Optimal Taxation, Informality and Welfare: Redistribution Costs and Efficiency Gains

Paweł Doligalski University of Bristol

Luis E. Rojas * UAB and Barcelona SE

2024

We characterize the welfare effects of the informal sector by proposing a decomposition into efficiency and redistribution components. We focus on an economy where a planner wants to redistribute income with taxation and sets the optimal tax scheme. Since the informal sector can limit the taxation possibilities for the government but at the same time provide a shelter against tax distortions for individuals we show that the net welfare effect can be positive or negative. We show that the relative advantage between informal and formal employment across different income levels is the key dimension that shapes the welfare costs of the informal sector. Using the model estimated with Colombian microdata, we show that, conditional on the optimal tax policy, the Colombian shadow economy benefits efficiency at the expense of redistribution. Consequently, the presence of the informal sector reduces welfare only when preferences for redistribution are strong.

Keywords: shadow economy, informal labor market, income taxation, redistribution.

JEL Codes: H21, H26, J46.

*Paweł Doligalski: School of Economics, University of Bristol, The Priory Road Complex, Priory Road, BS8 1TU, Bristol, the United Kingdom, email: pawel.doligalski@bristol.ac.uk, tel: +44 117 954 6930. Luis E. Rojas: Department of Economics and Economic History, Autonomous University of Barcelona, Building B - Campus UAB, 08193 Cerdanyola del Vallés, Barcelona, Spain, email: luis.rojas@uab.cat, tel: (+34) 93 581 47 39 Ext. 4739. Luis Rojas acknowledges financial support from the European Research Council under the European Union's Horizon 2020 research and innovation program (788547-APMPAL-HET). We are grateful for useful comments of an anonymous referee, Árpád Ábrahám, Charles Brendon, Antoine Camous, Hal Cole, Mike Golosov, Piero Gottardi, Ramon Marimon, Wojciech Kopczuk, Dirk Krueger, Humberto Moreira, Erwin Ooghe, Wojciech Paczos, Evi Pappa, Dominik Sachs, Julia Schmieder, Jon Temple, Yanos Zylberberg. All errors remain our own.

1. Introduction

Informality significantly impacts developing economies, where, on average, the shadow sector represents 30% of output (Schneider, Buehn, and Montenegro, 2011) and a larger portion of employment. In Colombia, for example, approximately 55% of workers are employed in the informal sector.[1] This prevalence of informality is often perceived as an obstacle to development, undermining government capacity for revenue collection and investment. Recognizing its importance, the United Nations has prioritized reducing informality among its sustainable development goals (UN General Assembly, 2017). Despite these challenges, informality serves as a crucial survival strategy for the impoverished, offering an escape from tax burdens and potentially reducing the efficiency costs of taxation for vulnerable populations. Thus, the welfare implications of the informal sector are complex and multifaceted.

This paper introduces a theoretical framework to assess informality's welfare impact, employing a model calibrated with Colombian microdata. Our analysis concentrates on how informality influences the optimal scope for redistribution. While informality may seem to undermine social welfare by facilitating tax evasion and compromising redistribution efforts, our findings suggest a nuanced reality. In certain contexts, the informal sector may actually enable greater redistribution or reduce tax distortions, improving overall welfare.

We model an economy with heterogeneous agents differentiated by their productivity in both formal and informal sectors. Given the existing income tax structure, agents choose their sector of employment (extensive margin) and work hours (intensive margin). A benevolent government optimizes a nonlinear income tax to maximize social welfare. Formally, we assume that individual productivities and earnings from the informal sector are unobserved by the planner, while only earnings from the informal sector are obervable. Thus, the planner chooses the allocation to maximise the social objective, subject to the adequate incentive-compatiblity constraints. Our evaluation considers two scenarios: one where agents can engage in informal work and another where the informal sector is artificially eliminated. This approach allows us to isolate the welfare effects attributable to informality.

Using a simple model with two types, we show that the impact of informality on social welfare is ambiguous—depending on the productivity profiles, it can be either negative or positive. To shed light on the mechanisms behind this result, we decompose the impact of informality into two welfare-relevant dimensions: the degree of redistribution and the efficiency of labour supply. Informality generates a redistribution gain if it allows the government to decrease the tax on or to increase the transfer to the agents favoured by the social welfare function (presumably, the poor). Informality generates an efficiency

[1] Own calculations based on the 2013 wave of the household survey run by the official statistical agency of Colombia (DANE). Alternative measurements of informality in Colombia yield very similar numbers (see, e.g., ILO 2014).

gain if it leads to the overall lower amount of labour supply distortions in the economy. Informality can enhance efficiency by sheltering agents from distortionary tax rates or by making the government choose a less distortionary tax to start with.

The simple model allows us to derive analytical comparative statics of these two dimensions of social welfare. Suppose that the government wants to redistribute from the rich to the poor. We find that the redistribution gain is strictly decreasing in the informal productivity of the rich, and is strictly positive if the informal productivity of the rich is sufficiently low. In turn, the efficiency gain is strictly increasing in the informal productivity of the poor, and is strictly positive if the informal productivity of the poor is above a threshold (that lays below the formal productivity of the poor). Thus, if the informal employment is sufficiently attractive to the poor and sufficiently unattractive to the rich, it will improve social welfare.

Further, we explore a more comprehensive model featuring a continuum of agent types, costs associated with informal employment, and the possibility of dual-sector employment. This model, parameterized with Colombian microdata and solved using recent advancements in optimal taxation theory, allows for a sophisticated comparison between scenarios with and without informality. We discover that the informal sector in Colombia contributes positively to labour efficiency while having a negative effect on redistribution. Importantly the net effect depends on the extent of redistribution that the planner wants to implement. When the planner's redistributive preferences are weak, the shadow economy facilitates a welfare improvement comparable to a 1% increase in consumption. However, when the redistributive preferences are strong, the shadow economy negatively impacts welfare, with a welfare loss equivalent to about 3% decrease in consumption.

Related Literature.

Kopczuk (2001) studies optimal redistribution with tax avoidance and shows that avoidance is always detrimental to welfare if individuals have homogeneous avoidance opportunities. Conversely, he provides an example with heterogeneous avoidance opportunities where tax avoidance is socially beneficial. da Costa and Lobel (2022) investigate the welfare implications of the shadow economy using Brazilian data and the social preferences implicit in the actual income tax schedule in Brazil. Our study diverges from theirs by considering the possibility for agents to concurrently engage in both formal and informal sectors. We further dissect the welfare impact of informality into two distinct aspects: efficiency and redistribution. Moreover, we delve into how varying redistributive preferences influence the perceived costs and benefits associated with informality, offering a nuanced understanding of its role within an economy. Other studies, noting a potential role of an informal sector in reducing distortions, include Cowell (1985) and Davidson, Martin, and Wilson (2007).

Our analysis leverages the theoretical advancements in the optimal nonlinear income taxation in the presence of an informal sector, as developed by Doligalski and Rojas

(2023). Other papers studying optimal tax systems with evasion or avoidance opportunities focus on the role of consumption taxes (Emran and Stiglitz 2005; Boadway and Sato 2009; Huang and Rios 2016), tax thresholds (Kanbur and Keen 2014), income shifting (Piketty, Saez, and Stantcheva 2014; Selin and Simula 2020), nonlinear costs of evasion (Grochulski 2007; Casamatta 2023) or non-expected-utility approach to decisions under uncertainty (Piolatto and Trotin 2016). Kanbur, Keen, and Tuomala (1994) study nonlinear income taxation under a non-welfarist objective of minimising poverty. Simula and Trannoy (2010) study optimal income taxation under the threat of migration to a different country, which shares some similarities with moving to an informal sector.

2. Two-type Model and Welfare Decomposition

This section characterizes the welfare implications of the shadow economy by dissecting its effects into redistribution and efficiency gains. We employ a simplified model that facilitates an analytical exploration of the comparative statics associated with these welfare components. Our model posits an economy populated by two distinct worker types-denoted as L (low productivity) and H (high productivity) without the complexities of shadow employment costs or the feasibility of simultaneous formal and informal sector employment.

The population is partitioned into L and H types, with respective shares μ_L and $\mu_H = 1-\mu_L$. These individuals derive utility from consumption c and labor supply n, modeled through a quasilinear utility function $U(c,n) = c - v(n)$, where v is an increasing, strictly convex, and twice differentiable function with $v'(0) = 0$. Although our analysis adopts a linear utility perspective for consumption to simplify the exposition, extending these findings to scenarios with concave utility from consumption remains straightforward.

Each agent is endowed with two linear production functions corresponding to two labor markets, enabling formal (f) and informal (s) sector employment. An agent of type $i \in \{L, H\}$ working in sector $j \in \{f, s\}$ exhibits productivity w_i^j. Income generated in each sector is defined as $y_i^j = w_i^j n_i^j$, where n_i^j represents the labor supplied to sector j. We designate type H as having superior formal sector productivity: $w_H^f > w_L^f$. Additionally, we assume higher formal sector productivity for all types: $\forall_i \; w_i^f > w_i^s$, suggesting an inherent inefficiency in the shadow economy that precludes its use in optimal scenarios where individual types are observable. This assumption is revisited and relaxed in discussions of the full model.

2.1. The planner's problem

The social planner observes only the formal income of each individual. Furthermore, the planner can transfer resources between agents with taxes T_i. We can think about y_i^f

and $y_i^f - T_i$ as a pre-tax and an after-tax reported income. It is convenient to express agents' choices of shadow income as a function of their formal income:

$$y_i^s\left(y^f\right) \equiv w_i^s v'^{-1}(w_i^s) \times \mathbf{1}(y^f = 0). \tag{1}$$

If agents have any formal earnings, their shadow earnings are zero. If instead they have no formal earnings, they are unconstrained in choosing their shadow income. Given this function, we can specify agents' consumption $c_i = y_i^f + y_i^s\left(y_i^f\right) - T_i$ and labor supply $n_i = y_i^f/w_i^f + y_i^s\left(y_i^f\right)/w_i^s$, conditional on a truthful revelation of types.

The social planner maximizes the sum of utilities weighted with Pareto weights λ_i

$$W = \max_{\left\{\left(y_i^f, T_i\right) \in \mathbb{R}_+ \times \mathbb{R}\right\}_{i \in \{L,H\}}} \lambda_L \mu_L U\left(c_L, n_L\right) + \lambda_H \mu_H U\left(c_H, n_H\right) \tag{2}$$

subject to a resource constraint

$$\mu_L T_L + \mu_H T_H \geq 0, \tag{3}$$

and incentive-compatibility constraints

$$U\left(c_i, n_i\right) \geq U\left(y_{-i}^f + y_i^s\left(y_{-i}^f\right) - T_{-i}, \frac{y_{-i}^f}{w_i^f} + \frac{y_i^s\left(y_{-i}^f\right)}{w_i^s}\right) \qquad i \in \{H, L\}. \tag{4}$$

The incentive compatibility constraints capture the limited information available to the planner. They imply that no agent can be better off by choosing formal income of the other type and, if this income level is zero, freely adjusting shadow earnings.

Lemma 1. *Suppose that $\lambda_i > \lambda_{-i}$. In the optimum,*

- *type $-i$ faces no labor distortions and does not work in the shadow economy.*
- *type i faces labor distortions and may work in the shadow economy.*

Proof. See Appendix A.

□

Lemma 1 extends the classical *no distortion at the top* principle to our context. When $\lambda_i > \lambda_{-i}$, there's an incentive for the social planner to redistribute income from type $-i$ to type i. Consequently, the incentive constraint for type $-i$ becomes binding, and the planner cannot improve the social objective by distorting the labor supply of type $-i$. Since an agent will opt for the shadow economy employment only if their formal sector labor supply is distorted downwards, type $-i$ will invariably abstain from shadow economy participation in the optimum. Conversely, introducing distortions into type i's

labor decisions mitigates the binding nature of the incentive constraint, thereby facilitating greater redistributive measures. As such, there exists a scenario where type i might engage in the shadow economy.

Note that the shadow economy may be utilised in the optimum due to the agents' private information about their productivity. If individual productivities were publicly known, the planner would rely on non-distortionary, individualised lump-sum taxes. Given that all agents are more productive in the formal sector (an assumption we will relax in the next section), no one would then find it optimal to work informally.

2.2. Welfare decomposition

Suppose that $\lambda_i > \lambda_{-i}$, such that the planner wants to redistribute resources from type $-i$ to type i. There are two candidate allocations for the optimum: a *Mirrleesian allocation* in which type i works formally (denoted with superscript M) and a *shadow economy allocation* in which type i works informally (denoted with superscript SE). Note that the *Mirrleesian allocation* is also the optimum in the setting without the shadow economy. We examine the welfare **gain** from the existence of the shadow economy by comparing these two allocations.

Proposition 1. *Suppose that $\lambda_i > \lambda_{-i}$. The welfare difference between the shadow economy allocation and the Mirrleesian allocation can be decomposed in the following way*

$$\underbrace{W^{SE} - W^{M}}_{\text{welfare gain}} = \underbrace{\lambda_i \mu_i \left(U\left(w_i^s n_i^{SE}, n_i^{SE}\right) - U\left(w_i^f n_i^M, n_i^M\right)\right)}_{\text{efficiency gain}} + \underbrace{(\lambda_i - \lambda_{-i})\mu_i \left(T_i^M - T_i^{SE}\right)}_{\text{redistribution gain}},$$

where

- *the efficiency gain is increasing with w_i^s and is positive when $w_i^s > \bar{w}_i^s$,*
- *the redistribution gain is decreasing with w_{-i}^s and is positive when $w_{-i}^s < \bar{w}_{-i}^s$,*
- *the productivity thresholds satisfy $\bar{w}_i^s < w_i^f$ and $\bar{w}_{-i}^s < w_{-i}^f$.*

Proof. See Appendix A. □

Proposition 1 decomposes the welfare gain from the shadow economy into an *efficiency gain*, measuring the difference in distortions imposed on type i, and a *redistribution gain*, capturing the change in the level of transfers received by type i.

Efficiency gain. In the shadow economy allocation, type i supplies the efficient level of labor to the inefficient shadow sector (an extensive margin distortion). In the Mirrleesian allocation, due to the distortions imposed by the planner, type i supplies an inefficient amount of labor to the efficient formal sector (an intensive margin distortion). The

relative inefficiency of the shadow sector depends on the productivity difference $w_i^f - w_i^s$. When this difference is sufficiently small ($w_i^s > \bar{w}_i^s$), distortions in the shadow sector are smaller than distortions in the formal sector and the shadow economy improves the efficiency of labor allocation. Intuitively, in this case the shadow economy provides a shelter against tax distortions. If instead the shadow economy distortions are large ($w_i^s < \bar{w}_i^s$), the efficiency gain of the informal sector will be negative.

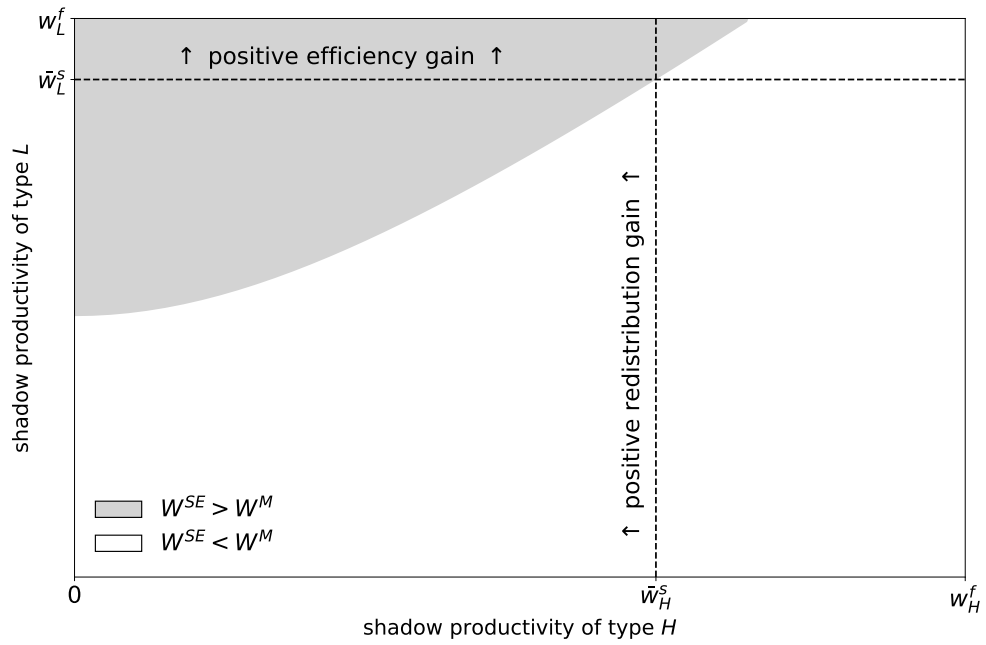
Redistribution gain. The shadow economy improves redistribution if the planner is able to provide type i with a higher transfer (or equivalently raise a higher tax from type $-i$). The scale of redistribution is determined by the payoff of type $-i$ from misreporting. In the Mirrleesian allocation the deviating worker works formally and can earn only as much as type i. In the shadow economy allocation the deviating worker cannot supply any formal labor, but is unconstrained in supplying shadow labor. As the shadow productivity of type $-i$ increases, the payoff from misreporting in the shadow economy allocation rises and the redistribution is reduced. On the other hand, when w_{-i}^s is sufficiently low ($w_{-i}^s < \bar{w}_{-i}^s$), the shadow economy deters the deviation of type $-i$, helping the planner to tell the two types of agents apart. In this case the informal sector is used as a screening device.

Proposition 1 is illustrated in Figure 1, where we assume that the planner maximizes the utility of type L: $\lambda_H = 0$. Intuitively, the shadow economy does not have to strengthen both redistribution and efficiency simultaneously to be welfare improving. Particularly interesting is the region where the redistribution gain is negative, but the efficiency gain is sufficiently high such that welfare is higher with the shadow economy. In this case the shadow economy allocation Pareto dominates the Mirrleesian allocation. Type L gains, since the welfare is higher with the shadow economy. Type H benefits as well, as the negative redistribution gain implies a lower tax burden on this type.

Kopczuk (2001) provides an example in which, starting from the allocation without tax evasion, a marginal increase in evasion yields welfare gains.[2] According to our decomposition, in his example welfare improves due to greater redistribution, but at the cost of efficiency. It may suggest that the shadow economy improves welfare by allowing for more even division of a smaller aggregate output. We show that such a scenario is only one of many possibilities. The shadow economy can reduce redistribution, while still being welfare-improving, in which case all agents benefit from the presence of the shadow economy.

[2] Kopczuk (2001) also presents a second example of welfare-improving tax evasion in which some agents have a distaste for paying taxes. We abstract from agents having preferences directly over tax payments.

Figure 1: Welfare impact of the shadow economy



3. Full Model and Quantitative Methodology

In this section, we present an extension of the model to quantitatively assess the significance of the informal economy on welfare, focusing on the gains related to efficiency and redistribution. This extension corresponds to the framework developed in Doligalski and Rojas (2023), which introduces comprehensive heterogeneity in productivity profiles and the costs associated with participation in the informal sector. There, we provide the optimal tax formula across various social welfare functions, utilizing Colombian data for empirical estimation.

Our analytical approach involves comparing the welfare outcomes under the optimal tax scheme—both in scenarios where the informal sector exists and in hypothetical situations where it is absent—with the optimal tax formula accordingly applied. To analyze the welfare impact, we employ our decomposition method across three distinct social welfare functions, each representing different redistributive preferences towards less productive individuals.

The empirical aspect of our analysis leverages the model calibrated with data from the Colombian economy, aiming to quantify the welfare implications of the informal sector under the implementation of an optimal tax scheme. This evaluation specifically considers how these welfare outcomes vary in relation to the redistributive priorities of the policy planner.

Following this introduction, we detail the model's structure, the derived tax formula, and our approach to welfare decomposition within the comprehensive framework. Sub-

sequently, we briefly discuss Doligalski and Rojas (2023) model's estimation process. The forthcoming section will then present our findings, shedding light on the impact of the informal sector on welfare.

3.1. Model Setup

The economy is populated by a continuum of heterogeneous workers that differ in their productivity $\theta \in [\underline{\theta}, \bar{\theta}] \subseteq \mathbb{R}_+$ and their participation cost in the informal sector $\kappa \geq 0$. There is no participation cost in the formal sector. Workers decide (i) whether to participate in the informal sector, which involves a fixed cost κ, and (ii) how much labour to allocate to each of the sectors they participate in, where the wages in the formal and the informal sector are functions of their productivity type and are given by $w^f(\theta)$ and $w^s(\theta)$, respectively. Note that workers have an option of working in both sectors simultaneously.

The formal sector income is taxed with a nonlinear income tax $T(y^f)$ that depends on formal income only. Shadow income is not observed by the tax authority. The worker has quasi-linear preferences in consumption and hours worked, and the problem they solve corresponds to

$$V(\theta, \kappa, T) = \max_{n^f \geq 0, n^s \geq 0} \; y^f - T\left(y^f\right) + y^s - v\left(n^f + n^s\right) - \kappa \cdot \mathbb{1}\left(n^s > 0\right), \qquad (5)$$
$$\text{where } y^j = w^j(\theta) n^j, \text{ for } j \in \{f, s\}.$$

Here, $v(\cdot)$ is assumed to be a convex function that captures the disutility of work, and the utility is linear in consumption that is given by the total after-tax income $y^f - T\left(y^f\right) + y^s$. The last term captures that if the worker decides to participate in the shadow sector, then they incur the cost κ. The function $V(\theta, \kappa, T)$ then represents the utility attained by a worker with productivity θ and participation cost κ under the tax scheme T.

3.2. Government's problem

The government decides on the nonlinear income tax schedule $T(y^f)$ to maximize a weighted utilitarian social welfare function subject to a fixed tax revenue requirement. The welfare objective is

$$W = \int_{\underline{\theta}}^{\bar{\theta}} \int_0^{\infty} \lambda(\theta) V(\theta, \kappa, T) dG_\theta(\kappa) dF(\theta), \qquad (6)$$

where $\lambda(\theta)$ signifies the Pareto weight for workers of productivity θ; $G_\theta(\kappa)$ represents the cumulative distribution of κ across productivity types θ, and $F(\theta)$ is the cumulative distribution function of θ.

We introduce a parameter r to embody the planner's redistributive inclinations, with Pareto weights defined as

$$\lambda(\theta) = r(1 - F(\theta))^{r-1}, \tag{7}$$

following the framework in Rothschild and Scheuer (2013). A r value of 1 implies equal weights across the population, while $r > 1$ accentuates the focus on lower productivity individuals, amplifying the redistributive objective.

Doligalski and Rojas (2023) tackle this optimization problem, deriving a sufficient statistics formula for the optimal tax schedule. This formula extends beyond the Diamond-Saez framework by incorporating elasticity of taxable income for workers engaged part-time in informal activities, and elasticity concerning extensive transitions out of formality.

In our analysis, W^{SE} denotes the welfare under the optimal tax scheme in the presence of an informal sector, and W^{M} represents the welfare in a hypothetical scenario devoid of the shadow economy, applying the corresponding optimal tax schedule. Here, $T(y)$ and $T_M(y)$ symbolize the optimal tax schedules in the presence and absence of the shadow economy, respectively.

3.3. Generalized welfare decomposition

We extend our decomposition in the following proposition:

Proposition 2. *First, consider an allocation from the model with an informal sector (y^f, T) with an associated schedule of shadow earnings y^s. Denote total income by $y(\theta,\kappa) = y^f(\theta,\kappa) + y^s(\theta,\kappa)$ and total labor supply by $n(\theta,\kappa) = \frac{y^f(\theta,\kappa)}{w^f(\theta)} + \frac{y^s(\theta,\kappa)}{w^s(\theta)}$. Second, consider the Mirrleesian allocation (y_M, T_M) and denote the Mirrleesian labor supply by $n_M(\theta) = \frac{y^M(\theta)}{w^f(\theta)}$. Define the welfare gain of the shadow economy as*

$$\begin{aligned}\mathcal{WG} = \int_{\underline{\theta}}^{\overline{\theta}} \int_0^{\infty} \lambda(\theta)[U(y(\theta,\kappa) - T(y^f(\theta,\kappa)), n(\theta,\kappa)) - \kappa \mathbb{1}_{y^s(\theta,\kappa)>0} \\ - U\left(y_M(\theta) - T_M(y_M(\theta)), n_M(\theta)\right)] dG_\theta(\kappa) dF(\theta).\end{aligned} \tag{8}$$

The welfare gain can be represented as $\mathcal{WG} = \mathcal{EG} + \mathcal{RG}$, where the efficiency gain $\mathcal{EG}$ and the redistributive gain $\mathcal{RG}$ are given by

$$\mathcal{EG} = \int_{\underline{\theta}}^{\overline{\theta}} \int_0^{\infty} \lambda(\theta) \left[U(y(\theta,\kappa), n(\theta,\kappa)) - \kappa \mathbb{1}_{y^s(\theta,\kappa)>0} - U\left(y_M(\theta), n_M(\theta)\right)\right] dG_\theta(\kappa) dF(\theta), \tag{9}$$

$$\mathcal{RG} = \int_{\underline{\theta}}^{\overline{\theta}} \int_0^{\infty} \lambda(\theta) \left[T_M(y_M(\theta)) - T(y^f(\theta,\kappa))\right] dG_\theta(\kappa) dF(\theta). \tag{10}$$

Proof. It follows from linearity of preferences with respect to consumption: $U(y-T,n) = U(y,n) - T$. The decomposition can be easily generalized to utility functions which are strictly concave in consumption and non-separable between consumption and labor supply.[3] □

The efficiency gain $\mathcal{EG}$ is the difference in the social welfare between the optimal and the Mirrleesian allocations if each agent where to consume her total income. Intuitively, it is a measure of social welfare before redistribution. It captures the influence of the informal sector on the allocation of labor supply, including the fixed cost of shadow employment. Suppose that some workers are more productive in the informal sector than in the formal sector and that for the others the marginal tax rates are lower than in the Mirrleesian economy. Then, as long as the aggregate fixed cost of shadow employment of non-formal workers is not too large, the efficiency gain will be positive. In that case the informal sector enhances the efficiency of labor supply. Alternatively, if the shadow productivity is relatively low and the fixed cost of shadow employment is high, the efficiency gain can be negative.

The redistributive gain $\mathcal{RG}$ is the difference in welfare-weighted taxes and transfers. It captures the influence of the informal sector on the optimal allocation of tax burden among workers. If in the absence of the informal sector the planner is able to reduce taxes for individuals with high Pareto weights (e.g. low productivity individuals) and increase taxes for individuals with low Pareto weights (e.g. high productivity individuals), then the redistributive gain will be negative. In that case the informal sector restricts redistribution. Alternatively, if the shadow economy allows the planner to raise transfers at low income levels, the redistributive gain can be positive.

3.4. Parametrization

For our empirical calibration, we adopt the parameter estimates from Doligalski and Rojas (2023), utilizing data from the Colombian household survey of 2013. This dataset enables distinction between formal and informal incomes and includes various individual and job characteristics. The model is estimated using maximum likelihood, where key assumptions include a log-linear relationship between individual characteristics and sector-specific wages, and the distribution of κ is modeled using a generalized Pareto distribution. This methodology accounts for self-selection into sectors without imposing

[3] Suppose that preferences over consumption and labor supply are given by $U(c,n)$, where $U_{cc} \leq 0$ and U_{cn} can be non-zero. Define the redistributive gain as below, the efficiency gain is then given by the difference between the welfare gain and the redistributive gain.

$$\mathcal{RG} = \int_{\underline{\theta}}^{\bar{\theta}} \int_0^{\infty} \lambda(\theta) \Big\{ U\left(y(\theta,\kappa) - T(y^f(\theta,\kappa)), n(\theta,\kappa)\right) - U\left(y(\theta,\kappa), n(\theta,\kappa)\right) \\ - \left[U\left(y_M(\theta) - T(y_M(\theta)), n_M(\theta)\right) - U\left(y_M(\theta), n_M(\theta)\right)\right] \Big\} dG_\theta(\kappa) dF(\theta). \quad (11)$$

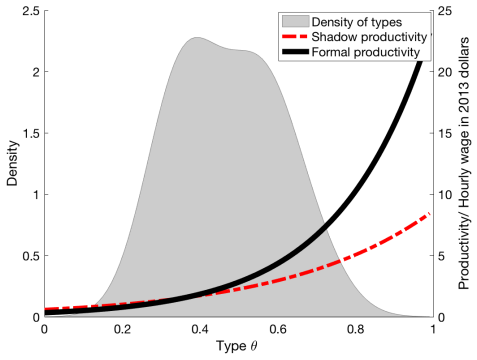
restrictions on the productivity distribution. Figure 2 shows the estimated productivity profiles and the density of productivity types in Doligalski and Rojas (2023).

The density of productivity types (θ) is approximately normal, excluding the tails, resulting in sectoral wages being distributed roughly log-normally with a Pareto tail. Central to our analysis is the relative productivity between shadow and formal work across the population. According to our calibration, the bottom quarter of workers are more productive in the shadow sector, while the median worker is 16% more productive in the formal sector. For the 90th percentile worker, productivity is 40% higher in the formal sector.

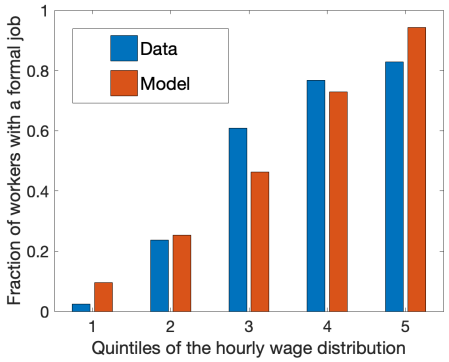
This widening of the relative advantage across sectors is depicted in Figure 2. Low productivity individuals face minimal or no productivity loss by participating in the shadow sector, while this loss increases with productivity. For each percentage point increase in the type distribution, shadow productivity increases on average by 1.2%, whereas formal productivity increases by 1.9%.

Figure 2: Estimation results

(a) Productivity profiles and type distribution



(b) Formality and hourly wages



(a) Kernel density estimate of the productivity type distribution obtained from the observed $X\beta$ in our sample (left axis). (b) Fraction of workers that are have formal employment at each quintile of the hourly wage distribution in the calibrated model and the data.

We want to examine the impact of the shadow economy under different assumptions about the government's redistributive preferences (controlled by the parameter r). To that end, we consider three scenarios that span the varying degree of government's concern for inequality, ranging from weak to strong concern. The scenarios are: *(i)* weakly redistributive, ($r = 1.1$), *(ii)* moderately redistributive ($r = 1.4$), and *(iii)* strongly redistributive ($r = 1.8$). To illustrate the three scenarios, let's compare the Pareto weights of the bottom type $\underline{\theta}$ and the type in the 90th percentile of the productivity distribution. In the weakly redistributive case, the Pareto weight of the bottom type is 26% higher than the weight of an agent in the 90th percentile of θ. In a moderately and strongly redistributive cases, the Pareto weight of the bottom type is 150% and 530% higher, respectively.

4. Results

This section quantifies the welfare impact of the informal sector by comparing the optimal allocation in our calibrated model with a counterfactual scenario where the informal sector is absent — termed the *Mirrleesian* economy.

In the left column of Figure 3, we present a comparative analysis of the optimal tax schedules in both economies. The Mirrleesian economy is characterized by sharply increasing marginal tax rates at lower income levels. At higher income levels the tax rates level out and, in more redistributive cases, exhibit a U-shape, whereby they first fall and then rise again. Conversely, in scenarios incorporating the informal sector, tax rates remain relatively stable at lower income brackets, escalating only near the median income level (approximately \$10,000). The tax schedules for economies with weak redistributive preferences align above this income threshold. With strong redistributive preferences, the informal sector's presence generally lowers tax rates across a broad income spectrum, excluding the highest earners.

Our model, calibrated with Colombian data, illustrates the nuanced welfare considerations introduced by informality. We observe a positive efficiency gain and a negative redistributive impact from the Colombian informal sector (refer to the right column of Figure 3). Under weak redistributive preferences, the informal sector's net effect is welfare-enhancing, equating to a 1% uptick in consumption. With moderately strong preferences, this welfare gain moderates to 0.25%. Conversely, strong redistributive preferences invert this effect, culminating in a welfare reduction equivalent to a 3% decrease in consumption.

The analysis, detailed in Table 1, elucidates that the informal sector bolsters efficiency but dampens redistributive potential. This dichotomy stems from two primary factors: first, the least productive individuals exhibit higher productivity in the informal versus the formal sector, amplifying efficiency gains. Second, the informal sector's influence on reducing tax rates contributes to lower labor distortions in the formal sector, enhancing efficiency but simultaneously curtailing redistributive capacity. Notably, the adverse impact on redistribution becomes particularly pronounced under strong redistributive preferences, where the shortfall in redistribution outweighs efficiency benefits.

Delving deeper, Figure 4 dissects these gains across different levels of productivity and the fixed cost of informality for each social preference scenario, revealing substantial heterogeneity among the population. Focusing first on the redistribution gain, we observe that low productivity individuals lose and high productivity individuals gain from the presence of the informal sector due to lower income redistribution. Both the magnitude of loses and gains, as well as the proportion of agents losing, increase with the strength of the government's redistributive preferences. Furthermore, the redistribution gain is particularly high for high productivity individuals with low fixed cost of informality (the right column), who further decrease their tax bill by working partially in the informal

Table 1: Welfare impact of the shadow economy

social preferences	efficiency gain	redistribution gain	(total) welfare gain
weakly redistributive	1.06%	-0.07%	0.99%
moderately redistributive	4.22%	-3.96%	0.25%
strongly redistributive	7.84%	-10.78%	-2.93%

The welfare gain is expressed as a proportional change of consumption in the Mirrleesian allocation required to reach the value of the social welfare function in the shadow economy allocation.

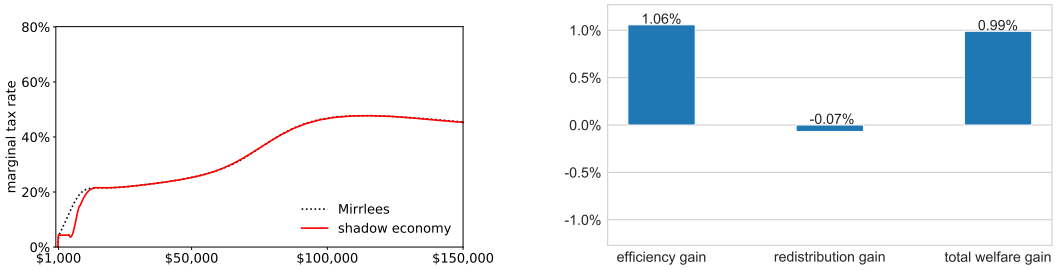
sector.

Consider now the efficiency gain. Agents with high fixed cost of informality who work only in the formal sector (the left column) experience small efficiency gains from the informal sector due to the lower marginal tax rates and, thus, lower labour supply distortions. In comparison, agents who face no fixed cost of working informally (right column) are affected on the efficiency dimension not only via the marginal tax rates in the formal sector, but also via their productivity in the informal sector. Among them, low productivity agents are relatively more productive informally and experience additional efficiency gains, while high productivity agents are less productive informally and, thus, suffer efficiency losses.

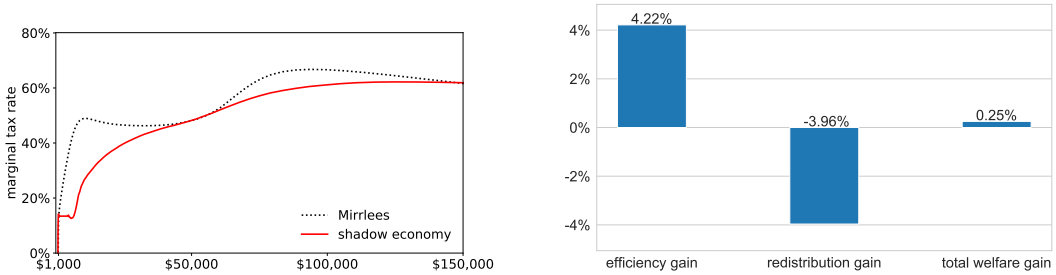
Finally, consider the total impact of the shadow economy on agents' utilities. When social preferences for redistribution are weak, both low productivity agents who can engage in informal activities and all high productivity agents gain from the existence of the informal sector. As a result, the shadow economy is, overall, welfare-improving. However, as the strength of the redistributive motive increases, the losses of low productivity types due to a restricted income redistribution become more severe, implying that all low productivity types lose because of the informal sector. Since the social welfare function places a higher weight on these individuals, the shadow economy becomes less beneficial (moderately redistributive scenario) or harmful (strongly redistributive scenario) for social welfare.

Figure 3: Optimal tax schedules with and without a shadow economy

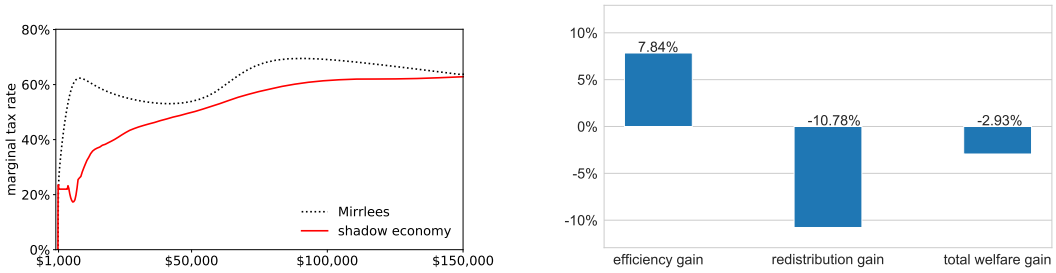
(a) Weakly redistributive social preferences



(b) Moderately redistributive social preferences



(c) Strongly redistributive social preferences

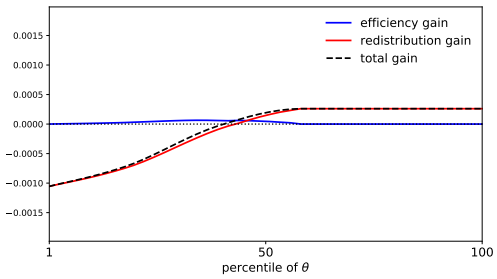


'Mirrlees' stands for a Mirrleesian tax schedule in the model without a shadow economy. 'Optimal' stand for the optimal tax schedules in the model with a shadow economy. With moderately redistributive social preferences and the optimal tax schedule the 50th, 95th and 99th percentiles of formal income in the model with a shadow economy are approx. $10,000, $41,000 and $79,000, respectively.

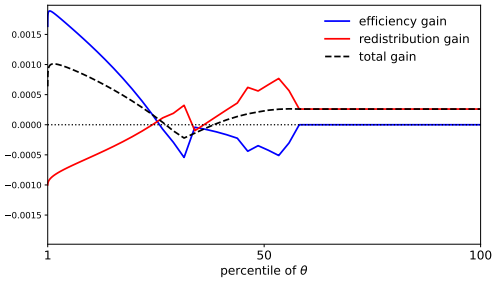
Figure 4: Efficiency and redistribution gains across the population

(a) Weakly redistributive social preferences

High fixed cost of informality

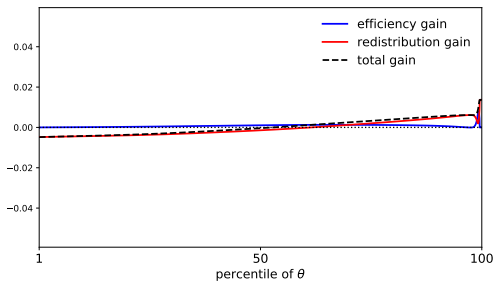


No fixed cost of informality

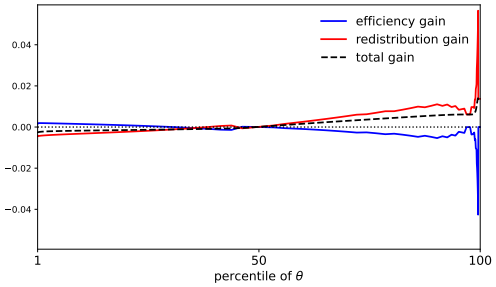


(b) Moderately redistributive social preferences

High fixed cost of informality

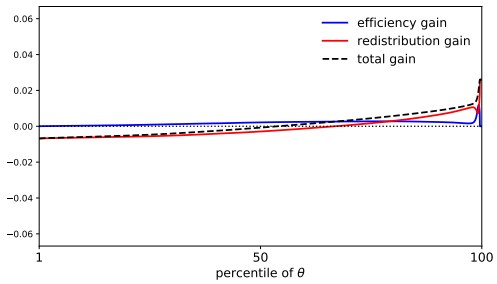


No fixed cost of informality

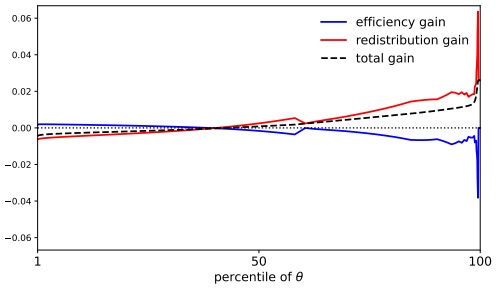


(c) Strongly redistributive social preferences

High fixed cost of informality



No fixed cost of informality



Note: Welfare decomposition for each productivity type θ and a fixed cost of informality κ that is either prohibitively high (the left column) or zero (the right column).

5. Policy Implications

The standard view is that the shadow economy harms welfare by reducing tax capacity and limiting the government's ability to redistribute. However, our analysis shows that this view is incomplete. The shadow sector also provides efficiency gains by allowing low productivity individuals to avoid tax distortions. These efficiency gains can offset the redistribution costs, which are mainly imposed by high productivity individuals who use the shadow sector to evade taxes.

The dominance of these effects depends on the participation costs in the shadow economy and the government's preferences regarding inequality. Low productivity individuals benefit from the shadow sector by avoiding tax distortions, aligning with the planner's redistribution goals. Conversely, high productivity individuals exploit the shadow sector to evade taxes, thereby limiting the government's ability to redistribute resources effectively.

Our analysis allows for a reevaluation of policy prescriptions, such as the UNâs proposed objective to reduce informality as a means to reduce inequality. It is not informality in general that can hurt inequality; rather, our paper provides a more nuanced understanding of the role of the informal sector in the welfare implications of optimal redistribution policy. The possibility of informality among high productivity workers harms the ability to redistribute and finance public expenditures, and should thus be a priority in policy considerations. In contrast, tax avoidance by low productivity individuals can even lead to higher welfare from a redistributive perspective.

Extending this view to tax enforcement or audit design, efforts should aim to make it more costly for high productivity workers to participate in the shadow sector. For example, penalties could be highly progressive concerning the amount of income concealed in the shadow sector and could further increase based on any signals of high productivity, such as formal income.

Our quantitative exercise with the model calibrated to Colombian data illustrates how the balance between the gains and costs of the informal sector can shift depending on the strength of a government's preferences for redistribution. For a government with weak redistributive desires, the presence of the shadow economy can lead to better outcomes, characterized by minimal redistribution costs and significant efficiency gains. However, as the desire to redistribute increases, the efficiency gains are overshadowed by the redistribution costs.

Regarding tax enforcement and audits, if targeting enforcement across the productivity distribution is costly, then for a government with ambitious redistribution objectives, enforcement is highly complementary to a progressive tax scheme. Conversely, for a government with more moderate redistribution objectives, low enforcement, even if it has low cost, can be optimal.

This nuanced approach suggests that policy measures should be tailored to the specific redistributive goals of the government. High enforcement costs may necessitate targeted strategies that maximize compliance among high productivity individuals, while more moderate goals might be achieved with a less aggressive enforcement stance. This flexibility allows for a more effective balance between efficiency and equity in tax policy.

Finally, our analysis reveals that the relationship between the size of the shadow economy and redistribution capacity is more complex than it appears. The impact of the shadow economy on redistribution is not solely a function of its size but also of the demographic and economic characteristics of those who participate in it. Understanding which sectors of the population are overrepresented in the shadow sector is crucial to grasping the extent to which it limits redistribution.

Studies have shown that the composition of the shadow economy varies widely across countries and even within regions of a country. For instance, Schneider and Enste (2000) highlight that in many developing countries, the shadow economy often includes a large proportion of low-skilled workers, whereas in some developed countries, high-skilled professionals tend to engage more in informal activities. According to our analysis, it is the latter scenario–where high-skilled professionals are more involved in the shadow economy–that poses a greater limitation to redistribution.

An important caveat is that this discussion focuses on the taxation aspect of informality. It does not address potential regulatory avoidance, such as food safety standards, which could increase the costs of informality, or dynamic concerns related to differences in human capital acquisition between the informal and formal markets.

6. Conclusions

This work delves into the welfare implications of the informal sector, introducing a novel approach to decomposing the change in social welfare into *efficiency gain* and *redistributive gain*. Traditional methods, such as evaluating the social welfare function or examining the Pareto frontier, conflate efficiency with redistribution, obscuring the specific drivers of welfare changes. Our decomposition method clarifies this ambiguity, distinguishing between welfare improvements due to more efficient labor allocation and those arising from equitable tax redistribution.

Applied to the Colombian context, our findings reveal that the shadow economy strengthens efficiency of labor supply at the expense of possible redistribution. When preferences for redistribution are weak, the former channel dominates and the existence of the shadow economy is welfare improving. These results highlight the non-trivial welfare implications of informality.

Looking forward, our analysis opens several pathways for extending this research. First, considering government interventions like audits and penalties could refine our under-

standing of how differentiated enforcement strategies influence the shadow economy's welfare effects. Such measures could potentially align tax evasion opportunities with welfare optimization rather than merely curbing evasion (Cremer and Gahvari, 1996). Moreover, our welfare decomposition technique can illuminate the efficiency and redistributive outcomes of various economic changes, from structural shifts to educational advancements, independent of informality's context.

References

Boadway, R., and M. Sato (2009): "Optimal tax design and enforcement with an informal sector," *American Economic Journal: Economic Policy*, 1(1), 1–27.

Casamatta, G. (2023): "Optimal income taxation with tax avoidance and endogenous labour supply," *Canadian Journal of Economics/Revue canadienne d'économique*, 56(3), 913–939.

Cowell, F. A. (1985): "THE ECONOMIC ANALYSIS OF TAX EVASION.," *Bulletin of Economic Research*, 37(3).

Cremer, H., and F. Gahvari (1996): "Tax evasion and the optimum general income tax," *Journal of Public Economics*, 60(2), 235–249.

da Costa, C. E., and F. Lobel (2022): "The Cost of Informality: An Optimal Taxation Approach," *Available at SSRN 4231413*.

Davidson, C., L. Martin, and J. D. Wilson (2007): "Efficient black markets?," *Journal of Public Economics*, 91(7-8), 1575–1590.

Doligalski, P., and L. E. Rojas (2023): "Optimal redistribution with a shadow economy," *Theoretical Economics*, 18(2), 749–791.

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

Grochulski, B. (2007): "Optimal nonlinear income taxation with costly tax avoidance," *FRB Richmond Economic Quarterly*, 93(1), 77–109.

Huang, J., and J. Rios (2016): "Optimal tax mix with income tax non-compliance," *Journal of Public Economics*, 144, 52–63.

ILO (2014): "Trends in informal employment in Colombia: 2009 - 2013," Discussion paper, International Labour Organization.

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

Kanbur, R., M. Keen, and M. Tuomala (1994): "Optimal non-linear income taxation for the alleviation of income-poverty," *European Economic Review*, 38(8), 1613–1632.

Kopczuk, W. (2001): "Redistribution when avoidance behavior is heterogeneous," *Journal of Public Economics*, 81(1), 51–71.

Piketty, T., E. Saez, and S. Stantcheva (2014): "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities," *American Economic Journal: Economic Policy*, 6(1), 230–271.

Piolatto, A., and G. Trotin (2016): "Optimal income tax enforcement under prospect theory," *Journal of Public Economic Theory*, 18(1), 29–41.

Rothschild, C., and F. Scheuer (2013): "Redistributive taxation in the Roy model," *The Quarterly Journal of Economics*, 128(2), 623–668.

Schneider, F., A. Buehn, and C. E. Montenegro (2011): "Shadow Economies all over the World: New Estimates for 162 Countries from 1999 to 2007," in *Handbook on the shadow economy*, ed. by F. Schneider, pp. 9–77. Edward Elgar Cheltenham.

Schneider, F., and D. H. Enste (2000): "Shadow economies: Size, causes, and consequences," *Journal of Economic Literature*, 38(1), 77.

Selin, H., and L. Simula (2020): "Income shifting as income creation?," *Journal of Public Economics*, 182, 104081.

Simula, L., and A. Trannoy (2010): "Optimal income tax under the threat of migration by top-income earners," *Journal of public Economics*, 94(1-2), 163–173.

UN General Assembly (2017): "Work of the Statistical Commission Pertaining to the 2030 Agenda for Sustainable Development (a/res/71/313)," *UN General Assembly: New York, NY, USA*.

A. Proofs from Section 2

Proof of Lemma 1. The planner can increase social welfare by transferring consumption from type $-i$ to type i, so at the optimum the incentive constraint of $-i$ will binds and the incentive constraint of i will be slack. Denote the undistorted level of formal income of type $-i$ by $y_{-i}^{f*} \equiv w_{-i}^f v'^{-1}(w_{-i}^f)$. If $y_{-i}^f \neq y_{-i}^{f*}$, the planner can extract more resources without violating the incentive constraint by setting $y_{i-}^f = y_{-i}^{f*}$ and increasing T_{-i} to keep the utility level of type $-i$ constant. Since $y_{-i}^{f*} > 0$, type $-i$ will not work in the shadow economy.

To see that the planner optimally distorts the labor supply of type i, notice that a marginal adjustment of y_i^f, starting from the undistorted level y_i^{f*}, has no direct impact on welfare of type i by the Envelope Theorem. However, the distortion in a correct direction will reduce the payoff of $-i$ from misreporting, relax the incentive constraint and, hence, allow for more redistribution. In particular, if $w_i^f < w_{-i}^f$ ($w_i^f > w_{-i}^f$), a marginal decrease (increase) of y_i^f will relax the incentive constraint. □

Proof of Proposition 1. The difference in the utility level of type i between the two allocations is

$$U\left(c_i^{SE}, n_i^{SE}\right) - U\left(c_i^M, n_i^M\right) = U\left(w_i^s n_i^{SE}, n_i^{SE}\right) - U\left(w_i^f n_i^M, n_i^M\right) + T_i^M - T_i^{SE}. \quad (12)$$

The difference in utility level of type $-i$ is

$$U\left(c_{-i}^{SE}, n_{-i}^{SE}\right) - U\left(c_{-i}^M, n_{-i}^M\right) = T_{-i}^M - T_{-i}^{SE} = -\frac{\mu_i}{\mu_{-i}}\left(T_i^M - T_i^{SE}\right), \quad (13)$$

where the first equality follows from Lemma 1, since in the two allocations the labor supply of $-i$ is undistorted, and the second equality follows from the resource constraint. Using both utility differences, we can decompose $W^{SE} - W^M$ as stated in the proposition.

Define a function $\Psi(w) \equiv U\left(wv'^{-1}(w), v'^{-1}(w)\right)$, equal to the utility level of an individual with productivity w who supplies labor efficiently and receives no transfers. The efficiency term can be restated as $\lambda_i \mu_i \left(\Psi(w_i^s) - U\left(w_i^f n_i^M, n_i^M\right)\right)$. Since Ψ is an increasing function, the efficiency gain is increasing in w_i^s and changes sign at $\bar{w}_i^s \equiv \Psi^{-1}\left(U\left(w_i^f n_i^M, n_i^M\right)\right)$. To see that $\bar{w}_i^s < w_i^f$, note that since n_i^M is distorted, $\Psi(w_i^f) > U\left(w_i^f n_i^M, n_i^M\right)$.

To characterize the redistribution term, note that, due to the binding incentive constraints, we have

$$U\left(c_{-i}^{SE}, n_{-i}^{SE}\right) - U\left(c_{-i}^M, n_{-i}^M\right) = \Psi(w_{-i}^s) - T_i^{SE} - U\left(w_i^f n_i^M, w_i^f n_i^M / w_{-i}^f\right) + T_i^M. \quad (14)$$

Combining it with (13), we find that $T_i^M - T_i^{SE} = \mu_{-i}\left(U\left(w_i^f n_i^M, w_i^f n_i^M / w_{-i}^f\right) - \Psi(w_{-i}^s)\right)$. It implies that the redistribution term is decreasing in w_{-i}^s and changes sign at $\bar{w}_{-i}^s \equiv \Psi^{-1}\left(U\left(w_i^f n_i^M, w_i^f n_i^M / w_{-i}^f\right)\right)$. $\bar{w}_{-i}^s < w_{-i}^f$ holds, since $U\left(w_i^f n_i^M, w_i^f n_i^M / w_{-i}^f\right) < \Psi(w_{-i}^f)$ due to the optimal distortion of n_i^M. □