

Ensayos

sobre POLÍTICA ECONÓMICA



THE DESIGN OF OPTIMAL EDUCATION POLICIES WHEN INDIVIDUALS DIFFER IN INHERITED WEALTH AND ABILITY

DARÍO MALDONADO

ENSAYOS SOBRE POLÍTICA ECONÓMICA,
VOL. 25, NÚM. 55,
EDICIÓN DICIEMBRE 2007
PP. 84-108

Los derechos de reproducción de este documento son propiedad de la revista Ensayos Sobre Política Económica (ESPE). El documento puede ser reproducido libremente para uso académico, siempre y cuando no se obtenga lucro por este concepto y además, cada copia incluya la referencia bibliográfica de ESPE. El(los) autor(es) del documento puede(n) además poner en su propio website una versión electrónica del mismo, pero incluyendo la referencia bibliográfica de ESPE. La reproducción de esta revista para cualquier otro fin, o su colocación en cualquier otro website, requerirá autorización previa de su Editor de ESPE.

EL DISEÑO DE POLÍTICAS EDUCATIVAS ÓPTIMAS CUANDO LOS INDIVIDUOS DIFIEREN EN SU RIQUEZA HEREDADA Y SU HABILIDAD

DARÍO MALDONADO*

En éste artículo considero el papel de las políticas educativas dentro de un esquema de redistribución del ingreso cuando los individuos difieren en dos parámetros: habilidad y riqueza heredada. Discuto la diferencia entre las reglas que emergen cuando los individuos difieren en sólo uno de estos parámetros y cuando difieren en los dos. La conclusión principal es que la forma de la regla que emerge cuando los individuos difieren sólo en su habilidad es la misma que cuando difieren en habilidad y riqueza heredada. La diferencia entre las reglas que emergen en las dos situaciones es en la manera de implementarlas y no en la forma de la diferencias entre las asignaciones de educación.

Clasificación JEL: H21, H23, H52, I28, J31.

Palabras clave: tributación óptima, educación, provisión pública, *screening* multidimensional.

* Agradezco los comentarios de Helmuth Cremer, Pierre Pestieau, Guido Friebel y Juan Daniel Oviedo y la ayuda financiera del Banco de la República.

Facultad de Economía,
Universidad del Rosario.
Correo electrónico: dario.maldonadoca@urosario.edu.co

Documento recibido el
10 de agosto de 2007;
versión final aceptada el
20 de noviembre de 2007.

THE DESIGN OF OPTIMAL EDUCATION POLICIES WHEN INDIVIDUALS DIFFER IN INHERITED WEALTH AND ABILITY

DARÍO MALDONADO*

In this paper I consider the role of education policies in redistribution of income when individuals differ in two aspects: ability and inherited wealth. I discuss the extent to which the rules that emerge in unidimensional settings apply also in the bidimensional setting considered in this paper. The main conclusion is that, subject to some qualifications, the same type of rules that determine optimal education policies when only ability heterogeneity is considered apply to the case where both parameters of heterogeneity are considered. The qualifications pertain to the implementation of the optimal allocation of resources to education and not the way the optimal allocations first- and second-best differ.

JEL Classification: H21, H23, H52, I28, J31.

Keywords: optimal taxation, education, public provision, multidimensional screening.

* The author thanks Helmuth Cremer, Pierre Pestieau, Guido Friebel, Juan Daniel Oviedo and the anonymous referee for their comments and contributions; and also Banco de la República for funding.

Department of Economics,
Universidad del Rosario.

E-mail: dario.maldonadoca@urosario.edu.co

Document received 10 August 2007; final version accepted 20 November 2007.

I. INTRODUCTION

The reform of education policy forms part of the policy agendas of most countries today. Efficiency and redistributive concerns motivate the need to reform education systems. Proposals for reform aim at an appropriate allocation of resources to education; most reform agendas want to guarantee that education resources are placed where they are most efficient. For some proponents of these reforms, this has two implications. First, the most able individuals should receive the biggest part of the education budget since that guarantees its most efficient use. Second, education systems must be designed in order to minimize the effect that differences in pure wealth may have on access to education.¹ This concern can be rephrased in terms of two parameters: the ability to learn and the ability to pay for access to education. The first parameter affects the efficiency of education investments; the second determines the amount of individual resources available for consumption and investment in education. Heterogeneous learning abilities motivate policies in which the bulk of education resources are allocated to the individuals with the greatest learning abilities in order to maximize the returns from these investments. Heterogeneity in exogenous wealth motivates policies in which, on grounds of equality of opportunity, public resources for education are allocated to all individuals regardless of their ability to pay.

In this paper I address the design of education policies when individuals differ in both of these parameters in order to examine the extent to which optimal taxation

¹ However, if investments in education and other forms of wealth —like non-monetary family resources— are complementary, this conclusion may not apply.

models support the previous arguments. A proper examination of this question requires the consideration of other aspects of policy design. The first is the fact that education policies never appear as isolated policy instruments. Almost all modern economies are managed by governments that use other policy instruments; probably the most important of them is the tax code. The second aspect is the non-linearity of welfare-improving policies, which is another characteristic of most policies implemented in reality. These considerations follow from the Second Theorem of Welfare Economics and from traditional results found in the literature on optimal non-linear income taxation (Cremer, Pestieau and Rochet, 2001; Stiglitz, 1987).

Consequently, in this paper I consider the optimal design of education policies when the government optimally sets the income tax code and taxes are non-linear. My main concern in this paper is how education policy depends on the information available to a utilitarian government. The main conclusion is that the same type of rules for allocating resources for education apply when wealth heterogeneity is considered as when it is not considered. In this respect the conclusion is the same as in Maldonado (2007): the education gap between high- and low-ability individuals is wider in the first-best than in the second-best allocation if the education elasticity of wage is decreasing in ability. This conclusion has the important implication that heterogeneity in exogenous wealth does not affect the way education policy differs in first- and in second-best. Heterogeneity in ability to pay, however, does have implications for optimal education policies: it affects the way education policy is implemented.

Traditionally, economists have recognized that income taxes affect labor supply, however, they also affect education choices.² Education taxes or subsidies can be used to reinforce or reduce the effect of income taxes on education expenditure. Thus, depending on whether one wants to reinforce or reduce this effect, one can tax or subsidize resources devoted to education. It is in this respect that the implementation of the optimal allocation of resources to education when the double heterogeneity is considered differs from the case when it is not considered. In a two-class economy where individuals differ only in ability, even if ability and education are complements, low-ability individuals may receive an education subsidy.³ This will

² The effects of income taxes on education choice are quantified by Blöndal, Field and Girouard (2002).

³ The traditional “no-distortion at the top” result holds in the two-class economy version of this model. Consequently marginal tax rates on labor income and on education for high-ability individuals are equal to zero.

happen if the education elasticity of wage decreases with ability. This result holds as well when individuals differ in ability and in exogenous wealth as long as the marginal tax rate on labor income for low-ability individuals is positive. Additionally, in the multidimensional case, high-ability individuals may face a negative marginal income tax (a subsidy). Under the stated condition on the education elasticity of wage, these individuals will also face a marginal tax on education. Otherwise, if low-ability individuals face a subsidy on labor market income, the implementation of the optimal allocation of education resources implies a marginal education tax on low-ability individuals and a marginal education subsidy for high-ability individuals.

Related problems have been analyzed by Arrow (1971), Ulph (1977), Hare and Ulph (1979), Toumalala (1986), De Fraja (2002); and Bovenberg and Jacobs (2005). However, either they make important restrictions on functional forms (particularly the wage function) and the instruments used by the government (for example linear taxes) or they do not take into account the double heterogeneity discussed in this paper. In this paper, as in Maldonado (2007), I depart from these assumptions. In terms of conclusions, this paper may be best contrasted with Bovenberg and Jacobs (2005), where the authors argue that the Atkinson-Stiglitz theorem holds for the education policy problem. In other words, education policy only has the role of restoring efficiency, by undoing the effect of income taxes on education choices. In this paper I argue that if one allows for a more general wage function, the role of education subsidies is to restore efficiency when the only source of heterogeneity is exogenous wealth; otherwise if individuals also differ in ability, the Atkinson-Stiglitz efficiency theorem does not hold. The mechanism behind this result highlights a different role for education policies: by affecting relative wages education policy can relax the incentive constraints and make redistribution easier.

The rest of the paper is organized as follows. Section II sets up the model. Section 3 studies the optimal education policies. Section IV shows numerical simulations of the model studied in Section III. The last section concludes.

II. THE MODEL AND THE LAISSEZ-FAIRE SOLUTION

The model I analyze is an extended version of the one used in Maldonado (2007), which is a modified version of the optimal taxation model in Stiglitz (1982). This paper modifies the model used in those papers to account for investments in education and to introduce exogenous wealth heterogeneity. This second parameter of

heterogeneity does not affect the benefits from investments in education but does affect the resources available for consumption and education.

Individuals derive utility directly from consumption and labor supply. Labor market income and inherited wealth are used to pay for consumption and for investment in education; prices of consumption and education are normalized to one. Individuals differ in their ability to learn and their inherited wealth. The first of these two parameters, together with labor supply and investment in education, determine labor market income. In the absence of government intervention, the decision problem of an individual of type i can be expressed as follows:

$$\begin{aligned} \max_{c^i, l^i, q^i} \quad & u(c^i) - v(l^i) \\ \text{s.t.} \quad & q^i + c^i \leq \theta^i + Y^i, \\ & Y^i = \omega(\phi^i, q^i)l^i. \end{aligned}$$

where $u(c^i) - v(l^i)$ is the utility the individual gets from consumption, c^i , and labor supply, l^i . Following the tradition in optimal tax models, I assume that labor market income, Y^i , is linear in labor supply; labor supply must be understood as the time devoted to work and $\omega(\phi^i, q^i)$ is the productivity (or the wage) per unit of time of an individual with ability ϕ^i and who has devoted resources q^i to investment in education. Individuals differ in their ability to learn, ϕ^i , and ability to pay, θ^i . Utility is increasing in consumption and decreasing in labor supply, and labor market productivity, $\omega(\phi^i, q^i)$ is increasing in both arguments. Utility and labor market productivity are concave in choice variables.⁴

The model assumes that there are two possible uses of monetary resources: consumption or investment in education. Consequently the opportunity cost of investing in education is given by foregone consumption instead of the possible alternative of modeling education investments in terms of time and regarding the opportunity cost of investing in education as foregone earnings.

Throughout this paper I will address myself to an analysis of the general tax function $T(Y^i, q^i)$ when it is optimally set by a utilitarian government. Given the tremendous

⁴ The assumption that prices are normalized to one can be the result of assuming that production technologies are linear with equal coefficients for education and consumption. The assumption that the coefficients are equal has no impact on the results of the paper as long as technologies continue to be linear.

degrees of freedom that this formula gives to the planner, any policy that may be implemented by the government can be understood with its use.⁵ When subject to this tax function, the problem of individuals becomes

$$\begin{aligned} \max_{c^i, Y^i, q^i} \quad & u(c^i) - v\left(\frac{Y^i}{\omega(\phi^i, q^i)}\right) \\ \text{s.t.} \quad & q^i + c^i \leq \theta^i + Y^i - T(Y^i, q^i) \end{aligned}$$

where the traditional change of variables in optimal tax literature has been employed.⁶ The first order conditions of the individual problem yield the following arbitrage conditions:⁷

$$MRS_{cl}^i = \omega(\phi^i, q^i)[1 - T_Y(Y^i, q^i)], \quad (1)$$

$$MRS_{cq}^i = [1 + T_q(Y^i, q^i)] \quad (2)$$

and

$$MRT_{lq}^i = \frac{1}{\omega(\phi^i, q^i)} \frac{1 + T_q(Y^i, q^i)}{1 - T_Y(Y^i, q^i)}. \quad (3)$$

In these formulae MRS and MRT stand for Marginal Rate of Substitution and Marginal Rate of Transformation, as usual. For the following analysis it is important to note that the MRT between labor supply and investment in education depends on both marginal tax rates, T_Y and T_q . These formulae will also be used to understand the implementation of optimal policies further along in this paper.

⁵ To more clearly see that this formula can subsume many different policies, note that I am not making any assumption about the slope of this formula. Consequently, this formula includes the case of linear subsidies (T_q constant and negative) or the case where education policy takes the form of in-kind transfers (T_q taking infinite values in all q except in the education level that the government wants to achieve); for more details on the relation of non-linear taxation and in-kind transfers see Cremer and Gahvari (1997) and for more details on the tax function see Maldonado (2007).

⁶ This change of variables simply implies replacing l^i with its equivalent Y^i/ω .

⁷ For the sake of simplicity, in these expressions I make some abuse of notation. It is well known that the optimal non-linear tax functions are not differentiable; however, these expressions enable us to see the wedge between the MRS and the MRT introduced by taxes. Introducing non-differentiability at this stage will only complicate the argument without giving any advantage.

III. THE OPTIMAL EDUCATION POLICY

Before addressing the optimal design of the tax function $T(Y^i, q^i)$ under asymmetric information conditions, a few words about the optimal policies under first-best conditions are in order. The first-best problem of a utilitarian planner is

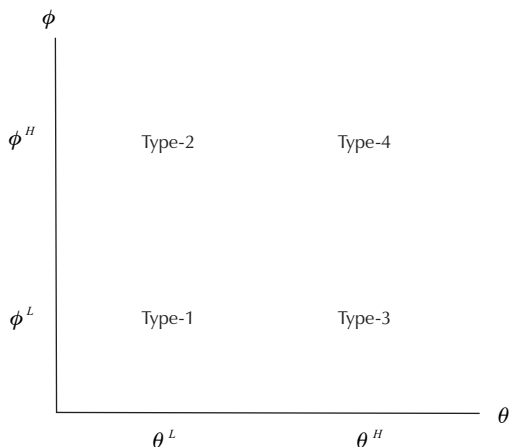
$$\begin{aligned} \max_{\{T^i, Y^i, q^i\}_i} \quad & \sum_i \pi^i \left[u(\theta^i + Y^i - T^i - q^i) - v\left(\frac{Y^i}{\omega(\phi^i, q^i)}\right) \right] \\ \text{s.t.} \quad & \sum_i \pi^i T^i \geq 0. \end{aligned} \tag{4}$$

In the first-best the government directly observes all parameters and decision variables. Consequently it can set any tax-transfer scheme and faces no restrictions on the policies it uses to maximize welfare. A utilitarian government would set the tax function $T(Y^i, q^i)$ so that all individuals receive the same consumption level regardless of their ability or inherited wealth. However, low-ability individuals would work less and receive less education than high-ability individuals. Moreover, as argued by the Second Theorem of Welfare Economics, the government does not need to introduce any distortions of individual choices. This means that marginal tax rates (the derivatives of $T(Y^i, q^i)$) will be equal to zero. The balance between labor supply and education investment for each type of individual will be set according to the same first order conditions found in the laissez-faire solution.

Under asymmetries of information things are different. As in Maldonado (2007), my main concern in this paper is the characterization of the function $T(Y^i, q^i)$ when the government has limited information on individuals' characteristics and choice variables. In this paper I deal with qualifications of the conclusions obtained in my previous paper which result from consideration of the double heterogeneity. As in most of the literature on optimal non-linear taxation, this paper adopts the mechanism-design approach to characterize $T(Y^i, q^i)$.

In line with the literature on optimal taxation (Stiglitz, 1987), I will assume that the government observes labor market income $Y^i = \omega(\phi^i, q^i) \times l^i$ but it does not observe ability, productivity, or labor supply. I will also assume that the government observes investments in education and consumption level. Inherited wealth is not observed either. The tax function $T(Y^i, q^i)$ will be characterized for a four-class economy where ϕ^i and θ^i take only two values. Specifically,

Figure 1
Types of Individuals



$\phi^i \in \{\phi^L, \phi^H\}$ with $\phi^L < \phi^H$ and $\theta^i \in \{\theta^L, \theta^H\}$ with $\theta^L < \theta^H$. Individuals will be labeled according to Figure 1.

As a consequence of asymmetric information, the government must rely on direct revelation mechanisms; the Revelation Principle guarantees that there is no loss of generality associated with this choice and that any optimal indirect mechanism (such as a tax function) would be equivalent to the direct mechanism. The key elements of this analysis are the incentive constraints that induce truthful revelation. Let (R^i, Y^i, q^i) be the second-best allocation implemented by the planner, where R^i is after-tax income income ($R^i = Y^i - T^i$), Y^i is labor market income and q^i is the education investment of an individual of type i . Note that the allocation only includes variables observable to the government. Because of asymmetric information the government must design the mechanism (R^i, Y^i, q^i) so that the optimal choice for any individual i is the allocation designed for his type instead of the allocations designed for individuals of other types. This means the planner will face 12 incentive constraints that take the following form:

$$u(\theta^i + R^i - q^i) - v\left(\frac{Y^i}{\omega(\phi^i, q^i)}\right) \geq u(\theta^i + R^k - q^k) - v\left(\frac{Y^k}{\omega(\phi^i, q^k)}\right). \quad (5)$$

The problem of the planner is problem (4) with the additional incentive constraints, (5). Note that there are twelve incentive constraints restricting the allocations that the government can choose.⁸ Letting λ be the Lagrange multiplier of the resource constraint and μ^{ik} the multiplier of the constraint which deters type- i individuals from mimicking type- k individuals, the Lagrangian of this maximization problem is:

$$\begin{aligned} \Lambda = & \sum_{i=1}^4 \pi^i \left[u(\theta^i + R^i - q^i) - v\left(\frac{Y^i}{\omega(\phi^i, q^i)}\right) \right] + \lambda \left\{ \sum_{i=1}^4 \pi^i [Y^i - R^i] \right\} \\ & + \sum_{i,k:i \neq k} \left[\mu^{ik} u(\theta^i + R^i - q^i) - v\left(\frac{Y^i}{\omega(\phi^i, q^i)}\right) \right. \\ & \quad \left. - u(\theta^i + R^k - q^k) + v\left(\frac{Y^k}{\omega(\phi^i, q^k)}\right) \right] \end{aligned} \quad (6)$$

The first order conditions of this problem are:

$$\begin{aligned} [Y^i]: & \pi^i \left[-v' \left(\frac{Y^i}{\omega(\phi^i, q^i)} \right) \frac{1}{\omega(\phi^i, q^i)} + \lambda \right] \\ & + \sum_{k:k \neq i} \left[-\mu^{ik} v' \left(\frac{Y^i}{\omega(\phi^i, q^i)} \right) \frac{1}{\omega(\phi^i, q^i)} + \mu^{ki} v' \left(\frac{Y^i}{\omega(\phi^k, q^i)} \right) \frac{1}{\omega(\phi^k, q^i)} \right] = 0, \end{aligned} \quad (7)$$

$$\begin{aligned} [R^i]: & \pi^i \left[u'(\theta^i + R^i - q^i) - \lambda \right] \\ & + \sum_{k:k \neq i} \left[\mu^{ik} u'(\theta^i + R^i - q^i) - \mu^{ki} u'(\theta^k + R^i - q^i) \right] = 0 \end{aligned} \quad (8)$$

and

⁸ The technical difficulties involved in solving multidimensional screening problems are well known. A general analysis of the mechanism design problem in a multidimensional setting has been done by Armstrong and Rochet (1999). The results they derive cannot be applied directly here since the optimal tax problem requires the introduction of a budget constraint and I am not assuming quasi-linearity of the utility function. The difficulty in analyzing the optimization problem in this paper is that if one wants to do the full Kuhn-Tucker analysis, one would have to compare the optimal solutions of the 144 different optimization problems that emerge for all the possible combinations of binding incentive constraints. The treatment I give to this problem is similar to that in Cremer, Pestieau and Rochet (2001).

$$\begin{aligned}
 [q^i]: \pi^i & \left[-u'(\theta^i + R^i - q^i) + v' \left(\frac{Y^i}{\omega(\phi^i, q^i)} \right) \frac{Y^i \omega_q(\phi^i, q^i)}{(\omega(\phi^i, q^i))^2} \right] \\
 & + \sum_{k:k \neq i} \mu^{ik} \left[-u'(\theta^i + R^i - q^i) + v' \left(\frac{Y^i}{\omega(\phi^i, q^i)} \right) \frac{Y^i \omega_q(\phi^i, q^i)}{(\omega(\phi^i, q^i))^2} \right] \\
 & - \sum_{k:k \neq i} \mu^{ki} \left[-u'(\theta^k + R^i - q^i) + v' \left(\frac{Y^i}{\omega(\phi^k, q^i)} \right) \frac{Y^i \omega_q(\phi^k, q^i)}{(\omega(\phi^k, q^i))^2} \right] = 0. \tag{9}
 \end{aligned}$$

To rearrange these expressions, I first define

$$\eta(\phi^i, q^i) \equiv \frac{q \times \omega_q(\phi^i, q^i)}{\omega(\phi^i, q^i)};$$

where $\eta(\phi^i, q^i)$ is the education elasticity of the wage function.⁹

The expressions for the marginal tax rates are obtained by combining and rearranging (7), (8) and (9). Using the notation $c^i = \theta^i + R^i - q^i$, $c^{ki} = \theta^i + R^k - q^k$, $l^i = Y^i / \omega(\phi^i, q^i)$, and $l^{ki} = Y^i / \omega(\phi^k, q^i)$. These expressions are

$$\begin{aligned}
 -MRS_{cl}^i & = \omega(\phi^i, q^i) \frac{1 + \sum_{k:k \neq i} \frac{\mu^{ik}}{\pi^i} - \sum_{k:k \neq i} \frac{\mu^{ki}}{\pi^i} \frac{u'(c^{ki})}{u'(c^i)}}{1 + \sum_{k:k \neq i} \frac{\mu^{ik}}{\pi^i} - \sum_{k:k \neq i} \frac{\mu^{ki}}{\pi^i} \frac{\omega(\phi^i, q^i)}{\omega(\phi^k, q^i)} \frac{v'(l^{ki})}{v'(l^i)}} \tag{10}
 \end{aligned}$$

and

$$\begin{aligned}
 MRT_{lq}^i & = \frac{1}{\omega(\phi^i, q^i)} \frac{1 + \sum_{k:k \neq i} \frac{\mu^{ik}}{\pi^i} - \sum_{k:k \neq i} \frac{\mu^{ki}}{\pi^i} \frac{\omega(\phi^i, q^i)}{\omega(\phi^k, q^i)} \frac{v'(l^{ki})}{v'(l^i)}}{1 + \sum_{k:k \neq i} \frac{\mu^{ik}}{\pi^i} - \sum_{k:k \neq i} \frac{\mu^{ki}}{\pi^i} \frac{\eta(\phi^k, q^i)}{\eta(\phi^i, q^i)} \frac{\omega(\phi^i, q^i)}{\omega(\phi^k, q^i)} \frac{v'(l^{ki})}{v'(l^i)}} \tag{11}
 \end{aligned}$$

⁹ Note that one can see $\omega(\phi^i, q^i)$ as a Mincer equation. Consequently, $\eta(\phi^i, q^i)$ corresponds to the parameter that accompanies years of education in this type of equation. The literature on education has called this parameter the return to education; even if, with the specification adopted in this paper, it were correct to give this name to $\eta(\phi^i, q^i)$, given the fact that the term return to education has been widely misused, I prefer to use for $\eta(\phi^i, q^i)$ its more direct meaning. I could also call $\eta(\phi^i, q^i)$ the growth rate of income with education. This point is elaborated by Heckman, Lochner and Todd (2005).

Equations (10) and (11) define the marginal tax rates on Y^i and q^i . In this case, because of the impossibility of knowing which of the incentive constraints are binding, it is difficult to determine a priori the signs of the marginal tax rates (some restrictions on marginal tax rates are provided in the appendix). However, it is possible to characterize the overall distortion on education with respect to labor supply.

From equation (11), it can be seen that the key to the analysis of the distortions on the allocation of education resources is the form of $\eta(\phi^i, q^i)$. The direction of the overall distortion on education with respect to labor supply will be downwards, flat or upwards as $MRT_{lq}^i \stackrel{\geq}{\leq} 1/\omega(\phi^i, q^i)$. From (11) it is seen that this happens according to the sign of

$$\sum_{k:k \neq i} \frac{\mu^{ki}}{\pi^i} \frac{\omega(\phi^i, q^i)}{\omega(\phi^k, q^i)} \frac{v'(l^{ki})}{v'(l^i)} \left[1 - \frac{\eta(\phi^k, q^i)}{\eta(\phi^i, q^i)} \right]. \quad (12)$$

If this expression is negative, the marginal rate of transformation between education and labor is greater than the inverse of the wage rate; if it is positive, it is smaller. Notice that for each individual i the ratio $\eta(\phi^k, q^i)/\eta(\phi^i, q^i)$ can only take two values: if the individual is a low-ability one, the ratio can be one or $\eta(\phi^H, q^i)/\eta(\phi^L, q^i)$; if he or she is a high-ability individual, the ratio is one or $\eta(\phi^L, q^i)/\eta(\phi^H, q^i)$. Accordingly, the distortion on the allocation of education resources will be related to whether $\eta(\phi^H, q^i) \stackrel{\geq}{\leq} \eta(\phi^L, q^i)$. This means that, if distorted at all, individuals with different ability parameters will face opposite distortions on education.¹⁰

The effect of the assumptions about the form of $\omega(\phi^i, q^i)$ on optimal education policy has already been introduced in Maldonado (2007). The three possible cases analyzed there are

- A1** $\eta(\phi^i, q)$ increasing in ϕ^i for all q ,
- A2** $\eta(\phi^i, q)$ decreasing in ϕ^i for all q , and
- A3** $\eta(\phi^i, q)$ is independent of ϕ^i for all q .

Consider the case in which A1 holds ($\eta(\phi^i, q)$ is increasing in ϕ^i). Individuals with a low-ability parameter will have $MRT_{lq}^i \geq 1/\omega(\phi^L, q^i)$ and individuals

¹⁰ By "distorted at all" I mean that for a type- i individual at least one of the multipliers μ^{ki} with k is such that $\phi^k = \phi^i$ is different from zero.

with a high-ability parameter will have $MRT_{lq}^i \leq 1/\omega(\phi^H, q^i)$.¹¹ This means that the education level of low-ability individuals will be distorted downward (if distorted) and that of high-ability individuals upward (if distorted). If it is A2 that holds, the opposite pattern is found. Similarly, when $\eta(\phi^L, q^i) = \eta(\phi^H, q^i)$ the expression in (12) will be equal to zero and there will be no distortion on the education level. The presence of a distortion requires that one of the incentive constraints that links individuals who differ in ability binds. If the only binding constraints are those linking individuals of equal abilities, there would be no distortions on education. These results are summarized in Proposition 1, which restates Proposition 1 in Maldonado (2007) to take into account the double-heterogeneity of individuals set forth in this paper.

Proposition 1. Suppose A1 (A2) holds. Consider the low-ability individuals, i.e. $i = 1,3$. If $\mu^{2i} \neq 0$ or $\mu^{4i} \neq 0$ then type- i individuals will face a downward (upward) distortion of education with respect to labor supply. Consider the high-ability individuals, i.e. $i = 2,4$. If $\mu^{1i} \neq 0$ or $\mu^{3i} \neq 0$, then type- i individuals will face an upward (downward) distortion on education with respect to labor supply. If A3 holds, education will not be distorted with respect to labor supply.

The results in Proposition 1 hold if at least one of the incentive constraints linking individuals differing in ability binds. If none of the incentive constraints is binding, the solution would be identical to that of the first-best problem. If the only incentive constraints that bind are those linking individuals differing in exogenous wealth, there are no distortions on education with respect to labor supply. If at least one of the incentive constraints that links individuals of different ability binds, and $\frac{\partial}{\partial \phi^i} \eta(\phi^i, q^i) > 0$ (A1), the education gap between high- and low-ability individuals will be wider than in the first-best. If $\frac{\partial}{\partial \phi^i} \eta(\phi^i, q^i) < 0$ (A2) holds, the same gap will be narrowed. In the case in which $\frac{\partial}{\partial \phi^i} \eta(\phi^i, q^i) = 0$ (A3), the gap will be kept to its

¹¹ I do not use strict inequality to also consider the case in which that type of individual is not distorted.

first-best level.¹² Moreover, if A1 holds, the optimal education policy is input-regressive, but if A2 holds, it is marginally input-progressive.¹³

The reason for this result is that the government uses distortions of education levels to deter mimicking behavior. If the indifference curves of individuals of different abilities in the (l^i, q^i) plane are parallel, distortions on this plane are not useful to deter mimicking behavior. This happens if $\eta(\phi^i, q^i)$ does not depend on ϕ^i . Note that this is not changed by the multidimensional heterogeneity, since these indifference curves do not depend on θ^i . However, if $\eta(\phi^i, q^i)$ depends on ϕ^i , distorting education choices becomes a useful method for separating individuals of different types.

Notice the effect of putting together the two parameters of heterogeneity (ϕ^i and θ^i). Under bidimensional heterogeneity it is not any more evident that marginal taxes are positive for low-ability individuals and zero for high-ability ones, as it would be in the result in a model in which individuals only differ in ability. Now, the marginal income tax rate can have negative or positive signs and high-ability individuals may face non-zero marginal taxes. An analysis of the relationships between the marginal tax rate on labor market income and on education is presented in the appendix. It is interesting to note here that positive and negative marginal tax rates on labor income can emerge in this model. Moreover, a marginal tax on labor income of a given sign is not always coupled with an education marginal tax rate on of the same sign.

A particularly interesting case emerges if A2 holds. In this case, the optimal allocation is implemented with a positive tax on labor market income and a marginal subsidy on education for low-ability individuals, together with a negative tax on labor income and a positive tax on education for high-ability individuals. This leads to the possibility of having something that can be called over-qualification (or under-activity): low-ability individuals acquire more education than it may be optimal for them

12 Whether A1, A2 or A3 holds in data is still a controversial question. Some researchers have sought estimates of this parameter but there is no consensus; see Arias, Hallock and Sosa (1999), Ashenfelter and Rouse (1998), Girma and Kedir (2005) and Tobias (2003).

13 The results in Proposition 1 cannot be easily extended to a more than two-type case. In such a case any individual different from the one with the highest or lowest abilities could be simultaneously mimicked by individuals with higher and lower abilities. This makes the term $1 - \frac{\eta(\phi^k, q^i)}{\eta(\phi^l, q^i)}$ take positive and negative values for a given individual, and (12) to have an ambiguous sign.

to have in the first-best situation. However, they will provide a lower labor supply than in the first-best situation.

As argued above, in Maldonado (2007) I show that if individuals differ only in learning ability, i.e. if $\pi^3 = \pi^4 = 0$, $\pi^1 > 0$ and $\pi^2 > 0$, the optimal tax structure satisfies zero marginal tax rates on labor and education for high-ability individuals and a positive marginal tax rate on labor supply for low-ability individuals. The marginal taxes on education for low-ability individuals follow the same logic as in the model with the double heterogeneity: whether they receive a subsidy or a tax depends on the sign of the education elasticity of wage.

A different polar case is that in which individuals differ only in inherited wealth, i.e. $\pi^2 = \pi^4 = 0$ and $\pi^1 + \pi^3 = 1$. In the first-best solution of this problem the planner will set equal levels of labor supply and education and a lump-sum transfer from type-3 to type-1 individuals. Since the planner is a pure utilitarian he will want to redistribute income from type-3 to type-1 individuals. This means that there is only one binding constraint: the one preventing type-3 individuals from mimicking type-1 individuals. Consequently, among the Lagrange multipliers of the incentive constraints, only μ^{31} will differ from zero, and the first order conditions (7) to (9) yield

$$-MRS_{cl}^1 = \omega(\phi, q^1) \frac{1 - \frac{\mu^{31} u'(c^{31})}{\pi^1 u'(c^1)}}{1 - \frac{\mu^{31}}{\pi^1}},$$

$$MRT_{lq}^1 = \frac{1}{\omega(\phi, q^1)},$$

$$-MRS_{cl}^3 = \omega(\phi, q^3) \text{ and}$$

$$MRT_{lq}^3 = \frac{1}{\omega(\phi, q^3)}.$$

The above equations imply undistorted choices for type-3 individuals and an upward distortion on labor supply for type-1 individuals. Consequently, the marginal rate on labor income is zero for type-3 individuals and negative for type-1 individuals. Simultaneously, there will be a positive marginal tax rate on education equal to the negative of the marginal income tax to avoid distorting education levels. Thus, the structure of the returns to labor will have no effect on optimal education policy. In this case, education should be taxed for high-ability individuals and the only role for

this tax will be to restore efficiency between labor supply and education. The reason for this is that, since individuals differ only in exogenous wealth, indifference curves in the (l^i, q^i) plane are parallel, thus, there is no need to distort education with respect to labor supply.

Recapping, the main result in this section is that the second-best education gap between high- and low-ability individuals will be wider or narrower than the first-best gap depending on which of the two possibilities helps to deter mimicking behavior. If the education elasticity of the wage function is increasing in ability, mimicking behavior will be deterred by a widening of the education gap between high- and low-ability individuals. In the case where the education elasticity of the wage function is decreasing in ability it will be a narrowing of the education gap that will deter mimicking behavior. The effect of this tax on individual education levels may possibly be reversed by using the marginal tax on education. This is true if the education elasticity of wage decreases with ability. This rule implies that the education of low-ability individuals must be subsidized in some cases, particularly if they face a positive marginal income tax.

IV. NUMERICAL SIMULATIONS

In this section I show some numerical simulations for the problem stated in (6). The simulations have three objectives: they shed light on the role of the size-differences between the parameters, show some specific cases where we can see which of the incentive constraints are binding, and illustrate the workings of the qualitative features of the model previously discussed.

All the simulations share some features: the proportion of each type of individual and the utility function of consumers. I assume that there are equal numbers of each type of individual in the economy ($\pi^i = 0.25$ for all $i \in \{1, 2, 3, 4\}$) and that the utility function is

$$u(c) - v(l) = \sqrt{c} - l^2.$$

I show simulations for three different types of wage functions that were chosen according to the behavior of their education elasticity with respect to the ability parameter. According to the theoretical results in the previous sections, this amounts to the question of whether the education elasticity of the wage function increases, decreases or does not change with the ability parameter.

Table 1 shows results for the wage function $\omega(\phi, q) = 10q^\phi$ which satisfies A1, i.e. $\eta(\phi, q)$ is increasing in ϕ . Table 2 shows results for the wage function $\omega(\phi, q) = \phi + q - q^2$ which satisfies A2, i.e. $\eta(\phi, q)$ is decreasing in ϕ . Table 3 shows the results for the function $\omega(\phi, q) = 10\log(q)$ which satisfies A3. For each of the wage functions I show two different simulations: in the first, the difference between the exogenous incomes is “small” and in the second, the difference is “big”. For each set of simulations sharing the same wage function, the only thing that changes is the value of θ^H . In each case I report consumption, labor supply, and education level for the first- and second-best allocations. I additionally report the marginal tax rates, the compound effect of both marginal

Table 1
Education Elasticity Increasing in $\phi, \omega(\phi, q) = 10q^\phi$

Parameters: $\theta^L = 5, \theta^H = 10, \phi^L = \frac{1}{4}, \phi^H = \frac{1}{3}$ (small difference in θ^i)									
Type	First-best			Second-best					
	c	l	q	c	l	q	T_Y	T_q	$\frac{1+T_q}{1-T_Y}$
1	18.19	0.710	2.148	13.60	0.861	2.605	0.0004	0.049	1.049
2	18.19	1.146	7.462	18.79	1.109	7.107	0.0000	0.000	1.000
3	18.19	0.710	2.148	17.20	0.738	2.265	0.0010	-0.001	1.000
4	18.19	1.146	7.462	21.33	0.968	5.793	0.0050	-0.005	1.000

Binding incentive constraints: 1 \rightarrow 3, 2 \rightarrow 4 and 4 \rightarrow 1

Parameters: $\theta^L = 5, \theta^H = 30, \phi^L = \frac{1}{4}, \phi^H = \frac{1}{3}$ (big difference in θ^i)									
Type	First-best			Second-best					
	c	l	q	c	l	q	T_Y	T_q	$\frac{1+T_q}{1-T_Y}$
1	24.62	0.566	1.587	14.56	0.924	2.510	-0.120	0.297	1.158
2	24.62	0.846	4.738	20.41	1.087	6.896	-0.032	0.032	1.000
3	24.62	0.566	1.587	32.20	0.462	1.213	0.000	0.000	1.000
4	24.62	0.846	4.738	35.79	0.652	3.426	-0.034	-0.011	0.956

Binding incentive constraints: 3 \rightarrow 4, 4 \rightarrow 1 and 4 \rightarrow 2

Source: Authors' calculations.

Table 2
Education Elasticity Decreasing in $\phi, \omega(\phi, q) = \phi + q - q^2$

Parameters: $\theta^L = 30, \theta^H = 60, \phi^L = 100, \phi^H = 200$ (small difference in θ^i)

Type	First-best			Second-best					
	c	l	q	c	l	q	T_Y	T_q	$\frac{1+T_q}{1-T_Y}$
1	241.6	1.611	0.190	204.2	1.708	0.211	0.026	-0.038	0.987
2	241.6	1.933	0.241	253.1	1.888	0.235	0.000	0.000	1.000
3	241.6	1.611	0.190	229.0	1.653	0.197	0.001	-0.001	1.000
4	241.6	1.933	0.241	274.2	1.812	0.224	0.002	-0.002	1.000

Binding incentive constraints: 1 \rightarrow 3, 2 \rightarrow 4 and 4 \rightarrow 1

Parameters: $\theta^L = 30, \theta^H = 300, \phi^L = 100, \phi^H = 200$ (big difference in θ^i)

Type	First-best			Second-best					
	c	l	q	c	l	q	T_Y	T_q	$\frac{1+T_q}{1-T_Y}$
1	332.5	1.373	0.136	207.3	1.769	0.232	-0.017	-0.036	0.948
2	332.5	1.647	0.196	269.5	1.892	0.230	-0.034	0.057	1.022
3	332.5	1.373	0.136	410.7	1.235	0.095	0.000	0.000	1.000
4	332.5	1.647	0.196	474.8	1.455	0.143	-0.056	0.096	1.038

Binding incentive constraints: 1 \rightarrow 2, 3 \rightarrow 4 and 4 \rightarrow 1

Source: Authors' calculations.

tax rates on education and the binding incentive constraints for the second-best allocation.¹⁴

Some of the standard features of the optimal income tax problem can be seen in these tables. Most important, because of the separable utility function in the first-best allocation, consumption is equalized among all types of individuals. However,

¹⁴ Quantitative comparisons of the results of sets of simulations with different wage functions should not be made here. This explains why the numerical values for the parameters do not coincide. There is nothing to be gained from choosing the same values and, due to computational constraints, there could be considerable costs in terms of the time needed to obtain interior solutions for each example.

Table 3
Education Elasticity Constant in $\phi, \omega(\phi, q) = \phi \log(q)$

Parameters: $\theta^L = 30, \theta^H = 60, \phi^L = 100, \phi^H = 120$ (small difference in θ^L)									
Type	First-best			Second-best					
	c	l	q	c	l	q	T_y	T_q	$\frac{1+T_q}{1-T_y}$
1	2,070	3.163	316.3	1,733	3.407	340.7	0.0270	-0.0270	1.000
2	2,070	4.084	490.1	2,320	3.815	457.9	0.0000	0.0000	1.000
3	2,070	3.163	316.3	1,763	3.407	340.7	0.0190	-0.0190	1.000
4	2,070	4.084	490.1	2,337	3.797	455.7	0.0004	-0.0004	1.000

Binding incentive constraints: 1 \rightarrow 3, 3 \rightarrow 4, 3 \rightarrow 1, 4 \rightarrow 1 and 4 \rightarrow 3

Parameters: $\theta^L = 30, \theta^H = 2100, \phi^L = 100, \phi^H = 120$ (small difference in θ^L)									
Type	First-best			Second-best					
	c	l	q	c	l	q	T_y	T_q	$\frac{1+T_q}{1-T_y}$
1	2,725	2.677	267.7	1,752	3.527	352.7	-0.007	0.007	1.000
2	2,725	3.465	415.9	2,447	3.744	449.3	-0.011	0.011	1.000
3	2,725	2.677	267.7	3,085	2.482	248.2	0.000	0.000	1.000
4	2,725	3.465	415.9	3,822	2.939	352.7	-0.033	0.033	1.000

Binding incentive constraints: 1 \rightarrow 4, 3 \rightarrow 1, 3 \rightarrow 4, 4 \rightarrow 1 and 4 \rightarrow 2

Source: Authors' calculations.

labor supply and education are higher for high-ability individuals than for low-ability ones. This prevents the implementation of first-best allocations since utility decreases with ability and is independent of exogenous wealth. Moreover, utility in the first-best allocation is independent of exogenous income, which also works against the implementation of the first-best allocation. As a consequence, in the second-best allocation consumption must depend on ability and exogenous wealth, and distortions appear in the picture.

In the second-best allocation marginal taxes stop being equal to zero. The tables show the binding incentive constraints that cause this. These patterns of binding incentive constraints show that the intuitive result of having only “downward” incentive constraints and the no-bunching that holds in one-dimension cases no longer

apply when individuals differ in more than one parameter. In all the simulations there is at least one “upward” incentive constraint that binds.¹⁵ Table 3 shows two cases where there is bunching, i.e. when the upward and downward incentive constraints that link two given types of individuals bind simultaneously.

With respect to marginal taxes, the main regularity across the tables is that there always exists a type of individual who is not distorted. It turns out that this is either a type-2 or type-3 individual. This was expected, since these are the types of individuals with higher and lower marginal rates of substitution. Moreover, when facing a non-zero tax, it is negative for type-2 individuals and positive for type-3 individuals. This is in line with the one-dimensional cases. In the case where individuals differ only in ability, low-ability individuals face a positive marginal income tax and in the case where they differ only in exogenous wealth, low-exogenous wealth individuals face negative marginal income taxes.

The signs of marginal education taxes are consistent with what has been argued throughout this paper and particularly the discussion in the appendix. Low-ability individuals may receive subsidies on their education expenditure. In all cases at least one low-ability individual (type-1 or type-3) has $T_q < 0$. If A2 holds (Table 2), whenever a low-ability individual faces a positive marginal income tax he also faces a marginal subsidy on education. If A1 holds, a positive marginal income tax can go together with a subsidy or a tax on education (respectively type-1 and type-3 in top panel of Table 1). When A3 holds (Table 3), marginal taxes on education are always equal to the marginal tax on labor income.

The education levels and the education gap between high- and low-ability individuals also confirmed the analytical results in this paper. The last column of Table 2 shows that, under assumption A2, type-1 individuals face an upward distortion on education, while the bottom panel shows that type-2 and type-4 individuals face downward distortions on education. In Table 3 the opposite is found. It is only in Table 3, where $\eta(\phi, q)$ does not depend on ϕ , that there are no distortions of education with respect to labor supply. Finally, note that the simulation exercises constitute examples of policies that can be input-regressive or input-progressive not only in marginal terms but also in absolute terms, (in the sense of Arrow, 1971). In Table 1, type-1 individuals

¹⁵ Upward incentive constraints are those that prevent an individual with low ability or low exogenous wealth from mimicking an individual with high ability or high exogenous wealth.

have higher levels of education than type-3 and type-4 individuals. In Table 3, type-1 individuals have exactly the same level of education as type-4 individuals.

V. CONCLUDING COMMENTS

In this paper I have discussed the design of optimal education policies when income taxation is also designed optimally and when individuals differ in learning ability and in inherited wealth. I have treated the problem carefully, in accordance with insights provided by the traditional theories of welfare economics and optimal taxation. This meant introducing, not only education policy, but also income taxation in a setting where education and labor supply are non-separable.

I have shown that, in a model where the government observes directly the individual's education level, the distortion of the education level may not have the same sign as the distortion of labor supply. Education subsidies may complement income taxes. But this is just one of the possible cases that can emerge in this model. I have shown the condition under which it emerges, namely, if the education elasticity of the wage function is decreasing in ability, and it is optimal to set a positive marginal tax on low-ability individuals, these individuals will also face a subsidy on education.

The main purpose of the paper has been to contrast the optimal policies that emerge when the two dimensions of heterogeneity are considered with those that emerge when only one dimension is considered. The main conclusion is that the way the second-best education gap (between high- and low-ability individuals) differs from the first-best one does not depend on whether one considers the bidimensional heterogeneity or only heterogeneity in ability. However, important differences between the ways to implement this education gap emerge between the two models. There are important differences between the model where heterogeneity is limited to inherited wealth and the one where double-heterogeneity is considered. In the first, the education gap is kept to its first-best level (conditional on the levels of labor supply that do differ in both situations). The model where heterogeneity is limited to inherited wealth highlights the nature of the use of education policy in the models where heterogeneity in ability is considered. In the first, the only use for education policy is to correct the inefficiency caused by the effect of income taxes on the labor-supply-education margin; in the second, education policy has a real role in the redistribution of income.

REFERENCES

1. Arias, O.; Hallock, K. F.; Sosa, W. "Individual Heterogeneity in the Returns to Schooling: Instrumental Variables Quantile Regression using Twins Data", *Empirical Economics*, vol. 26, pp. 7-40, 1999.
2. Armstrong, M.; Rochet, J. C. "Multi-dimensional Screening: A User's Guide", *European Economic Review*, vol. 43, pp. 959-979, 1999.
3. Arrow, K. J. "A Utilitarian Approach to the Concept of Equality in Public Expenditures", *Quarterly Journal of Economics*, vol. 85, pp. 409-415, 1971.
4. Ashenfelter, O.; Rouse, C. "Income, Schooling, and Ability: Evidence from a New Sample of Identical Twins", *Quarterly Journal of Economics*, vol. 213, pp. 253-284, 1998.
5. Blöndal, S.; Field, S.; Girouard, N. "Investment in Human Capital through post-Compulsory Education and Training: Selected Efficiency and Equity Aspects", working paper no. 333, Organization for Economic Co-operation and Development (OECD), Economics Department, 2002.
6. Bovenberg, A. L.; Jacobs, B. "Redistribution and Education Subsidies", *Journal of Public Economics*, vol. 89, 2005-2035, 2005.
7. Cremer, H.; Gahvari, F. "In-kind Transfers, Self-selection and Optimal Tax Policy", *European Economic Review*, vol. 41, pp. 97-114, 1997.
8. Cremer, H.; Pestieau, P.; Rochet, J. C. "Direct versus Indirect Taxation: the Design of the Tax Structure Revisited", *International Economic Review*, vol. 42, pp. 781-799, 2001.
9. De Fraja, G. "The Design of Optimal Education Policies", *Review of Economic Studies*, vol. 69, pp. 437-466, 2002.
10. Girma, S.; Kedir, A. "Heterogeneity in Returns to Schooling: Econometric Evidence from Ethiopia", *The Journal of Development Studies*, vol. 242, pp. 1405-1416, 2005.
11. Hare, P. G.; Ulph, D. T. "On Education and Distribution", *Journal of Political Economy*, vol. 87, pp. S193-S212, 1979.
12. Heckman, J.; Lochner, L.; Todd, P. "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond" (mimeo), *NBER*, working papers, no. 11544 2005.
13. Maldonado, D. "Education Policies and Optimal Taxation", *International Tax and Public Finance* (on edition), 2007.
14. Stiglitz, J. E. "Self-selection and Pareto Efficient Taxation", *Journal of Public Economics*, vol. 17, pp. 213-240, 1982.
15. Stiglitz, J. E. "Pareto Efficient and Optimal Taxation and the New Welfare Economics", in A. J. Auerbach and M. Feldstein (eds.), *Handbook of Public Economics*, Amsterdam: North-Holland, 1987.
16. Tobias, J. L. "Are Returns to Schooling Concentrated among the Most Able? A Semi-parametric Analysis of the Ability-earnings Relationships", *Oxford Bulletin of Economics and Statistics*, vol. 65, pp. 1-29, 2003.
17. Toumala, M. "Optimal Income Taxation and Educational Decisions", *Journal of Public Economics*, vol. 30, pp. 183-198, 1986.
18. Ulph, D. T. "On the Optimal Distribution of Income and Educational Expenditure", *Journal of Public Economics*, vol. 8, pp. 341-356, 1977.

APPENDIX 1 THE MARGINAL TAX RATES

The specific signs of the marginal tax rates on labor income and education expenses are ambiguous in this model because the multidimensionality assumption makes it difficult to know which of the incentive constraints are binding. However, some restrictions on the possible marginal tax rates can be obtained.

From (10) and (11), the two conditions needed to understand these relations are:

$$C1 \quad \sum_{k:k \neq i} \frac{\mu^{ki}}{\pi^i} \frac{u'(c^{ki})}{u'(c^i)} - \sum_{k:k \neq i} \frac{\mu^{ki}}{\pi^i} \frac{\omega(\phi^i, q^i)}{\omega(\phi^k, q^i)} \frac{v'(l^{ki})}{v'(l^i)} \geq 0$$

and

$$C2 \quad \sum_{k:k \neq i} \frac{\mu^{ki}}{\pi^i} \frac{u'(c^{ki})}{u'(c^i)} - \sum_{k:k \neq i} \frac{\mu^{ki}}{\pi^i} \frac{\eta(\phi^k, q^i)}{\eta(\phi^i, q^i)} \frac{\omega(\phi^i, q^i)}{\omega(\phi^k, q^i)} \frac{v'(l^{ki})}{v'(l^i)} \geq 0.$$

Conditions C1 and C2 define four cases under which different configurations of marginal tax rates can appear. Notice that the two conditions differ due to the presence of the relative education elasticity of the wage function of the mimicked individual i and the mimicker k in the second of them. This makes it possible for the left-hand side of both conditions to have different signs.

From equation (1) it can be seen that $T_y(Y, q) \geq 0$, depending on whether $-MRS_{cl}^i \geq \omega(\phi^i, q^i)$. This means that the marginal tax rate on labor supply faced by individual i will be positive if the left-hand side of C1 is strictly greater than zero, and negative if it is strictly less than zero. Similarly, (1) and (3) imply $T_q(Y, q) \geq 0$, depending on whether $-MRS_{cl}^i \times MRT_{lq}^i \geq 1$. This means that marginal tax rates on education will be negative if the left-hand side of C2 is strictly greater than zero, and positive if it is strictly smaller than zero.

Generally, the marginal tax rates faced by a type- i individual will be zero when the corresponding multipliers μ^{ki} (i.e., the multiplier of the constraint that prevents a type- k individual from mimicking a type- i individual) are all zero. Therefore if one of the tax rates is zero, the other one will also be zero. Only in very special cases (depending on $\eta(\phi, q)$) will one but only one of the marginal tax rates be different from zero.

Table A1.1 shows the possible pattern of marginal tax rates in each of these cases; it also labels the cases for further reference. In the table $CI (>)$, $CI (<)$, refers to cases when the left-hand side of conditions C1 or C2 are strictly greater or strictly smaller than zero, respectively.

From Proposition 1 and equation (3), it can be seen that if $\frac{\partial}{\partial \phi} \eta(\phi, q) > 0$ then $T_q^i < -T_Y^i$ for low-ability individuals and $T_q^i > -T_Y^i$ for high-ability individuals. The opposite pattern will be found if $\frac{\partial}{\partial \phi} \eta(\phi, q) < 0$. Assuming $\frac{\partial}{\partial \phi} \eta(\phi, q) = 0$, it will always be that $T_q^i = -T_Y^i$. Thus, depending on which of these assumptions hold, the possible cases resulting from Table A1.1 are restricted. Table A1.2 shows the resulting possible relations between marginal tax rates under the different assumptions on $\eta(\phi, q)$.

Table A1.1
Possible Combinations of Marginal Tax Rates

	C1 (>)	C1 (<)
C2(>)	Case 1: $T_Y^i > 0, T_q^i < 0$	Case 2: $T_Y^i < 0, T_q^i < 0$
C2(<)	Case 3: $T_Y^i > 0, T_q^i > 0$	Case 4: $T_Y^i < 0, T_q^i > 0$

Source: Authors' calculations.

Table A1.2
The Relation between the Marginal Tax Rates

	$\frac{\eta(\phi^k, q^i)}{\eta(\phi^i, q^i)} < 1$	$\frac{\eta(\phi^k, q^i)}{\eta(\phi^i, q^i)} > 1$	$\frac{\eta(\phi^k, q^i)}{\eta(\phi^i, q^i)} = 1$
$T_Y^i < 0$	$T_q^i \leq 0$	$T_q^i > 0$	$T_q^i = -T_Y^i$
$T_Y^i > 0$	$T_q^i < 0$	$T_q^i \leq 0$	$T_q^i = -T_Y^i$
$T_Y^i = 0$	$T_q^i = 0$	$T_q^i = 0$	$T_q^i = 0$

Source: Authors' calculations.

It is difficult to know a priori which of these possibilities will occur. This depends on whether the multipliers of the incentive constraints are or are not different from zero, as well as their size and the differences between the marginal rates of substitution between labor and consumption of mimickers and mimicked. The only thing that can be said is that, when it is different from zero, the marginal tax rate on labor income will be non-positive for type-2 individuals and non-negative for type-3 individuals, since the marginal rate of substitution will always be greater than or equal to that of its mimickers for the former, whereas for the latter it will always be less or equal.