 8. Let $\tilde{q}_1(z)$, $\tilde{e}(z)$, and $\tilde{\eta}_\rho(z)$ denote q_1 , e , and η_ρ expressed in formal-income space. Define upper-tail averages by $s^+(z) \equiv \frac{\int_z^\infty s(x) g(x) dx}{1 - G(z)}$.

As in the Jacquet and Lehmann (2021) perturbation method, the planner considers arbitrary smooth directions R and imposes the government budget constraint through a multiplier rather than by requiring each direction to be mechanically budget neutral. Along the path $T_x = T - xR$, define the Lagrangian

$$\mathcal{L}(x; R) = W(T_x) + \lambda[\mathcal{N}(x) - \mathcal{R}]$$

Where $\mathcal{N}(x)$ is net government revenue along the perturbation and $\rho_x(n) \equiv \rho(n; T_x)$

$$\mathcal{N}(x) \equiv \int_{\mathcal{N}} \rho_x(n) T_x(z_{1,x}(n)) h(n) dn - \gamma \int_{\mathcal{N}} \rho_x(n) \mathbf{1}_{\{z_{1,x}(n) > \bar{z}\}} h(n) dn \quad (\text{B24})$$

At an interior optimum, every admissible perturbation must have zero first-order effect on the Lagrangian:

$$0 = \left. \frac{d\mathcal{L}(x; R)}{dx} \right|_{x=0} \quad (\text{B25})$$

Dividing by λ ,

$$0 = \frac{1}{\lambda} \left. \frac{dW(T_x)}{dx} \right|_{x=0} + \left. \frac{d\mathcal{N}(x)}{dx} \right|_{x=0} \quad (\text{B26})$$

Here we use the result

$$\left. \frac{d\mathcal{N}(x)}{dx} \right|_{x=0} = \int \rho_R(n) [T(z_1) - \gamma \mathbf{1}_{\{z_{1,x} > \bar{z}\}}] + \rho(n) [-R(z_1) + T'(z_1) z_R] dn \quad (\text{B27})$$

Here, for a given exemption threshold $\bar{z}$, the interior perturbation R holds $\bar{z}$ fixed and induces no crossing of the threshold, so the derivative of $\gamma \mathbf{1}_{\{z_{1,x}(n) > \bar{z}\}}$ contributes no interior term. The other effect, workers crossing $\bar{z}$, is the separate exemption-boundary margin in (B35)–(B38), where it enters as the administrative-cost term $\gamma g(\bar{z})$. We use the welfare derivative in (B23), the Roy response (B20) and defining the Roy correction $\Gamma(z; \bar{z}) \equiv [T(z) - \gamma \mathbf{1}_{\{z > \bar{z}\}}] \tilde{\eta}_\rho(z)$ we see

$$0 = \int_{\mathcal{N}} \rho(n) \{ [q_1(n) - 1 + \Gamma(z_1(n); \bar{z})] R(z_1(n)) + T'(z_1(n)) z_R(n) \} h(n) dn \quad (\text{B28})$$

Substitute the intensive-margin response (B15) then changing variables from skill to formal earnings,

$$0 = \bar{\rho} \int \left\{ [\tilde{q}_1(z) - 1 + \Gamma(z; \bar{z})] R(z) + \frac{T'(z)}{1 - T'(z)} \tilde{e}(z) z R'(z) \right\} g(z) dz \quad (\text{B29})$$

To solve this optimality condition, define

$$\mathcal{A}(z) \equiv [\tilde{q}_1(z) - 1 + \Gamma(z; \bar{z})] g(z) \quad \text{and} \quad \mathcal{B}(z) \equiv \frac{T'(z)}{1 - T'(z)} \tilde{e}(z) z g(z)$$

Then (B29) is equivalent to

$$\bar{\rho} \int [\mathcal{A}(z) R(z) + \mathcal{B}(z) R'(z)] dz = 0$$

Now apply integration by parts to the term involving $R'(z)$. For any admissible perturbation in this interior unconstrained problem, the transversality condition implies that the IC does not bind at the endpoints. Thus

$$[\mathcal{B}(z)R(z)]_0^\infty = 0 \quad (\text{B30})$$

The integration-by-parts identity is $\int_0^\infty \mathcal{B}(z) R'(z) dz = [\mathcal{B}(z)R(z)]_0^\infty - \int_0^\infty \mathcal{B}'(z)R(z) dz$. The first-order condition becomes

$$0 = \int_0^\infty [\mathcal{A}(z) - \mathcal{B}'(z)] R(z) dz$$

If $\mathcal{A}(z) - \mathcal{B}'(z)$ were positive on some small interval, one could choose a smooth positive perturbation R supported on that interval to make the integral positive. If it were negative on some small interval, one could choose the opposite bump. Hence the integral can be zero for all smooth interior perturbations only if the bracketed term is zero pointwise on the interior:

$$\mathcal{B}'(z) = \mathcal{A}(z) = [\tilde{q}_1(z) - 1 + \Gamma(z; \bar{z})]g(z) = -[1 - \tilde{q}_1(z) - \Gamma(z; \bar{z})]g(z) \quad (\text{B31})$$

This differential equation pins down $\mathcal{B}(z)$ by integrating down from $\mathcal{B}(z_1) = 0$ such that

$$-\mathcal{B}(z) = - \int_z^\infty [1 - \tilde{q}_1(x) - \Gamma(x; \bar{z})] g(x) dx \quad (\text{B32})$$

Inserting the definition of $\mathcal{B}(z)$ into (B32)

$$\frac{T'(z)}{1 - T'(z)} \tilde{e}(z) z g(z) = \int_z^\infty [1 - \tilde{q}_1(x) - \Gamma(x; \bar{z})] g(x) dx \quad (\text{B33})$$

The right-hand side of (B33) can be written in terms of upper-tail averages.

$$\begin{aligned} \frac{T'(z)}{1 - T'(z)} \tilde{e}(z) z g(z) &= [1 - G(z)][1 - \tilde{q}_1^+(z) - \Gamma^+(z; \bar{z})] \\ \frac{T'(z)}{1 - T'(z)} &= \frac{1 - G(z)}{z g(z)} \frac{1 - \tilde{q}_1^+(z) - \Gamma^+(z; \bar{z})}{\tilde{e}(z)} \end{aligned}$$

Using the local Pareto parameter $\frac{1-G(z)}{z g(z)} = \frac{1}{\alpha(z)}$

$$\frac{T'(z)}{1 - T'(z)} = \frac{1 - \tilde{q}_1^+(z) - \Gamma^+(z; \bar{z})}{\alpha(z) \tilde{e}(z)}$$

Finally, rewrite to

$$T'(z) = \frac{1 - \tilde{q}_1^+(z) - \Gamma^+(z; \bar{z})}{1 - \tilde{q}_1^+(z) - \Gamma^+(z; \bar{z}) + \alpha(z)\tilde{e}(z)} \quad (\text{B34})$$

This characterizes the optimal unconstrained marginal tax rate for a given $\bar{z}$. Thus, Roy selection changes the optimal marginal tax formula through the net fiscal value of additional formal workers

$$\Gamma^+(z; \bar{z}) = \frac{1}{1-G(z)} \int_z^\infty [T(x) - \gamma \mathbf{1}_{\{z(x) > \bar{z}\}}] \tilde{\eta}_\rho(x) g(x) dx.$$

Exemption Boundary

The unconstrained first-order condition in (B34) takes the exemption boundary $\bar{z}$ as given. However, because of the presence of administrative costs, the exemption boundary is a separate margin. To characterize the ex-post choice of $\bar{z}$, hold the marginal-rate schedule above the boundary fixed and compare two effects: raising the boundary gives all taxpayers above $\bar{z}$ an inframarginal tax cut, but it also saves administrative costs on the marginal group removed from the administered tax base. Write the schedule as

$$T_z(z) = \begin{cases} 0, & z \leq \bar{z}, \\ \int_{\bar{z}}^z T'(s) ds, & z > \bar{z}, \end{cases} \quad (\text{B35})$$

Raising the exemption boundary by dz gives current taxpayers above $\bar{z}$ an inframarginal tax cut. This is equivalent to the reform direction

$$R_{\bar{z}}(z) = T'(\bar{z}) \mathbf{1}_{\{z > \bar{z}\}} dz \quad (\text{B36})$$

It also directly saves administrative costs $\gamma \bar{\rho} g(z) dz$. Consider then the optimal boundary using the Lagrangian $\mathcal{L}(x; R) = W(T_x) + \lambda[\mathcal{N}(x) - \mathcal{R}]$. Normalizing by $\lambda \bar{\rho}$ and find the optimum

$$\frac{1}{\lambda \bar{\rho}} \frac{d\mathcal{L}}{d\bar{z}} = \gamma g(\bar{z}) - T'(\bar{z})[1 - G(\bar{z})][1 - \tilde{q}_F^+(\bar{z}) - \Gamma^+(\bar{z}; \bar{z})] = 0 \quad (\text{B37})$$

A regular interior ex-post boundary satisfies

$$\gamma g(\bar{z}) = T'(\bar{z})[1 - G(\bar{z})][1 - \tilde{q}_F^+(\bar{z}) - \Gamma^+(\bar{z}; \bar{z})] \quad (\text{B38})$$

This condition characterizes the ex-post exemption boundary for a given schedule above the threshold.

2.5 Effects of Productivity Differentials

Conditional on the sufficient-statistics objects that appear in (B34), the value of κ has no separate direct effect on the optimal marginal tax formula.

Proposition. Fix the formal earnings distribution $g(z)$, the formal earnings response $\tilde{e}$, the formal-share semi-elasticity $\tilde{\eta}_\rho$, the marginal social welfare weights $\tilde{q}_1$, and the administrative cost γ . Then changing the primitive productivity wedge $\kappa(n)$ leaves the optimal tax formula (B34) unchanged.

Proof. The interior formula depends on primitives only through

$$\alpha(z), \tilde{e}(z), \tilde{q}_1^+(z), \Gamma^+(z; \bar{z})$$

The Roy correction is

$$\Gamma^+(z; \bar{z}) = \frac{1}{1 - G(z)} \int_z^\infty [T(x) - \gamma \mathbf{1}_{\{z(x) > \bar{z}\}}] \tilde{\eta}_\rho(x) g(x) dx$$

Thus, no remaining term in (B32) contains κ . The intuition for this is that the effect the productivity differential has on welfare already is captured by the MSWWs. Simultaneously, while the differential does affect the counterfactual informal earnings of a marginal worker changing sectors this does not change the fiscal cost of the change which remains $T(z)$ whilst the welfare cost of the worker in question will be 0 by definition of indifference.

Note that the same holds for the exemption boundary condition.

2.6 Non-negative Marginal Tax Rates

Denote the unconstrained optimal tax $T'_u(z)$. If the planner also imposes a shape constraint such as $T'(z) \geq 0$, then the same perturbation logic gives a Kuhn-Tucker. Let $t_c(z) \equiv T'_c(z)$ denote the constrained optimum where all sufficient statistics in this section are evaluated at T_c .

Define the constrained formal-plus-Roy social weight

$$\mathcal{Q} \equiv \tilde{q}_1^+(z) + \Gamma^+(z; \bar{z}) \quad (\text{B39})$$

Let $\omega(z)$ be the normalized Kuhn-Tucker multiplier on the constraint $t_c(z) \geq 0$. Since the perturbation is $T_x = T_c - xR$, the induced perturbation of the marginal tax rate is $t_x(z) = t_c(z) - xR'(z)$. The multiplier therefore enters the first variation through the $R'(z)$ term. The constrained analogue of (B29) is

$$0 = \bar{\rho} \int \left\{ [\tilde{q}_1(z) - 1 + \Gamma(z; \bar{z})] R(z) + \frac{t_c(z)}{1 - t_c(z)} \tilde{e}(z) z R'(z) - \omega(z) R'(z) \right\} g(z) dz \quad (\text{B40})$$

with

$$\omega(z) \geq 0 \quad \text{and} \quad t_c(z) \geq 0 \quad \text{and} \quad \omega(z) t_c(z) = 0$$

Integrating by parts as above gives

$$\frac{t_c(z)}{1 - t_c(z)} = \frac{1 - \mathcal{Q}(z; \bar{z})}{\alpha(z)\tilde{e}(z)} + \frac{\omega(z)}{\tilde{e}(z)z} \quad (\text{B41})$$

Where the constraint is slack, $\omega(z) = 0$, and the interior formula (B32) applies at the constrained optimum. Where the constraint binds, $t_c(z) = 0$, and

$$\mathcal{Q}(z; \bar{z}) \geq 1 \quad (\text{B42})$$

Define the unconstrained right-hand side evaluated at the constrained allocation by $t_{u,c}(z)$. Equivalently, the nonnegative constrained marginal tax schedule is

$$t_c(z) = \max\{0, t_{u,c}(z)\} \quad (\text{B43})$$

This is the analogue of the clipping argument in appendix section A9, but it must be read as a fixed-point condition evaluated at the constrained optimum.

2.7 Extension: Sector-Specific Administrative Costs

The main proof assumes that administrative costs are paid only for formal taxpayers above the exemption. Suppose instead that the fiscal administrative cost can differ by sector. Let γ_F be the cost of a formal taxpayer and γ_I the cost of an informal worker. The admin cost term in the budget constraint then becomes

$$\int_{\mathcal{N}} \mathbf{1}_{\{z_1(n; T) > \bar{z}\}} [\gamma_F \rho(n; T) + \gamma_I (1 - \rho(n; T))] h(n) dn \quad (\text{B44})$$

A marginal Roy switcher changes the cost from γ_I to γ_F . The interior Roy correction is therefore

$$\Gamma^{\text{alt}}(z; z) = [T(z) - (\gamma_F - \gamma_I) \mathbf{1}_{\{z > \bar{z}\}}] \tilde{\eta}_\rho(z) \quad (\text{B45})$$

Thus formal-only costs are the special case $\gamma_I = 0$, while an equal cost for everyone in the covered range is the special case $\gamma_F = \gamma_I$. In addition to impacting optimal rates the administrative costs also impacts the boundary condition controlling the placement of $\bar{z}$.

The direct cost saving from increasing the exemption by one unit is now $(\gamma_F \rho(\hat{n}(z); T) + \gamma_I (1 - \rho(\hat{n}(z); T))) h(n)$. Raising the exemption boundary removes all workers at the margin from the administered range, not only formal workers. Relative to the formal-only case, the direct administrative-cost saving must therefore be inflated by the average covered cost per formal worker at the boundary:

$$\gamma_F + \gamma_I \frac{1 - \rho(z; T)}{\rho(z; T)}$$

The ex-post boundary condition therefore becomes

$$g(z) \left[\gamma_F + \gamma_I \frac{1 - \rho(z; T)}{\rho(z; T)} \right] = T'(z)[1 - G(z)][1 - \tilde{q}_1^+(z) - \Gamma^{alt}(z; z)] \quad (\text{B46})$$

3 Survey Methodology

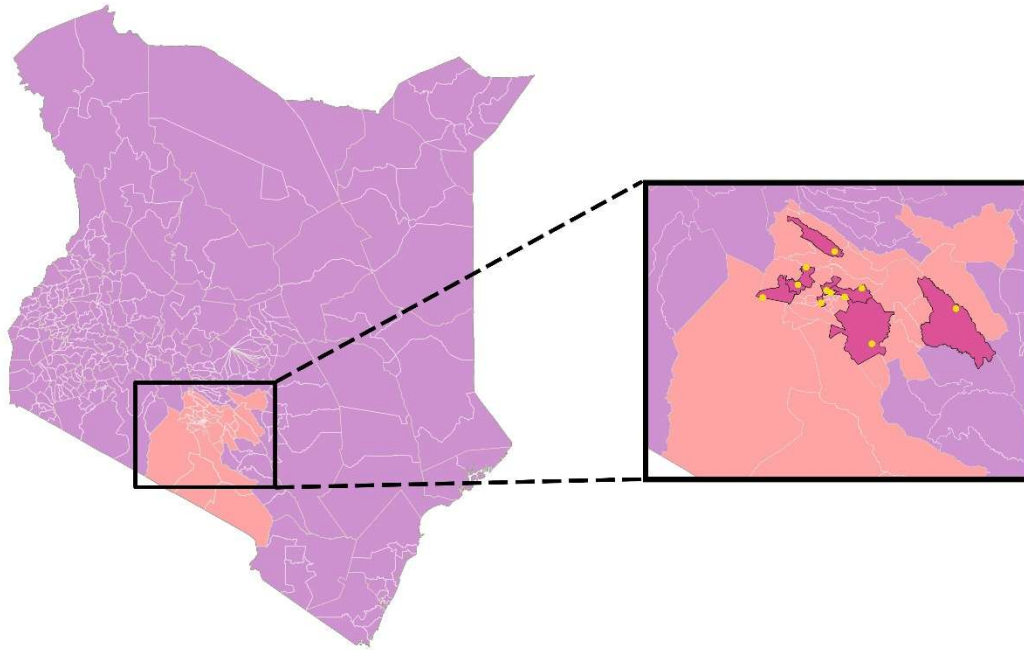
3.1.1 Sampling

The sampling frame consists of all subcounties within Nairobi, Kiambu, Machakos, and Kajiado counties. Subcounties were selected with probability proportional to population size, using population counts from the 2019 Kenya Population and Housing Census (KNBS, 2019). The selected subcounties were (w. number of responses in parenthesis): Gatundu South (50), Kabete (51), Kasarani (n/a)²⁵, Kajiado (43), Kiambaa (51), Kikuyu (42), Mavoko (73), Mwala (46), Roysambu (51), Ruaraka (69), and Starehe (51).

Within each selected subcounty, random geographic coordinates were drawn using QGIS and adjusted to the nearest identifiable settlement to ensure feasibility. All locations were reviewed by the field supervisor prior to enumeration. When a location was deemed unsafe, coordinates were either shifted to a nearby safe settlement within a one-kilometer radius or redrawn at random following the same procedure. For safety reasons, fieldwork in Kasarani subcounty was discontinued.

²⁵ Kasarani was originally drawn for the sample but due to violent unrest at the time of surveying it was dropped.

Figure O1: Sampling Framework



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

Upon arrival at the approved starting location, enumerators selected a random initial walking direction using a spun-pin method. Enumerators followed this direction until reaching an intersection, at which point a new direction was chosen using the same procedure. If a direction was deemed unsafe, the pin was re-spun until a safe direction was identified. This protocol was designed to preserve randomness while prioritizing enumerator safety.

Enumerators followed a skip-one sampling rule, interviewing every second establishment encountered along their route. Sampling was conducted at the building level rather than the household level, reflecting the prevalence of home-based economic activity and shared dwellings in the study areas. This approach avoids requiring enumerators to classify establishments as residential or commercial and reduces ambiguity in household selection.

Enumerators were assigned to subcounties with a target of approximately 50 completed interviews per subcounty. However, enumerators completed full field days even if the target was reached before the end of the day, resulting in modest variation in sample sizes across subcounties. This variation is explicitly accounted for through the construction of base sampling weights, which adjust for differential sampling intensity across locations.

While enumerators were instructed to record all refusals, refusal reporting was incomplete in practice and is therefore not systematically observed across the full sample. Partial refusal data collected by a subset of enumerators, together with direct field supervision by the author and the

field coordinator, indicate that refusal rates were moderate and did not threaten field completion targets. There is no evidence that refusal was systematically concentrated in particular subcounties or time periods. Remaining differences in contact and participation rates across locations are addressed through the construction of base sampling weights.

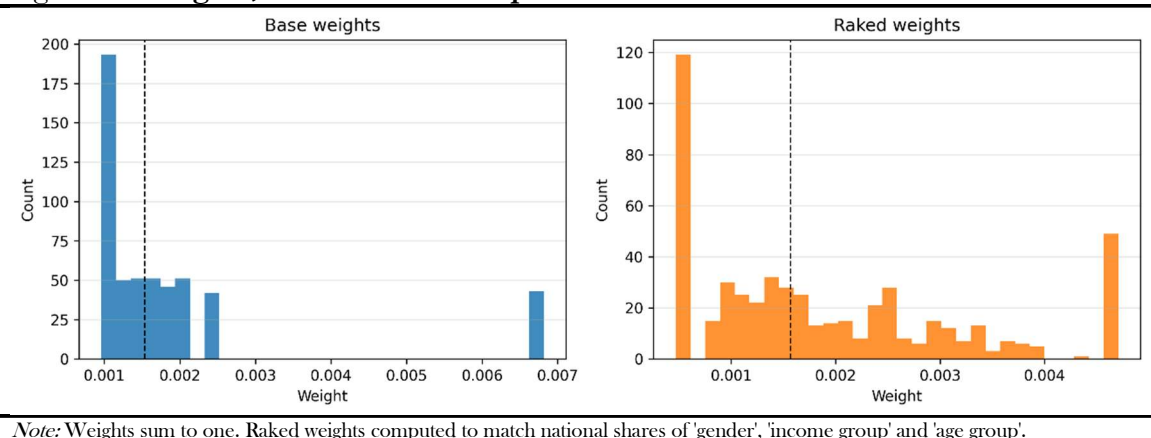
3.1.2 Weights

Survey weights are constructed in two steps. First, I compute base weights to account for unequal sampling intensity across subcounties. Second, I adjust these base weights using raking so that the weighted sample aligns with external population benchmarks for key demographics.

Base weights correct for the fact that subcounties differ both in population size and in the number of completed interviews. Respondents from subcounties with relatively few interviews per resident receive higher weight, while respondents from more intensively sampled subcounties receive lower weight. Operationally, the base weight assigned to a respondent in subcounty h is proportional to the ratio of the subcounty population to the number of completed interviews in that subcounty. Base weights are normalized to sum to one in the analysis sample.

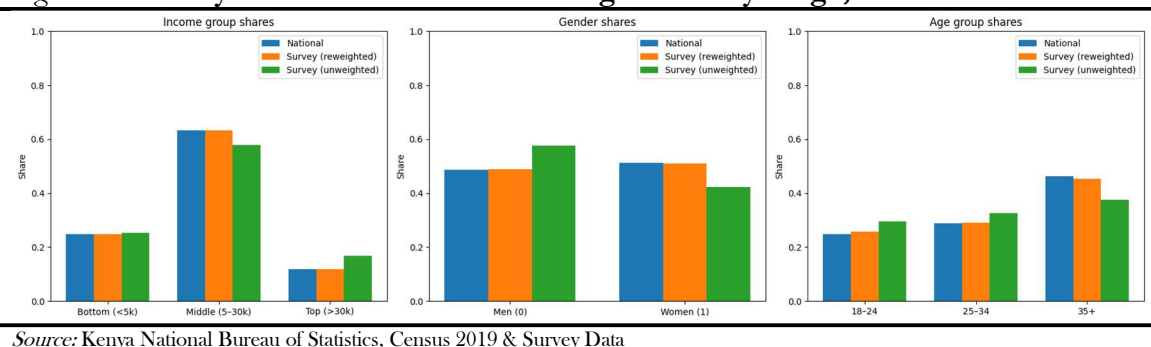
Raking (iterative proportional fitting) then adjusts these base weights so that the weighted sample matches national population shares along three dimensions: income, gender, and age. Income margins are constructed using external data that combine formal- and informal-sector income distributions to obtain population shares in three broad income groups. Gender and age margins are taken from the 2019 Kenya Population and Housing Census, with age grouped into three categories (18–24, 25–34, and 35+). The raking algorithm iteratively rescales the base weights until the weighted marginal distributions of the sample match each of these population benchmarks while preserving the relative differences implied by the base weights.

Figure O2: Weights, Baseline Full Sample



After applying raking weights, the sample composition more closely resembles the national composition.

Figure O3: Survey Means and National Averages in Kenya: Age, Gender & Income



Unless otherwise stated, all weighted results in the paper use raked weights.

3.1.3 Implementation and Framing Assignment

The survey is built in oTree (Chen et al., 2016) which allows for interactive animations and graphics to build understanding. For example, when changing allocations, respondents see the size of piles of coins shrink building understanding even for respondents with limited numeracy.

When a new survey instance is initialized, the potential respondent is immediately assigned to either the gain-gain (80 pct.) or the loss-gain (20 pct.) framing by random draw in the backend of the software. This draw is not affected by respondent characteristics nor by any previous draws. The framing assignment only becomes visible to the enumerator at the third screen, and it is as such not possible for the enumerator to exert influence over the framing assignment.

Below I show that the sample was balanced across the two framings. For the final sample 20.4 percent received the "loss-gain" framing-

Table O4: Balance Table: Framing Assignment

Variable	Gain-Gain (%)	Loss-Gain (%)	Difference (pp)	Std. Diff.	p-value
Female	43.6	38.9	-4.7	-0.1	0.375
Age 18--24	29.4	30.6	1.2	0.03	0.814
Age 25--34	34.1	26.9	-7.3	-0.16	0.137
Age 35+	36.3	42.6	6.3	0.13	0.236
Income up to 5k	21.3	14.8	-6.5	-0.17	0.103
Income 5k--30k	46.4	41.7	-4.8	-0.1	0.373
Income above 30k	12.8	15.7	2.9	0.08	0.449
Income not reported	19.4	27.8	8.3	0.20	0.080

Note: All percentages are measured for the unweighted cleaned sample consisting of 530 observations. Of these responses, 422 had the "gain-gain" framing and the remaining 108 had the "loss-gain" framing. Differences are evaluated using Welch two-sample *t*-tests on indicator variables; standardized differences are reported as a scale-free measure of covariate balance.

In the survey, income levels are randomly drawn from a pre-specified set covering approximately the 4th to the 99th percentile of the income distribution. For the vertical elicitation, one income is first drawn at random from the full set of monthly incomes, excluding the two lowest values, and assigned to the high-income recipient. A second income is then drawn for the low-income recipient from the subset of incomes strictly below the high-income level, excluding both the high-income value itself and the immediately adjacent lower income level. For monthly incomes above the exemption income taxes are computed as average remittances at that income level.

In the subsequent horizontal elicitation, income assignment depends on the income drawn in the vertical round. If the high-income draw exceeds the exemption threshold, that same income level is assigned to both recipients. If it lies below the exemption, a new income is drawn from the subset of incomes above the exemption and assigned to both recipients.

Table O5: Income Sets, Monthly Ksh

Monthly Income (Below Exempt)	Monthly Income (Above Exempt)	Income Taxes
1,650	25,000	750
3,350	29,150	1,500
5,000	33,350	2,500
6,650	37,500	4,000
8,350	41,650	5,000
10,000	45,850	6,500
11,650	50,000	7,500
13,350	54,150	8,500
15,000	58,350	9,000
16,650	62,500	11,000
18,350	66,650	11,500
20,000	83,350	17,500
21,650	125,000	28,500
23,350	208,350	52,000

Note: Monthly income below exempt indicates earnings below the 2021 exemption threshold. Monthly income indicates earnings above threshold. Income taxes are average tax remittance for the specified income level as observed in the KRA 2021 administrative tax data.

After being assigned to a framing the survey proceeds as follows:

1. Consent page
2. Background information
3. Comprehension module
4. Allocation task (repeats three times)
 - a. Vertical distribution task
 - b. Horizontal distribution task
5. Questions on opinions

For the allocation task the respondents face the following set of allocation options.

Table O6: Allocation Set

Budget	Recipient A	Recipient B	MRS
1000	1000	0	1.0
995	950	45	1.1
990	925	65	1.2
970	855	115	1.4
943	785	158	1.6
918	720	198	1.6
891	655	236	1.7
866	595	271	1.7
839	535	304	1.8
814	480	334	1.8
787	425	362	1.9
762	375	387	2.0
736	325	411	2.1
711	280	431	2.2
684	235	449	2.5
658	195	463	2.7
631	155	476	3.2
605	120	485	3.5
577	85	492	5.2
551	55	496	8.1
525	27	498	11.8
500	0	500	13.5

Note: Amounts are experimental points. Recipient B equals Budget – Recipient A by construction. “MRS” reports the target relative price (marginal rate of substitution) used to construct the allocation menu; because payoffs are integer-valued, the implied discrete tradeoff between adjacent options may differ slightly from the target MRS, particularly at low budgets.

3.1.4 Cleaning & Enumerator Quality

Enumerator compliance was evaluated using three criteria: *(i)* average time per response screen, *(ii)* variance in observed allocation choices, and *(iii)* direct in-field observation by the author and the field coordinator. Extremely low variance in responses was treated as a potential indicator of enumerator interference or respondent disengagement. Enumerator performance was monitored in real time throughout the field period.

Based on these criteria, two enumerators failed all three compliance benchmarks and were dismissed from further data collection. All responses collected by these enumerators outside periods of direct supervision were discarded. Responses collected under direct supervision were retained. All remaining five enumerators met all compliance benchmarks and were retained in the analysis sample.

These procedures resulted in the exclusion of 114 observations due to enumerator non-compliance. After applying these exclusions, the cleaned dataset comprises 538 completed interviews.

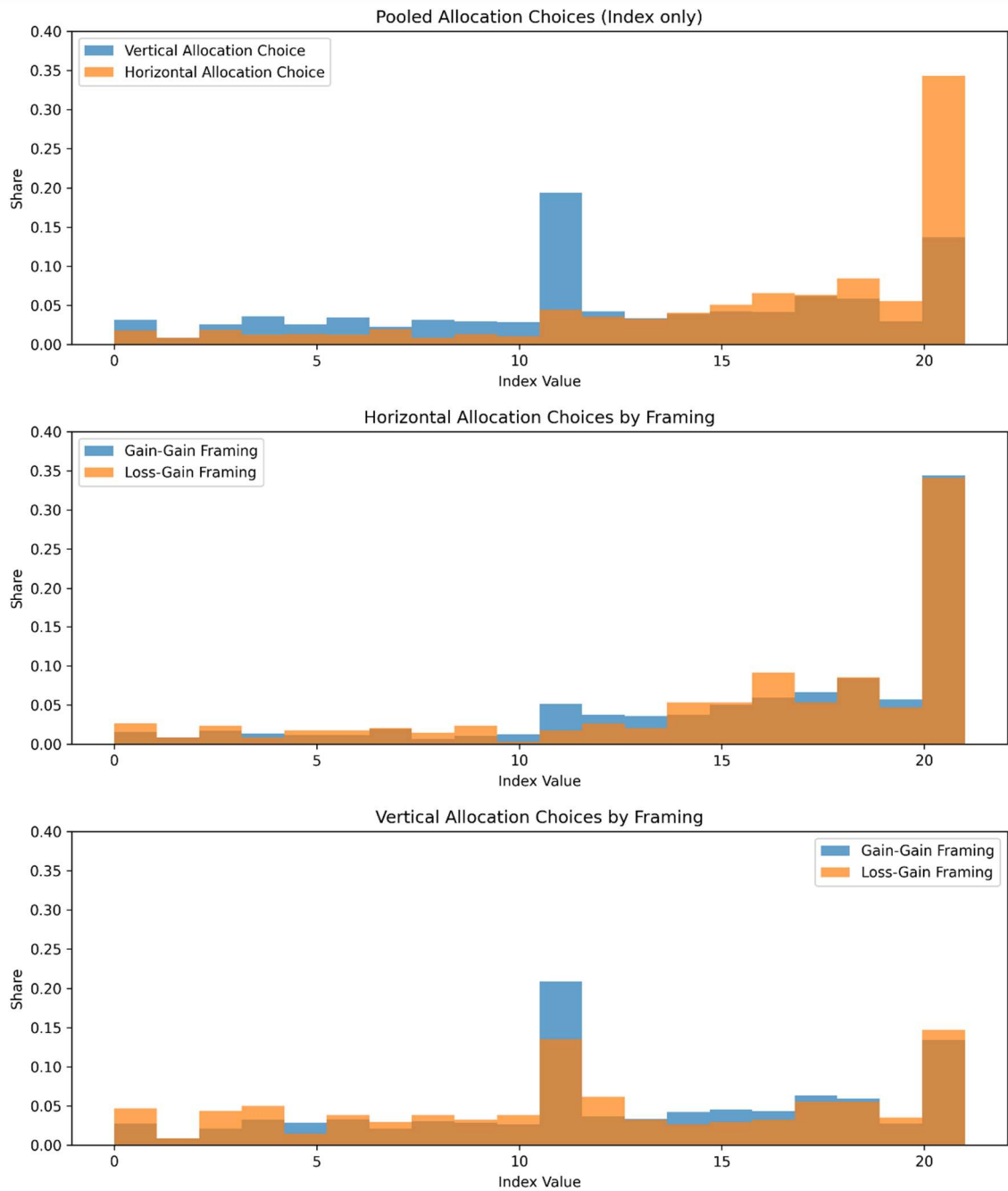
To further ensure respondent engagement, an additional response-time filter was applied. Observations with implausibly short completion times were excluded, yielding a final analysis sample of 530 respondents.

All enumerators received a full day of training prior to fieldwork. The training focused on understanding and explaining the trade-offs inherent in the survey. Enumerators were not informed about the specific hypotheses being tested in the paper. In addition, each enumerator was supervised in the field for at least half a day by both the author and the field coordinator as part of ongoing quality assurance.

3.1.5 Pooled Responses:

Figure O7 presents histograms of pooled allocation choices. For vertical comparisons, there is noticeable excess mass at index 11, corresponding to an equal split of the endowment. This pattern is consistent with a fairness heuristic favoring equal division. To the extent that the analysis targets marginal rates of substitution, such focal choices can blur interpretation. However, choosing an equal split implies that this fairness ideal dominates both alternatives that increase total surplus (lower indices) and those that further equalize outcomes at a higher efficiency cost (higher indices). Absent this auxiliary preference for exact equality, the underlying marginal valuation can therefore be reasonably inferred to lie close to this point. Consistent with this interpretation, median vertical preferences are nearly unchanged under the alternative loss–gain framing, where focal equal-split behavior is substantially reduced.

Figure O7: Histograms of Redistribution Choices



Note: The x-axis represents the index of the chosen allocation, where 0 is the least and 20 is the most redistributive. The corresponding implied MRS ratios can be gleaned from Table O6. The histograms pool responses across all income comparisons.

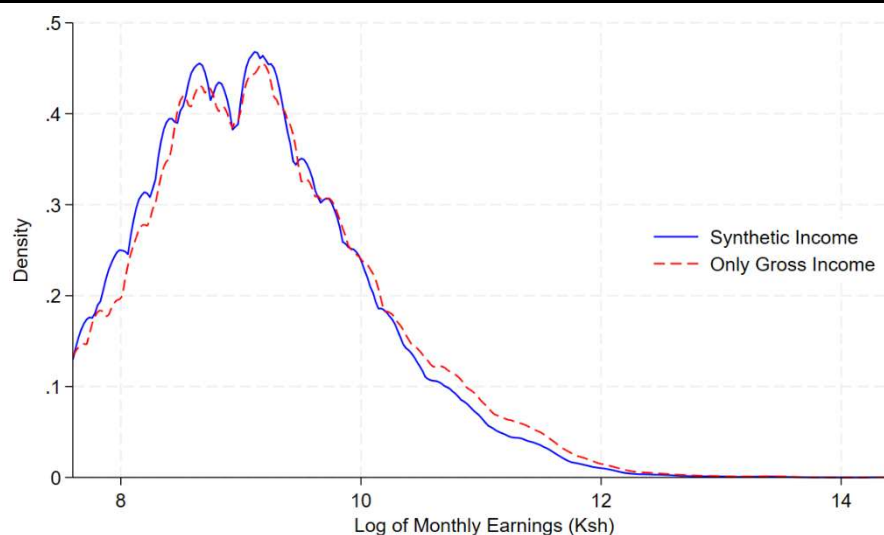
3.1.6 Deviations from Pre-registration

- Pre-registration states that responses will be aggregated using a bottom-up estimation procedure. During analysis, I discovered that the currently employed non-directional aggregator strictly dominates the bottom-up by using all relevant information from the data and imposing less structure. The two estimators yield nearly identical results, but the bottom-up estimator has higher variance.

- The pre-registration states that, as a secondary outcome, an analysis of the preferences only of formal workers would be conducted. However, given issues of misreporting of sectoral affiliation in the data, it is not meaningful to compute this measure.
- The intended gross sample for the survey was 900 respondents. However, due to political unrest at the time of the survey, the enumeration period was shortened to ensure the safety of the enumerators, leading to the final sample of 652 completed interviews.
- The pre-registration notes that in addition to weights for national representativity, weights would also be computed for regional representativity. However, since these weights have no use in subsequent analysis, they have not been computed.

4 Income Imputation

Figure O8: Synthetic and Observed Income Distribution

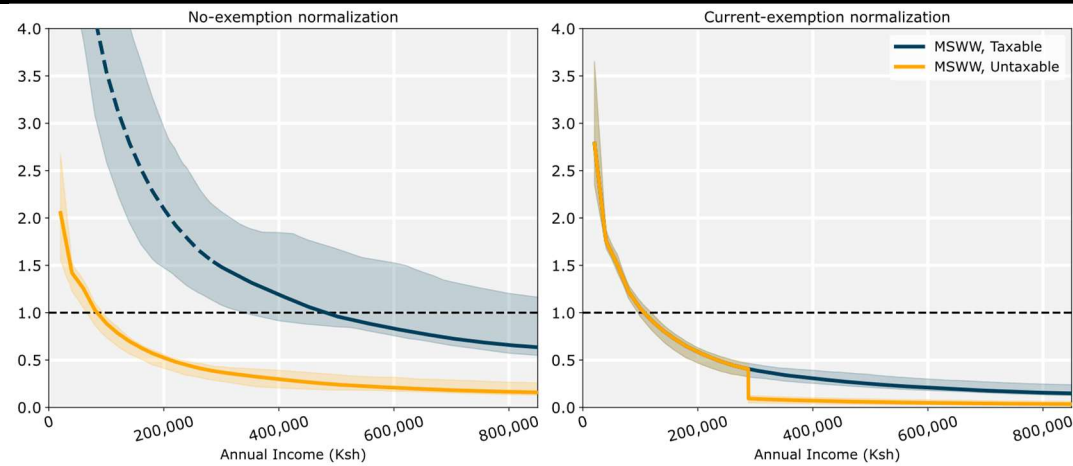


Notes: Observed income refers to reported monthly earnings in the KCHS 2021 microdata. Synthetic income replaces missing or implausibly small values (< KSh 1,500) using predicted earnings from a log-income regression estimated separately for employees and the self-employed. The model uses consumption, demographics, education, location, and household employment composition as predictors. A stochastic residual draw is added to preserve the observed variance of earnings.

5 Ratios Normalized to MSWWs

Moving from observed ratios to Marginal Social Welfare Weights requires normalizing the weights to one. Below I illustrate two such normalizations. The left-hand panel shows MSWWs for the counter-factual of no exemption. Unlike the structural approach from the optimal tax simulations, this illustrative plot merely extrapolates a fixed ratio between the taxable and untaxable workers below the threshold. The right-hand panel uses the current threshold for normalizing.

Figure O9: Empirical Estimates of Marginal Social Welfare Weights, Kenya 2025



Note: Marginal social welfare weights (MSWWs) are shown under two normalization conventions. In the left panel, weights are scaled so that the schedule integrates to unity under the counterfactual of no exemption threshold, i.e., if informal workers were assigned the same relative weight as above the threshold. In the right panel, weights integrate to unity given the current exemption, so the reference point is the observed tax schedule. Shaded bands indicate 95% bootstrap confidence intervals. The suffix "(fixed ratio)" indicates that the MSWWs are plotted as if $z=n$ in order to be able to visually compare to the fixed ratio we see when plotting across incomes, z . As noted in the pre-registration, only workers who received the gain-gain framing are included.

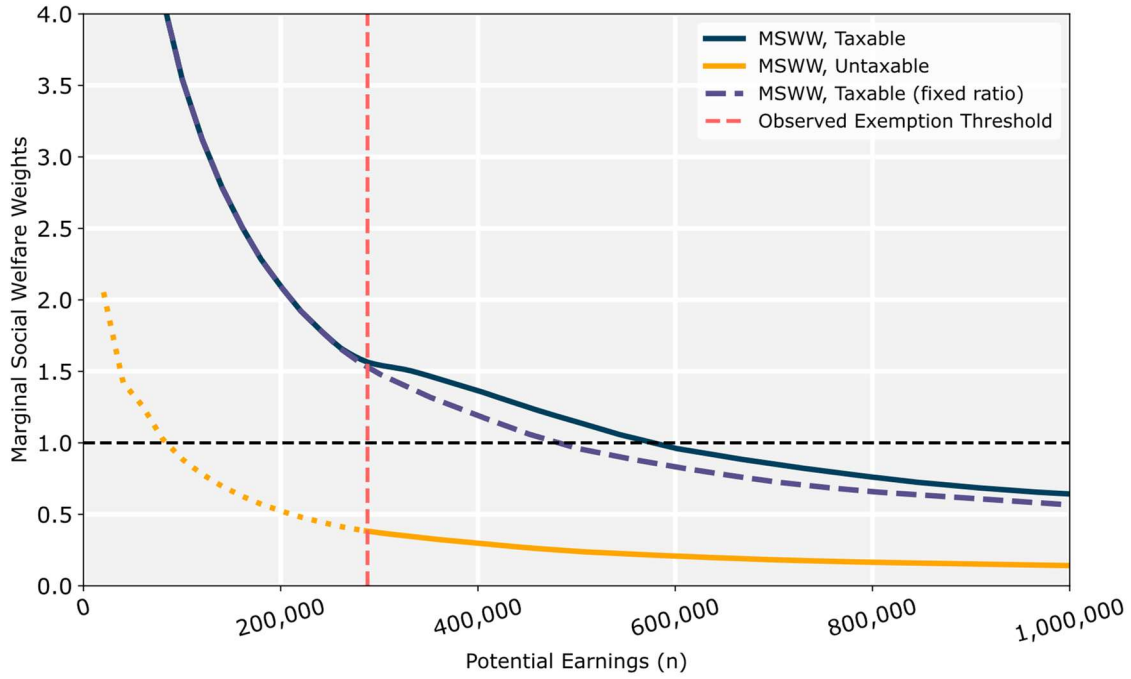
Source: Survey data, Kenya 2025 and author's calculations.

The difference between the left and right panels illustrates the importance the placement of the exemption threshold has on the MSWWs of taxpayers. The imputed ratio below the threshold is only illustrative (the full structural fit is reported in A11) but shows how lowering the threshold increases the number of disadvantaged taxpayers thus increasing the average MSWW of taxpayers.

6 Empirical Marginal Social Welfare Weights across Skills

For a cleaner comparison of horizontal inequality aversion, I plot the MSWWs from the survey across skills, n , rather than incomes. Here the ratio is directly indicative of horizontal inequality.

Figure O10: Survey Marginal Social Welfare Weights across Skills



Note: Schedules are scaled according to counterfactual of no exemption. That is, if all informal workers are assigned one sixth of the weight of formal workers, then MSWWs integrate to one. The suffix "(fixed ratio)" indicates that the MSWWs are plotted as if $z=n$ in order to be able to visually compare to the fixed ratio we see when plotting across incomes, z .

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

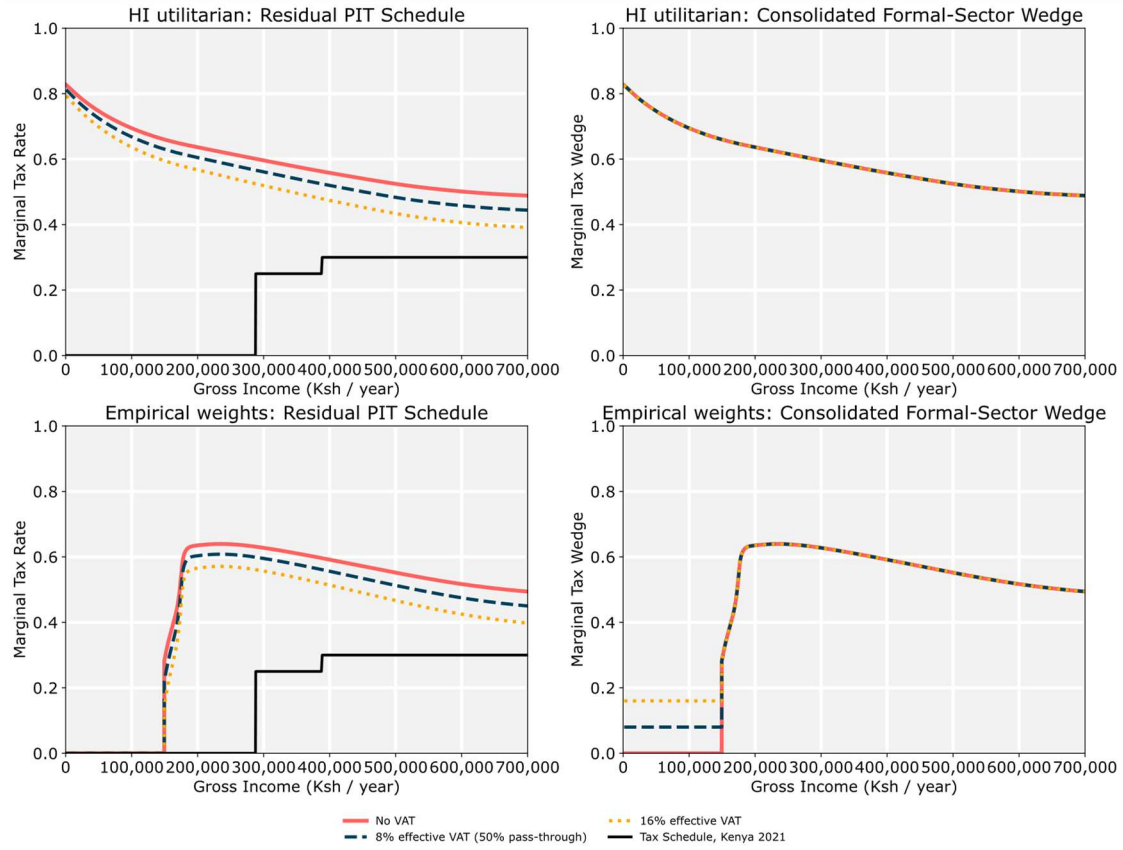
7 Sensitivity of Exemptions to VAT

For an effective common VAT wedge s and a residual PIT rate $P'(z)$, the consolidated formal-sector wedge is $m_F(z) = s + (1 - s)P'(z)$. The residual PIT rate required to implement a target wedge $m_F^*(z)$ is therefore

$$P'^*(z) = \max \left\{ 0, \frac{m_F^*(z) - s}{1 - s} \right\}$$

The non-negativity constraint matters as VAT can provide the full desired wedge at the bottom, while horizontal-inequality considerations can still extend the residual PIT exemption. However, in the simulations below I show that the VAT rate has no impact on the statutory exemption. The experiment uses effective VAT wedges of 8% and 16%. The 8% case corresponds to 50% pass-through of the statutory 16% rate.

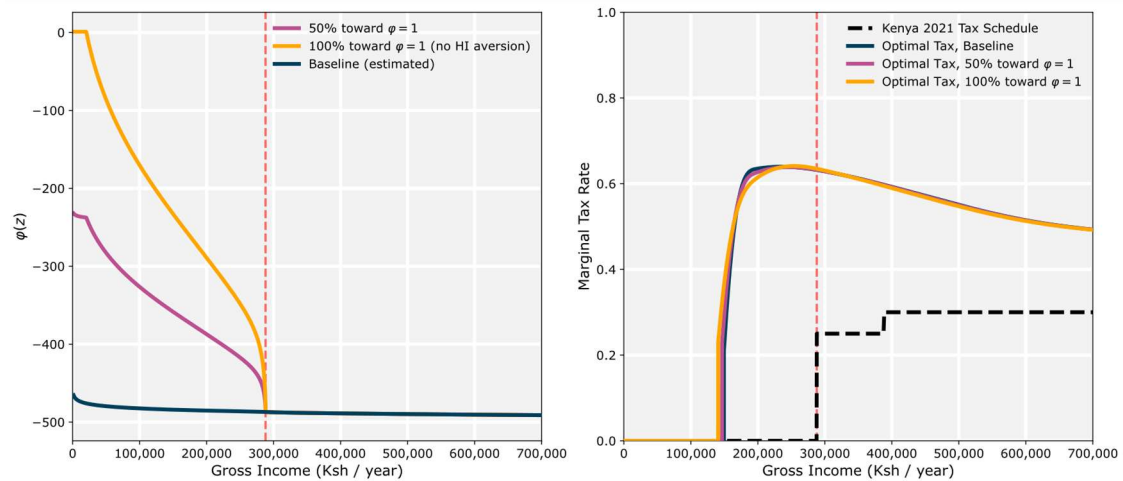
Figure O11: Sensitivity of Exemptions to VAT



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

8 Robustness: Structural SWF Parameters

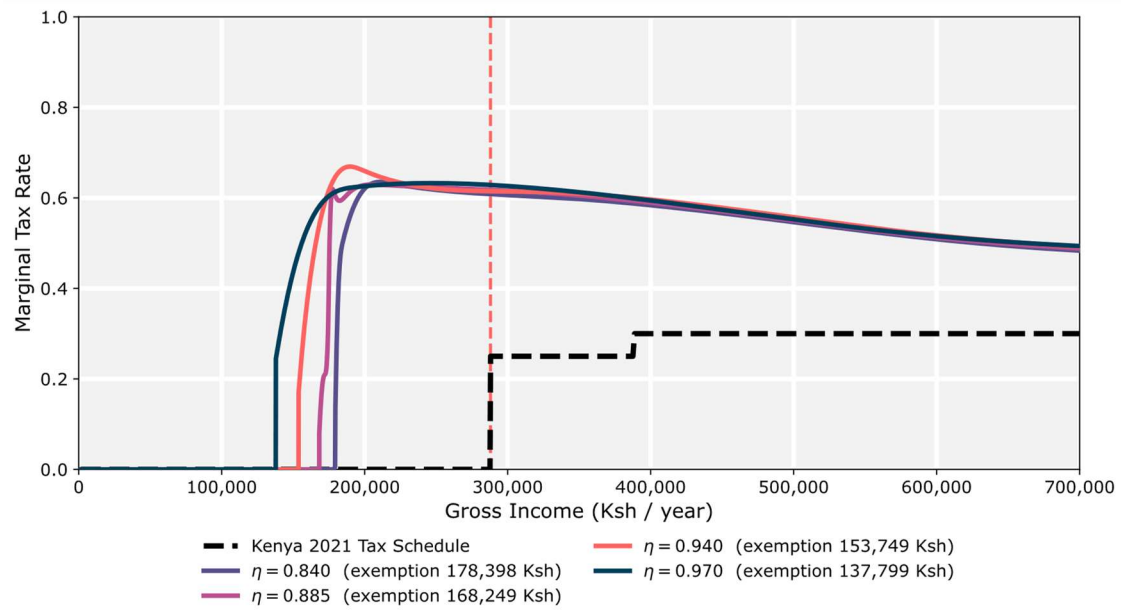
Figure O12: Horizontal Inequality Aversion Extrapolation



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

Note: The left panel plots the horizontal preference parameter $\varphi(z)$ against gross income. φ is identified from elicited preferences above the 288,000 Ksh bracket (dashed vertical line); below this threshold it is unobserved and the baseline extrapolates the estimated quadratic form in log-income. The two robustness schedules set $\varphi_f(z) = \varphi_{baseline}(z) + f \cdot \alpha(z) \cdot (1 - \varphi_{baseline}(z))$, where $f \in \{0.5, 1\}$ is the fraction of the distance toward $\varphi = 1$ and $\alpha(z) = \sqrt{((\log \bar{n} - \log n(z)) / (\log \bar{n} - \log n_{min}))}$ is a square-root taper in log-income that equals 1 at the bottom and 0 at the 288,000 Ksh boundary $\bar{n}$. The right panel shows the corresponding ex-post optimal tax rates.

Figure O13: Vertical Preferences and Optimal Taxes

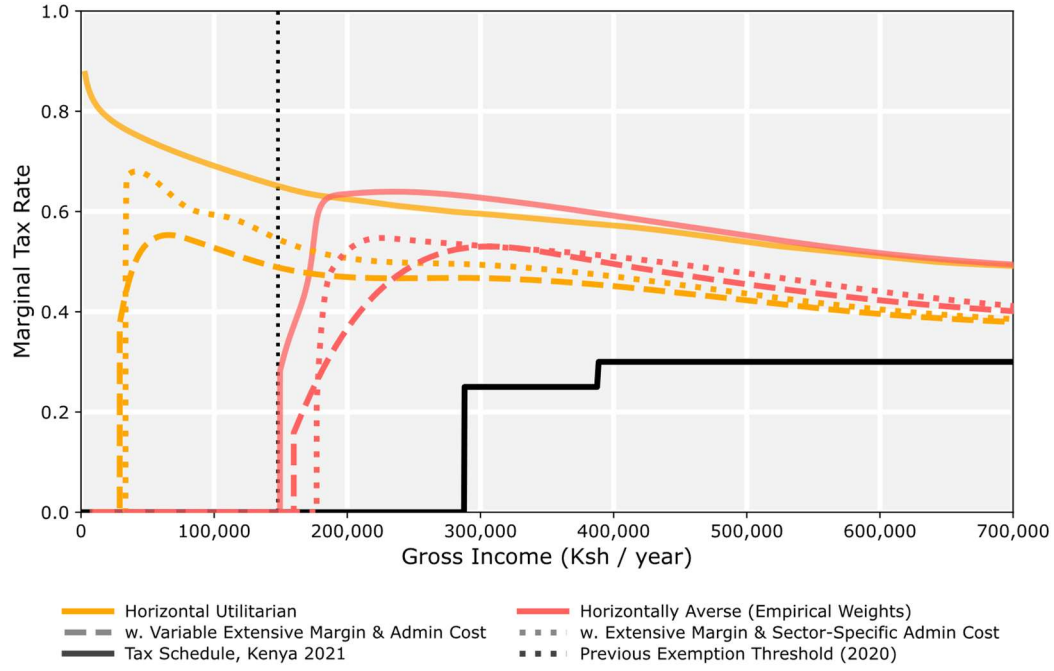


Source: Simulation results based on data from KCHS and KRA.

Note: Each schedule plots the ex-post optimal marginal tax rate for a given value of the vertical preference parameter $\tilde{\psi}$, holding the horizontal preference profile ϕ fixed at the estimated baseline $\hat{\phi}$. The four values span the estimation band around the point estimate.

9 Additional Simulations

Figure O14: Optimal Taxation with Admin Costs and Extensive Margin



Note: Schedules denoted "w. Extensive Margin" feature a participation elasticity of $\tilde{\eta}_p = 0.5$ while schedules with "Variable Extensive Margin" have $\tilde{\eta}_p = 1 - 0.5F_1(z)$. Schedules with "Adm. Cost" and "Sector-Specific Admin Cost" have marginal inclusion costs of an additional worker in the taxbase of $\gamma_F = 5,000$ Ksh and $\gamma_I = 0$ and $\gamma_F = 5,000$ Ksh and $\gamma_I = 1,500$ Ksh respectively. All series plot ex-post optimal marginal tax rates using the empirically estimated vertical preference parameter $\tilde{\psi}$. "Horizontal Utilitarian" incorporates informality but imposes, $\phi = 1$, implying no horizontal inequality aversion. 'Empirical Weights, Kenya' accounts for the presence of an informal sector and for the observed aversion to horizontal inequality embodied in $\tilde{\phi}$.