Models of optimal taxation typically assume that governments can perfectly observe taxpayers’ income. There is plenty of evidence to suggest otherwise. Taxpayers can use income falsification to reduce how much taxes they pay by manipulating the information the government receives about their income through legal (tax avoidance) and illegal means (tax evasion). Additionally, participating in these activities is not costless and it involves more than just lying or sending an incorrect signal. This paper relaxes the aforementioned standard assumptions and answers the question of how the optimal tax schedule changes when a normative model of taxation allows for costly income falsification. The outcome is a tax function that encourages agents to display all of their income (non-falsification), allowing a social planner to increase tax revenue and aggregate welfare while eliminating agents’ wasteful income falsification activities. These results suggest that there are gains to reforming the tax schedule to encourage taxpayers to report their income in full. Moreover, I explore policy implications when the model is calibrated to realistic parameters of taxpayer behavior in the U.S. and show that marginal tax rates under costly income falsification are lower than those prescribed by a similarly defined static Mirrleesian model. Ultimately, this paper shows that assuming that income is perfectly observables is not innocuous.

J.E.L Codes: H21, H24, H31, J22.
Keywords: Optimal taxation, income tax, tax avoidance, tax evasion, non-linear marginal taxes, Mirrleesian optimal taxes.

*I am grateful to Marek Kapicka, Peter Rupert, Henning Bohn, Julian Neira, Liming Chen, Brian Thomas and for their valuable feedback. I am also in debt to the AEA CSMGEP Summer Fellowship Program and researchers in the Federal Reserve Bank of Chicago and the Board of the Federal Reserve for their comments and assistance.
1 Introduction

Models of optimal taxation assume that governments can observe taxpayers’ income perfectly. There is plenty of evidence to suggest otherwise. Taxpayers can use income falsification—the act of manipulating the income figures they provide to the government—to reduce how much taxes they pay. For instance, taxpayers can report income that otherwise would be considered labor income as capital gains or dividend income (tax avoidance), or simply report less income than they earned (tax evasion). Another standard assumption is that taxpayers communicate with the government through messages or signals that can be misrepresented at no cost, but taxpayers’ activities show that engaging in income falsification consists of more than just lying or sending an incorrect signal. Exploiting weaknesses in the tax code or finding legal ways to alter reported income can require the services of a tax accountant or a tax attorney, and reducing taxable income by illegal means imposes, in the least, an expected cost of being discovered and punished.

This paper assesses the consequences of relaxing these assumptions on the optimal tax schedule. Specifically, how does the policy prescriptions of normative models of taxation change when they allow for strategic actions like tax avoidance and undetectable tax evasion? To answer this question, I generalize the optimal non-linear income taxation model built by Mirrlees [1971] to allow agents to engage in costly income falsification; agents can choose the amount of income that are going to “display”, and pay a cost to show an income display that is different from the true income. These income displays are part of an alternative structure for communicating information between agents and the social planner. This informational structure lets agents take strategic actions to reduce their tax liability that does not involve changing their labor supply, which allows the model to captures additional margins of adjustment to taxation used by taxpayers.

The outcome is a mechanism where agents display all of their income (non-falsification). The prescribed tax schedule allows the social planner to increase tax revenue and aggregate welfare, while eliminating agents’ wasteful income falsification activities. This is achieved by designing it to provide agents with the incentives to increase their work effort through marginal taxes, and to self-select into non-falsification through lump sum taxes. Starting with the latter, lump sum taxes work to make agents indifferent between falsification and fully displaying their income. On the other hand, marginal taxes direct agents to the allocation where these lump sum taxes maximize tax revenue. In cases where the cost of income falsification is low, marginal taxes encourage agents to work “harder” so that the social planner can tax a larger income pool. Under these circumstances, this need to increase tax revenues drives marginal taxes below the level of classical Mirrleesian marginal taxes. However, as this cost increases, the redistributive limits imposed by it are lifted, and the social planner can redistribute more income across agents until it reaches the Mirrleesian solution. This convergence occurs because as the cost of income falsification increases, it slowly reaches the point where agents are better off avoiding taxes by paying the cost of changing their labor supply rather than paying the cost of income falsification.

These results suggest that there are gains to reforming the tax schedule to encourage
taxpayers to report all of their income. This argument corroborates the old adage that “low tax rates over a broad base is good tax policy”: taxes under income falsification are always lower than in the standard Mirrleesian case, and higher income by high ability agents translates into a large tax base, by size even if not by the number of taxpayers. Taking into account income falsification also strengthens the implementability of the optimal tax schedule; by taking into account other ways that taxpayers use to avoid taxes, its normative prescriptions could be more closely compared to the current tax schedule. Ultimately, this paper shows that assuming that income is perfectly observable is not innocuous, and shows that the optimal tax function is limited by the available income falsification opportunities found within the tax system.

Just like costly state falsification in contract theory, this paper complements the costly state verification models of tax evasion in public economics. Starting with Allingham and Sandmo (1972), Yitzhaki (1974), and Sandmo (1981), these models features a tax collecting institution that pays a cost to acquire information about taxpayers’ true income through tax auditing. This institution can impose punishment when they detect a deviation in reported income. In contrast to this approach, my model proposes an optimal tax designed to prevent agents’ income falsification, thereby eliminating compliance costs and the need for punishment or enforcement. As in costly state falsification contracts, the social planner in this paper doesn’t pay a cost to acquire information, but sets taxes with the understanding that taxpayers can manipulate their income information at a cost. This approach is appropriate, first, in situations when tax authorities can’t punish or enforce report deviations because of some technological constraint (like in the case of undetectable tax evasion), or second, when income manipulation is legal (tax avoidance) and therefore not punishable.

This paper also contributes to the study of taxpayers’ responses to marginal taxes. Specifically, it builds on insights from Saez (2004) and Piketty, Saez and Stantcheva (2014), who point out that taxpayers use other alternatives rather than changing their labor supply to reduce taxable income. This notion is also closely related to the large literature on taxable income elasticities started by Lindsey (1987) and Feldstein (1995), which recognizes and estimates taxpayers response to taxation. This paper applies their observations into a normative model of taxation. On the theoretical side, this paper also builds on Grochulski (2007) and Lacker and Weinberg (1989). It adds to the first by relaxing his assumption that income is received as an exogenous endowment, allowing agents to choose the amount of income they are going to produce. In doing so, it incorporates tax payers’ labor supply and income reporting responses jointly and changes the policy prescriptions of the model. And finally, it incorporates the second’s informational structure into the non-linear taxation framework.

1 For some examples see Cowell (1990), Cremer and Galvani (1994), Kaplow (1990), Slemrod (1994)
2 Income falsification in the U.S.

The academic literature contains ample evidence of taxpayers’ income falsification in the United States. In the case of (legal) tax avoidance, research shows that individuals (including those represented by closely held corporations) are able to legally minimize their tax liabilities by strategically choosing what kind of debt instruments and business structures they use for their economic activities. For example, Maki (1996) and Scholz (1994) record that after the implementation of the Tax Reform Act of 1986 there was a large shift in the debt instruments used by households, from no-longer-deductible consumer interest into a still-deductible mortgage or home equity loans. As a product of the same tax reform, the marginal tax rate for personal income was lowered below the corporate marginal tax. Gordon and MacKie-Mason (1990, 1997) report that this change generated a large shift in the legal structure of firms, from C-Corporations to S-Corporations, to take advantage of lower personal income marginal taxes available to firms established as the latter. Similarly, Gordon and Slemrod (1988) estimated that due to tax arbitrage the US collected approximately zero tax revenue from taxes levied on capital income in 1983. Tax arbitrage and tax exemptions are also responsible for differences between statutory and effective tax rates. Cagetti and De Nardi (2009) calculated that around the 1990’s the effective average estate tax was 16% when the statutory rate was 40%-60%.

On the illegal side, the IRS periodically publishes their estimates of aggregate tax evasion called the “Tax Gap”: the difference between what tax payers should pay (according to tax legislation), and what they actually pay. In their 2011 report, the IRS estimates that this gap was 16.9% of total tax liability. Tax Gap measures of tax evasion are even incorporated into the calculations for the National Income and Product Accounts. The Bureau of Economic Analysis uses an adjustment factor that gets multiplied by the reported income of households and businesses to calculate total US income. In the past 40 years, for some classifications like nonfarm proprietor’s income, this factor of adjustment was as high 2.9, implying that these taxpayers only reported 34% of their income.

These estimates show that legal and illegal income manipulation are not easily measured; we rely on changes in the tax schedule for evidence of changes in tax avoidance behavior, and estimate tax evasion based on data obtained from randomized tax audits. The former are not a good indicator of the overall extent of tax avoidance, for which we have no estimates, and the latter suffers from inherent weaknesses of the tax auditing process caused by the variation in the detected non-compliance results of tax auditors (which gives the Tax Gap the interpretation of a measure calculated by an “average tax official” in contrast to the obviously superior “fully informed tax official”). However, as they are, they describe income falsification and its uses, which I now incorporate into a model of optimal taxation.

---

3 This is measuring the Gross Tax Gap, which is based on voluntary tax compliance. The Net Tax Gap is a measurement of the Gross Tax Gap minus the tax receipts from enforced and late payments. The Net Gap was 14.5% in 2011.
3 Optimal taxation with incomplete information and income falsification

Certain features of the tax system, like tax loopholes or imperfect tax evasion monitoring, are easy opportunities for moral hazard problems to appear; their exploitation proves that taxpayers can manipulate their income to obtain an advantageous tax position. This means that they have power to control the flow of information used for tax purposes, so the government has to look through a murky glass when designing and collecting taxes. To model these frictions, I need to modify the standard Mirrleesian model of optimal taxation. As mentioned previously, it’s not enough for agents to send messages about their income; falsifying income in the real work takes a lot more effort than just lying about how much income a person receives. Furthermore, the information shared with the government represents the reality upon which taxes are calculated. So, I use the costly state falsification optimal contracts approach of Lacker and Weinberg (1989), where agents have control of publicly available information—in this case official income figures—and can deviate from the truth by paying a falsification cost. I incorporate this informational structure first. Then I show how the social planner adapts to make sure that the optimal mechanism is successful when implemented in a decentralized economy, where agents can act according to their opportunities for income falsification.

3.1 Economic environment and agents’ behavior

Agents’ preferences and production

Agents are privately informed about their ability which can be either high or low, \( \Theta \equiv \{\theta_L, \theta_H\} \). From the perspective of a social planner choosing an individual at random, an agent is of high type with probability \( \pi \) and of low ability with probability \((1 - \pi)\), according to their proportion in the population. For \( i \in \{H, L\} \), agents’ preferences and production functions are represented by:

\[
W(c_i, l_i) = U(c_i) - V(l_i)
\]

\[
y_i = \theta_i l_i.
\]

(3.1)

\(W(c_i, l_i)\) is the utility derived from consumption and labor effort and is additively separable, consisting of two parts: \(U(c_i)\) and \(V(l_i)\). I assume that \(U'(c_i), -U''(c_i), V'(l_i), V''(l_i)\) are positive, exist, and meet the usual regularity conditions: \(\lim_{c_i \to 0} U'(c_i) = \infty\), \(\lim_{c_i \to 0} V'(l_i) = 0\). An agent’s income, \(y_i\), is produced by multiplying their ability and labor effort.

Income display and agents’ problem

To formalize the idea that the government can’t observe taxpayers income perfectly, I assume that produced income, \(y_i \geq 0\), is also private information. Agents, just like taxpayers, are required to disclose information about their income. They do so by choosing an income
"display", \( x_i \geq 0 \), which comes from a set of believable income displays:

\[
x_i \in \text{argmax } \tilde{x}, \quad \bar{X} \equiv \{ \tilde{x} | \text{ for any } j, \tilde{x} \in \text{argmax } x_j \}.
\] (3.2)

This means that any agent’s display belongs to the set of utility maximizing display choices of all other types of agents, \( \bar{X} \). Taxes, \( T(x_i) \), are based on their displayed income according to the tax schedule previously appointed by the social planner. This gives agents the option of displaying a level of income that gives them an advantageous tax position, pay the costs associated with making this display and keep the remainder as private consumption. The agent’s budget constraint reflects these possibilities:

\[
c_{i}^{\text{displayed}} = x_i - T(x_i)
\]

\[
c_{i}^{\text{private}} = (y_i - x_i) - g(y_i, x_i)
\]

\[
c_i = c_i^{\text{private}} + c_i^{\text{displayed}} = y_i - g(y_i, x_i) - T(x_i).
\]

The cost of falsification function, \( g(y_i, x_i) \), is the cost of displaying \( x_i \) while privately producing \( y_i \). This cost could stand for the risk premium paid by a risk-averse agent that is engaging in tax evasion. It could also represent expenditures on tax accountants or tax attorneys, or time spent in learning the tax code and filling more complicated tax forms. I assume that these activities are unproductive because they represent a reallocation of resources to tasks that wouldn’t exist had the opportunities for income falsification never existed. In addition, I assume that \( g(\cdot) \) is the same for both type of agents, that it takes non-negative values \( g: \mathbb{R}_+^2 \to \mathbb{R}_+ \), is zero when the agents’ displayed income matches realized income, \( g(z, z) = 0 \) for any \( z \in \mathbb{R} \), and

\[
g_x(y, x) < 0, \quad g_y(y, x) > 0 \text{ if } x < y
\]

\[
g_x(y, x) > 0, \quad g_y(y, x) < 0 \text{ if } x > y.
\]

The social planner knows the functional form of the income falsification cost, but does not observe individual realizations; which prevents the social planner from backing out information about agents’ types from the cost they pay to display their income. Agents act optimally given their preferences, production technology, and their opportunities for income falsification by choosing the amount of income they are going to produce and display:

---

4 An agent that displays income outside of this set reveals himself as a deviator, to whom the social planner can impose very large tax penalty. This does not mean that the social planner has the capacity to enforce truth-telling or that it can observe agent’s ability or income realizations. Rather it means that the social planner can detect single deviators from equilibrium strategies and has a contingent plan for them, which could involve complete income expropriation.
\[
\max_{y_i, x_i} U(y_i - g(y_i, x_i) - T(x_i)) - V(y_i / \theta_i)
\]
\[
s.t.
\]
\[
c_i = c_i^{\text{private}} + c_i^{\text{displayed}} = y_i - g(y_i, x_i) - T(x_i)
\]
\[
c_i^{\text{displayed}} = x_i - T(x_i)
\]
\[
c_i^{\text{private}} = (y_i - x_i) - g(y_i, x_i)
\]
\[
x_i \in \arg\max \tilde{x}, \tilde{X} \equiv \{ \tilde{x} | \text{for any } j, \tilde{x} \in \arg\max x_j \}
\]

Income displays, messages, and the Revelation Principle

This income display structure is consistent with the Revelation Principle and is without loss of generality. First, its consistent with the Revelation Principle because in this environment any arbitrary mechanism can be replicated by a direct revelation mechanism where agents truthfully report their ability types via messages. Truth-telling, however, does not rule out income falsification. An agent can send a message stating to be of a particular ability type while displaying an income level that is not equal to her produced income. As long as the message sent and income displayed are in accord with the income displays of all other agents of her type (those who sent the same message), the social planner can’t verify whether the income displayed is the entirety of her produced income. Second, the structure of income displays, as previously described, is without loss of generality because the message space is extraneous. To prove this, suppose that instead of imposing taxes solely on displayed income, \( T(x_i) \), taxes are based on both the displayed income and a message sent to the social planner \( T(x_i, m_i) \), \( T : X \times M \rightarrow \mathbb{R} \). In this mechanism, \( m \in M \) is an element of the message space \( M \). Assume that for two different income realizations \( y \) and \( y' \) the income displayed is the same, or \( x(y) = x(y') = x \), but the message sent is different, \( m(y) = m, m(y') = m' \neq m \). For the pair \((x, m)\) to be optimal when \( y \) realized it must be that
\[
y - g(y, x) - T(x, m) \geq y - g(y, x) - T(x, m')
\]
\[
T(x, m) \leq T(x, m').
\]
Similarly, the optimality of \((x, m')\) when \( y' \) is realized requires that
\[
y' - g(y', x) - T(x, m') \geq y' - g(y', x) - T(x, m)
\]
\[
T(x, m') \leq T(x, m).
\]
So that two income realizations with the same displayed income must lead to the same allocation, making messages extraneous.

Additionally, I want to make an important distinction between reporting and displaying income. Agents are not sending a message about their level of income, they are “displaying

---

5 Similar claims were first made by Lacker and Weinberg (1989) in the context of optimal contract theory.
it”. They are taking an action to make information publicly available and readily verifiable. In contrast, in the context of mechanism design, a message can be any form of communication between agents and the social planner. This communication may include information about income levels, but is not verifiable and does not necessarily reveal private information. Also, an income display doesn’t necessarily have to equal total produced income, it may not; displaying income determines how much income is observed by the social planner, and implies the possibility that a fraction of agents’ income that remains out of its reach.

3.2 Social Planner’s problem with costly income falsification

A social planner is tasked with finding the mechanism that will deliver the greatest expected aggregate welfare. This mechanism is meant to be implemented in a decentralized economy using marginal and lump-sum taxes. For this reason, the social planner is bound by the limitations imposed by the informational structure of the economic environment, in particular, that agents have the capacity to falsify their income and change their labor effort in response to taxation. This translates to asking the social planner to limit the optimal mechanism to be incentive compatible and feasible under expected tax receipts. Since the implementation of this mechanism is simply setting agents’ marginal rates of substitution between consumption and labor equal to the net of tax rate, I will incorporate the tax function directly into the following equations.

Feasibility and incentive compatibility

There are three incentive compatibility constraints for each type of agent $i \neq j \in \{H, L\}$. Note that each equation takes into account the agent’s possible income display strategies outlined in equation (3.2). The first set of constraints, the income falsification constraints (IFC’s), reduces the set of mechanism to those where agents don’t profit by displaying an income level different from their realized income:

$$U(y_i - g(y_i, x_i) - T(x_i)) - V(y_i / \theta_i) \geq U(y_i - g(y_i, x_j) - T(x_j)) - V(y_j / \theta_i)$$

or

$$g(y_i, x_j) - g(y_i, x_i) \geq T(x_i) - T(x_j).$$

From the perspective of a high ability agent, this constraint prevents the social planner from choosing a mechanism where this agent could display the same income as a low ability agent, receive a tax transfer from other high ability agents, pay a cost of income falsification and keep the difference as private consumption.

The second set, the labor supply constraints (LSC’s), is the same found in many Mirrleesian models of optimal taxation,

$$U(y_i - g(y_i, x_i) - T(x_i)) - V(y_i / \theta_i) \geq U(y_j - g(y_j, x_j) - T(x_j)) - V(y_j / \theta_i).$$

It restrains the social planner from choosing a mechanism where agents alter their labor
supply in response to taxation. The common analogy used to explain this constraint is that it prevents high ability agents from “shirking”.

The third set of equations contains the double deviation constraints. These expressions eliminate mechanisms where an agent could use both methods (income falsification and changes in labor effort) to appear to be an agent of the other type:

\[
U(y_i - g(y_i, x_i) - T(x_i)) - V(y_i/\theta_i) \geq U(y_j - g(y_j, x_i) - T(x_i)) - V(y_j/\theta_i).
\]  

(3.5)

This equation rules out mechanisms where, for example, a low ability agent works like a high ability agent while displaying the income level corresponding to low ability agents. This is a profitable deviation if the low ability agent’s costs of income falsification and exerting more labor effort are more than compensated by the additional consumption from private income and tax transfers.

Finally, the feasibility of this mechanism is guaranteed by the following equation, which depends on consumption and expected displayed income:

\[
\pi^c_H \text{ reported} + (1 - \pi)c^L \text{ reported} = \pi x_H + (1 - \pi)x_L \\
\text{or} \\
\pi T(x_H) + (1 - \pi)T(x_L) = 0.
\]  

(3.6)

Social Planner’s optimization

The social planner chooses a mechanism \( \Gamma^* = \{y^*_i, x^*_i, T^*(x^*_i)\}_{i=H,L} \) that consists of an income allocation, a recommendation for how much income each type of agent should display, and the lump-sum tax transfers that correspond to each income display:

\[
\max_{y_i, x_i, T(x_i)} \pi [U(c_H) - V(l_H)] + (1 - \pi) [U(c_L) - V(l_L)] \\
s.t. \\
x_i \in \arg\max \hat{x}, \ \bar{X} \equiv \{\hat{x} \mid \text{for any } j, \ \hat{x} \in \arg\max x_j\} \quad (3.2) \\
g(y_i, x_j) - g(y_i, x_i) \geq T(x_i) - T(x_j) \quad (3.3) \\
U(y_i - g(y_i, x_i) - T(x_i)) - V(y_i/\theta_i) \geq U(y_j - g(y_j, x_i) - T(x_i)) - V(y_j/\theta_i) \quad (3.4) \\
U(y_i - g(y_i, x_i) - T(x_i)) - V(y_i/\theta_i) \geq U(y_j - g(y_j, x_i) - T(x_i)) - V(y_j/\theta_i) \quad (3.5) \\
\pi T(x_H) = -(1 - \pi)T(x_L) \quad (3.6)
\]

**Theorem 3.1** (Non-Falsification Theorem). *The solution to the social planner’s problem is a non-falsification mechanism*.\(^6\)

The search for the solution to this problem can be reduced to non-falsification mechanisms where agents display all of their produced income. The intuition behind this result

\(^6\) The proof of this theorem and the following corollaries are in the appendix.
rests on understanding the assumption that the cost of income falsification is unproductive. Income falsification has negative effects on the social planner’s budget constraint because it reduces tax revenues. This reduction is not directly offset by an increase in agents’ private consumption, which are paying an unproductive cost that reduces their welfare and the net benefit of falsification. These two forces contribute to welfare losses that could be averted if there was no income falsification. The optimality of the solution $\Gamma^*$ implies that the social planner should reduce these welfare losses by constructing a tax schedule that incentivizes agents to display all of their income. This is accomplished by making agents indifferent between income falsification and complete income displays through lump-sum taxes. These new lump-sum taxes will collect more tax revenue and increase welfare since they capture the unproductive cost of tax avoidance.

In addition to reducing deadweight losses, non-falsification reduces the number of decision variables in the social planner’s problem. Now the optimization problem can be restricted to those mechanisms where the income display recommendation is equal to agent’s income allocation. Accordingly, the social planner’s constraints should be re-written and simplified.

**Corollary 3.2.** The income falsification constraint, \((3.3)\), can be simplified to
\[
g(y_i, y_{-i}) \geq T(y_i) - T(y_{-i}).
\]

**Corollary 3.3.** The double deviation constraint is redundant if these constraints are satisfied
\[
g(y_i, y_{-i}) \geq T(y_i) - T(y_{-i}) \quad \text{(IFC’s)}
\]
and
\[
(y_i - T(y_i)) - V\left(y_i/\theta_i\right) \geq U\left(y_j - T(y_j)\right) - V\left(y_j\theta_H\right) \quad \text{(LSC’s)}.
\]

**Corollary 3.4** (Downward Binding Income Falsification Constraint). The income falsification constraint on low ability agents is always non-binding,
\[
g(y_i, y_h) \geq 0 \geq T(y_i) - T(y_h).
\]

Rewriting the incentive compatibility constraints brings some interesting insights. Corollary 3.2 shows that income falsification sets an upper bound on tax revenues by imposing limits on total redistribution. As will be explained later, this limitation is a major driving force in the calculation of optimal taxes. Corollary 3.3 points out that a double deviation strategy is more costly than a single deviation strategy. Because agents pay costs associated with each deviation, its clear that making two deviations, rather than one, will make the agent worse off. Finally, Corollary 3.4 follows directly from the functional form of the objective function. A concave, utilitarian objective function guarantees that the income falsification constraint is only binding for the high ability type agents, since $T(y_H) \geq 0 \geq T(y_L)$. This result, together with the “single crossing property” of these kind of problems, reduces the set of relevant incentive compatibility constraints to those imposed on the allocation and taxes of high ability agents.
Simplified Social Planner’s Problem

After reducing the number of incentive compatibility constraints, the social planner’s problem can be restated as searching for the mechanism \( \Gamma^* = \{y^*_H, y^*_L, T^*(y^*_H), T^*(y^*_L)\} \) that solves:

\[
\max_{y_i, T(y_i)} \pi \left[U(y_H - T(y_H)) - V(y_H/\theta_H)\right] + (1 - \pi) \left[U(y_L - T(y_L)) - V(y_L/\theta_L)\right]
\]

\[
s.t.
\begin{align*}
\lambda & \pi T(y_H) + (1 - \pi) T(y_L) = 0 \\
\gamma & g(y_H, y_L) \geq T(y_H) - T(y_L) \text{ (IFC)} \\
\mu & U(y_H - T(y_H)) - V(y_H/\theta_H) \geq U(y_L - T(y_L)) - V(y_L/\theta_H) \text{ (LSC)}
\end{align*}
\]

where \( \lambda, \gamma, \mu \) are the multipliers on the feasibility, income falsification, and labor supply constraints respectively. The Kuhn-Tucker theorem characterizes the possible solutions to this problem. The optimality of \( \Gamma^* \) means that no resources are wasted, or \( \lambda > 0 \), eliminating one possibility. The complementary slackness conditions on the incentive compatibility constraints allows for a total of four additional alternatives. However, one of these can be eliminated by noting that the first best allocation can’t be achieved in this economic environment.

**Theorem 3.5.** The social planner’s allocation are always constrained by at least one incentive compatibility constraint.\(^7\)

The logic behind this result can be illustrated by making a small change to the way the incentive compatibility constraints are written. Instead of inequality constraints, I can add two slack variables, \( \chi_{labor} \) and \( \chi_{falsification} \), to the system of incentive compatibility constraints and rewrite them as equalities. These variables are defined as the difference between high ability agents’ utility level if they deviate from their income allocation and taxes. For the sake of clarity, assume that the utility functions are quasi-linear, \( W(c_i, l_i) = c_i - V(l_i) \):

\[
\frac{y_H - y_L}{\Delta \text{Consumption}} - \frac{T(y_H) + T(y_L)}{\Delta \text{Taxes}} - \frac{V(y_H/\theta_H) - V(y_L/\theta_H)}{\Delta \text{Labor}} = \chi_{labor} \geq 0
\]

\[
\frac{\eta |y_H - y_L|}{\Delta \text{Cost}} - \frac{T(y_H) + T(y_L)}{\Delta \text{Benefit}} = \chi_{falsification} \geq 0
\]

The first variable, \( \chi_{labor} \), is the cost of changing high ability agents’ labor supply choice to display the same amount of income as a low ability type agent. The first term in this equation is the change in consumption of a high ability agent that changes his labor supply

\(^7\) The proof is found in the appendix.
to mimic low ability agents. This difference is positive and is the cost that he pays for this deviation strategy. The next two terms are the benefit of this strategy, and is the tax benefit (not paying taxes and receiving transfers) and extra leisure gained by working less. Similarly, the second variable, $\chi^{\text{falsification}}$, balances the utility cost of income falsification with the tax benefit of displaying the income level of low ability agents. For rational agents, the optimal deviation strategy is the one which cost the least. Therefore, the binding constraint for the social planner for any level of $g(\cdot)$ is

$$\min \{ \chi^{\text{falsification}}, \chi^{\text{labor}} \}.$$  

This means that the social planner will be at least constrained with one incentive compatibility constraint. If not, the marginal benefit of tightening the constraints is positive—when the Lagrange multipliers $\gamma$ and $\lambda$ are positive—and the social planner could increase welfare by changing the income allocation and taxes until high ability agents are indifferent between deviating through the cheapest option.

This logic narrows the solution to three complementary slackness alternatives, either:

Region 1: $\gamma > 0$, $\mu = 0$ - only the income falsification constraint binds

Region 2: $\gamma > 0$, $\mu > 0$ - the income falsification and labor supply constraints bind.

Region 3: $\gamma = 0$, $\mu > 0$ - only the labor supply constraint binds

These three conditions outline three incentive compatibility regions, each with a different solution to the social planner’s problem depending on the cost of income falsification. As this cost rises, the solution moves through Region 1 to Region 2, and eventually to Region 3. For each of these regions there is a closed form solution for marginal taxes, and a numerical solution for the entire mechanism. The complete numerical solution needs some further assumptions, so I will discuss marginal taxes first.

### 3.3 Marginal Taxes and Redistribution

In a model with two types of agents the tax schedule is a stepwise function in the C-Y space with diagonal sides rising at a slope of the net of tax rate of the agent to the right, immediately followed by flat sections, forming a kink at the income and consumption bundle of low and then high ability agents. This encourages agents to self-select into their respective income and consumption bundles per the proper implementation of the mechanism. When I refer to marginal tax rates, $\tau$, I refer to one minus the slope of the indifference curve (the next of tax rate) of each type of agent at the kink, which is not defined as a marginal rate in the tax schedule since the first derivative of the tax schedule function is not defined at this point. However, this is how the marginal tax rate is defined in a model with a continuous distribution of ability types. Please note that I follow this convention throughout this paper.

For $g_{yH}', g_{yL} \in (0,1)$ and marginal taxes are:

---

8. The way marginal tax rates $\tau$ are defined here are more akin to an average tax rate

9. If the marginal cost of deviation is less than zero, there is no way the social planner could prevent income falsification. On the other hand, values above one mean would mean it’s too costly, and income
Region 1:
\[
\begin{align*}
\tau_H &= - (1 - \pi) g_{yH} \left( \frac{U_L'}{U_H'} - 1 \right) < 0 \\
\tau_L &= \pi g_{yL}' \left( 1 - \frac{U_H'}{U_L'} \right) > 0
\end{align*}
\]

Region 2:
\[
\begin{align*}
\tau_H &= - (1 - \pi) g_{yH} \left( \frac{(1 - \pi - \mu) \pi U_L'}{(\pi + \mu)(1 - \pi) U_H'} - 1 \right) < 0 \\
\tau_L &= 1 - \left( \frac{1 - \pi - \mu}{1 - \pi - \mu \frac{V_H}{V_L}} \right) \left[ 1 + \pi g_{yL}' \left( 1 - \frac{(\pi + \mu)(1 - \pi) U_L'}{(1 - \pi - \mu) U_H'} \right) \right] > 0
\end{align*}
\]

Region 3:
\[
\begin{align*}
\tau_H &= 0 \\
\tau_L &= 1 - \left( \frac{1 - \pi - \mu}{1 - \pi - \mu \frac{V_H}{V_L}} \right) > 0
\end{align*}
\]

In Region 1, marginal taxes on high ability agents are negative to encourage them to display all of their income. A reduction in lump-sum taxes at the margin means that high ability agents gain by reporting any additional unit of income, while on the other hand, hiding that income would cost them resources. Low ability agents face positive marginal taxes to align their actions with the incentive-feasible mechanism; they don’t have need any incentives to display their income in full since they are receiving tax transfers upon their identification.

As the cost of income falsification increases, the solution reaches Region 2. This region also features negative marginal taxes on high ability agents for the same reason as before. However, in this region, the two incentive compatibility constraints bind, meaning that high ability agents are indifferent between using income falsification or labor supply responses to reduce their displayed income. Throughout this region, as the cost of income falsification increases marginal taxes for high ability agents also increase, converging to zero.

In this region marginal taxes for low ability agents are converging to the term
\[
1 - \left( \frac{1 - \pi - \mu}{1 - \pi - \mu \frac{V_H}{V_L}} \right).
\]

Finally, the expression directly above term and zero are the Region 3 solutions of marginal taxes for low ability agents and high ability agents respectively. This last region corresponds to the solution to a static Mirrleesian model with two agents, where the income falsification constraint is irrelevant. In this region all agents face the highest marginal taxes.

---

falsification would not be relevant.
Graphical Analysis

To analyze optimal taxes graphically, assume that \( g(y - x) = \eta|y - x| \) for any \( y > x \in \mathbb{R} \), and \( \eta \in (0, 1) \). This functional form models the cost of income falsification as the absolute distance between produced and displayed income with a constant slope representing the cost as fixed proportion of hidden income. This is a very natural representation of the cost of income falsification.

Starting with a familiar graph, Figure 1, the solution for Region 3 is characterized by sloping indifference curves; the high ability agent’s curve is always flatter than the low ability agent’s curve for any level of income because \( V'(y_i/\theta_i)/U''(c_i) = \theta_i \). The binding labor supply constraint is illustrated by the intersection of the indifference curves, or Point A. At this intersection high ability agent are indifferent between their consumption-income bundle and that of the low ability agents. In this region, this is the incentive compatibility constraint limiting redistribution; the vertical distance from Point A and the 45 degree line is the total lump-sum transfer given to the low ability agent. Consequently, the tax to the high ability agent is the vertical distance between the 45 degree line and Point B where the slope of the indifference curve of high ability agents is one. Point B maximizes tax revenue coming from high ability agents (or the distance between the 45 degree line and the high ability type’s curve) and maintains incentive compatibility. If high ability agents’ lump-sum taxes were any larger, ceteris paribus, their consumption-income bundle would be below the optimal allocation, giving them the incentives to deviate and produce the income of low ability agents. Similarly, any income allocation to the right or left of Point B on the indifference curve would yield less tax revenue, making it suboptimal. As mentioned before, marginal taxes are defined as one minus the slope of the indifference curves at Points A and B. For high ability agents, this implies zero marginal taxes. For low ability agents the slope is well below one, resulting in positive marginal taxes.
Figure 1: Region 3 - Binding labor supply constraint.

Figure 2: Region 3 - Non-binding income falsification constraint.
In Figure 2, I superimpose on Figure 1 a green line representing the IFC. This line, with a slope of \(1 - \eta\), proposes an allocation based on the distributional constraints of income falsification as an alternative to the optimal allocation. Points C and D mark an allocation that would allow more income redistribution than Points A and B—as evidenced by their greater vertical distance from the 45 degree line to the level of consumption of both agents at the same level of income. However, choosing this alternative allocation would result in a violation of incentive compatibility—high ability agents would be better off by producing and consuming the income-consumption bundle of low ability type agents. Therefore, this alternative allocation does not meet all of the requirements of the optimal mechanism. This is what I meant when I asserted that the income falsification constraint is irrelevant in Region 3; the high ability agents’ cost of income falsification is higher than the cost of changing their labor supply, implying that the binding constraint for the optimal mechanism is the LSC.

Figure 3: Region 2 - Binding income falsification and labor supply constraints.

For Region 2, illustrated in Figure 3, the green line (IFC) is steeper (representing a lower cost of income falsification) and runs through the intersection of the agents’ indifference curves, Point E. This intersection marks the equivalence of either constraint in how they limit the redistribution efforts of the social planner. The vertical distance from Point E to the 45 degree line marks the transfers to low ability agents, just like the vertical distance from Point F marks the amount of taxes high ability agents need to pay. Note that the
slope of the indifference curve at Point F is clearly above one, implying negative marginal taxes for high ability agents. They need this incentive to increase the amount of income they produce. Since all tax revenue comes from them, increasing the amount of income they produce relaxes the limits to redistribution set by the IFC. This is illustrated by the growing difference between the green line and the 45 degree line as income increases. This could be used as an analogy for the tax base: as high ability agents’ income increases, there is a larger pool of income to be taxed, which in turn is redistributed to low ability agents.

Figure 4: Region 1 - Binding income falsification constraint.

In the last region, Figure 4, the green line is below the intersection of the indifference curves making it the binding constraint on redistribution. As in the previous cases, the intersections of this line, Points G and I, with the indifference curves of both high and low ability agent marks the incentive-feasible allocations for both types. Just like in Region 2, the slope of the indifference curves at the intersection points defines marginal taxes: high ability agents face negative marginal taxes since Point I is almost always above the point with a slope of one, and low ability agents almost always face positive marginal taxes. I say almost always because throughout this region as the cost of income falsification tends towards zero the slope of this line increases towards one, making income transfers more difficult. At the extreme, when $\eta = 0$, the slope of the green line is one (overlapping the 45 degree line) resulting in zero optimal lump-sum and marginal taxes for both types of agents.
4 Numerical Solution

The complete solution for this model requires the numerical computation of lump-sum taxes and multipliers. For this purpose, I assume the same functional form for the cost of avoidance as in the graphical analysis in the previous section, $\eta|y-x|$, and the utility from consumption and labor takes the form:

$$U(c_i) - V(l_i) = \frac{c_i^{1-\gamma} - 1}{1-\gamma} - \frac{\alpha}{\sigma} y^\sigma.$$  

The parameter $\gamma$ determines the degree of curvature of the utility from consumption, $\alpha$ is the relative weight of leisure to consumption utility, and $\sigma$ determines the agents’ compensated labor supply elasticity, $\epsilon_c = \frac{1}{1-\sigma}$. The values for these parameters are $\gamma = 1.5$, $\alpha = 2.55$ and $\sigma = 3$ respectively. The value of sigma implies an elasticity of labor supply somewhere in between the upper limit of micro-estimates, and conventional macro estimates, $\epsilon_c = 0.5$. The value of alpha is chosen to match the average percentage of hours that are spent on work from the NLSY - approximately 8 hours. Finally, workers ability or productivity, will be represented by hourly wages: $\theta_L = 40$ and $\theta_h = 100$. These values of $\theta$ are the 90th and 99th percentiles of the US wage distribution respectively.

Optimal mechanism as a function of the cost of income falsification

Under these assumptions the optimal mechanism is a function of $\eta$. Figure 5 shows marginal taxes for different values of this parameter. When $\eta = 0$, marginal taxes for high and low ability agents start at zero. For high ability agents, as $\eta$ increases through the range $(0,1)$ marginal taxes become negative and then return to zero. This path is shaped by two forces working in opposite directions: the marginal cost of income falsification $g'_y = \eta$, and the utility inequality among the agents $(U'_L/U'_H - 1)$. Low ability agents face monotonically increasing marginal taxes through this range of values for $\eta$, reaching their height at Region 3, shaped by the same forces influencing marginal taxes for high ability agents. Note that after $\eta$ reaches the point where $\chi^\text{labor} \geq \chi^\text{falsification}$ (Region 3) there is no change to the optimal allocation, and $\eta$ becomes irrelevant. In the same figure, graphed along with marginal taxes is a line representing $\eta$. Its purpose is to illustrate that high ability agents’ marginal tax may never exceed the marginal cost of income falsification. The logic is simple, if the social planner imposes marginal taxes above the marginal cost of income falsification, agents would have incentives to hide their income. As stated by the Non-falsification Theorem (3.1), income falsification is not optimal, and the social planner could improve the welfare of at least one type of agents by incentivizing complete income displays.
Figure 5: Marginal taxes for different values of $\eta$.

Figure 6: Indirect utility for different values of $\eta$. 
Figure 6 shows the path of aggregate welfare and high and low ability agents’ indirect utility. As $\eta$ increases the agents’ indirect utility gets closer to each other—as expected when using a utilitarian objective function. Aggregate welfare slowly increases to reach its peak in Region 3. This welfare equalization process hinges on the social planner’s capacity to redistribute income, which is in itself a function of the cost of income falsification. Throughout Region 1, welfare increases as the agents’ indifference curves shift away from the 45 degree line in the consumption-income space. This process is a combination of how the income allocation and taxes change as $\eta$ increases. As the green line in Figure 4 becomes flatter, the income falsification constraint is relaxed, and the social planner can impose larger lump-sum taxes on high ability agents. To increase the size of these lump-sum taxes the social planner also encourages more labor effort from them. The result of this incentive can be seen in Figure 7. Because high ability agents are encouraged to increase their labor effort for low values of $\eta$, income increases monotonically throughout Region 1, reaches its peak while passing through Region 2, and then retreats to the level in Region 3—precisely as negative marginal taxes approach zero. This figure also shows the opposite effect on low ability agents, who receive transfers and work less as $\eta$ increases.

Figure 7: Income-consumption bundle for different values of $\eta$. 

![Figure 7: Income-consumption bundle for different values of $\eta$.](image)
This process has a direct impact on aggregate output, which is illustrated in Figure 8. Aggregate output increases monotonically throughout Region 1, and stays above its level in autarky through most of Region 2. As \( \eta \) increases towards Region 3 output falls below autarky levels. Another consequence of the welfare equalization process is the path of income and consumption inequality. As \( \eta \) moves in the \((0, 1)\) range, income inequality between high and low ability agents increases, see Figure 9. This is a result of the incentives provided by marginal taxes, and shows that income inequality does not translate directly into welfare inequality. Similarly, consumption inequality is almost always decreasing. Although it’s difficult to see in the graph, consumption inequality increases for very low values of \( \eta \) and then starts to descend until its lowest level in Region 3. For these values, high ability agents have to be compensated to report all of their income, which might result in greater consumption inequality as compared to autarky (or \( \eta = 0 \)).

Figure 8: Aggregate expected income for different values of \( \eta \).
Optimal mechanism as a function of the ability dispersion

The previous exercise takes the agent’s productivity levels as given. The analysis would be incomplete without studying how the incentive compatibility regions change as differences in agents’ productivity increase. Figure 10 fixes the level of productivity of low ability agents, and increases the level of high ability agent’s productivity. The shaded section between the two black lines represents the incentive compatible Region 2. This is a correspondence between $\theta_H$ and $\eta$ for which high ability agents would be indifferent between income falsification and labor supply changes as a way to avoid taxation. In this graph, as the difference in ability between agents grows, so does the proportional cost of income falsification ($\eta$) the agents would be willing to pay. In this region, as the high ability agents’ income rises, the option to reduce how much tax they pay using income falsification opportunities becomes more attractive compared to the cost of changing their labor supply—as indicated by the larger cost they are willing to pay. This model characteristic is consistent with a low elasticity of labor supply to tax changes, and positive (and possibly large) elasticity of tax avoidance to changes in taxes at the top of the income distribution. This result can also provide

\[ \text{See Saez (2004)} \]
an idea of the deadweight losses incurred by taxpayers who engage in income falsification. Since evidence points out that income falsification is an important factor at the top of the income distribution, wealthy taxpayers are spending a large proportion of their falsification gains in unproductive activities. These expenditures increase the deadweight loss caused by a tax system that allows for income falsification. Note that Figure 10 was constructed using conservative values for the elasticity of labor supply. Accordingly, these estimates of $\eta_e$ are conservative as well. It is possible that these costs are much higher for some taxpayers. For instance, reducing the labor supply of workers that are highly attached to the workforce or are subject to mandatory work week laws can be very costly. For these taxpayers, the lack of opportunities to reduce how much tax they pay through their labor supply can make income falsification very attractive.

Figure 10: Aggregate expected income for different values of $\eta$.

5 Discussion

Income as private information and exogenous income falsification

Taking agents’ ability as private information and treating their labor supply decisions as independent from the social planner is a standard practice. Compulsory labor has been outlawed throughout the world, and using an individual’s ability as the basis for taxation is either difficult due to informational constraints or not appropriate under tax laws. The
argument for private income information and income falsification is not as established. One may argue that the social planner should be able to “fix” the problem first by eliminating income falsification, and then search for the optimal tax schedule. Can’t a social planner just get rid of tax loopholes and exceptions? Or impose a system of income reporting that does not give agents the opportunities for income falsification? This position argues that the social planner should have power to manipulate all the structure of the tax system, including the tax code, when making normative prescriptions. If this argument is valid, then taking portions of the structure of the tax system (like tax loopholes and opportunities for tax evasion) as exogenous is inconsistent with a description of how the world should be.

Curiously, this is the same argument used to question whether ability should be modeled as private information after the work of Mirrlees. I cite the subsequent defense of his work to justify income as private information. The idea was explained succinctly by A.B. Atkinson, who noted that there exists a “status of different types of information” that need to be taken into account when in designing optimal income taxes. Information varies in the difficulty that tax authorities have in acquiring and using it for tax purposes. If information about a person’s ability is either very difficult to collect or is not admissible for tax design purposes, then it’s justified as exogenous when designing the optimal tax schedule. If information about taxpayers’ income meets similar criteria, then it’s justified by the same argument. I divide these criteria into two categories: technical and technological difficulties of acquiring information about taxpayers’ income, and the legality or appropriateness of using to levy taxes. Like determining the ability of a taxpayer, gathering accurate information about taxpayers’ income can be difficult: governments need to obtain, store, and service large amounts of data, which can be technically and technologically demanding—like establishing effective institutions and methods of collecting information, or enacting proper compliance statutes and auditing organizations. There are also legal and political barriers. The tax code, with all of its regulations and clauses, is generally designed by legislative bodies under the influence of many political, social, and economic forces; it’s only appropriate to take some of its features as exogenous if these forces are not explicitly modeled. Ultimately, the right categorization of the different types of information depends on the context. The normative focus of this paper is not to describe the optimal tax system, rather it’s to find optimal tax schedule and outline some of the consequences of implementing it. Designing the optimal tax system (with all the laws and features of tax collection and compliance) would require political and technological considerations outside of the scope of this paper.

Implementing the optimal tax schedule

By modeling alternative ways in which taxpayers respond to taxation, this paper answers some of the questions about implementing Mirrleesian optimal taxes. A successful implementation of a tax mechanism in a decentralized economy should take into consideration all the factors influencing agents’ relevant tax related decisions. Otherwise, taxpayers in real life would not self-select into the optimal mechanism, choosing instead another alternative over-

\[\text{11} \text{Tuomala (1990), pg. 58.}\]
looked by the mechanism’s designers. Hence, optimal taxes have to conform to the reality and incentives motivating taxpayers’ behavior. These limitations suggest that contrary to one of the tenets of the Mirrleesian literature, the optimal tax function should be disciplined by the characteristics of the tax system, namely, tax laws and income reporting structure that give taxpayers’ opportunities for income falsification. This is precisely the case in the model presented in this paper. In Regions 1 and 2 the tax function is limited by the cost of income falsification which represents agents’ income falsification opportunities. For this reason, the tax schedule developed here has significant policy implications since it incorporates these important aspects of taxpayers’ behavior observed in the real world.

6 Conclusion

Evidence shows that taxpayers respond to the incentives created by the tax system, both by the tax schedule and the opportunities for income falsification found in the tax code. Yet the normative tax literature does not model income falsification in their design of the optimal tax schedule. This article fills this gap by developing a costly income falsification framework for the normative analysis of taxation, and explores its policy prescriptions.

The inclusion of income falsification in these models emphasizes the role of incentives. The optimal tax schedule should provide taxpayers the incentives for complete income reports. By doing so, a government could increase their tax revenue while at the same time reducing unproductive activities in their economy. In cases where income falsification is the most affordable way to reduce income tax liability (IC Regions 1 and 2), marginal taxes are lower than those found in similar models of optimal taxation—low marginal taxes give agents the incentives to produce and display all of their income. Otherwise, taxpayers would find the cheapest way to avoid or evade taxes while causing dead weight losses.

There are some logical extensions of the model that would enhance its ability to make relevant policy recommendations. First, there are groups in the population, such as those self-employed, who are able to use income falsification at lower costs than other taxpayers. In the context of the model this means that taxpayers are heterogeneous in their ability to use income falsification technologies. This heterogeneity might have an important impact on the tax burden placed on different types of agents, welfare, and redistribution. Another issue that was not covered is the functional form of the cost of income falsification. In a model with only two types of agents, there is no need to worry about the functional form of the cost of income falsification. This is not true when there three or more types of agents. In this case, the characteristics of the optimal mechanism depends directly on the shape of $g(\cdot)$. Unfortunately, there are no estimates of the cost of income falsification function. However, even without fully characterizing the tax schedule, the current model presents basic principles that are not going to change whether there are two, three, or an infinite number of types of agents.
References


Review of Economic Studies, 38(114): 175–208. 2


7 Proofs

Definition 1. For any mechanism \( \Gamma = \{y_i, x_i, T(x_i)\}_{i=H,L} \) the net benefits of deviation through income falsification (IF), labor supply (LS), or both choices (DD) for \( i \in \{H, L\} \) are:

\[
\begin{align*}
\chi_{i}^{IF} &= g(y_i, x_j) - g(y_i, x_i) - \geq T(x_i) + T(x_j) \\
\chi_{i}^{LS} &= U(y_i - g(y_i, x_i) - T(x_i)) - V(y_i/\theta_i) - U(y_j - g(y_j, x_j) - T(x_j)) + V(y_j/\theta_i) \\
\chi_{i}^{DD} &= U(y_i - g(y_i, x_i) - T(x_i)) - V(y_i/\theta_i) - U(y_j - g(y_j, x_i) - T(x_i)) + V(y_j/\theta_i)
\end{align*}
\]

Lemma 7.1. Claim: The tax function in the optimal mechanism is increasing in income displayed and satisfies \( \chi_{i}^{j} > 0 \) for \( i \in \{H, L\} \) and \( j \in \{IF, LS, DD\} \).

Proof. As a baseline, define \( c_{i}^{a} = y_{i}^{a} \) to be the autarky solution to the social planner, where \( T(x_i) = T'(x_i) = 0 \) for \( i \in H, L \). For any \( \pi \), the first order conditions of this problem and the fact that \( \theta_H > \theta_L \) imply that \( y_H^a > y_L^a \). Given that preferences and productivity are not affected by taxation, \( y_H > y_L \) is the optimal relationship between income for agents of high and low ability in the following proposed mechanisms. First, suppose the optimal mechanism is \( \tilde{\Gamma} = \{\tilde{y}_i, \tilde{x}_i, T(\tilde{x}_i)\}_{i=H,L} \) features a decreasing tax function, so that \( T(\tilde{x}_H) < 0 < T(\tilde{x}_L) \). Any such mechanism should satisfy \( y_H = x_H \), since any other choice for income display would reduce the consumption of high ability agents, who receive a tax transfer upon identification. Additionally, for any such mechanism to induce a separating equilibrium between the agent-types, it must be that \( \chi_{i}^{j} > 0 \) for \( j \in \{IF, LS, DD\} \). Otherwise low ability agents would deviate to the allocation of high ability agents and not tax revenue could be collected (if no tax revenue is collected, the optimum is the autarky solution and \( \tilde{\Gamma} \) is ruled out as the optimum). However, even if these conditions are met, any such mechanism would reduce aggregate welfare versus the autarky solution due to the concavity of the agent’s utility function:

\[
\pi W(\tilde{c}_H, \tilde{y}_H/\theta_H) + (1-\pi)W(\tilde{c}_L, \tilde{y}_L/\theta_L) < \pi W(c_{H}^{a}, y_{H}^{a}/\theta_H) + (1-\pi)W(c_{L}^{a}, y_{L}^{a}/\theta_L),
\]

which is contrary to the welfare maximization objectives of the social planner. This requires the optimal mechanism to utilize a tax function that is increasing in displayed income, or \( T(\tilde{x}_H) > 0 > T(\tilde{x}_L) \). However, for symmetrical reasons as to the case of \( \tilde{\Gamma} \), \( \chi_{i}^{j} > 0 \) for \( j \in \{IF, LS, DD\} \), otherwise no tax revenue can be collected and agents are better off under autarky.

Proof of the Non-falsification Theorem (3.1). Starting with high ability agents, suppose the falsification mechanism \( \Gamma_f = (x_H, x_L, y_H, y_L, T(x_H), T(x_L)) \), where \( x_H \leq y_L \), is a solution to the social planner’s problem. The optimality of non-falsification requires that there exists a non-falsification mechanism \( \Gamma_n = (\hat{x}_H, x_L, y_H, y_L, \hat{T}(\hat{x}_H), T(x_L)) \), where \( \hat{x}_H = y_H \) and
\begin{align*}
g(y_H, \hat{x}_H) = g(y_H, y_H) &= 0, \text{ that satisfies the feasibility and incentive compatibility constraints and improves aggregate welfare compared to } \Gamma_f. \\
\text{In order for } \Gamma_{nf} \text{ to satisfy the feasibility constraint (3.6), it must collect at least the same amount of revenues, therefore it must be that}
\hat{T}(\hat{x}_H) \geq T(x_H). \\
\text{Furthermore, in order for } \Gamma_{nf} \text{ to satisfy the left hand side of the incentive compatibility constraints (3.3), (3.4) and (3.5) for the high ability agent } (i = H, j = L), \text{ it has to provide at least the same level of utility as } \Gamma_f:\nU(y_H - g(y_H, \hat{x}_H) - \hat{T}(\hat{x}_H)) \geq U(y_H - g(y_H, x_H) - T(x_H)),
\end{align*}

which means that taxes must satisfy
\begin{equation}
\hat{T}(\hat{x}_H) \leq g(y_H, x_H) + T(x_H). 
\end{equation}

Both conditions on \( \hat{T}(\hat{x}_H) \) satisfy the right hand side of equations (3.3) and (3.4) of high ability agents trivially. Similarly, it satisfies equations (3.3) and (3.5) for low ability agents \( (i = L, j = H) \) since by lemma 7.1 taxes are an increasing function of income reports. This deviations would unambiguously make the low ability agent worse off by reducing consumption through increasing the cost of income falsification and a larger positive tax.

Lastly, to satisfy the double deviation constraint (3.5) of high ability agents we need that the level of utility of the low type deviator under \( \Gamma_{nf} \) be at most the amount of utility received in \( \Gamma_f \), which is satisfied if:

\begin{align*}
U(y_H - g(y_H, \hat{x}_H) - \hat{T}(\hat{x}_H)) \leq U(y_H - g(y_H, x_L) - T(x_H)) \\
y_H - g(y_H, \hat{x}_H) - \hat{T}(\hat{x}_H) \leq y_H - g(y_H, x_L) - T(x_H) \\
\hat{T}(\hat{x}_H) \leq g(y_H, x_H) + T(x_H) 
\end{align*}

Equations (7.1), (7.2) and (7.3) together imply that taxes under non-falsification for high ability agents must be
\begin{equation}
\hat{T}(\hat{x}_H) = g(y_H, x_H) + T(x_H)
\end{equation}

which satisfies incentive compatibility and feasibility constraints as well as collect more tax revenue. This tax revenue can be redistributed to the agents in a lump sum, increasing their overall utility levels. Hence, \( \Gamma_{nf} \) Pareto dominates \( \Gamma_f \), a contradiction of our assumptions and proving the claim. A symmetric argument can be made for the low type agents, but its worth noting that since they receive a tax transfer upon identification they have no incentive for falsification and their optimal income display is equal to their true income. The case against
any falsification mechanism where \( x_H \geq y_L \) relies in noting that all the social planner needs to do is change the income display recommendation to a non-falsification allocation. No change in taxes is needed to make them better off since displaying income larger than their true income reduces their consumption unambiguously without gaining any tax advantage.

\[ \]

**Proof of Corollary 3.2.** The IFC can be simplified by eliminating the utility of labor \( V(\cdot) \) from both sides of the inequality, and setting \( g(y_i, x_i) = g(y_i, y_i) = 0 \) using theorem 3.1.

\[ U(y_i - T(y_i)) \geq U(y_i - g(y_i, y_{-i}) - T(y_{-i})) \]
\[ y_i - T(y_i) \geq y_i - g(y_i, y_{-i}) - T(y_{-i}) \]
\[ -T(y_i) \geq y_i - g(y_i, y_{-i}) - T(y_{-i}) \]
\[ T(y_i) - T(y_{-i}) \leq g(y_i, y_{-i}) \]

**Proof of Corollary 3.3.** Subtract equation (3.4) from (3.5), and apply theorem 3.1. Eliminating the right hand side of both equations and the utility from labor effort which are shared amongst both constraints we have the following:

\[ U(y_i - T(y_i)) \geq U(y_i - g(y_i, y_{-i}) - T(y_{-i})) \]
\[ y_i - T(y_i) \geq y_i - g(y_i, y_{-i}) - T(y_{-i}) \]
\[ -T(y_i) \geq y_i - g(y_i, y_{-i}) - T(y_{-i}) \]
\[ g(y_i, y_i) \geq T(y_i) - T(y_{-i}) \]

which is precisely the simplified expression for the report constraint. As long as equations (3.3) and (3.4) are constraints to the social planner’s problem there is no need to include the double deviation constraint.

**Proof of Proposition 3.4.** Lemma 7.1 guarantees that \( T(y_H) \geq 0 \geq T(y_L) \). Together with the non-negativity restriction on the cost of avoidance, this condition implies that the report constraint never binds for the low ability agent.

\[ T(y_i) - T(y_h) \leq 0 \leq g(y_i, y_h) \]

The second claim in this proposition follows directly from the the single crossing property. It guarantees that only the high ability type’s labor constraint could bind at the optimum.

**Proof of Theorem 3.5.** Update the net-benefit of deviation definition at the beginning of this section to reflect theorem 3.1: define a cost of deviation for the high ability agent using the
labor and income falsification constraints:
\[ U(y_h - T(y_h)) - V(y_h/\theta_h) - U(y_h - g(y_h, y_l) - T(y_l)) + V(y_h/\theta_h) = \chi_{IF} \geq 0 \]
\[ U(y_h - T(y_h)) - V(y_h/\theta_h) - U(y_l - g(y_l, y_l) - T(y_l)) + V(y_l/\theta_h) = \chi_{LS} \geq 0 \]

Replace this slack constraints in place of the inequality incentive compatibility constraint and set the latter to equality constraints. The social planner’s problem is now constrained by these costs of deviation. Suppose that there exists an welfare-maximizing incentive-feasible mechanism \( \Gamma^* = (y_{H*}, y_{L*}, T(y_H)^*, T(y_L)^*) \) that solves the social planner’s problem when the incentive compatibility constraints are defined as above, and the multipliers in each of these constraints are zero. The complementary slackness conditions imply that the costs of deviation are positive or \( \chi_{LS} > 0 \), and \( \chi_{IF} > 0 \). However, the social planner can increase aggregate welfare by at least one of the values of the non-zero multipliers by revising the optimal mechanism until
\[ \min \{\chi_{IF}, \chi_{LS}\} = 0. \]
This contradicts the optimality of \( \Gamma^* \).