Taxation and Informality

Felipe Lobel

January 30, 2017

Abstract

This paper studies an optimal non-linear tax problem in the presence of informality. Empirical evidence suggests that informality is costly and this cost increases in the worker’s skill. I show that, in this case, it is possible to improve welfare by limiting access to the formal sector. I also consider a framework in which workers of the same skill may face different costs. I derive a modified Mirrlees Formula for an optimal non-linear income tax schedule in the presence of informality that accounts for the effects of this additional heterogeneity. I prove that, in this richer model, the second best requires positive sales taxes. This demonstrates that Atkinson-Stiglitz theorem fails when workers can become informal.

1 Introduction

Informality is a broad phenomenon that plagues many economies worldwide. To avoid high levels of taxation and burdensome government regulation, many economic agents flee the formal job market and take refuge in sectors that are outside the government’s reach. This article intends to look upon informality from the viewpoint of optimal tax theory.

Individuals make decisions along two margins. The labor supply operates on the intensive margin, while the sectorial choice operates on the extensive margin. Income taxation affects the intensive margin according to the elasticity of labor supply. The individual extensive margin choice is based on the trade-off between tax avoidance and informality’s participation cost, which has a diverse nature - it includes lack of social security, working environment, etc. An alternative approach is to consider a productivity loss associated to informal jobs, instead of participation costs. These approaches are equivalent, but the former is chosen because it dialogues with Taxation and Migration literature.

The goal of this article is to investigate the role of taxation on the potential evasion effects towards informality. I am interested in how the extensive choices affect social welfare. For a given tax schedule, the extensive margin choices depend on the compared indirect utilities across sectors. I show that individual choices not necessarily match the optimal social choices.

To tackle these issues I develop two frameworks. First, I assume homogeneous participation cost, which means that individuals with the same productivity always incur in the same informality’s participation cost. Furthermore, as highlighted by Ulyssea (2010), "as we move from the bottom to the top of the skill distribution, the rate of change in the utility derived from a formal job must be higher than the one from informal jobs". Using this finding, it is shown that a social welfare improvement is possible by limiting access to formal jobs.
Second, to represent informalization process in a realistic way, I consider that participation cost may differ across equally skilled workers, according to a generic density function. This feature allows more heterogeneity of types across sectors. Solving the planner’s problem, I derive the optimal marginal income tax rates for each type, extending the Diamond-Saez formula. This extended formula informs the role played by evasion effect in a tax reform, casting light on how optimal nonlinear income tax schedule differs from Mirrlees (1971) and Diamond (1998) in the presence of informality.

A major contribution of this article is that Atkinson-Stiglitz theorem fails when workers can possibly move to informality. It is always feasible for the government to eliminate informality, for instance, setting up a zero income tax for all skill types. I show that is never optimal to eliminate it completely, and this result is a sufficient condition for the failure of Atkinson-Stiglitz theorem, according to which, when preferences over consumption and leisure are separable (as they are in my model) then second-best can be implemented with zero sales taxes. Taxation problem becomes particularly appealing when we ponder informality, because the planner have very limited information about informal participants. As the planner can observe only income earned by formal workers, if she screens agents using only nonlinear income taxation, as suggested by Atkinson-Stiglitz results, she is disregarding information about informal workers that could be obtained through other policy instruments, such as sales taxes. Then, nonlinear income taxation by itself is not able to implement a second-best solution. Differently from income tax, sales taxes can be collected from informal workers, as well. So, in order to design the optimal mechanism, sales taxes constitute a valuable device for the planner. For this reason, when workers can possibly move to informality, optimal tax schedule should combine income and sales taxes.

Sales tax play an important role, not only because it allows taxing informal workers, but also because it makes informality less attractive as a refugee for income tax evasion. As I show in the Modified Mirrlees Formula, there is an evasion effect taking place after an income tax reform. Therefore, sales taxes soften this distorting effect that plagues nonlinear optimal income taxation. This discussion seems appealing because sales tax is an instrument widely used by governments in developed and developing economies, but most of the literature set sales taxes aside and focus on income tax schedule.

Related Literature. This paper is related to the literature on Taxation Under the Threat of Migration and Multisectorial Taxation. Lehmann et al. (2013) investigates in what respects potential migrations across countries affect the nonlinear income tax schedules. Although, they also deal with an extensive margin choice affecting optimal taxation, there are substantial differences from our focus. By the planner’s perspective, an individual that seeks for an informal job based on tax-driven forces is different from an agent that leaves the country. The former worker is still contemplated by the social welfare function. So it seems that the social justice criteria more appropriate for this case is a concave utilitarian planner, instead of Rawlsian. Empirical evidence brings up another distinction from the migration literature. While Lehmann et al. (2013) suggests that skilled workers are more willing to migrate, Ulyssea (2010) points that qualified agents are less susceptible in moving to informality. These distinctions on the problems’ essence lead to variations on the tax scheme design.
Rothschild and Scheuer (2012) and Gomes, Lozachmeur, and Pavan (Gomes et al.) study the design of optimal tax systems in a Roy model, where productivity is sector-specific. Their approach provides us insights to deal with taxation and endogenous sector choice. However, our model does not focus on sector-specific productivity, which rules out general equilibrium effects that play an important role on this referred literature. Even so, resembling results are found. For instance, Gomes, Lozachmeur, and Pavan (Gomes et al.) show that optimal multisectorial tax schedules imply failure of the Atkinson-Stiglitz theorem. The reason why this theorem does not hold in their economy is that when occupational choice is endogenous, the informational costs of redistribution are also endogenous and can be effectively manipulated by differential taxation. Taking informality into account, I also contrast Atkinson-Stiglitz, but for another driven force, which is the lack of informal workers information provided by an income based tax schedule.

2 Model

2.1 Homogeneous Cost

I consider a continuum of agents and two sectors indexed by $j \in \{F, I\}$, where F extends for the formal sector and I extends for the informal sector. Each agent chooses in which sector to work - extensive margin choice - and how many hours to supply in the chosen sector - intensive margin choice. It is not possible for an individual to work part time in each sector. Once one sector is chosen, all the hours supplied will be restricted to that sector. Workers have the same productivity in both sectors, so an agent’s type is given by $\theta_j = \theta_F = \theta_I = \theta \in [\hat{\theta}; \bar{\theta}]$. The types are private information, but the cumulative distribution function of $\theta$, denoted $F(\cdot)$, is common knowledge. It admits a continuous and strictly positive density $f$.

The planner can observe the formal workers’ income. So, the nonlinear income taxation $T(\cdot)$ is based on the $y_F(\theta)$. The planner cannot observe the informal workers’ income, so it is not possible to tax the informal earnings. I assume quasilinear preferences:

$$U(x_j, y_j, \theta) = x_j - \psi(y_F)$$

where $x_j$ is the consumption and $\psi(\cdot)$ is the disutility in labor supply. When $j = F$, the consumption is given by $x_F = y_F(\theta) - T(y_F(\theta))$. Analogously, when $j = I$, then $x_I = y_I(\theta)$.

As usual, this disutility is convex and isoelastic, given by:

$$\psi(I) = \frac{I^{1+\frac{1}{e}}}{1 + \frac{1}{e}}$$

where $e$ is the elasticity of labor supply.

The income $y_j$ comes from the individual problem that differs according to the sector chosen. Formal agents face:

$$\max_{y_F} y_F(\theta) - T(y(\theta)) - \psi\left(\frac{y_F(\theta)}{\theta}\right) \equiv V_F(\theta)$$

While, informal workers solve:
\[
\max_{y_i} y_i(\theta) - \phi \left( \frac{y_i(\theta)}{\theta} \right) 
\equiv V_I(\theta)
\]

There is a cost of participating in the informal sector, and in this first model the cost is type dependent. For a given productivity \( \theta \), the cost is given by \( c(\theta) > 0 \). Hence, all workers with skill \( \theta \) will have the same intensive and extensive choices. Based on Ulyssea (2010) and contrasting with Lehmann et al. (2013), when \( \theta \) increases the cost increases more than the informal sector’s indirect utility, which means \( c'(\theta) > V_I'(\theta) \). Note that \( V_I(\theta) \) is exogenous, so I am not making an assumption over endogenous variables.

As productivity is private information it is needed incentive compatible allocations. As usual, incentive compatibility is equivalent to envelope theorem and monotonicity:

\[
\dot{V}_F(\theta) = \frac{y(\theta)}{\theta^2} \psi' \left( \frac{y(\theta)}{\theta} \right) > 0
\]

\( y(\theta'') > y(\theta'), \theta'' > \theta' \)

**Definition 1.** The location rent of a \( \theta \)-individual is the excess of his formal sector indirect utility over his informal indirect utility, net of participation cost.

\[
R(\theta) = V_F(\theta) - \left[ V_I(\theta) - c(\theta) \right]
\]

(1)

An individual stays in the formal sector voluntarily if and only if \( R(\theta) \geq 0 \). She thus becomes informal when \( R(\theta) < 0 \).

**Lemma 1.** The location rent is increasing with the productivity.

**Proof.** This proof is straight forward because differentiating totally 1, it follows:

\[
R'(\theta) = V_F'(\theta) - V_I'(\theta) + c'(\theta)
\]

Envelope theorem assures that \( V_F'(\theta) > 0 \) and by hypothesis, \( c'(\theta) - V_I'(\theta) > 0 \). So, \( R'(\theta) > 0, \forall \theta \in [\underline{\theta}; \bar{\theta}] \).

In other words, more productive workers tend to be formal and there is at most one type that is indifferent across the two sectors.

**Definition 2.** For a given tax schedule, the natural threshold is a type \( \theta^*(T) \) which is indifferent between the two sectors.

By Definition 1, \( R(\theta^*) = 0 \), so all types above \( \theta^* \) chooses to be formal and types below \( \theta^* \) prefer to be informal.

**2.2 Heterogeneous Cost**

The previous model provides interesting insights, but by the other hand suggests a clear cut-off across sectors, which does not seem much realistic. In order to allow more heterogeneity across
sectors, I consider heterogeneous participation cost. The general set up is identical to the model already presented, but in this second model, each worker is bidimensional. One dimension is the skill $\theta \in [\underline{\theta}, \bar{\theta}]$ and the second dimension is the participation cost $s \in \mathbb{R}^+$ that she supports if decides to become informal. I do not make any assumption on the correlation between skills and migration costs. Differently from the previous model, I am not assuming that the cost increases with the skill, or any similar hypothesis in this way. I consider for a fixed skill, a generic density function for the migration costs, which may vary from $[s; \bar{s}]$. As in the first model, the type is private information and the government observes only the income.

Contrasting with the previous model, there is not anymore a clear threshold that segregates the economy between formal and informal workers. What comes up instead, is a skill dependent threshold function $c(\theta)$. It means that for each skill level $\theta$ there is a cost given by $c(\theta)$ that makes the individual indifferent between the two sectors. This threshold is the smallest cost that keeps the individual in the formal sector. Since the cost can vary from $s$ to $\bar{s}$, this model allows more heterogeneity of types across sectors. The cumulative distribution function $G(c(\theta) | \theta) = Pr(s \leq c(\theta) | \theta)$ captures the probability of an individual with skill $\theta$ to become informal.

3 Optimal Extensive Margin Choices

3.1 Homogeneous Model

In the first model, we have seen that there is a type $\theta^*$, which is a clear threshold between the two sectors. For a given type $\theta$, the individual extensive margin choice depends only if $\theta$ is below or above $\theta^*$. But suppose for a moment that the planner could implements a policy that limits access to the formal sector. Would that be socially optimal? To tackle this question, the planner should chooses an optimal social threshold $\tilde{\theta}$ that solves the following problem:

$$\max_{V_F, y, \tilde{\theta}} \int_{\underline{\theta}}^{\tilde{\theta}} \phi(V_I(\theta) - c(\theta)) f(\theta) d\theta + \int_{\tilde{\theta}}^{\bar{\theta}} \phi(V_F(\theta)) f(\theta) d\theta$$

s.a. : $\int_{\tilde{\theta}}^{\bar{\theta}} (y(\theta) - (V(\theta) + \psi(y(\theta)/\theta)) f(\theta) d\theta \geq 0$

$$V_F(\theta) = \frac{y(\theta)}{\theta^2} \psi(y(\theta)/\theta)$$

Proposition 1. The socially optimal threshold $\tilde{\theta}$ is strictly greater than the natural threshold ($\theta^*$).

Proof. Appendix.

A sketch of this proof can be provided supposing by contradiction that $\tilde{\theta}$ was below $\theta^*$. Then $\tilde{\theta}$ worker would prefer to be informal, but the planner would force her to be formal. Clearly, this type would be worse off. By the other hand, since the new threshold $\tilde{\theta}$ is the smallest formal type, the government is redistributing resources to $\tilde{\theta}$. So, the government budget constraint is thither after choosing the optimal social threshold $\tilde{\theta}$. Therefore, both the individual and the
planner would be worse off, then \( \tilde{\theta} \) could not be an optimal threshold. Thus, \( \tilde{\theta} \) must be strictly greater than \( \theta^* \).

In other words, if there is a policy able to influence the informal individual’s extensive margin choices, then limiting access to the formal sector would be welfare improving. The intuition for this result is based on two factors. First, according to hypothesis, the outside option of becoming informal is more attractive to lower types relatively to higher types. Second, there is a cost for the planner to keep a low type formal. This cost comes from the informational cost to keep the tax system incentive compatible and from resources redistributed to low skilled workers. Therefore, despite of the goal for smoothing utility across types, the cost to keep lower types formal is not worth the utility gain associated to this effort.

### 3.2 Heterogeneous Model

Following the same steps from the previous model, I want to analyze the threshold shape and then possible welfare improvements attained by extensive margin interventions.

**Definition 3.** For a given tax schedule, the natural threshold is a function \( c^*: [\theta; \bar{\theta}] \rightarrow \mathbb{R}^+ \), such that \( c^*(\theta) \) assigns the smallest cost that makes the individual-\( \theta \) indifferent between the two sectors.

**Proposition 2 (Threshold Function Slope).** The natural threshold function \( c(\theta)^* \) is increasing with the type.

**Proof.** Appendix.

This result is intuitive because since we are not assuming anything about the correlation between skill and costs, it is expected that workers who pay more taxes will be more willing to become informal. As I showed that marginal taxes are positive, the skilled workers are paying more taxes. Thus, for a fixed informality cost, skilled agents will be more susceptible to flee the formal job market.

As in the previous section, now I want to investigate if it is possible to reach a welfare improvement by distorting the natural threshold function. This result is proved in the next proposition.

**Proposition 3 (Social Optimal Threshold Shape).** If the planner could distort the natural threshold function \( c(\theta)^* \), she would chooses a flatter curve, as \( c(\theta)^* \). The two curves intercept in only one point.

**Proof.** Appendix

The main interpretation from Proposition 3 is that optimally, besides the optimal tax schedule, the planner would implement policies in order to formalize high skilled workers and informalize low skilled workers. This result is consistent to the homogeneous model’s result, where we have seen that the planner would avoid lower types in the formal job markets.

The intuition behind this result is that optimal nonlinear income taxation that gives rise to \( c(\theta)^* \) is not a second-best solution. There is a possibility to obtain social welfare improvement through the use of other policies that affect extensive margin choices.
Example 1. Based on the graph, consider the two agents \( \theta_1 \) and \( \theta_2 \), where \( \theta_1 < \theta_2 \). Let the cost \( s \in (c(\theta_2), c(\theta_2)^*) \) and the cost \( s' \in (c(\theta_1)^*, c(\theta_1)) \). For this given \( s \), the agent with skill \( \theta_2 \) would prefer to be informal under \( c(\theta_2)^* \) and would prefer to be formal under \( c(\theta_2) \). By the other hand, fixing the cost \( s' \), the worker \( \theta_1 \) would prefer to be formal under \( c(\theta_1)^* \) and would prefer to be informal under \( c(\theta_1) \).

4 Optimal Nonlinear Income Taxation

4.1 Homogeneous Model

In the model with homogeneous cost, as we have seen in Lemma 1 there is a clear cut-off across sectors. So, there is a smaller mass of agents subjected to the tax system. But the exclusion of lower types from the tax system does not affect the formal type’s incentives. The order across the formal workers remain the same, as well as their incentives to fake a different type. Thus, the optimal nonlinear income taxation in the formal sector, follows exactly the Mirrleesian taxation.

Proposition 4 (Mirrlees Formula). When workers have the possibility to avoid the tax system, incurring in a homogeneous informalities’s participation cost, the optimal marginal tax rates are:

\[
\frac{T'(y(\theta))c}{1 - T'y(\theta)} = \frac{M(\theta)}{\theta f(\theta)}
\]  

4.2 Heterogeneous Model

Taking into account the heterogeneous participation cost, the order among formal workers might change according to the endogenous extensive choices. For instance, a high skilled worker might be attracted to informality due to a low realization of participation cost. By the other hand, a low skilled worker may be associated to a high cost and then would prefer to become formal. Therefore, in this model some modification in the standard tax formulas will be present.
To deal with the extensive margin choices involved in this new mechanism designed when agents can flee the formal job market, I introduce a participation constraint to the Mirrlees model. In this new environment, the planner must solve the following problem in order to obtain the optimal nonlinear income tax schedule:

\[
\max_{V_F, g, c(\theta)} \int_{\theta}^{\bar{\theta}} \int_{2}^{c(t)} \phi(V_I(\theta) - s)g(s | \theta)f(s)d\theta d\theta + \int_{\theta}^{\bar{\theta}} \phi(V_F(\theta))(1 - G(c(\theta) | \theta))d\theta
\]

s.a.: \[
\int_{\theta}^{\bar{\theta}} T(y_F(\theta))(1 - G(c(\theta) | \theta)f(\theta)d\theta \geq 0
\]

\[\dot{V}_F(\theta) = \frac{y(\theta)}{\theta^2} \psi'(\frac{y(\theta)}{\theta})\]
\[c(\theta) = V_I(\theta) - V_F(\theta)\]

**Proposition 5** (Modified Mirrlees Formula). *When workers have the possibility to avoid the tax system, incurring in a heterogeneous informality’s participation cost, the optimal marginal income tax rates are:*

\[
\frac{T'(y(\theta))(1 - G(c(\theta) | \theta)e)}{1 - T'y(\theta)} = \frac{\hat{M}(\theta)}{\theta f(\theta)} - \int_{\theta}^{\bar{\theta}} T(y(t))g(c(t) | \theta)f(t)dt
\]

*Proof. Appendix.*

This optimal tax formula differs from (Piketty, 1997), (Diamond, 1998) and (Saez, 2000) in three main aspects:

**Evasion Effect.** This is captured by the last term on the right hand side of 4, and accounts for the agents that are escaping from the formal jobs when marginal tax rates are raised. Note that, once a worker move to informality, the planner reduces it tax collection in level. In other words, the marginal increase in the tax rates that culminate with the evasion of certain type, provoke a loss to the government budget correspondent to all the taxes that this type used to pay, while formal. The evasion effect restricts the government’s ability to tax.

**Efficiency Effect.** The classical trade-off between redistribution and efficiency remains, but it becomes less intense. Note by the left hand side of 4 that when the marginal tax rates are raised, the labor supply reduces according to its elasticity. But, only a mass of formal agents, given by \(1 - G(c(\theta) | \theta)\) is affected by the tax increase. So, the consequence in terms of production loss is also restricted to a smaller mass of agents. Compared to the standard case, this efficiency effect enhances the government’s potential to tax and redistribute.

**Modified Mechanical Effect.** Typically, when the planner raises marginal tax rates, there is a mechanical effect associated to it. This is due to the increase in funds collected by the government. In our model, the funds are collected only from a restricted mass of formal agents \(1 - G(c(\theta) | \theta)\). Thus for the same perturbation, the revenue collection is smaller compared
to the standard case. The modified mechanical effect weakens the government’s redistributive capability relative to the standard mechanical effect.

This last effect is given by the term $\hat{M}(\theta)$ in 4, which is:

$$\hat{M}(\theta) = \int_{\bar{\theta}}^{\delta} \left[1 - G(c(\theta) \mid \theta)\right] \left(1 - \frac{\phi'(V_F(\theta))}{\lambda}\right) f(\theta)d\theta$$

(5)

that is exactly the mechanical effect, but restricted to the mass $1 - G(c(\theta) \mid \theta)$.

5 Sales Taxes

So far, I have investigated the optimality of the tax schedule restricted only to income taxes. This methodology seems plausible because it follows the rationale presented in Mirrlees (1971), Atkinson and Stiglitz (1976), Diamond (1998) and Saez (2000). But in our economy, income taxes are restricted to only one sector, so it is worth investigating the effects of considering sales taxes. While the income tax is enforced only to the formal side, the sales tax is imposed to both sectors.

In this section, I want to cast light on the possibility of obtaining a welfare improvement through the use of positive sales taxes. I focus on the model with heterogeneous participation cost. Actually, I show that when workers have the possibility of fleeing the formal job market, it is optimal to consider the use of sales taxes. This result contrasts with the Atkinson-Stiglitz theorem, which says that when preferences over consumption and leisure are separable (as they are in our economy), the second-best can be implemented with zero sales taxes. The reason why this theorem does not hold in our environment is that the sales taxes plays an informational role.

I will proceed this argumentation in two steps. First, I show that it is never be optimal for the government to eliminate the informality. Then, considering the existence of the informal sector, I demonstrate that sales taxes should be positive.

Note that it is always feasible for the government to eliminate the informal sector. For instance, if the government set up a zero income tax for all the skill types, there is no reason for any type to become informal. This policy is compatible with the government’s budget, but it is not optimal because the planner is concave and aims to smooth utility across types. The question that stands is: Since it is always possible to eliminate informality, would be desirable under any circumstance to eliminate it? The answer is provided by the next proposition.

Definition 4. The full support hypothesis assures that $\forall \theta \in [\theta; \bar{\theta}]$, there is a positive measure set of individuals with type $(\theta; s)$. The $\bar{s}$ represents the minimum possible cost of becoming informal.

Proposition 6. Considering full support, it is always socially optimal to ensure informality’s existence.

Proof. Appendix. □

The intuition for this proof is that if the government intends to eliminate informality, it would be necessary to provide incentives for all types to become formal workers. In particular, the type $(\theta'; \bar{s})$ should become formal. Let $\theta'$ be a sufficiently high skill. To keep this type in the formal
side, the taxation enforced to her should be moderate because the cost $\bar{s}$ indicates the lowest cost of becoming informal. But, the tax enforced to $(\theta'; s')$ is the same collected from $(\theta'; s')$, where $s'$ is a sufficiently high cost. For instance, note that $(\theta'; s')$ had potential to pay much more taxes if it was not the concern to keep $(\theta'; s')$ formalized. The extra taxes potentially collected from $(\theta'; s')$ would allow more distribution across types and would be welfare improving.

Considering the existence of informality provided by Proposition 6, the planner will solve the following problem, which will lead us to Proposition 7:

$$\max_{V_F, y, t} \int_\theta^{c(t)} \int_\theta^c \phi[(1 - t)(V_I(\theta) - s)]g(s \mid \theta)f(\theta)d\theta d\theta$$

$$+ \int_\theta^{c(t)} \phi[(1 - t)V_F(\theta)] - t\psi(y(\theta))\phi(1 - G(c(\theta) \mid \theta))f(\theta)d\theta$$

s.a.: $\int_\theta^{c(t)} T(y_F(\theta))(1 - G(c(\theta) \mid \theta) + t[y_F(\theta) - T(y_F(\theta))])$

$$(1 - G(c(\theta) \mid \theta)) + y_F(\theta)G(c(\theta) \mid \theta)]f(\theta)d\theta \geq 0$$

$$V_F(\theta) = \frac{y(\theta)}{\theta^2} \psi'(\frac{y(\theta)}{\theta})$$

**Proposition 7.** When workers have the possibility of fleeing the formal job market, the optimal tax schedule should have positive sales taxes.

**Proof.** Appendix

## 6 Concluding Comments

The next steps for this project is to collect empirical data to address applied conclusions. In the model with heterogeneous costs I want to calibrate the density function $F(\cdot)$ to Brazilian economy parameters. Then I can do applied exercises, such as evaluate the optimality of the given Brazilian tax schedule compared to optimum when informality is accounted. Considering that informality is tax-responsive, I want to compare Brazilian informal sector dimension to the optimal size suggested by the model and the calibrated density function.
References


Lehmann, E., L. Simula, and A. Trannoy (2013). Tax me if you can! optimal nonlinear income tax between competing governments.


Appendix

Proof of Proposition 2

Proof. We know that \( c(\theta) = V_I - V_F \). So using the envelope theorem we have:

\[
c'(\theta) = \frac{y_I(\theta)}{\theta} \psi'(\frac{y_I(\theta)}{\theta}) - \frac{y_F(\theta)}{\theta} \psi'(\frac{y_F(\theta)}{\theta})
\]

But from formal and informal workers problem’s we have the following first order conditions:

\[
f(y_I) = 1 - \psi'(\frac{y_I}{\theta}) \frac{1}{\theta} = 0
\]

\[
f(y_F) = 1 - \psi'(\frac{y_F}{\theta}) \frac{1}{\theta} = T'(y_F(\theta))
\]

Suppose for a moment that \( y_I(\theta) \) that satisfies 9 is less or equal to \( y_F(\theta) \) that satisfies 10. But, \( f(y_I) \) and \( f(y_F) \) are both decreasing in the income. So it is straight forward that:

\[
f(y_F) \leq f(y_I) = 0 < T'(y_F(\theta)) = f(y_F(\theta))
\]

The first equality on 11 comes from 9, while the second inequality derives from lemma 2 and the last equality comes from 10.

The expression 11 leads to a contradiction, so \( y_I(\theta) > y_F(\theta) \). We assumed that \( \psi \) is convex, what guarantees that:

\[
\psi'(\frac{y_I}{\theta}) y_I(\theta) \frac{1}{\theta^2} > \psi'(\frac{y_F}{\theta}) y_F(\theta) \frac{1}{\theta^2}
\]

Thus, substituting 12 into 8, we can finally conclude that \( c'(\theta) > 0 \).

\[\square\]

Proof of Proposition 3

Proof.

Lemma 2. \( T'(y_F(\theta)) > 0 \)

Proof. From the planner’s budget constraint we have:

\[
F(V_F) = \int_{0}^{\theta} T(y_F(\theta))(1 - G(c(\theta) \mid \theta)f(\theta))d\theta = 0
\]

The derivative of 13 with respect to \( V_F \), must be negative, otherwise the planner could raise \( V_F \) infinitely and the budget constraint would never bind. In this sense we have that,
\[
\int_{0}^{\bar{\theta}} -T(y(\theta))g(c(\theta) \mid \theta)f(\theta)d\theta > 0
\]

We know that the modified mechanical effect $\hat{M}(\theta)$ is greater than zero. Now, considering the signal obtained from 14 and applying it on the Modified Mirrlees Formula 4, it concludes our proof.

\[
\left[ c(\theta) \right] : \phi(V_I(\theta) - \tilde{c}(\theta))g(\tilde{c}(\theta) \mid \theta)f(\theta) - \phi(V_F(\theta))g(\tilde{c}(\theta) \mid \theta)f(\theta) = 0
\]

\[
\phi(V_I(\theta) - \tilde{c}(\theta)) - \phi(V_F(\theta)) - \lambda T(y(\theta)) = 0
\]

For $\theta$ sufficiently high such that $T(y(\theta)) > 0$, 15 gives us:

\[
\phi(V_I(\theta) - \tilde{c}(\theta)) > \phi(V_F(\theta))
\]

Thus, $\tilde{c}(\theta) < c(\theta) = V_I(\theta) - V_F(\theta)$ because $\phi(\cdot)$ is increasing.

Analogously we might argue that for $\theta$ sufficiently low such that $T(y(\theta)) < 0$:

\[
\phi(V_I(\theta) - \tilde{c}(\theta)) < \phi(V_F(\theta))
\]

Therefore, $\tilde{c}(\theta) > c(\theta) = V_I(\theta) - V_F(\theta)$

It is still remaining to prove that there is only one productivity level in which the optimal threshold function is equal to the natural one. This is straightforward from the fact that $T'(y(\theta)) > 0$, as we proven in Lemma 2. So there is only one $\hat{\theta}$, such that $T(y(\hat{\theta})) = 0$. Then, by 16 and 17 follows the proof.

\textbf{Proof of Proposition 5}

\textit{Proof.} I will use optimal control techniques to solve problem 3. Initially I will substitute the last constraint into the hamiltonian's objective function:
\( H = \int_{\theta}^{\bar{\theta}} \int_{s}^{V_{T}(\theta) - V_{F}(\theta)} \phi(V_{I}(\theta) - s)g(s \mid \theta)f(\theta)dsd\theta + \int_{\theta}^{\bar{\theta}} \phi(V_{F}(\theta))(1 - G(c(\theta) \mid \theta))d\theta \\
+ \lambda \int_{\theta}^{\bar{\theta}} T(y_{F}(\theta))(1 - G(c(\theta) \mid \theta))f(\theta)d\theta + \mu \left( \frac{y(\theta)}{\theta^2} \right) \psi'(\frac{y(\theta)}{\theta}) \geq 0 \) (18)

Taking the first order conditions:

\[ \frac{\partial H}{\partial V_{F}} = \phi'(V_{F}(\theta))(1 - G(c(\theta) \mid \theta))f(\theta) + \phi(V_{F}(\theta))g(V_{I}(\theta) - V_{F}(\theta))f(\theta) - \phi(V_{F}(\theta))g(V_{I}(\theta) - V_{F}(\theta))f(\theta) \\
+ \lambda \int_{\theta}^{\bar{\theta}} T(y_{F}(\theta))(1 - G(c(\theta) \mid \theta) + T(y(\theta)))g(V_{I}(\theta) - V_{F}(\theta) \mid \theta)]f(\theta) = -\hat{\mu} \] (19)

\[ \frac{\partial H}{\partial y} = \lambda T'(y(\theta))(1 - G(V_{I}(\theta) - V_{F}(\theta)))f(\theta) + \mu(1 - T'y(\theta))) \frac{1}{c\theta} = 0 \] (20)

Manipulating 19, we have:

\[ \hat{\mu} = \lambda f(\theta)[(1 - G(V_{I}(\theta) - V_{F}(\theta)))(1 - \frac{\phi'V_{F}(\theta)}{\lambda}) - T(y(\theta)))g(V_{I}(\theta) - V_{F}(\theta) \mid \theta)] \] (21)

Integrating both sides from \( \theta \) to \( \bar{\theta} \), it follows:

\[ \mu = \lambda \hat{M}(\theta) \int_{\theta}^{\bar{\theta}} -T(y(\theta)))g(V_{I}(\theta) - V_{F}(\theta) \mid \theta)]f(\theta)d\theta \] (22)

Where, \( \hat{M}(\theta) = \int_{\theta}^{\bar{\theta}} [1 - G(V_{I}(\theta) - V_{F}(\theta)))(1 - \frac{\phi'(V_{F}(\theta))}{\lambda})f(\theta)d\theta \) is the modified mechanical effect.

Substituting 22 in 20, we have:

\[ \lambda T'(y(\theta))(1 - G(V_{I}(\theta) - V_{F}(\theta)))f(\theta) = -(1 - T'y(\theta))) \frac{1}{c\theta} \lambda \hat{M}(\theta) \int_{\theta}^{\bar{\theta}} -T(y(\theta)))g(V_{I}(\theta) - V_{F}(\theta) \mid \theta)]f(\theta)d\theta \] (23)

Dividing both sides by \( \lambda \), using that \( c(\theta) = V_{I}(\theta) - V_{F}(\theta) \) and reorganizing terms, we have the modified Mirrlees formula:

\[ \frac{T'(y(\theta))(1 - G(c(\theta) \mid \theta)\epsilon}{1 - T'y(\theta)} = \frac{\hat{M}(\theta)}{\theta f(\theta)} - \int_{\theta}^{\bar{\theta}} T(y(\theta)))g(c(\theta) \mid \theta)f(\theta)d\theta \] (24)
Proof of Proposition 6

Proof. Suppose by contradiction that it is optimum for the planner to eliminate informality. So every worker is formal. In particular, the type \((\theta', 0)\) is formal.

Let, \(\theta'\) be below \(\bar{\theta}\), but sufficiently high such that \(T(y(\theta')) > 0\). I want \(\theta' < \bar{\theta}\) to make sure that \(T'(y(\theta')) > 0\).

Lemma 3. Every type that faces a positive income marginal tax would work more hours if informal compared to the labor supply choice if formal, which implies that, \(y_I(\theta') > y_F(\theta')\).

Proof. From the \(\theta'\)-worker intensive margin problems:

\[
\max_{y_F} y_F(\theta') - T(y_F(\theta')) - \psi\left(\frac{y_F(\theta')}{\theta'}\right)
\]

\[
[y_F] : 1 - T'(y_F(\theta')) - \psi\left(\frac{y_F(\theta')}{\theta'}\right) \frac{1}{\theta'} = 0
\]

\[
T'(y_F(\theta')) = 1 - \psi\left(\frac{y_F(\theta')}{\theta'}\right) \frac{1}{\theta'}
\] (25)

This same worker would face the following problem to choose his informal labor supply:

\[
\max_{y_I} y_I(\theta') - \psi\left(\frac{y_I(\theta')}{\theta'}\right)
\]

\[
[y_I] : 1 - \psi\left(\frac{y_I(\theta')}{\theta'}\right) \frac{1}{\theta'} = 0
\]

(26)

Now, suppose by contradiction that \(y_I(\theta') \leq y_F(\theta')\). Then, using that \(\psi(c\cdot \text{dot})\) is convex, we have that:

\[
\psi\left(\frac{y_I(\theta')}{\theta'}\right) \frac{1}{\theta'} \leq \psi\left(\frac{y_F(\theta')}{\theta'}\right) \frac{1}{\theta'}
\] (27)

Putting this results together we find:

\[
0 = 1 - \psi\left(\frac{y_I(\theta')}{\theta'}\right) \frac{1}{\theta'} \geq 1 - \psi\left(\frac{y_F(\theta')}{\theta'}\right) \frac{1}{\theta'} = T'(y_F(\theta')) > 0
\] (28)

Where the first and the last equality comes from 26 and 25, respectively. The first inequality comes from 27 and the last inequality comes from construction of \(\theta'\).

Therefore, expression 28 leads to a contradiction and, thus we conclude that \(y_I(\theta') > y_F(\theta')\).

As \((\theta', 0)\) is formal, then by construction:

\[
c(\theta') < s = 0
\]
\[
V_I(\theta') - V_F(\theta') < 0 \\
V_I(\theta') < V_F(\theta')
\]

\[
y_I(\theta') - \psi(y_I(\theta')) < y_F(\theta') - T(y_F(\theta')) - \psi(y_F(\theta'))
\]  \hspace{1cm} (29)

By Lemma 3, I know that \(y_I(\theta') > y_F(\theta')\), so the informal worker could have chosen \(y_F(\theta')\), but preferred \(y_I(\theta')\). So, we can infer that:

\[
y_I(\theta') - \psi(y_I(\theta')) > y_F(\theta') - \psi(y_F(\theta'))
\]  \hspace{1cm} (30)

Therefore, using that \(T(y(\theta')) > 0\), equations 29 and 30 lead to a contradiction and the proof follows. \(\Box\)

**Proof of Proposition 7**

*Proof.* From the first order condition of problem 6, with respect to sales taxes "t":

\[
[t] = - \int_{\theta}^{\hat{\theta}} \int_{\theta}^{c(\theta)} \phi'[(1-t)(V_I(\theta) - s)](V_I(\theta) - s)g(s | \theta)f(\theta)d\theta d\theta \\
+ \lambda \int_{\theta}^{\hat{\theta}} [(y_F(\theta) - T(y_F(\theta)))(1 - G(c(\theta) | \theta)) + y_I(\theta)G(c(\theta) | \theta)]f(\theta)d\theta \\
- \int_{\theta}^{\hat{\theta}} \phi'((1-t)V_F(\theta) - t\psi(y_F(\theta)))(1 - G(c(\theta) | \theta))[V_F(\theta) + \psi(y_F(\theta))]d\theta = 0
\]  \hspace{1cm} (31)

Now, denote the first order condition equation from 31 as \(W(\theta)\). So, considering \(t = 0\) and \(t^*\) the optimal sales taxes, we have three possible cases:

- \(W(\theta) > 0\), then \(t^* > 0\)
- \(W(\theta) < 0\), then \(t^* < 0\)
- \(W(\theta) = 0\), then \(t^* = 0\)

I show that in fact \(W(\theta) > 0\), which proves that \(t^* > 0\). Evaluating the first term in 31, and considering that \(s\) can assume values in the continuum from \(\underline{s}\) to \(c(\theta)\), we can find a lower bound for the first term in 31:
\[-\int_\theta^\tilde{\theta} \int_\theta^\infty \phi'[(1-t)(V_I(\theta) - s)](V_I(\theta) - s)g(s \mid \theta)f(\theta)dsd\theta \geq -\int_\theta^\tilde{\theta} \phi'[(1-t)(V_F(\theta))G(c(\theta) \mid \theta)](1-G(c(\theta) \mid \theta))\int_\theta^\infty \phi'[(1-t)(V_I(\theta) - s)](V_I(\theta) - s)f(\theta)dsd\theta \geq -\int_\theta^\tilde{\theta} \phi'[(1-t)V_F(\theta)](V_I(\theta) - s)G(c(\theta) \mid \theta)f(\theta)d\theta \] (32)

Considering \( t = 0 \) and substituting the lower bound obtained by (32) in (31):

\[-\int_\theta^\tilde{\theta} \phi'[(1-t)V_F(\theta)](\bar{x}_I(\theta))G(c(\theta) \mid \theta)f(\theta)d\theta - \int_\theta^\tilde{\theta} \phi'(V_F(\theta))(1-G(c(\theta) \mid \theta))\{[V_F(\theta) + \psi(y_F(\theta) \theta)]d\theta + \lambda \int_\theta^\tilde{\theta} [V_F(\theta) + \psi(y_F(\theta)) \theta](1-G(c(\theta) \mid \theta)) + G(c(\theta) \mid \theta)y_I(\theta))]f(\theta)d\theta \leq W(\theta) \] (33)

Rearranging terms and considering \( t = 0 \), we have:

\[\int_\theta^\tilde{\theta} G(c(\theta) \mid \theta)(\lambda y_I(\theta) - \phi'(V_F(\theta)))\bar{x}_I(\theta)]f(\theta)d\theta + \int_\theta^\tilde{\theta} (1-G(c(\theta) \mid \theta))\{[V_F(\theta) + \psi(y_F(\theta) \theta))]f(\theta)d\theta - \phi'(V_F)f(\theta)d\theta \leq W(\theta) \] (34)

Therefore we conclude that \( W(\theta) > 0 \) and thus \( t^* > 0 \).