Efficient redistribution through the reduction of compliance costs

Mauro Bux, Università del Salento

Abstract

This paper analyzes in a simple two-type model (skilled and unskilled workers) the redistributive properties of reducing taxpayers' compliance costs, when an optimal nonlinear income tax scheme is implemented. It is shown that, depending on the distribution of skills, the government can efficiently redistribute income from the more able to the less able individual by providing services helping taxpayers to comply with tax. The results offer theoretical support to existing programs targeted to simplify income reporting and tax payment for low income taxpayers.

Keywords: compliance costs, tax administration, redistribution

JEL: H21, H23, H83
Introduction

This study analyzes the redistributive properties of reducing the taxpayers’ compliance costs, when an optimal nonlinear income tax scheme is implemented. Compliance costs are defined as the costs taxpayers bear in order to be compliant with their fiscal duties. They can have different nature, perhaps being psychological, monetary, related to time, and so on. The attention to this type of cost stems from the theoretical and empirical recognition that the negative consequences associated with taxation do not involve only direct distortion of economic decisions. Indeed, it is widely agreed that taxation causes four different categories of deadweight loss: namely, substitution, active non-compliance, administration and passive compliance. This work focuses on the latter.

Of course, taxpayer compliance costs are strictly related to the level of complexity of the tax system. Moreover the expected compliance costs might contribute to affect the individual decision to totally or partially dodge taxation, that is, to evade and/or avoid. For this reason, it is important to investigate how these costs affect taxpayers’ behavior, and how the latter, in turn, affects the government’s objectives.

During the last decades, a large literature has been produced on this topic, both from a theoretical and from an empirical point of view. In particular, economists have gradually integrated the notion of compliance costs within standard models of optimal taxation. Moreover, more comprehensive analyses have been conducted, in order to explore the interaction between players with contrasting objectives such as the taxpayers and the government. Empirical works have documented the magnitude of the phenomena at hand, and have established links on the causes that affect taxpayers compliance costs.

The growing interest on the topic under consideration is testified also by the numerous experiences of public policies implemented all over the world, in order to reduce taxpayers compliance costs. Indeed, recently, governments have been using different tools at their disposal (e.g., pre-populated returns)
with the twofold objective of making it easier for taxpayers to be tax compliant and, at the same time, decreasing tax dodging. One of the aim of this paper is showing how lowering compliance costs can be an effective and efficient way for the government to redistribute resources.

An introduction to the most relevant literature is presented in the first section, dealing with a comprehensive view on the impact of taxation on taxpayers, the government, and the economic activity. A particular emphasis is devoted to the description of the standard model of optimal taxation with compliance costs, since this is at the base of the model I present later in the paper. Finally, I present some of the empirical research conducted on the compliance costs problem. In particular, these papers focus on the measurement of these costs for the taxpayers, as well as on the impact of specific public policies aimed at reducing them.

The second section is dedicated to the theoretical analyses, and a series of models is described. First, I consider a framework where the government is restricted to levying only a linear income tax. From the second model on, I introduce the more realistic assumption that the government can set a nonlinear income tax schedule, which brings about the well known adverse selection problem. The latter turns out to be decisive in the very last model, where the government is able to relax the incentive compatibility constraint by introducing the compliance-cost-reducing measure. By doing so, the government is able to implement a redistributive tax policy, while at the same time, enhancing social welfare. Therefore, this result loosens the renowned tension between efficiency and equity, which is at the base of modern public finance analyses on the optimal tax problem. Of course, this conclusions heavily rely on some assumptions on taxpayers compliance costs, which are duly stated in the section dedicated to the models.

The last section concludes with a summary of the results and a discussion related to their relevance on the implementation of specific public policies.
Review of the literature

A relatively recent strand of literature has posed attention on the necessity to consider a broad range of costly economic consequences due to taxation. This line of research recalls Adam Smith’s classification of the direct and indirect costs generated by taxation, described in his notable work "An Inquiry into the Nature and Causes of the Wealth of Nations". From Smith’s lines we can identify four major costs associated to taxation: substitution, active non-compliance, administration and passive compliance. The first two can be considered to be indirect costs, whereas the others are direct: this is merely because substitution and active non-compliance require some individual action in order for the cost to be born. The substitution effect has to do with the impact taxation has on individual economic decisions. It might involve consumption, investments, labor offer, and more generally any situation where the taxpayer has to decide how to allocate a scarce resource at her disposal, which is subject to a specific form of taxation. Active non-compliance refers to taxpayers attempt to dodge taxation as much as they can. In particular, it is commonly agreed that evasion is about deliberately (and illegally) deciding not to pay taxes, whereas avoidance refers to the (legal) exploitation of tax code flaws in order to reduce one’s tax liability. Administrative costs concerns the public expenditure needed to enforce the tax code. Finally, compliance costs are the ones taxpayers bear in order to obey their fiscal duties.

Yitzhaki (1979) represents an early attempt to conduct a theoretical research on the economic consequences of taxation with a more comprehensive approach. In particular, the author concentrates on commodity taxation when administrative costs are taken in consideration. He stresses how the excess burden and the administrative cost can be viewed as substitutes, which results in the optimality condition that equates the marginal cost of raising the administrative cost with the marginal cost of raising the tax rates. A similar approach is followed in Azienman (1987) where it is found that administrative
costs in the collection of revenue can justify the application of both tariffs and inflation tax. However, the author assumes that enforcement costs of tariffs and inflation tax are smaller than alternative taxes, which might not be the case for all the countries depending on the frequency of use of the two above mentioned tax instruments.

One of the most influential paper in the same is Mayshar (1991), where the concept of tax technology is used in order to study the economic implications of both administrative activity by the government and tax-resisting behavior by the taxpayers. Mayshar introduces new elements in the tax technology, eventually affecting the individual budget constraint as follows. First, the individual utility function is assumed to be defined over consumption and total work effect, the latter being made of three components: namely, hours worked $L$, tax-shielding effort $S$, and effort in imposed tax compliance $m$. Therefore, the taxpayers maximize their utility function

$$U(C, L + S + m(\theta)),$$

given the constraints

$$X = wL$$
$$C = X - T(X, S, \theta).$$

The tax technology is represented by the function $T(X, S, \theta)$, defined over the tax base $X$, the taxpayer tax-shielding activity $S$ (i.e., evasion and avoidance), and the vector of tax instruments $\theta$. The tax technology function is assumed to be increasing in the tax base and the tax instrument, while decreasing in the tax-shielding activity. The (passive) compliance cost $m = m(\theta)$ is a function of the tax instrument chosen by the social planner. Mayshar highlights that this specification of the tax technology also suits standard tax models. Indeed, if we take the vector $\theta$ to be a single variable chosen by the government, perhaps the tax rate $t$, then $T(X, S, t) \equiv tX$ and $m(\theta) \equiv 0$. Moreover, if we represent by $D(\theta)$ the average public direct cost when adopting the vector of tax instruments $\theta$, then the net tax revenue can be defined as
\[ R(X, S, \theta) = T(X, S, \theta) - D(\theta) \]. From the last definition, it is clear that what the social planner collects in terms of revenues is systematically less than what the implemented tax schedule and enforcement activity could potentially yield. This is precisely related to the concept of costly administration, which gave birth to Okun's renowned metaphoric expression of the "leaky bucket".

The individual maximization problem is then solved as follows. The optimal choice on labor supply for the representative agent is given by the first order condition:

\[
-\frac{U_L(C^*, L^* + S^* + m(\theta))}{U_C(C^*, L^* + S^* + m(\theta))} = w[1 - T_x(X^*, S^*, \theta)] = w^*.
\]

This is the familiar condition where the marginal rate of substitution between consumption and labor equates the net-of-tax wage rate, \( w^* \). The optimal tax-shielding activity, instead, is given by the following first order condition:

\[
w^* = w[1 - T_x(X^*, S^*, \theta)] = -T_s(X^*, S^*, \theta).
\]

This condition states that the net-of-tax marginal benefit from working should be equal to the marginal benefit from tax-shielding.

By solving the above cited maximization problem, we get the optimal values for all the individual selected variables - namely, \( L^*, S^*, X^*, C^*, U^* \) - as a function of \( \theta \) and \( w \). Mayshar then derives the reduced form tax technology function, \( T(\theta, w) \equiv T(X^*(\theta, w), S^*(\theta, w), \theta) \). As a consequence, the reduced form of the net tax revenue function is \( R(\theta, w) \equiv T^*(\theta, w) - D(\theta) \). In order to study the overall effect of marginally increasing the vector of tax instrument used by the social planner, we take the derivative of the tax technology function with respect to the tax instrument \( \theta \):

\[
T_i^*(\theta, w) = T_x(X^*(\theta, w))X^*_\theta + T_s(S^*(\theta, w))S^*_\theta + T_\theta.
\]

---

It is easy to see that the overall effect is composed of three distinct components. The first one, $T \times X^\star_\theta$, represents the distortive substitution effect and is expected to be negative whenever increasing the tax instrument discourages production effort, that is, $X^\star_\theta < 0$. The second term, $T \times S^\star_\theta$, is the effect on tax-shielding activity, expected to be negative. Finally, $T_\theta$ is the direct positive impact on gross tax revenue, the so called “mechanical effect”. Mayshar believes that the dominance of the positive direct effect for low levels of $\theta$, in conjunction with subsequent dominance of the indirect substitution and tax-shielding effect, might explain the typical hump shape of the Laffer curve (i.e., fiscal revenues as a function of the tax rate). The further step in Mayshar’s analysis is the use of the envelope theorem in order to express the impact of an increase in $\theta$ on individual utility, in terms of consumption units:

$$U^\star_i(\theta, w) = -[T_\theta(X^\star, S^\star, \theta) + w^\star m_\theta(\theta)].$$

On the other hand, by differentiating the reduced form net revenue function with respect to the tax instrument, we have:

$$R^\star_\theta(\theta, w) \equiv T^\star_\theta(\theta, w) - D_\theta(\theta).$$

Consequently, Mayshar proposes the notion of marginal cost of funds (MCF) for the $i$th instrument, defined as the taxpayer’s marginal welfare loss per dollar of marginal net revenue:

$$MCF_i = -\frac{U^\star_i / U_C}{R^\star_\theta} = \frac{T^\star_\theta(X^\star, S^\star, \theta) + w^\star m_\theta(\theta)}{T^\star_\theta(\theta, w) - D_\theta(\theta)}.$$
budget constraint. So, taxpayers rational choices result in the indirect utility function \( V(U^*(\theta, w), G) \), where \( G \) represents a given amount of public good which enters the individual indirect utility function in a weekly separable manner. Then, the social planner aims at maximizing the social welfare function \( W(U^*(\theta, w), G) \) under the budget constraint \( G = \frac{NR^*(\theta, w)}{p} \), where \( p \) is the unit cost of the public good and \( N \) is the total number of taxpayers. The first order condition for the optimal choice with respect to \( \theta \) equals the marginal cost of public funds, \( \text{MCF}_i \), to the marginal benefit of public funds, \( \text{MBF} \), defined as follows:

\[
\text{MCF}_i = \text{MBF} = \frac{NW_U/NU_C}{N}.
\]

This condition states that the increased social utility due to the production of a public good (MBF) should be equal to the social utility loss in collecting that unit of tax revenue. Since there are several ways the government can raise extra revenue (perhaps, applying different tax instruments from the vector \( \theta \)), the optimal policy includes equalizing all the marginal costs of funds related to every possible tax instrument. This, ultimately, implies collecting a given amount of revenue in the most efficient way.

There are at least three reasons why the work by Mayshar, described above, is of fundamental importance. First, it provides a very clear framework to analyze a broader set of economic consequences of taxation, transcending the mere substitution effect. Secondly, the paper under examination provides helpful policy advices, for example with respect to cost-benefit analysis on publicly funded projects. Lastly, Mayshar’s framework is also useful in order to examine specific phenomena which are related to tax administration (e.g., compliance, evasion, avoidance, complexity of the tax system). Slemrod and Yitzhaki (1996) provide some further contribution to the same topic. In particular, they argue that the definition of the marginal cost of public funds very much relies on the assumption that the cost borne by taxpayers when
they try to reduce their tax liability is exactly equal to the social cost. This assumption does not hold whenever we consider a taxpayer who is at a corner solution (for example, she has already evaded every tax she can) or when the agent hires an accountant whose cost is deductible from taxable income. Indeed, in both cases, the private cost would not be identical to the social cost.

Another major contribution to the field of tax enforcement is found in Kaplow (1990), which focuses on the relation between optimal taxation and tax enforcement. The basic question of Kaplow’s research is whether it is more efficient to raise tax rates or increase enforcement activity, in order to collect a given amount of revenue. In particular, the paper extends Ramsey’s analysis on optimal commodity taxation to include the presence of administrative and evasion costs. The author starts from recognizing that greater enforcement, not only distorts agents optimal economic choices by raising the effective marginal tax rates (EMTR), but it also implies greater administrative costs. Therefore, the intuitive conclusion should be that raising tax rates is the most efficient way for the government to collect extra revenue. However, Kaplow shows that when distortion rises disproportionally with the EMTR, relying on greater enforcement is more efficient than increasing the nominal tax rates. One of the keys of Kaplow’s framework is that enforcement makes evasion more costly to the taxpayer, eventually reducing this phenomenon and, consequently, raising tax revenue. This positive effect must then be confronted with the distortive impact brought about by enforcement, namely, the substitution effect. It is hard to predict which of the two contrasting effects overcomes the other; however, the main conclusion is that an optimal amount of enforcement is indeed efficient, which is in contrast with what one may intuitively think.

Keen and Slemrod (2016) contribute largely to this literature by introducing the concept of “enforcement elasticity of tax revenue”, that is, a sufficient statistic for taxpayers behavioral response to administration. The authors argue that this is an even more robust statistic with respect to the well known
elasticity of taxable income, since it fits also the case where concealment costs are partly transfers. The study at hand provides an interesting result: the optimal enforcement elasticity should be equal to the marginal compliance and administration costs relative to revenue raised. Moreover, on the same topic as in Kaplow (1990), the paper concludes that, the choice on whether it is more efficient to raise a given amount of revenue by increasing the tax rates or the enforcement activity, heavily depends on the elasticities of enforcement and taxable income, and on the shapes of compliance and administration costs. Finally, and most importantly, Keen and Slemrod call for more and better empirical knowledge on the topic, especially oriented to measuring the two above cited fundamental elasticities.

All the previous cited works, highlight the links between the tax instruments, the administrative costs to implement and enforce them, the compliance costs for the taxpayer, and the evasion and avoidance phenomena. With respect to compliance costs, both empirical and theoretical literature has been produced in the last decades. An econometric analysis on the Finnish experience with pre-populated returns is conducted in Kotakorpi and Laamanen (2014). The authors analyze data from the Finnish tax proposal experiment implemented between 1995 and 2004. This consisted of the same procedure explained in the previous paragraph, that is, a proportion of wage earning taxpayers received a prefilled return, based on third party information, and had the option to make adjustments or claim for discretionary deductions; otherwise, they could just confirm the prefilled return. The aim of the paper was measuring the effect of this policy on taxpayers reporting behavior. The authors find that receiving a partially prefilled income tax return increased taxpayers' tendency to report deduction items that are prefilled in the tax return, and reduced the tendency to report both income and deduction items that are not prefilled. This is explained by the function of compliance costs, assumed to be composed of a fixed part and a variable part which depends on the number of items reported by the taxpayer. This features make the compliance cost function first concave
and then convex, which might explain the results found in Kotakorpi and Laamanen (2014).

Prefilled returns are at the centre of the proposal for the United States tax system, described in Goolsbee (2006). Once again, the idea is for the IRS to exploit all the relevant information it already receives from third parties, in order to decrease taxpayers burden in being tax compliant. The proposed measure, which would involve nearly 40 percent of U.S. taxpayers, is estimated to potentially reduce the burden of tax compliance by more than $4 billion.

The usage of pre-filing by tax administrations around the world is becoming more and more frequent, in conjunction with the availability of appropriate online services. This is very well described in the 2015 biennial Tax Administration report by the OECD, where it is stated that 40 percent of revenue bodies, so far, report some use of pre-filing services.

To the best of my knowledge, this is the first study to theoretically analyze how the government can efficiently achieve its redistributive objectives, by levying an optimal nonlinear income tax scheme and introducing a measure aimed at reducing taxpayers’ compliance costs.

**The models**

In this section, I present a series of models analyzing the welfare consequences of the introduction of a public measure aimed at reducing taxpayers compliance costs. The models are different from one another, in that they assume the availability of various tax schemes at the government's disposal, as well as different features of the taxpayers compliance costs.

The first model represents an extension of the standard optimal linear income tax framework, where taxpayers are assumed to bear a cost - in terms of time
spent in filing their tax return - in order to be tax compliant. In order to study a more realistic framework, the second model extends the range of tax instruments at the government disposal to non-linear income taxation. The third model introduces an innovative element on the compliance costs borne by the differently skilled taxpayers. In this framework, a series of assumptions is stated, which clarifies how individuals with different abilities would optimally behave, and what is the consequent outcome in terms of social welfare. In this case, public intervention to reduce taxpayers compliance costs is found to be efficient, due to the relaxation of the incentive compatibility constraints. However, it should be stressed that the above mentioned result relies on specific assumptions which will be clarified during the description of the model.

**Optimal linear income taxation with fixed compliance costs**

In this model, the economy is composed of two types of representative agents differing solely in their abilities, as measured by their respective wage per unit of time, $w$. Since it is assumed that $w^G > w^B$, I shall refer to the type G agents as the "skilled", whereas the type B agents are the "unskilled". As we know, the presence of this heterogeneity among taxpayers makes government intervention distortive, since levying an income tax is expected to affect individual respective labor and consumption decisions.

Both types of agents are assumed to have the same preferences over consumption and leisure, as represented by the following utility function:

$$U(c_i, L_i)$$  \( \tag{1} \)

where $c_i$ and $L_i$ are the vector of consumption goods and the labor supply affecting the utility of a type $i$ agent, respectively. Usual assumptions on the properties of the above cited utility function apply:
with subscripts representing partial derivatives.

The social planner is assumed to be able to implement a simple tax scheme which involves a linear income tax and a uniform lump-sum tax. On the taxpayers side, instead, agents are assumed to bear a constant compliance cost, represented in terms of time spent in filling in the tax returns. Of course, by multiplying this time cost by the individual wage rate, the compliance cost becomes monetary, and most importantly, it is higher for the skilled agents with respect to the unskilled. After mentioning these assumptions, the individual budget constraint can be written as follows:

\[ c^i = w^i (L^i - m(1 - \theta))(1 - t) + b \]  

(2)

where \( m \) is the constant compliance cost, \( \theta \in [0,1] \) is a scalar representing some public policy aimed at reducing that cost, \( t \) is the income tax rate and \( b \) is the lump-sum tax (transfer).

As it's common procedure, taxpayers are assumed to be completely rational individuals, maximizing their utility function subject to their respective budget constraint. This amounts to the following maximization problem:

\[ \Omega = U(c^i, L^i) + \lambda [w^i (L^i - m(1 - \theta))(1 - t) + b - c^i] \]  

(3)

where agents optimally choose their pair of \((c^i, L^i)\) based on the following first order conditions:

\[ \frac{\partial \Omega}{\partial c^i} = U_c(c^i, L^i) - \lambda = 0 \]  

(4)

\[ \frac{\partial \Omega}{\partial L^i} = U_L(c^i, L^i) + \lambda w^i (1 - t) = 0 \]  

(5)

leading to the optimality condition:
\[ \frac{\partial U}{\partial L} = \frac{\partial L}{\partial c} = w'(1-t) \] (6)

From (3) and (4) we see that \( \lambda \) can be considered both as the budget constraint multiplier and as the marginal utility of income. From (5), we see that the reasoning behind the optimal choice on labor offer has to do with the disutility agents feel when spending effort in working, as compared to the benefit coming from enjoying a greater after-tax income. Condition (6) combines (4) and (5) stating that the optimal labor offer and consumption decisions, considering a given tax rate \( t \), equate the marginal rate of substitution between consumption and labor to the individual net-of-tax wage rate.

The agents' optimal choices are a function of individual net-of-tax wage rate and governmental lump-sum transfer, as follows:

\[ (c^*,L^*) = (c'(w'(1-t),I'),\hat{L}(w'(1-t),I')) \] (7)

Therefore, a new indirect utility function can be defined:

\[ V'(w'(1-t),I') \equiv U(c^*,L^*) \] (8)

depending on the net-of-tax wage rate, \([w'(1-t)]\) and on the lump-sum payment, the latter defined as:

\[ I' = b - w'm(1-\theta)(1-t) \] (9)

In particular, \( b \) is the governmental lump-sum tax (transfer), whereas the second term amounts to the net-of-tax "extra income" the taxpayer enjoys after receiving back a share, \( \theta \), of the total time spent in filing her tax return, assumed to be the constant entity \( m \) for any agent, multiplied by the agent's own wage rate, \( w' \).
On the government's side, the problem is to maximize social utility under the revenue constraint (including the resources needed to finance the policy aimed at reducing the individual compliance costs):

$$\Gamma = W(V^G, V^b) + \rho [t^G (L^G - m(1-\theta)) + t^b (L^b - m(1-\theta)) - 2b - \bar{R} - \theta m(w^G + w^b)]$$

(10)

where $\rho$ multiplies the government revenue constraint and $W$ is the social welfare function (increasing in both its arguments, that is, the skilled and unskilled agents indirect utility functions). From (8) and (9), we know that the government’s tax instruments enter taxpayers indirect utility function. In other words, the government can directly affect the agents levels of utility, by implementing a given tax policy. Therefore, the social planner is assumed to maximize (10) by choosing the optimal tax rate, $t$, and lump-sum tax (transfer), $b$, based on the following first order conditions:

$$\frac{\partial \Gamma}{\partial t} = \frac{\partial W}{\partial V} \left[ \frac{\partial V^G}{\partial t} w^G m(1-\theta) - \frac{\partial V^G}{\partial w^G} \right] + \frac{\partial W}{\partial V^b} \left[ \frac{\partial V^b}{\partial t} w^b m(1-\theta) - \frac{\partial V^b}{\partial w^b} \right] +$$

$$+ \rho \left[ t^G \left( \frac{\partial L^G}{\partial I^G} w^G m(1-\theta) - \frac{\partial L^G}{\partial I^G} \right) + w^G \left( \frac{\partial L^G}{\partial t} w^G m(1-\theta) - \frac{\partial L^G}{\partial w^G} \right) \right] = 0$$

(11)

$$\frac{\partial \Gamma}{\partial b} = \frac{\partial W}{\partial V^G} \frac{\partial V^G}{\partial I^G} + \frac{\partial W}{\partial V^b} \frac{\partial V^b}{\partial I^b} + \rho [t^G \frac{\partial L^G}{\partial I^G} + t^b \frac{\partial L^b}{\partial I^b} - 2] = 0$$

(12)

Equation (11) is the familiar condition for optimality in the choice of the linear tax rate. In particular, it makes sure the benefits and disadvantages of marginally raising $t$ are equalized. Indeed, increasing the tax rate affects the agents' indirect utility functions via their respective lump-sum payments and wage rates, as it is clear from the first two terms in (11). Equation (12) states the condition to optimally set the governmental lump-sum tax (transfer). It is clear to see how marginally increasing $b$ causes two main effects. First, it impacts positively the indirect utility functions of both agents via a greater lump-sum payment. On the other hand, this has a negative effect on the
revenue constraint, since a greater lump-sum payment gives both agents less incentives to increase their labor supply.

Let us now turn the attention to the parameter $\theta$, which is the fiscal instrument at the government disposal for reducing taxpayers compliance costs. If we start from a situation where $\theta$ is zero (e.g., no cost-reducing policy is implemented by the government) and marginally increase it while keeping the other tax instruments at the optimal level, we get:

$$\frac{\partial \Gamma}{\partial \theta} = \frac{\partial W}{\partial \theta} \frac{\partial V^G}{\partial \theta} w^G m(1-t) + \frac{\partial W}{\partial \theta} \frac{\partial V^B}{\partial \theta} w^B m(1-t) + \rho \left[ tw^G \frac{\partial L^G}{\partial \theta} w^G m(1-t) + tw^B \frac{\partial L^B}{\partial \theta} w^B m(1-t) + m(w^G + w^B) \right]$$

(13)

In order to interpret the result in (13) in a more clear manner, further steps are needed. By applying the envelope theorem, we know that

$$\frac{\partial V^i}{\partial I^i} = \lambda^i$$

(14)

$$\frac{\partial V^i}{\partial t} = -\lambda^i w^i (L^i - m(1-\theta))$$

(15)

where $\lambda^i$ is the private marginal utility of income for a type $i$ agent. By substituting (14) and (15) into (11), (12), and (13), the results become:

$$\frac{\partial \Gamma}{\partial t} = -\lambda^G \left[ \frac{\partial W}{\partial V^G} w^G (L^G - m(1-\theta)) \right] - \lambda^B \left[ \frac{\partial W}{\partial V^B} w^B (L^B - m(1-\theta)) \right] + \rho \left[ tw^G \left( \frac{\partial L^G}{\partial V^G} w^G m(1-t) + \frac{\partial L^G}{\partial w^G} \right) + tw^B \left( \frac{\partial L^B}{\partial V^B} w^B m(1-t) + \frac{\partial L^B}{\partial w^B} \right) \right] = 0$$

(16)

$$\frac{\partial \Gamma}{\partial b} = \lambda^G \frac{\partial W}{\partial V^G} + \lambda^B \frac{\partial W}{\partial V^B} + \rho \left[ tw^G \frac{\partial L^G}{\partial V^G} + tw^B \frac{\partial L^B}{\partial V^B} - 2 \right] = 0$$

(17)
By rearranging the first condition to optimally set the governmental lump sum tax (transfer) in (17), I get:

\[
\frac{\partial T}{\partial \theta} = \left[ \frac{\lambda^G}{\rho} \frac{\partial W}{\partial V^G} + tw^G \frac{\partial L^G}{\partial l^G} \right] w^G m(1-t) + \left[ \frac{\lambda^B}{\rho} \frac{\partial W}{\partial V^B} + tw^B \frac{\partial L^B}{\partial l^B} \right] w^B m(1-t) + (w^G + w^B) m(1-t)
\]  

\[
(18)
\]

(18)

where \(\alpha^G\) and \(\alpha^B\) are the net social marginal valuation of income for the skilled and the unskilled agent, respectively. Using the definitions in (19) into (18), the result becomes:

\[
\frac{\partial T}{\partial \theta} = \alpha^G w^G m(1-t) + \alpha^B w^B m(1-t) - (w^G + w^B) m(1-t)
\]

\[
(20)
\]

(20)

This shows the welfare consequences of the introduction of a compliance cost-reducing measure by the government. Analyzing whether the policy under examination is welfare enhancing implies reaching the condition for (20) to be positive, which after rearranging is:

\[
\alpha^G w^G + \alpha^B w^B > w^G + w^B
\]

\[
(21)
\]

(21)

or, equivalently,

\[
\frac{\alpha^G w^G}{w^G + w^B} + \frac{\alpha^B w^B}{w^G + w^B} > 1
\]

\[
(22)
\]

(22)

Since from (19)

\[
\frac{\alpha^G + \alpha^B}{2} = 1
\]

\[
(23)
\]

(23)

that is, the average net social marginal utility is equal to one, then condition (21) is infeasible. Therefore, It can be concluded that when an optimal linear
tax and an optimal individual lump-sum transfer are levied, no measure aimed at reducing the taxpayers’ compliance costs is efficient, in terms of welfare enhancement.

**Optimal non-linear income taxation with fixed compliance costs**

In the previous model, it was clear that the inefficiency of the governmental measure to reduce taxpayers compliance costs stemmed from the heterogeneity in the individual levels of ability for the two agents. In other words, for the same amount of time spent in filing the tax return, the skilled individual values that time more than the unskilled does. Moreover, relying only on a linear income tax scheme and on a lump-sum transfer, does not allow the government to achieve any redistributive goal.

In this paragraph a new model is presented, which assumes the social planner is able to implement a non-linear tax system. By doing so, the government can levy a progressive tax scheme, aimed at redistributing income from the skilled to the unskilled agent. This model is an extension of Stiglitz (1987), where fixed compliance costs are internalized in both types of agents’ budget constraints. As in the previous case, the different levels of ability of the two representative agents are measured by their respective wage per unit of time, \( w^G > w^B \). Also, preferences over consumption and labor are represented by the same utility function as the one considered in the linear taxation case. Importantly, the government is not able to observe neither individual abilities nor labor supplies; the only variable it can observe is gross earned income.

When the social planner aims at redistributing from the skilled to the unskilled agent, and the individual abilities and labor supplies are private information, an adverse selection problem arises. In that case, the skilled agent is tempted to choose the optimal levels of consumption and leisure of an unskilled, in order to enjoy a higher utility. In other words, if the non-linear
tax scheme is progressive, the skilled agent has an incentive to behave as if she was unskilled, in order to be taxed more lightly. However, if a skilled agent acts like an unskilled, this results in lower labor supply by the skilled, which of course is inefficient in terms of wasted utility in the economy. Therefore, the optimal non-linear income taxation should be the solution to a maximization problem where the social planner succeeds in effectively preventing the skilled agent from mimicking the unskilled. Since the government is assumed to ignore the individual abilities, it cannot rely on them to reach the optimal non-linear schedule. Therefore, it will set up its problem taking into account the only measure it can observe, gross income, defined as:

\[ Y^i = w^i (L^i - m(1-\theta)) \]  

(25)

where \( m \) is the fixed compliance cost and \( \theta \in [0,1] \) is a vector of public policies aimed at reducing this cost. This could be seen as a governmental tax filing service or, equivalently, as a public grant to be used solely for private fiscal advisory services. Notice that the assumption of the social planner’s ignorance about individual labor supplies is crucial, since if this didn’t hold, then the government could infer individual abilities and easily set an optimal income tax scheme.

In order to set the mathematical problem with only observable variables, it is expedient to introduce the following individual utility function and its relation to the one cited above:

\[ U^i(c^i, Y^i) = U(c^i, (Y^i/w^i) + m(1-\theta)) = U(c^i, L^i) \]  

(26)

\[ U(c^i, L^i) > 0 \quad U_{c^i}(c^i, L^i) < 0 \quad U_{Y^i}(c^i, Y^i) < 0 \quad U_{YY^i}(c^i, Y^i) > 0 \]

Notice that the new function, \( U^i(c^i, Y^i) \), is defined over consumption and gross income, which have a positive and negative impact, respectively; the latter reflects the increasing effort agents have to put, in order to earn a greater income. Notice also that the new utility function \( U^i(c^i, Y^i) \) strictly refers to
type-i agent. Indeed, for the same level of gross income and consumption, a skilled taxpayers enjoys a greater level of utility, since the inequality \( w^G > w^B \) still holds. Notice also that, following a step-by-step way of conducting the analysis, the model hereby does not introduce any major innovation related to the compliance cost: it is assumed to be a fixed amount of time the taxpayers dedicates to filing her tax returns. However, the government is now capable of implementing a non-linear tax scheme which, at least in principle, should make it easier to meet its redistributive objective.

Back to the analytical aspects, the social planner problem can be set as follows:

\[
\psi = U^G(c^G, Y^G) + \mu[U^B(c^B, Y^B) - \bar{U}] + \gamma[U^G(c^G, Y^G) - U^G(c^B, Y^B)] + \\
+ \rho[(Y^G - C^G) + (Y^B - C^B) - R] - \theta m(w^G + w^B) \tag{27}
\]

where \( \gamma \) multiplies the self-selection constraint for the skilled agent, and \( \rho \) multiplies the government revenue constraint. The social planner chooses the vectors of consumption and the labor supplies for both agents, based on the following first order conditions:

\[
\frac{\partial \psi}{\partial c^G} = U^G_c(c^G, Y^G)(1 + \gamma) - \rho = 0 \tag{28}
\]

\[
\frac{\partial \psi}{\partial Y^G} = U^G_y(c^G, Y^G)(1 + \gamma) + \rho = 0 \tag{29}
\]

\[
\frac{\partial \psi}{\partial c^B} = \mu U^B_c(c^B, Y^B) - \gamma U^G_c(c^B, Y^B) - \rho = 0 \tag{30}
\]

\[
\frac{\partial \psi}{\partial Y^B} = \mu U^B_y(c^B, Y^B) - \gamma U^G_y(c^B, Y^B) + \rho = 0 \tag{31}
\]

Similarly to the linear framework, let us take the derivative of the Lagrangian with respect to \( \theta \) (the policy instrument), in order to simulate the exogenous implementation of a public measure aimed at reducing taxpayers’ compliance costs, and analyze its effect on social welfare:
By integrating in (32) the first order conditions regarding the individuals’ optimal choices on consumption and gross income in (29) and (31), the result becomes:

\[
\left. \frac{\partial \psi}{\partial \theta} \right|_{\theta=0} = -U_y^G(c^G, Y^G)w^Gm(1+\gamma) - \mu U_y^B(c^B, Y^B)w^Bm + \gamma U_y^G(c^B, Y^B)w^Gm + \rho m(w^G + w^B)
\]

(32)

which collapses to:

\[
\left. \frac{\partial \psi}{\partial \theta} \right|_{\theta=0} = \rho w^Gm - w^Bm[\gamma U_y^G(c^B, Y^B) - \rho] + \gamma U_y^G(c^B, Y^B)w^Gm - \rho m(w^G + w^B)
\]

(33)

The result in (34) is negative, since the partial derivative of the individual utility function with respect to gross income is negative. This leads to the conclusion that, when an optimal non-linear income tax schedule is adopted, implementing a measure aimed at reducing the fixed taxpayers’ compliance costs decreases social welfare. This is due to the incentive constraint, which operates inefficiently when such a policy is adopted. In other words, the skilled agent is even more incentivized to mimicking the unskilled, resulting in an overall loss of efficiency.

**Optimal non-linear income taxation with heterogeneous compliance costs**

In this last model, the idea that taxpayers bear different compliance costs depending on their levels of ability and income is introduced, by stating the following important assumptions. First of all, the two agents have heterogeneous constant compliance costs, \(m^i\). Second, agents also differ in their efficiency in being compliant with their fiscal duties: in particular, the
skilled agent is more efficient than the unskilled, as represented by the parameters $e^G$ and $e^B < e^G$. Third, agents can buy on the market two types of fiscal services, measured in terms of time saved by the taxpayer: $a^G$, suitable for higher incomes (typically earned by skilled individuals), and $a^B$, suitable for lower incomes (typically earned by unskilled individuals). The per-unit-prices are $p^G$ and $p^B < p^G$, respectively for the high and the low quality fiscal service. A fraction of time, $\theta \in [0,1]$ is "given back" to taxpayers from the government, as a result of the free fiscal service policy (paid by the government at price $p^B$). Justifications for choosing a lower price for the fiscal services granted by the government stem from two major lines of reasoning. First, it may be the case that the government only provides "low quality" fiscal services, such that higher income earners find this free service insufficient for their needs. Second, economies of scale might apply such that the government has access to fiscal services at the lowest possible cost.

The assumed relations between the above cited parameters make the following inequality hold:

$$w^G \frac{1}{e^{GB}} < p^B < w^B \frac{1}{e^B} < p^G < w^G \frac{1}{e^{GG}} ,$$  \hspace{1cm} (35)

The double superscripts on the skilled agent's efficiency parameter, $e^{GB}$ and $e^{GG}$, symbolize a mimicking and a non-mimicking skilled agent efficiency in filing a tax return, respectively. Of course, there is no need to use double superscripts for the low skilled agent, since she never finds it convenient to mimic the more able individual when choosing her optimal consumption and income levels.

Before interpreting assumption (35), for the sake of simplicity in the development of the analysis, I will simplify it in the following way:

$$w^G \frac{1}{e} < p^B < w^B < p^G < w^G \hspace{1cm} e > 1$$  \hspace{1cm} (36)
A straightforward interpretation of assumption (36) result in the derivation of the following optimal behaviors by the two representative agents:

- when the skilled agent does not mimic the unskilled, then the skilled prefers to buy high quality fiscal services on the market, rather than bearing the compliance cost; therefore, \( a^G_* = m^G \); that is, the skilled will buy as much high quality fiscal service as it is needed in order to totally cover her compliance cost;

- when the skilled mimics the unskilled agent, then the skilled prefers to bear the compliance cost, rather than buying low quality fiscal services on the market: \( a^B_* = 0 \); equivalently, in this case, the skilled agent will not make use of the compliance-cost-reducing measure implemented by the government;

- low skilled agents always prefer buying low quality fiscal services on the market, rather than bearing the compliance cost; of course, they have even more incentives to do so, if the (low quality) fiscal services they make use of are paid by the government.

The three optimal behaviors stated above, in turn, imply the following budget constraints for the different taxpayers:

when G does not mimic B: \( c^G + p^G a^G = w^G L^G \) \hspace{1cm} (37)

when G mimics B: \( c^G = w^G \left[ L^G - \frac{m^B}{e} (1 - \theta) \right] \) \hspace{1cm} (38)

for a low skilled agent: \( c^B + p^B a^B = w^B \left[ L^B + \theta m^B \right] \) \hspace{1cm} (39)

Notice that the highest income a type-B agent can reach does not require a high quality fiscal service \((a^G)\); therefore, this good is not included in her budget constraint.

After constructing the budget constraints, we are able to finally write the new utility functions, defined over \((c^i, Y^i)\), as follows:
when $G$ does not mimic $B$:  
\[ U^G(c^G, Y^G) \equiv U \left( c^G, \frac{Y^G}{w^G} \right) = U \left( c^G, L^G \right) \]  
(40)

when $G$ mimics $B$:  
\[ U^G(c^G, Y^G) \equiv U \left( c^G, \frac{Y^G}{w^G} + \frac{m^B}{e} (1 - \theta) \right) = U \left( c^G, L^G \right) \]  
(41)

for any low skilled agent:  
\[ U^B(c^B, Y^B) \equiv U \left( c^B, \frac{Y^B}{w^B} - \theta m^B \right) = U \left( c^B, L^B \right) \]  
(42)

As in the previous non-linear case, defining the new utility function, $U(c, Y)$, allows us to make the social planner solve optimally the individual maximization problem, while at the same time, setting the optimal non-linear taxes for both types of agents:

\[ \Lambda = U^G(c^G, Y^G) + \mu[U^B(c^B, Y^B) - \bar{U}] + \gamma[U^G(c^G, Y^G) - U^G(c^B, Y^B)] + \\
+ \rho[(Y^G - C^G) + (Y^B - C^B) - 2\theta p^B m^B] \]  
(43)

Once again, $\gamma$ multiplies the self-selection constraint for the skilled agent, and $\rho$ multiplies the government revenue constraint. Notice that, this time, the administrative cost of the policy instrument $\theta$ depends only on the unskilled taxpayers compliance costs. The social planner chooses the vectors of consumption and the labor supplies for both agents, based on the following first order conditions:

\[ \frac{\partial \Lambda}{\partial c^G} = U^G_c(c^G, Y^G)(1 + \gamma) - \rho = 0 \]  
(44)

\[ \frac{\partial \Lambda}{\partial Y^G} = U^G_y(c^G, Y^G)(1 + \gamma) + \rho = 0 \]  
(45)

\[ \frac{\partial \Lambda}{\partial c^B} = \mu U^B_c(c^B, Y^B) - \gamma U^G_c(c^B, Y^B) - \rho = 0 \]  
(46)

\[ \frac{\partial \Lambda}{\partial Y^B} = \mu U^B_y(c^B, Y^B) - \gamma U^G_y(c^B, Y^B) + \rho = 0 \]  
(47)
If we now increase marginally $\theta$, so as to compute the welfare effect of introducing the new public measure aimed at reducing taxpayers' compliance costs, we get the following result:

$$
\frac{\partial \Lambda}{\partial \theta_{\theta=0}} = -\mu U_y^G(c^B, Y^B)m^B w^B + \gamma U_y^G(c^B, Y^B) \frac{m^B}{e} w^G - 2 \rho p^B m^B \quad (48)
$$

It is now necessary to substitute within (48) the condition for the government to optimally choose the unskilled agent's income in (47). This is because the latter includes the so called "mimicking term", that is, the partial effect on utility generated by the skilled agent when she acts like an unskilled. By doing so, we obtain

$$
\frac{\partial \Lambda}{\partial \theta_{\theta=0}} = (\rho - \gamma U_y^G(c^B, Y^B)) m^B w^B + \gamma U_y^G(c^B, Y^B) \frac{m^B}{e} w^G - 2 \rho p^B m^B \quad (49)
$$

and rearranging

$$
\frac{\partial \Lambda}{\partial \theta_{\theta=0}} = \gamma U_y^G(c^B, Y^B) m^B \left[ \frac{w^G}{e} - w^B \right] + \rho m^B \left( w^B - 2 p^B \right) \quad (50)
$$

The first term in (50) represents the change in utility for the skilled agent when she behaves as an unskilled (notice the arguments in the function have "B" superscripts). This is multiplied by the fixed unskilled agent's compliance cost, $m^B$ (the only compliance cost contemplated in this last framework), and by the difference in "weighted" wage rates between the skilled and the unskilled agent (the weight being the individual level of efficiency in filing her tax return). This first term is surely positive, given the negative partial derivative between utility and income and the inequality stated in (36). Moreover, the assumption in (36) confirms that the second term in (50) is also positive, whenever the following condition holds:

$$
w^B > 2p^B. \quad (51)$$
In other words, for the public measure to be efficient and effective in reaching greater equality in the distribution of income between the two agents, it is necessary that low skilled individuals’ productivity be more than double that the price of the low quality fiscal advisors.

**Conclusions**

Over the last two decades, a strand of literature has focused on the issue of taxpayers’ compliance costs. At the same time, different governments have implemented measures in order to provide fiscal assistance to the taxpayers (e.g., prefilled tax returns) and lower the costs they bear to be tax compliant. However, no investigation has been conducted on the redistributive aspects associated with the reduction of compliance costs.

This paper aims at investigating how the government can efficiently redistribute income by implementing a compliance cost reducing policy. As a starting case, it is shown that when the social planner is restricted to the linear taxation, reducing the taxpayers’ compliance costs is not efficient. This is in line with previous research showing the inefficiency of rising administration costs. However, the theoretical analysis in this paper shows that, when the government levies an optimal non-linear tax scheme and agents bear different compliance costs, a well designed public measure aimed at redistributing income via a reduction of taxpayers’ compliance costs is welfare enhancing. The relaxation of the self selection constraint for the skilled agent is crucial for the welfare enhancing result to occur. In other words, after the government implements the new measure, the more able individual is less tempted from mimicking the unskilled agent, which attenuates the adverse selection problem. Even though this result is strong, in that it provides the government with an efficient tool to redistribute income from the more able to the less able taxpayers, it should be born in mind that it only holds when the following conditions occur: the skilled agent's efficiency in filling her tax return is such
that she finds convenient buying fiscal services only when her income is high \textit{(i.e., she does not mimic the unskilled agent's optimal behavior)}; the unskilled individual wage rate is more than double the hour price paid by the government to provide free low income fiscal services.
REFERENCES


