Poverty, tax evasion and the optimum general income tax

Enlinson Henrique Carvalho de Mattos[1]
Department of Economics - University of Illinois at Urbana-Champaign
Preliminary, please do not cite it

This paper combines poverty considerations in the social welfare function and tax evasion into an optimal general income tax problem. It investigates the optimal audit and tax structures using a model with two types of individual, endogenous labor supply and a hybrid social welfare function that captures the pluralism of the objectives of a full committed government. The results confirm the previous literature on tax evasion: (i) skilled households should never be audited and face any distortions in the labor supply, (ii) unskilled households have to be audited randomly (probability less than one) and (iii) Individuals should be rewarded if telling the truth about their income. In addition, the introduction of poverty concern as a negative externality brings new characteristics for the optimum income tax on poor: (iv) they may (or not) face negative marginal income tax. Also, a numerical example is provided to further explore the model.

Este artigo combina considerações sobre pobreza na função de bem-estar social e a possibilidade de evasão fiscal dos agentes econômicos em um problema de imposto de renda ótimo. Investiga-se a estrutura de fiscalização e imposto ótimos usando um modelo com dois tipos de indivíduos, oferta de trabalho endógena e função de bem estar social híbrida que captura o pluralismo dos objetivos do planejador central. Os resultados confirmam os resultados obtidos na literatura sobre evasão fiscal: (i) trabalhadores habilidosos nunca devem sofrer auditoria nem distorção na oferta de trabalho, (ii) trabalhadores pouco habilidosos tem que sofrer auditoria aleatoriamente (probabilidade menor que um) e (iii) indivíduos devem ser recompensados por não evadir o imposto. Além disso, a introducao da pobreza como externalidade negativa traz caracteríssticas novas para o imposto de renda ótimo que incide nas pessoas que estão abaixo da linha da pobreza: (iv) o imposto de renda marginal pode (ou nao) ser negativo. Um exemplo numérico é utilizado para melhor explorar os resultados do modelo teórico.

Keywords: optimum general income tax, poverty, tax evasion.

Palavras-chave: imposto de renda ótimo, pobreza, evasão fiscal.

JEL Classification Numbers: H42, H31, H21 and H23.

Área ANPEC: Area 4 - Economia do setor público.

---

[1] Capes Funded - Brasilia/Brazil. Address: 330 Wohlers Hall, 1206 South Sixth St., Champaign, IL 61820, USA, phone:\ 1-217-384-5221, e-mail: *dmcarval@uiuc.edu.*

1 Introduction

The implementation of income maintenance programs to combat poverty in many developing countries in the last decade has brought back the efficient redistribution policy into the discussion arena. The literature on optimal income maintenance program concerns the "targeting" aspect in the design of income transfers. The programs have to ensure that resources are concentrated on the poor, minimizing the leakages to the non-poor. This is an information problem, where non-targeted groups may benefit from the program and have incentives to mimic the behavior of the targeted ones.

Nichols and Zeckhauser (1982) first suggested that ordeals should be imposed on recipients of the aid in order to discourage those who would not otherwise claim the benefit. With respect to the guaranteed income scheme, a branch of the income maintenance schedule, Kesselman (1971) and Zeckhauser (1971) propose the wage subsidy as a way to alleviate the poverty and make the work more valuable to the low income household. This policy was severely criticized by Garfinkel (1973) in terms of disincentive to work and difficulty to be implemented. However as we see later in this paper, a wage subsidy may be the optimum solution to reduce poverty.

Besley and Coate (1992, 1995) following the above literature, examined the incentive effects of wage subsidies, welfare and workfare as alternative policies that aim to guarantee a certain income level in an economy where some individuals have earnings abilities that are greater than the minimum target. A wage subsidy can be a solution in the case that incomes are perfectly obsevable. In any other case, the superiority of the workfare scheme compared to the welfare relies on the fact that the first limits the benefited household's time to work in the private sector. Therefore, those whose private earnings would be greater than the minimum assured in the program would not choose the program.

On the other hand, Kanbur, Keen and Tuomala (1994) uses a non-welfarist objective function to introduce a model that captures the policy debate about income-based poverty index. The interesting result of the paper is that the optimal marginal tax rate on the very poorest is strictly negative. This is the first step before considering mixing the welfarist and the non-welfarist approach.

Wally (2001) observes that if the planner recognizes that it has to take into account the non-utility information, say income-poverty when setting a policy, welfarist and non-welfarist considerations must be traded off in the objective function. This approach captures the pluralism of the objectives of the government but assumes, as in Mirrlees (1971) that everybody can work. Poverty is considered a "public bad" or a negative externality to the welfare of the society and Wally (2001) derives the

optimal linear and non-linear tax for this hybrid social welfare function. He arrives at the conclusion that poor individuals can face negative or positive marginal tax rates and all the non-poor, except the most able, face a strictly positive marginal tax rate. These results reinforce the one obtained in Kanbur, Keen and Tuomala (1994) which suggests that the negative marginal tax can be used as optimal policy when the concern about reduction in poverty is introduced in the objective of the government.

However, none of the above papers considered the case that individuals can mimic their income and their type. As observed by Sandmo (2004), the theory of optimal taxation can be seen as a "recipe for minimizing the costs of taxation" under which the more direct costs of administration and compliance should play an important role into the analysis. More specifically, the possibility that the government (or tax administrator) may not observe the income of the households adds an additional cost to the tax administration and must be considered in the tax design. The literature on tax evasion is the way to bring these issues into the literature. Sandmo (1981) and Cremer and Gahvari (1996) present the optimal linear and general income tax respectively when tax evasion is introduced in the standard income tax model without considering poverty.

This paper combines income-based poverty consideration and tax evasion in the optimum income tax problem of Mirleess (1971). A hybrid social welfare function is used to capture both the equity-efficiency trade-off and the income based-poverty concern of the social planner. There are only two types of individuals but their incomes are observable only through an audit cost which brings the possibility of tax evasion into the picture. The paper characterizes a solution where individuals can mimic types (self-selection constraint) and income (moral hazard constraints) under the plurality role of the government. When compared to Wally (2001), this paper brings an additional informational problem to the social planner. The social planner does not observe the skill-type and the income of the agents.

The paper is organized as follows. Next section introduces the poverty concern into the general income tax model with two types of individuals where the incomes are observable. The section works as the benchmark case and an example is provided to ilustrate the results. Section 3 incorporates tax evasion possibility into the model and its solution is presented along with a numerical example. Section 4 concludes.

2 Benchmark Case

I start by providing the solution of the problem when the individuals' income are observable at no cost. This problem is similar to the one posed in Wally (2001)

however with only two types of persons. The solution works as a benchmark case and is compared to the case presented in the next section.

Consider an economy consisting of two types of individuals: N_r rich individuals and N_p poor individuals. Both types are risk averse with identical preferences but different skills. Agents of type p earn a lower wage than those of type r. Preferences are separable in the numeraire (C) and labor supply, L, the only two gods in this economy, and are represented by

$$\begin{aligned} U &= u(C) + v(1-L). \\ L &= Y/w_i, i = p, r. \end{aligned} \tag{1}$$

where U is twice continuos and differentiable, srictly increasing in C and srictly decreasing in L. Also, let $u(C)$ be stricly concave. Consumption is greater or equal to zero and $u(0) = 0$. Y is the individual's income and w_i denotes individual's wage (or skills) of type $i = p, r$.

The direct mechanism consists of four functions: $Y(w_i)$ and $T(w_i)$ where $i = p, r$. It works as follows: after the agent reports his type, w_i, the tax administrator assigns the income, $Y(w_i)$ and the amount of taxes to be paid, $T(w_i)$. For ease of notation, define $U_{i,k} = u(Y_i - T_i) + v_i(Y_k)$, $Y_i = Y(w_i)$, $v_i(Y_k) = v(1 - Y_k/w_i)$ and $T_i = T(w_i)$, where $k, i = p, r$.

The measurement of poverty is beyond the scope of the paper,[1] however I follow the literature and denote C^* as the poverty line. Any individual whose consumption is below this critical value is considered poor. First denote consumption of type i, $C_i = Y_i - T_i$ and then define generalized poverty gap, the sum of individual poverty as

$$\wp(C, C^*) = P(C_r, C^*) + P(C_p, C^*) \tag{2}$$

where the function $P(C_i, C^*)$ satisfies the conditions $P(C_i, C^*) \geq 0$, $P_c(C_i, C^*) < 0$ and $P_{cc}(C_i, C^*) > 0$ for all $C \in [0, C^*)$ and for $i = r, p$. Finally $P(C, C^*) = P_c(C, C^*) = 0$ for $C > C^*$. This measure of poverty takes into consideration not only the existence of poverty but also how deep it is.[2] Aggregate poverty is considered a public bad in this set up, as in Wally (2001), and it generates a negative externality. Therefore, its reduction improves social welfare.

The government (social planner) has the twofold objective: maximize social welfare and minimize aggregate poverty,

[1] See Ravallion (1994) for further details.

[2] Kanbur, Keen and Tuomala (1994) use the same formulation to define agregate poverty. However they consider the case that minimizing poverty is the only objective of the government.

$$W = U_{pp} + \delta U_{rr} - \beta \wp(C, C^*) \tag{3}$$

where δ and β are the relative social weights imposed on the skilled household and on the aggregate poverty. When the incomes are observable there is no uncertainty with respect to the net income of the agents, however the rich may still mimic the poor. Therefore the maximization must satisfy the self-selection constraint,

$$U_{rr} \geq U_{rp} \tag{4}$$

That is, in equilibrium the rich must be better off by not faking his type. I follow the literature and ignore the "upward" incentive constraints, ie, the constraint that the poor individual tries to mimic the rich does not bind. The maximization is also subject to the revenue constraint for the government

$$N_p T_p + N_r T_r = R \tag{5}$$

where N_i is the number of i's population, $i = r, p$. R stands for the necessary tax revenue.

Define $MRTS^i_{C,L} = 1 - \frac{v_i'(Y_k)}{u'(C)w_i}$, $i = r, p$. Notice that if $MRTS^i_{C,L} < (>)0$ then a negative (positive) marginal tax is imposed on individual i. In addition, I consider the interesting case that only the consumption of the poor is below the poverty line. Henceforth, the government has to maximize 3 with respect to Y_p, Y_r, T_p, T_r subject to 4 and 5.

The Lagrangian expression can be written,

$$\Lambda = U_{pp} + \delta U_{rr} - \beta P(C_p, C^*) + \lambda[U_{rr} - U_{rp}] + \mu[N_p T_p + N_r T_r - \bar{R}] \tag{6}$$

where λ and μ denotes the lagrangian multipliers for the self-selection and revenue constraints. The first order conditions are

$$\begin{aligned}
Y_p: \quad & \tfrac{\partial U}{\partial C_p} - \tfrac{1}{w_p} v_p'(Y_p) - \beta \tfrac{\partial P(C_p, C^*)}{\partial C_p} - \lambda[\tfrac{\partial U}{\partial C_p} - \tfrac{1}{w_r} v_r'(Y_p)] &= 0, & (7)\\
Y_r: \quad & \delta[\tfrac{\partial U}{\partial C_r} - \tfrac{1}{w_r} v_r'(Y_r)] - \lambda[\tfrac{\partial U}{\partial C_r} - \tfrac{1}{w_r} v_r'(Y_p)] &= 0, & (8)\\
T_p: \quad & -\tfrac{\partial U}{\partial C_p} + \beta \tfrac{\partial P(C_p, C^*)}{\partial C_p} + \lambda[\tfrac{\partial U}{\partial C_p}] + \mu N_p &= 0, & (9)\\
T_r: \quad & \delta[-\tfrac{\partial U}{\partial C_r}] - \lambda \tfrac{\partial U}{\partial C_r} + \mu N_r &= 0. & (10)
\end{aligned}$$

It may easily be shown from 7 and 8 that

$$MRTS^{p}_{C,L} = \frac{\lambda}{1-\lambda}[\frac{v'_p(Y_p)}{w_p} - \frac{v'_r(Y_p)}{w_r}] + \frac{\beta}{1-\lambda}[\frac{\partial P(C_p, C^*)}{\partial C_p}], \quad (11)$$

$$MRTS^{r}_{C,L} = 0. \quad (12)$$

Equation 11 reveals that the optimum marginal tax imposed on the poor individuals may be either positive or negative.[3] If $C_p > C^*$ ($\beta = 0$), the self-selection can be binding ($\lambda > 0$) or not ($\lambda = 0$), then a positive or zero marginal tax has to be imposed on the poor individual. The distortion (positive marginal tax on the poor) has to make less attractable for the rich to mimic the poor. When the self selection constraint is binding and $\beta = 1$, the final marginal tax may be positive, negative or zero depending on the relative effects of the terms in brackets. It is sufficient to show an example that clarifies the results.

Last, equation 12 reinforces one of the main important findings in the optimal tax literature: "no discrimination at the top". That happens because no one wants to imitate the rich, so there is no reason to distort his decisions.

2.1 Example

In order to investigate precisely the properties of the equilibrium marginal income tax on poor it is necessary to specialize the utility functions, poverty function and the parameters of the model. Therefore, assume that both households, rich and poor, face the same utility function, $U(C, L) = Log(Y - T) - Y/w$, the poverty measure is $P(C_p, C^*) = (C^* - (Y_p - T_p))^2)/C^*$ and let the economy be populated such that half of the population is poor and the other half is rich.[4] For normalization purpose assume $Np = Nr = 0.5$. For simplicity let R, the amount of revenue necessary for the government and β, the weight given by the social planner to the poverty measure, to be both equal to one.[5]

Now consider the following exercise. If we assume that the skill of the poor is equal to one ($w_p = 1$) while the skill of the rich goes from 1.5 to 10 ($w_r = aw_p, a \in [1.5, 10]$) and the poverty line is assumed to be equal to 1.5, then we can verify explicitly what

[3] Notice that $[\frac{v'_p(Y_p)}{w_p} - \frac{v'_r(Y_p)}{w_r}] > 0$ and $[\frac{\partial P(C_p, C^*)}{\partial C_p}] < 0$. To see that note that the rich consumes more leisure compared to the poor individual for the same amount of income. That implies a higher marginal utility of the leisure good for the poor agent. In addition note that $w_r > w_p$. The second inequality follows from the definition of $P(C, C^*)$.

[4] See Ebert (1992) for further details about this particular utility function.

[5] $\beta = 1$ is a consequence of the assumption of a utilitarian social planner where both types of households dislike aggregate poverty.

is and how the optimal choices of Y_p, Y_r, T_p, T_r change for the particular social welfare function $N_p U_{pp} + N_r U_{rr} - P(C_p, 1.5)$. The solution must also respect the self selection and the revenue constraints, 4 and 5 respectively.[6]

Figure 1 compares the optimal income tax on the poor households when the government is concerned with poverty ($\beta = 1$) and when it is not ($\beta = 0$). For the case that $\beta = 1$, the Figure 1 shows that it may be optimal for the poor households to face positive, negative or zero marginal income taxes depending on how distant are their skills compared to the rich households. The negative marginal income tax is the optimal solution for the poor if the income of the first (poor) is not too distant from the income of the last (rich). This result further qualify the conclusions derived in Waly's (2001) because they show under which conditions the lowest skilled household can face positive, negative or zero marginal tax.

Note also that the marginal tax on the poor individuals for the first case ($\beta = 1$) is always below the one presented for the last case ($\beta = 0$). This means that the poverty concern of the social planner introduces an extra-subsidy on the poor's earnings as an optimal income tax.

Figure 2 shows the consumption of the poor individuals when both cases are considered: $\beta = 0$ or $\beta = 1$. It shows that when the government is concerned with an income-based poverty ($\beta = 1$) the net income (consumption) of the poor households is much higher than the case when such issue is not considered ($\beta = 0$). This result is not surprising given the poverty concern of the government. However, as the Figure 3 shows, this increase in consumption occurs in part because of increase in the labor supplied by the poor households compared to the case that $\beta = 0$.

Figure 3 explains why the consumption of the poor individuals start to decrease with the skill distance and also shows the labor supply for the poor in both cases ($\beta = 0$ or $\beta = 1$). First, as the skill distance between the rich and the poor increases, the labor supply for the poor decreases. That leads to a lower level of income for the poor, which makes him to have a lower consumption level but a higher level of leisure. Second, under the scheme wthat $\beta = 1$ this effect is lower, since the poor are encouraged to "overwork" so that they can partially finance their consumption.

Figure 4 shows the amount of taxes paid by the poor individuals. It shows that when $\beta = 1$ the level of taxes imposed on the poor individuals is much higher than when $\beta = 0$. This can be done because the poors are working more in the first case. But as the level of taxes decreases with the skill distance between rich and poor eventually it becomes negative (transfer) and can finance higher levels of

[6]For lack of space, all figures contain three cases and are presented at the end of the paper. In the first two cases ($\beta = 1$ and 0) it is assumed that incomes are observable with no cost while the last one considers tax evasion. However they'll be explained according to the order of the text.

consumption (Figure 2) for the poor households.

The last figure (figure 5) confirms Waly's result. The agents below the poverty line are made worse-off with the introduction of income-based poverty consideration in the objective function of the social planner. This happens because these individuals (poor) "overwork" due to the marginal incentive (negative marginal tax) and the increase in consumption does not compensate the increase in the labor supplied by them. This reuslt only holds for the cases that the ability of the poor is not too distant from the ability of the rich household. For the cases that the this distance is high enough the utility does not change with the introduction of the poverty concern into the objective function of the government. This happens because the poor households are not working and consuming the same amount at the optimum in any circunstance ($\beta = 0$ or 1).

3 Incomes Observable through an audit cost

This section characterizes the optimal income tax, fines, auditing proabilities, consumption and leisure for both types of households, when poverty and tax evasion are combined in a genereal income tax model. When incomes are observable only through an audit cost the possibility of tax evasion arises. If the poor household receives a subsidy for having low income, he might have incentive to mimic his income. The rich one might as well to mimic his income level to obtain the poor's level of subsidy as well. These effects have to be taken into account when designing the optimal income tax.

While the possibility of tax evasion is appealing for most developing countries where the poor households usually works in the informal sector, that is not so obvious for developed economies. However as shown in Schneider and Enste (2000), informal economy is a relevant issue also for developed countries.

Therefore, let households' income be observable only with an audit cost, A which is stricly increasing in the number of people audited and respects, $0 < A' < +\infty$.[7] Penalties cannot exceed an indvidual's income and other punishments are excluded.

The direct mechanism consists of four functions: $Y(\bar{w})$, $p(\bar{w})$, $T(\bar{w})$, $F(\bar{w}, Y_A)$ where $\bar{w}$ is the reported type and Y_A is the income revealed through and auditing process. It works smilialrly to the previous section: after the agent reports his type, $\bar{w}$, the tax administrator assigns the income, $Y(\bar{w})$, the probability of auditing, $p(\bar{w})$,

[7] This model borrows a similar structure presented in Cremer and Gahvari (1996) to guarantee the existence of the revelation principle. More complicated structures could have been presented but this structure is sufficient to show the arguments.

the amount of taxes to be paid, $T(\bar{w})$, and the fines of $F(\bar{w}, Y_A)$ if he is audited and found to have a true income of Y_A.

The individuals can cheat in two different ways. First, by misreporting the type (poor and rich, p and r) and second by misreporting the income. It is assumed that only the income misreported action can be detected through an audit.

Also, define the expected utility of type i individual who reports to be of type k and earns income Y_k by EU_{ik},his expected utility is

$$EU_{ik} = (1 - p_k)u(Y_k - T_k) + p_k u(Y_k - F_k) + v_i(Y_k), i, k = p, r. \tag{13}$$

Since the individuals can be audited (or not) and pay fines (or taxes) two different states arise. In this case, assume the government wants to reduce the expected aggregated poverty ($E\wp(C, C^*)$) defined as $(1 - p)\wp(Y - T, C^*) + p\wp(Y - F, C^*)$.[8]

The government (social planner) has two obectives: maximize social welfare and minimize expected aggregate poverty,

$$W = EU_{pp} + \delta EU_{rr} - \beta E\wp(C, C^*) \tag{14}$$

where δ and β are the relative social weights imposed on the skilled household and on the expected aggregate poverty as in the previous section. The maximization is subject to the revenue and incentive compatibility constraints. Again, I follow the literature and ignore the "upward" incentive constraints, ie, the constraint that the poor individual tries to mimic the rich is not binding.

Before stating the constraints formally, notice that in equilibrium the revelation principle (truth-telling equilbrium) applies in this case [as in Mookherjee and Png (1989) and Cremer and Gavhari (1996)]. Therefore, any strategy out of the equilibrium path can be punished using the highest penalty (total income) without affecting the equilibrium utility of the individuals. In terms of policy implementation this implies.

$$F(\bar{w}, Y_A) = Y_A. \tag{15}$$

Also assume that the the minimum amount of tax and fine is $-C^*$. This states that if the optimal tax or fines call for a transfer from the government to the households, that must be at most by C^*. This assumption simplifies the proof of the existence of optimal mechanisms and also is realistic.[9] There is no reason for the government to redistribute more goods than the minimum to poverty line.

[8] An alternative definition for aggregate poverty would be $\wp((1 - p)(Y - T) + p(Y - F), C^*)$. This alternative does not change our results but introduces more algebra into the analysis.

[9] The proof is available upon request.

In addition, define the maximum utility of an individual with skill w, faces a audit probability of p and pay a fine equal to the maximum possible (his income) if audited and a tax T if not, as $V_{i,k} = V(w_i, p_k, T_k)$. This formulation corresponds to an agent of type i that claims to be of type k but declares a different income than the one assigned to type k. Since this type of cheating is detected, it is defined as

$$V_{i,k} = V(w_i, p_k, T_k) = (1 - p_k)u(\tilde{Y}(w, p, T) - T_k) + v_i(\tilde{Y}(w, p, T)),\ i, k = p, r. \quad (16)$$

where $\tilde{Y}(w, p, T)$ corresponds to the income that he chooses for himself under this situation and maximizes the above utility.

The problem of the government is to maximize the equation 14 with respect to Y_p, Y_r, p_p, p_r,T_p, T_r, F_p and F_r subject to the self-selection contraints

$$EU_{rr} \geq EU_{rp}, \quad (17)$$

$$EU_{rr} \geq V_{rr}, \quad (18)$$

$$EU_{rr} \geq V_{rp}, \quad (19)$$

$$EU_{pp} \geq V_{pp}. \quad (20)$$

and the revenue constraint

$$N_p[(1 - p_p)T_p + p_p F_p] + N_r[(1 - p_r)T_r + p_r F_r] - A(N_p p_p + N_r p_r) \geq \bar{R} \quad (21)$$

where $\bar{R}$ stands for the necessary tax revenue.

Self-selection constraints 17, 18 and 19 ensure that a rich household prefers a truthfull statement of his type and income than mimicking the poor person and his income, misreporting his income while declaring a rich person and misreporting his income and his type. Constraint 20 ensures that poor individual prefers a truthfull statement of his type and income than misreporting his income. Constraint 17 is the usual self-slection constraint when incomes are observable and constraints 18, 19 and 20 are the moral hazard conditions and have to be satisfied to avoid tax evasion.

As a preliminary, it may be shown that the rich individual is never audited and the poor household is audited randomly as stated int the next proposition.[10]

Proposition 1 *The optimal audit probabilities are* $p_r = 0$ *and* $p_p < 1$.

[10]The proofs of all the propositions are available upon request. They are not in this version due to the space constraint. In addition, calibration results for different countries are also available upon request.

The lagrangian set up becomes,

$$\begin{gathered}\Lambda = [(1-p_p)u(Y_p-T_p)+p_pu(Y_p-F_p)+v(1-\tfrac{Y_p}{w_p})]+\\ \delta[u(Y_r-T_r)+v(1-\tfrac{Y_r}{w_r})]-\beta[EP(C_p,C^*)]+\\ \lambda_1[u(Y_r-T_r)+v(1-\tfrac{Y_r}{w_r})-(1-p_p)u(Y_p-T_p)-p_pu(Y_p-F_p)+v(1-\tfrac{Y_p}{w_r})]+\\ \lambda_2[u(Y_r-T_r)+v(1-\tfrac{Y_r}{w_r})-V_{rr}]+\lambda_3[u(Y_r-T_r)+v(1-\tfrac{Y_r}{w_r})-V_{rp}]+\\ \lambda_4[(1-p_p)u(Y_p-T_p)+p_pu(Y_p-F_p)+v(1-\tfrac{Y_p}{w_p})-V_{pp}]+\\ \mu[N_p[(1-p_p)T_p+p_pF_p]+N_r[T_r]-A(N_pp_p)-\bar{R}] \end{gathered} \tag{22}$$

where the λ's denote the lagrangian multipliers for the self-selection constraints and μ is the marginal cost of an additional unit of revenue in utility terms. The first order conditions of this problem are

$$Y_p : (1-\lambda_1+\lambda_4)E[\frac{\partial u_p}{\partial C_p}]+(1+\lambda_4)[\frac{\partial v_p(Y_p)}{\partial Y_p}\frac{1}{w_p}]-\beta E\frac{\partial P(C_p,C^*)}{\partial C_p}-\lambda_1[\frac{1}{w_r}\frac{\partial v_r(Y_p)}{\partial Y_p}]=0, \tag{23}$$

$$Y_r : (\delta+\lambda_1+\lambda_2+\lambda_3)[\frac{\partial U}{\partial C_r}-\frac{1}{w_r}\frac{\partial v_r(Y_r)}{\partial Y_r}]=0, \tag{24}$$

$$T_p : -(1-\lambda_1+\lambda_4)(1-p_p)\frac{\partial u_p^T}{\partial C_p}+\beta(1-p_p)\frac{\partial P^T}{\partial C_p}-\lambda_3[\frac{\partial V_{rp}}{\partial T_p}]-\lambda_4[\frac{\partial V_{pp}}{\partial T_p}]+\mu N_p(1-p_p)=0, \tag{25}$$

$$T_r : -(\delta+\lambda_1+\lambda_2+\lambda_3)[\frac{\partial u_r^T}{\partial C_r}]-\lambda_2\frac{\partial V_{rr}}{\partial T_r}+\mu N_r=0. \tag{26}$$

$$F_p : -(1-\lambda_1+\lambda_4)p_p\frac{\partial u_p^F}{\partial C_p}+\beta p_p\frac{\partial P^T}{\partial C_p}+\mu N_pp_p=0, \tag{27}$$

$$p_p : (1-\lambda_1+\lambda_4)(u_p^F-u_p^T)+\beta(P^T-P^F)-\lambda_3[\frac{\partial V_{rp}}{\partial p_p}]-\lambda_4[\frac{\partial V_{pp}}{\partial p_p}]+\mu N_p(F_p-T_p-A')=0. \tag{28}$$

where $u_p^T = u(Y_p-T_p)$, $u_p^F = u(Y_p-F_p)$, $u_r^T = u(Y_r-T_r)$, $v_p(Y) = v(1-Y/w_p)$, $v_r(Y) = v(1-Y/w_r)$, $P^T = P(Y_p-T_p,C^*)$, $P^F = P(Y_p-F_p,C^*)$, $E(\partial u_p/\partial C_p) = (1-p_p)(\partial u(Y_p-T_p)/\partial C)+p_p(\partial u(Y_p-F_p)/\partial C)$ and $E(\partial P_p/\partial C_p) = (1-p_p)(\partial P(Y_p-T_p,C^*)/\partial C)+p_p(\partial P(Y_p-F_p,C^*)/\partial C)$.

The condition 24 implies that the skilled agent should face a zero marginal tax which is stated in the next proposition.

Proposition 2 *The optimal tax policy on rich individual requires a zero marginal tax.*

The intuition of this result is quite simple. The benefit of distorting a group's marginal tax is to relax the self-selection constraint. However, the poors individual do not have any interest in mimicking the rich and therefore this distortion is not necessary. This result confirms a previously result that no "distortion should be applied at the top".

Before analyzing the other conditions it is easier to exclude some results in order to narrow down the solution. The proposition 2 summarizes them.[11]

Proposition 3 *(i)Skilled households (rich individuals) are indifferent between truth-telling their type and income and mimicking their income ($\lambda_2 = 0$). (ii)Either skilled households are indifferent between truth-telling and mimicking their type and their income or the poor individuals are indifferent between truth-telling and mimicking their income ($\lambda_3 = 0$ or $\lambda_4 = 0$, both cannot be equal to zero). (iii)Individuals should be rewarded if caught telling the truth about their income ($F_p < T_p$).*

Now I can state the main findings of this section conditinal on the two possible set of solutions: (i) $\lambda_1 = 0, \lambda_2 = 0$, $\lambda_3 > 0$ or $\lambda_4 > 0$ or both λ_3 and $\lambda_4 > 0$. (ii)$\lambda_1 > 0, \lambda_2 = 0$, $\lambda_3 > 0$ or $\lambda_4 > 0$ or both λ_3 and $\lambda_4 > 0$. Using the definition of $MRTS^p_{C,L}$ and the condition 23 the following expression can be obtained

$$MRTS^p_{C,L} = \frac{\lambda_1[\frac{\partial v_r(Y_p)}{\partial Y_p w_p} - \frac{\partial v_p(Y_p)}{\partial Y_p w_r}] + \beta[E[\frac{\partial P(C_p, C^*)}{\partial C_p}]]}{(1 - \lambda_1 + \lambda_4)[E(\frac{\partial u_p}{\partial C_p})]} \tag{29}$$

Since $(1 - \lambda_1 + \lambda_4) > 0$, the first term of the right-hand side inside the brackets is positive and the derivative of the expected poverty with respect to the income (second term of the right-hand side also inside the brackets) is negative, the main conclusions can be summarized by the following proposition.

Proposition 4 *The optimal tax-audit policy implies that the marginal tax rate imposed on poor agents can be positive, negative or zero.*

[11] An important distinction between this model and the previous literature is that if I assume $\lambda_1 = 0$ that does not imply $\lambda_4 = 0$. That happens because even when λ_1 is set to zero, the social planner still has a reason to distort the behavior of the poor; the poverty concern. That in turn generates an "overwork" on poor due to the negative marginal tax imposed (see benchmark case). The poor individuals now have an incentive to mimic their income and work less which implies that λ_4 may be greater than zero.

This result is intuitive and not distant from the benchmark case. I discuss each possibility. If the self-selection constraint is not binding but the consumption of the poor is below the poverty line, then a negative marginal tax is assigned to the poor. He has now an extra incentive to work since he is being subsidized and the rich strictly prefers the truth-telling strategy. The alternative (ii) represents the case that the consumption of the poor is above the poverty line and the rich is indiferrent between mimicking the poor and telling the truth about his income. In that case, a positive marginal tax has to be imposed on the poor individual to distort his behavior making less attractable for the rich to mimic him. Turning to the third possibility the marginal tax assigned to the poor depends on how strong each effect is: the higher the poverty the lower (more negative) is the marginal tax imposed on the poor. Similarly, the stronger the self-selection constrain binds, the higher (more positive) is the marginal tax assigned to the poor. The only term added in comparison to the benchmark case is the lagrange multiplier λ_4. This term forces the marginal tax faced by the poor towards zero. It says that if the moral hazard constraint on the poor is (strongly) binding then no distortion should be assigned on poor's marginal tax. In other words, if he (the poor) is considering to cheat his income it is better to not impose distortion in his decision at the margin.

3.1 Numerical example

Similarly in this section, the properties of the equilibrium are explored in the numerical example with the additional characteristic that poverty and tax evasion are combined in the general optimum income tax model.

Adapting the previous example to the additional tax evasion structure, assume that both households, rich and poor, face the same utility function, $EU(C, L) = (1-p)Log(Y-T) + pLog(Y-F) - Y/w$, the poverty measure is $EP(C_p, C^*) = (1-p)((C^*-(Y_p-T_p))^2)/C^*)+p((C^*r-(Y_p-F_p))^2)/C^*)$ while all the other parameters is similar. Let the audit cost function $A(N_p p_p)$ to be equal to $N_p p_p$, that is $A = 1$.

Again, consider the same exercise. If we assume that the skill of the poor is equal to one ($w_p = 1$) while the skill of the rich goes from 1.5 to 10 ($w_r = aw_p, a \in [1.5, 10]$) and the poverty line is assumed to be equal to 1.5, then we can verify explicitly what is and how the optimal choices of $Y_p, Y_r, T_p, T_r, F_p, p_p$ change for the particular social welfare function $EU_{pp} + U_{rr} - EP(C_p, 1.5)$. The solutions must also respect the self selection and the moral hazard constraints 17, 18, 19 and 20.

Figure 1 shows that the introduction of tax evasion into the model smoothes out the optimal marginal tax on the poor individual compared to the case where $\beta = 1$ and incomes are observable with no cost. This happens because the marginal tax

cannot be too negative because the poor households can mimic income an request such marginal benefit. It also cannot be too positive since it is associated with a transfer from the government to the household (lump-sum) and then the rich could claim the benefit by mimicking their income.

Figure 2 presents the expected the expected optimal consumption for the poor ($E[C_p] = (1-p)*(Y_p-T_p)+p*(Y_p-F_p)$) assuming that the skill of the rich individual increases compared to the poor ones. It shows that the expected consumption can increase or decrease compared to the case where $\beta = 1$ and incomes are perfectly observable. When the distance of the skills between the rich and the poor is either small or too large than the expected consumption for the poor households under the first case is lower than their consumption for last one ($\beta = 1$ and observable income). This happens because in the first case, the rich has a strong incentive to misreport his type while the poor has high incentive to mimic his income in the second alternative. This makes the transfer from the government to the poor to decrease and consequently a lower level of expected consumption.

Figure 3 explicits that the optimal labor supply for the poor increases with the introdution of tax evasion (except when the abilities of the poor and rich are similar or when they are too far apart). This is surprising since it reveals that the poor are encouraged to "overwork" even more compare to the case that $\beta = 1$ and incomes are perfectly observable. This happens because the rich has a stronger incentive to mimic the type and/or income than the poor has to mimic their income for these skill levels.[12]

Figure 4 shows that the amount of taxes paid by the poor individuals also has increased compared to the previous section. This is the case not only as a consequence of the increase in the poor's labor supply but also due to the increase in the auditing cost. Once tax evasion possibility is introduced the government incurs in a extra cost. But similarly as the level of taxes decreases with the skill distance between rich and poor eventually it becomes negative (transfer) but the level of transfer is still lower in this case than in the one obtained in the previous section.

Figure 5 reveals that the poor households are worse off in the case that government cannot observe income without cost than in the case that incomes are observed with no cost. This is the case because these individuals (poor) "overwork" even more in the first case than in the second and it represents the extra cost in utility terms of the income observation.

Last, Figure 6 shows the three lambda's that reveal how strong each constraint is binding. Lambda 1 represents how the self-selection constraint is binding while lambdas 3 and 4 show the strengh of the hazard constraints. It shows that lambda 4

[12]See Figure 6.

(poors that want to misreport his income) is binding only when incomes of the rich and the poor households are similar or too far apart. As expected, lambdas 1 and 3 (riches that want to misreport their type and income)only bind when rich's income is similar to the poor's.

4 Conclusion

The paper has introduced poverty consideration and tax evasion into an optimum general income tax problem with endogenous labor supply. Under this set up the optimal audit and tax structures is investigated. The results confirm the previous literature on tax evasion: (i) skilled households are never audited and should not face any distortion in the labor supply, (ii) unskilled households are audited randomly (probability less than one) and (iii) Individuals should be rewarded if telling the truth about their income. In addition, the introduction of poverty concern brings new characteristics for the optimum income tax on poor: they may (or not) face negative marginal income tax depending on how distant (close) is their income compared to the poverty line and the rich's income.

A numerical example is provided to further explore the characteristics of the model. The example shows that, depending on how skilled are the two groups of agents, a positive, negative or zero marignal income tax can be imposed on the poor individual. But the agents below the poverty line have to overwork and that makes their utility to decrease compared to the case where incomes are observable with no cost.

Figure 1 - MRTS - Poor household

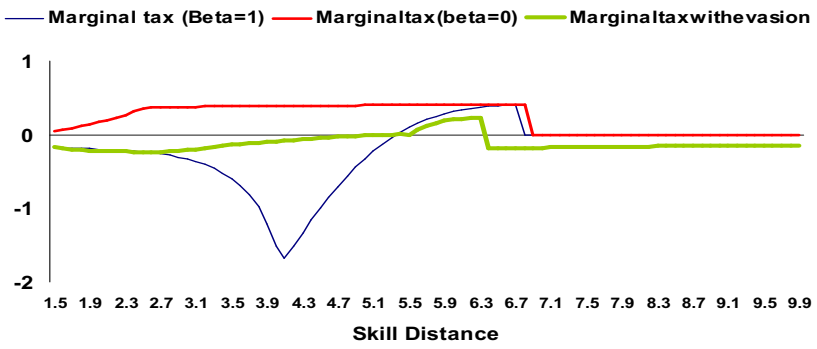


Figure 2 - Consumption of the Poor

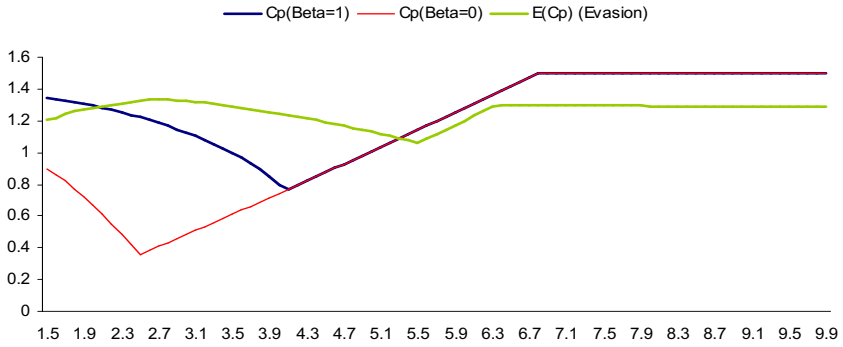


Figure 3 - Labor supply of the poor

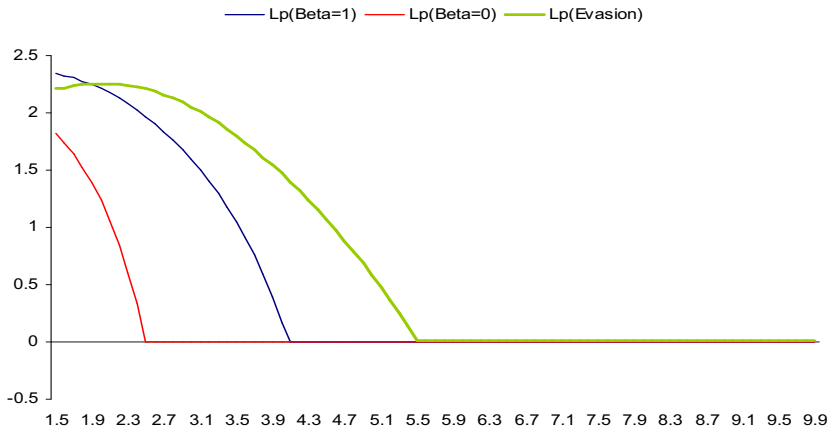


Figure 4 - Tax on Poor

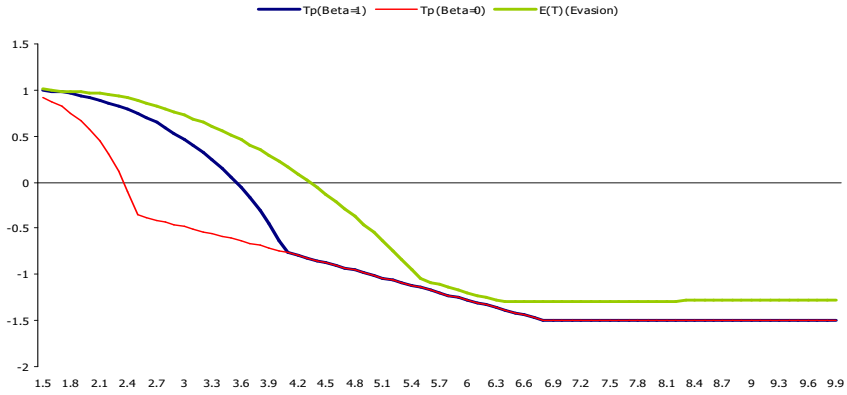
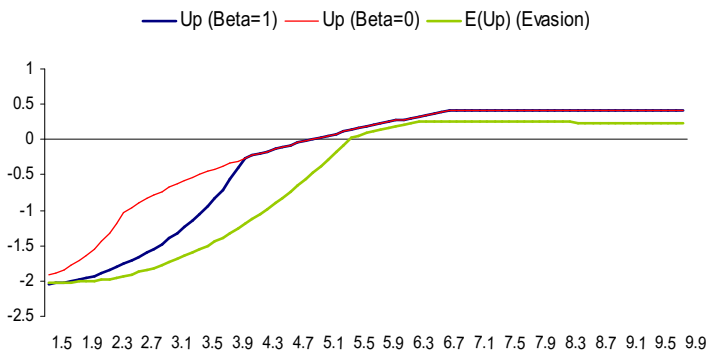
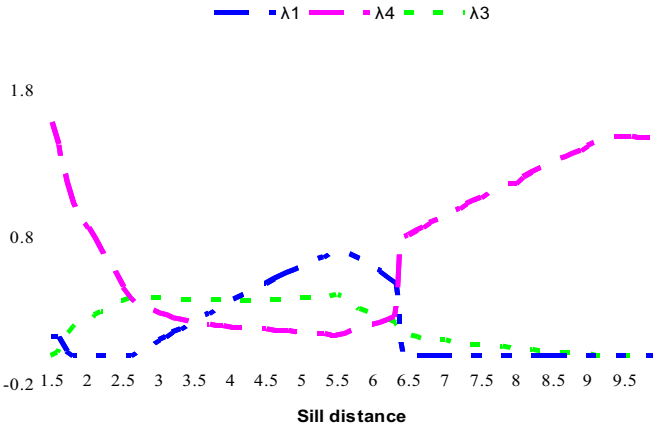


Figure 5 - Utility of the poor



Self-selection and moral hazard constraints



References

[1] Besley, T. and Coate, S.(1992), "Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs", American Economic Review, vol 82, 249–261.

[2] Besley, T. and Coate, S.(1995), "The design of income maintantenance programs", Reveiw of Economic Studies, 62, 187-221.

[3] Cremer, H. and Gahvari, F. (1996),"Tax evasion and the optimum general income tax", Journal of Public Economics, 60, 235-249.

[4] Garfinkel, I. (1973), "A skeptical note on "The optmality" of wage subsidy programs", American Economic Review, 63, 447-453.

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

[6] Kesselman, J. (1971) "Conditional Subsidies in Income Maintenance", Western Economic Journal, 9,1-20.

[7] Melumad, N. and Mookherjee D. (1989), "Delegation as a commitment:the case of income tax audits", Rand Journal of Economics, 20, 139-163.

[8] Mookherjee D. and PNG, I. (1989), "Optimal auditing, insurance, and redistribution", Quarterly Journal of Economics, 103, 399-415.

[9] Nichols, A. L. and Zeckhauser, R.(1982),"Targeting Transfer through Restrictions on Recipient", American Economic Review, vol 78, 372–377.

[10] Ravallion, M. (1994), "Poverty Comparisons" in Harwood Academic press fundamentals of pure and applied economics, 56.

[11] Sandmo, A. (1981), "Income tax evasion, labour supply, and the equity-efficiency tradeoff", Journal of Public Economics, 16, 265-288.

[12] Sandmo, A. (2004), "The theory of tax evasion: A retrospective review", mimeo.

[13] Schneider, F. and Enste, H. (2000): "Shadow Economies: Size, causes and consequences.". Journal of Economic Literature, 38:77-114.

[14] Sen, A. (1982), "Choice Welfare and Measurement". John Hopkins University press.

[15] Tanuri-Pianto, M. and Pianto, D. (2002), "A choice at the top and segmentation at the bottom: a quantile regression approach", mimeo.

[16] Wally, W. (2001), "Optimal general income tax and poverty as a public bad", Journal of Public Economics, vol 82, 2: 271-299.

[17] Zeckhauser, R.(1971), "Optimal mechanisms for income transfer", American Economic Review, 61, 324-334.