Int Tax Public Finance (2008) 15:131–143
DOI 10.1007/s10797-006-9015-7

Education policies and optimal taxation

Darío Maldonado

Received: 1 June 2005 / Accepted: 6 December 2006 /
Published online: 22 February 2007

Abstract This paper studies the design of education policies in a setting in which optimal redistributive labor taxation is available. It is usually argued that the crucial parameter to guide education policy is the complementarity of education and ability. This is true only when labor supply is fixed or when income taxes are not allowed. I show that, if labor supply is endogenous and if the government can tax income in a nonlinear way, the crucial parameter is how the education elasticity of wage changes with ability. Taking the elasticity criterion into account, education subsidies are optimal in cases in which, under the complementarity criterion, education taxes would be optimal. To do this, I use an asymmetric information setting that motivates nonlinear taxation of income and education.

Keywords Optimal taxation · Education

JEL Classification H21 · H23 · H52 · I28

1 Introduction

For a long time, economists have recognized that individuals adjust their behavior along several margins when facing tax schedules. However, most models used to analyze optimal taxes allow for only one of them. Following the seminal work of Mirrlees (1971), the usual margin considered is labor supply, but participation in the labor market (Diamond, 1980), education (Hare and Ulph, 1979), or retirement (Diamond and Mirrlees, 1978) have also been studied. It is only recently that models allowing simultaneous adjustment of several margins have been considered (Mardener and Rochet, 1995; Cremer, Lozachmeur, and Pestieau, 2004).

D. Maldonado (✉)
Universidad del Rosario, Calle 14 # 4-69, Bogotá, Colombia
e-mail: dario.maldonadoca@urosario.edu.co

My paper is related to this research. I consider the optimal tax when individuals are allowed to adjust two specific margins, namely, education and labor supply. My objective is to study the optimal design of education policies when the government can use the income tax schedule to redistribute. The model analyzed in this paper builds on the optimal tax literature as formulated by Stiglitz (1982), and on its further developments by Cremer and Gahvari (1997). These authors have stated the informational assumptions that justify the use of commodity taxes and in-kind transfers to supplement an optimal nonlinear income tax scheme. Under their assumptions on asymmetries of information, government intervention in education is justified either through taxes (or subsidies) or through public provision. Education subsidies and public provision of education are second-best instruments in the presence of particular asymmetries of information. However, these results say little about the particular design of education policy.

To study the design of education policy, I use a model with the following features. Individual labor market income is determined by a function in which labor supply, education, and innate ability interact. The first two are choice variables while the last is the parameter in which individuals differ. In this framework, I examine optimal policies under the assumption that the government taxes (or subsidizes) income and education in a nonlinear way. The assumptions that justify these policy instruments are the lack of observability of individual innate ability, wage, and labor supply (Stiglitz, 1982; Cremer and Gahvari, 1997); labor market income and education are observed. As usual, in the first-best, only lump-sum taxes are needed.

The main result in this paper is that the form of optimal education policy depends on the education elasticity of the wage function. Low-ability individuals face a positive marginal tax on labor income (as usual), but they may face a tax or subsidy on education. There will be a subsidy when the education elasticity of wage decreases with ability. As a result, low-ability individuals have more education (with respect to labor supply) in the second-best than in the first-best. When the education elasticity of wage increases with ability, a tax or a subsidy can emerge; in any case, it will be larger than the negative of the marginal income tax. In other words, low-ability individuals will have less education (with respect to labor supply) in the second-best than in the first-best. As usual, high-ability individuals face no distortions on labor supply or education.

A similar problem is analyzed by Bovenberg and Jacobs (2005). However, they assume that the wage function is homothetic. As a result, education elasticity of wage does not depend on ability. As a consequence, the optimal education subsidy has the same size of the tax on labor income in order to restore efficiency of education expenditure. In this paper, I show that the role of education subsidies is not to restore efficiency, and that the Atkinson–Stiglitz efficiency theorem does not hold once one allows for a more general form of the returns to education. I highlight a different role for education policies. By affecting relative wages, education policy can relax the incentive constraints and make redistribution easier.

Importantly, in my paper, even if ability and education are complements, low-ability individuals may receive more education in the second-best than in the first-best. This is the case if the education elasticity of wage is decreasing in ability. This contrasts with the result when income taxes are not allowed, as in most previous studies, where a downward distortion on education of the less able is optimal if education and ability are complements (Arrow, 1971; Hare and Ulph, 1979; Toumala, 1986; De Fraja, 2002). It is important to note that the education elasticity may be decreasing in ability even if education and ability are complements.[1] This means that elasticity criterion can yield a very different policy recommendation than the complementarity criterion.[2]

The above claim is even more important when one considers the empirical literature on education. For several decades, econometricians have estimated the education elasticity of wage. However, to the best of my knowledge, the relevance of this parameter has not appeared in any formal analysis of education policies. While most theoretical studies show that education policy depends on whether ability and education are complements, in this paper I argue that the crucial parameter is how the education elasticity of wage varies with ability. As argued earlier, this turns out to be a crucial distinction, since policy recommendations under both criteria can be very different. Some empirical studies have addressed the issue of whether education elasticity is increasing or decreasing with ability; however, there is no consensus among them (Ashenfelter and Rouse, 1998; Arias, Hallock, and Sosa, 1999; Tobias, 2003; Girma and Kedir, 2005).

The rest of the paper is organized as follows. Section 2 sets up a model where individuals choose consumption, labor supply, and education. Section 3 studies the optimal second-best policies. The last section concludes.

2 The model

In this paper, I use a modified version of the Stiglitz (1982) optimal taxation model. The size of the population is normalized to 1, individuals are indexed by i, and π^i is the proportion of individuals of type i. Utility depends positively on consumption, c^i, and negatively on labor supply, l^i; it is given by $u(c^i) - v(l^i)$, and satisfies the usual concavity assumptions.

The main difference with the model in Stiglitz (1982) is that individual labor productivity is endogenous. Productivity is given by $\omega(\phi^i, q^i)$, where q^i represents education, and ϕ^i is the *ability to benefit from education*, which is exogenously given. Productivity is increasing in both arguments, concave in q, and satisfies

$$\omega_{\phi q}(\phi^i, q^i) > 0 \tag{1}$$

which means that *ability and education are complements*. Therefore, high-ability individuals are more effective in drawing benefits from additional units of education,

[1] The education elasticity of wage is decreasing in ability if the complementarity between ability and education is not too strong.

[2] Ulph (1977) and Krause (2006) study education policies in optimal taxation frameworks. The first paper makes a first attempt to link education policy to the education elasticity of the wage function. However, his model turns to be too complicated to yield meaningful results. The second author highlights a different mechanism than the one in this paper. In his paper, in the second-best, it is also optimal to set a wider education gap of high-ability and low-ability individuals than in the first-best. As in the previous contributions to the subject, this results from the complementarity between ability and education.

than low-ability individuals.[3] The education elasticity of wage,

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

will be crucial in the subsequent analysis.[4] It is worth noticing that Eq. (1) does not imply how $\eta(\phi^i, q^i)$ varies with ϕ; under Eq. (1), $\eta_\phi(\phi^i, q^i)$ can be positive or negative. If $\eta_\phi(\phi^i, q^i)$ is positive (negative), the *proportional* benefit of an additional unit of education is higher (lower) for high-ability individuals than for low-ability individuals.

Let $Y^i = \omega(\phi^i, q^i)l$ represent labor market income and suppose that individuals are subject to a nonlinear tax on labor income and education, $T(Y^i, q^i)$. Normalizing the price of consumption and education to 1, and making the usual change of variables in optimal tax theory, the problem solved by individuals can be written as follows:

$$\max_{Y^i, q^i} \quad u(Y^i - T(Y^i, q^i) - q^i) - v\left(\frac{Y^i}{\omega(\phi^i, q^i)}\right). \tag{2}$$

This problem allows us to understand individual decisions under alternative assumptions about the form of government intervention. I turn to this issue in the rest of this section; for that I will make use of alternative assumptions on $T(Y^i, q^i)$.

2.1 The Laissez-faire

In the absence of government intervention, that is when $T(Y^i, q^i) = 0$, the optimal choice of individuals will satisfy the following arbitrage conditions (they follow from the first-order conditions of problem (2))

$$\mathrm{MRS}^i_{cl} \equiv \frac{v'(l^i)}{u'(c^i)} = \omega(\phi^i, q^i), \tag{3}$$

$$\mathrm{MRS}^i_{cq} \equiv \frac{v'(l^i)}{u'(c^i)} \times \frac{l^i \omega_q(\phi^i, q^i)}{\omega(\phi^i, q^i)} = 1 \tag{4}$$

[3] The results in this paper also hold if ability and education are substitutes. Assumption (1) is used in this paper because it allows a sharper contrast with the usual results in the literature on education (Arrow, 1971; Hare and Ulph, 1979; De Fraja, 2002).

[4] The *education elasticity of wage* is also the *return to education*. I use the first terminology because using the second may cause some confusion. Indeed, in the formal models used in most empirical papers, the return to education appears to be related to the relation between the demand for education and ability (see, for example, Arias, Hallock, and Sosa, 1999). However, this results from assuming a logarithmic utility function and separability between costs and benefits of education as in most empirical papers. It is not clear why monetary benefits and costs of education should enter the utility function in a separable way. In this section, I argue that it is $\omega_{\phi q}$ which determines this relation and not η. Moreover, this paper shows the importance of η without the need to make restrictive assumptions on the utility function.

and

$$\mathrm{MRT}^i_{lq} \equiv \frac{l^i \omega_q(\phi^i, q^i)}{\omega(\phi^i, q^i)} = \frac{1}{\omega(\phi^i, q^i)}. \tag{5}$$

The first is the traditional condition on the marginal rate of substitution between labor supply and consumption. In the absence of government intervention, it will be equal to the wage rate. The second is a condition on the marginal rate of substitution between consumption and education. In the absence of government intervention, it must equal 1. The third is a condition on the marginal rate of transformation between education and labor in the individual production plan. In the absence of government intervention, it will be equal to the inverse of the wage rate.

Together with conditions (3)–(5), individual choices must satisfy the individual budget constraint, which is the same constraint as in problem (2) with $T(Y^i, q^i) = 0$. Consequently, consumption, labor supply, and education depend on the ability parameter. Standard calculations show that the complementarity of ability and education (1) is a *sufficient* condition for education and labor supply to be increasing in ability.

2.2 The first-best

All along the paper, I will suppose that the government is utilitarian. This means that its objective function will be given by the sum of the utility of the individuals in the economy. The problem of the government is

$$\begin{aligned} \max_{\{T^i, Y^i, q^i\}_i} \quad & \sum_i \pi^i \left[u(Y^i - T^i - q^i) - v\left(\frac{Y^i}{\omega(\phi^i, q^i)} \right) \right] \\ \text{s.t.} \quad & \sum_i \pi^i T^i \geq 0. \end{aligned} \tag{6}$$

where π^i represents the proportion of individuals with ability ϕ^i.

The first-best allocation will satisfy the same arbitrage conditions as the *laissez-faire* (Eqs. (3) and (5)). The difference with the *laissez-faire* is that consumption will be equalized among individuals of different types. However, the complementarity of ability and education (1) is still a *sufficient* condition for labor supply and education to be increasing in ability.

2.3 Implementation

The first-best can be implemented with a tax schedule, $T(Y^i, q^i)$, under the condition that $T_Y(Y^i, q^i) = T_q(Y^i, q^i) = 0$ for all i. Thus, only lump-sum transfers are needed to implement the first-best (this is precisely the Second Theorem of Welfare Economics).

Outside the first-best, under asymmetric information, marginal tax rates may differ from zero and will distort individual choices. With nonzero marginal tax rates, relations (3)–(5) do not hold. Generally, the optimal tax function $T(Y, q)$ is not differentiable, so marginal tax rates must be interpreted as the differences between the ratio of the right and left-hand sides of Eqs. (3) and (4) and 1, at the point chosen by the individual.[5] Abusing notation, the conditions that describe individual optimal choices can be written as follows:

$$\mathrm{MRS}^i_{cl} = \omega(\phi^i, q^i)[1 - T_Y(Y^i, q^i)] \tag{7}$$

and

$$\mathrm{MRS}^i_{cq} = [1 + T_q(Y^i, q^i)]. \tag{8}$$

The effect of the marginal tax on labor is standard. The main interest in this paper is the overall effect of the tax function $T(Y, q)$ on the labor-education margin. It is easy to see that both marginal tax rates will affect education choice. To make this more clear, combine Eqs. (7) and (8) to write

$$\mathrm{MRT}^i_{lq} = \frac{1}{\omega(\phi^i, q^i)} \frac{1 + T_q(Y^i, q^i)}{1 - T_Y(Y^i, q^i)}. \tag{9}$$

If both marginal tax rates are positive ($T_Y > 0$ and $T_q > 0$), education will be distorted downward with respect to labor supply (this is seen comparing Eqs. (5) and (9)). However, both marginal tax rates do not necessarily have the same sign. A negative marginal tax on education that offsets the effect on education of a positive marginal tax on labor income is possible.

3 Government intervention under asymmetric information

If the government does not observe individual types, it must rely on the observation of individual choice variables to distinguish among them. This makes the first-best allocation not implementable. Therefore, the government will have to rely on instruments different from lump-sum transfers to redistribute. The instruments that must be used by the government have been studied by Atkinson and Stiglitz (1976) and Cremer and Gahvari (1997).

The famous Atkinson-Stiglitz result says that if labor supply and demand for other goods are separable, a nonlinear income tax is enough to redistribute. However, there are strong reasons to believe that education and labor supply are not separable. In such a case, the optimal income tax should be supplemented with an education tax. Whether public provision of education is also desirable has been answered by Cremer and Gahvari (1997). They show that if nonlinear education taxes are available, there is no need for public provision of education. This means that the tax schedule $T(Y^i, q^i)$

[5] Nondifferentiability of the optimal tax function is a result of the requirement of incentive compatibility. A smooth tax schedule would make the mimicked individual change its behavior with respect to that obtained with the nondifferentiable schedule. Even if the tax function is not smooth, we can talk about marginal taxes relating them to the slope of the indifference curve in the point chosen by the individual. Thus, marginal tax rates in Eqs. (7) and (8) would be those faced by the individual under a smooth optimal tax schedule (if this could be made smooth) that does not change individual behavior.

introduced in Eq. (2) is enough to implement the second-best allocation if education is observed.[6]

In this section, I characterize $T(Y^i, q^i)$ for a two-class economy. Specifically, $\phi^i \in \{\phi^L, \phi^H\}$, with $\phi^L < \phi^H$. In line with the optimal taxation literature, the government observes labor income and education expenditure of each individual. Labor supply, wage, and ability are not observed by the government. To study marginal tax rates, I use the mechanism design approach as is commonly done in the recent literature on optimal taxation.

Let (R^i, Y^i, q^i) be the second-best allocation implemented by the planner, where R^i is disposable income ($R^i = Y^i - T^i$), Y^i is labor income, and q^i is the education level of an individual of type i. Since the planner does not observe directly the type of individual, it must design the mechanism (R^i, Y^i, q^i) to induce individuals to self-select themselves through the choice of the allocation designed for their type. This means that for $i, k \in \{L, H\}$, the planner will face two incentive constraints of the form

$$u(R^i - q^i) - v\left(\frac{Y^i}{\omega(\phi^i, q^i)}\right) \geq u(R^k - q^k) - v\left(\frac{Y^k}{\omega(\phi^i, q^k)}\right). \tag{10}$$

The problem of the planner is problem (6), with the additional incentive constraints (10) and $c^i = R^i - q^i$. Following the insight given by the first-best allocation, and the usual practice in optimal taxation problems, I will suppose that the binding incentive constraint will be the one preventing high-ability individuals from mimicking low-ability ones. Consequently, letting λ be the Lagrange multiplier of the resource constraint, and μ the multiplier of the relevant incentive constraint, the Lagrangian of this maximization problem is:

$$\begin{aligned}\Lambda = &\sum_{i=L,H} \pi^i \left[u(R^i - q^i) - v\left(\frac{Y^i}{\omega(\phi^i, q^i)}\right)\right] + \lambda \sum_{i=L,H} \pi^i [Y^i - R^i] \\ &+ \mu\left[u(R^H - q^H) - v\left(\frac{Y^H}{\omega(\phi^H, q^H)}\right) - u(R^L - q^L) + v\left(\frac{Y^L}{\omega(\phi^H, q^L)}\right)\right]. \end{aligned} \tag{11}$$

From the first-order conditions of Eq. (11) (provided in the Appendix), the marginal rates of substitution and transformation can be expressed as follows:

$$\mathrm{MRS}_{cl}^H = \omega(\phi^H, q^H), \tag{12}$$

$$\mathrm{MRT}_{lq}^H = \frac{1}{\omega(\phi^H, q^H)}, \tag{13}$$

[6] If education is not directly observable at the individual level, nonlinear taxes on education are not available and direct provision of education is desirable. The same allocation that can be implemented under observability of education is easily implementable under the lack of observability of education. It is enough to use a large tax on supplemental privately provided education and public provision at the same levels as those in the allocation under observability of education.

$$\mathrm{MRS}^L_{cl} = \omega(\phi^L, q^L)\frac{1 - \dfrac{\mu}{\pi^L}}{1 - \dfrac{\mu}{\pi^L}\dfrac{\omega(\phi^L, q^L)}{\omega(\phi^H, q^L)}\dfrac{v'(l^{HL})}{v'(l^L)}} \tag{14}$$

and

$$\mathrm{MRT}^L_{lq} = \frac{1}{\omega(\phi^L, q^L)}\frac{1 - \dfrac{\mu}{\pi^L}\dfrac{\omega(\phi^L, q^L)}{\omega(\phi^H, q^L)}\dfrac{v'(l^{HL})}{v'(l^L)}}{1 - \dfrac{\mu}{\pi^L}\dfrac{\eta(\phi^H, q^L)}{\eta(\phi^L, q^L)}\dfrac{\omega(\phi^L, q^L)}{\omega(\phi^H, q^L)}\dfrac{v'(l^{HL})}{v'(l^L)}}. \tag{15}$$

As argued in Section 2, efficiency requires the equalization of the marginal rate of substitution between labor and consumption to the wage rate, and the marginal rate of transformation between education and labor to the inverse of the wage rate. Equations (12)–(15) show to what extent this is the case under the stated asymmetric information conditions. In other words, they allow us to know the signs and sizes of $T_Y(Y^i, q^i)$ and $T_q(Y^i, q^i)$.

First, take the marginal income tax rates. From Eq. (12), it follows that the labor-consumption relation is set at the efficient level for high-ability individuals, $T_Y(Y^H, q^H) = 0$. From Eq. (14) and $l^{HL} < l^L$, it follows that low-ability individuals will face a positive marginal tax on labor income, $T_Y(Y^L, q^L) > 0$. This is the standard result in optimal income taxation in a two-class economy (Stiglitz, 1982).

Regarding the marginal tax on education, there are two relevant questions. First, whether it is positive or negative. Second, its relation to the marginal tax rate on labor income, and therefore, the direction of the overall distortion on education induced by both marginal tax rates. From Eq. (13), it must be clear that high-ability individuals face no distortion on education, $T_q(Y^H, q^H) = 0$.

For low-ability individuals, things are different. From inspection of Eq. (15), one sees that the crucial element is how $\eta(\phi^i, q^i)$ varies with ϕ^i. Specifically, it will depend on which of the following assumptions hold:

A1 $\eta(\phi^H, q) > \eta(\phi^L, q)$ for all q,
A2 $\eta(\phi^H, q) < \eta(\phi^L, q)$ for all q, and
A3 $\eta(\phi^H, q) = \eta(\phi^L, q)$ for all q.

Any of these assumptions can hold if ability and education are complements, as assumed in Eq. (1). A1 holds if the complementarity is strong, and A2 holds if the complementarity is weak.[7]

[7] The derivative of $\eta(\phi^i, q^i)$ with respect to ϕ is

$$\eta_\phi = q\frac{\omega_{\phi q} - \omega_\phi \omega_q}{\omega^2},$$

where one sees that if $\omega_{\phi q} > (<)\ \omega_\phi\omega_q$, i.e., $\omega_{\phi q}$ is big (small), $\eta_\phi > (<)\ 0$. Additionally, note that A3 holds if $\omega(\phi^i, q^i) = f(\phi^i)g(q^i)$. Note also that only A2 can hold if ability and education are substitutes, and Eq. (1) is not satisfied.

Suppose A1 holds. From Eq. (15), $\mathrm{MRT}_{lq}^{L} > 1 \big/ \omega(\phi^L, q^L)$. This means (from Eq. (9)) that

$$\frac{1 + T_q(Y^L, q^L)}{1 - T_Y(Y^L, q^L)} > 1.$$

Consequently, low-ability individuals will face a downward distortion on education with respect to labor supply.

Consider whether this is the result of a tax or subsidy on education. Since $T_Y(Y^L, q^L) > 0$, the only restriction this places on $T_q(Y^L, q^L)$ is that it must be greater than $-T_Y(Y^L, q^L)$. Thus, since $T_Y(Y^L, q^L) > 0$, either a subsidy or a tax on education is possible.

Assume now that A2 holds. In this case, $\mathrm{MRT}_{lq}^{L} < 1 \big/ \omega(\phi^L, q^L)$, and thus

$$\frac{1 + T_q(Y^L, q^L)}{1 - T_Y(Y^L, q^L)} < 1.$$

This means, first, that education of low-ability individuals is distorted upward with respect to labor supply and, second, that education is subsidized. Moreover, the subsidy must be bigger than the negative of the marginal tax rate on labor income faced by low-ability individuals: $(T_q(Y^L, q^L) < -T_Y(Y^L, q^L))$. The following proposition summarizes the main result in this paper.

Proposition 1. *Low-ability individuals face a positive marginal tax rate on labor income,* $T_Y(Y^L, q^L) > 0$. *The distortion on education of low-ability individuals depends on the relation between* $\eta(\phi, q)$ *and* ϕ. *If A2 holds, low-ability individuals face an upward distortion on education, and a marginal education subsidy,* $T_q(Y^L, q^L) > -T_Y(Y^L, q^L)$. *If A1 holds, low-ability individuals face a downward distortion on education, and a marginal education subsidy or tax,* $T_q(Y^L, q^L) < -T_Y(Y^L, q^L)$. *High-ability individuals face zero marginal tax rates on labor income and education,* $T_q(Y^H, q^H) = -T_Y(Y^H, q^H) = 0$.

This proposition says that the education gap of high-ability and low-ability individuals in second-best is smaller than in the first-best if A2 holds. It will be larger if A1 holds.

The conclusion that the only role for education subsidies is to restore the distortions induced on education choice by the tax on income is valid only when $\eta(\phi^i, q^i)$ does not change with ability. In all other cases, there is a different role for education subsidies or taxes.

This result is explained by the same argument used by Stiglitz (1982) to explain when are commodity taxes useful. From Eq. (5), it can be seen that the marginal rate of transformation between education and labor income is related to $\eta(\phi^i, q^i)$; it can be written as

$$\mathrm{MRT}_{qY} = \frac{Y^i}{q^i}\eta(\phi^i, q). \tag{16}$$

If $\eta(\phi^i, q^i)$ does not depend on ϕ^i, the indifference curves of different individuals in the (q^i, Y^i) plane will be parallel. Thus, individuals that differ in ability will react in parallel ways to $T(Y^i, q^i)$. This makes useless distorting the choice of education with respect to labor supply since it does not help separate individuals of different types. Conversely, when $\eta(\phi^i, q^i)$ depends on ϕ^i, individuals of different ability will react in different ways to the taxation of education. Consequently, using an education tax or subsidy that differs from $-T_Y(Y^i, q^i)$ becomes useful to separate individuals according to ability.

The direction of the distortion on education with respect to labor supply is also related to Eq. (16). If the indifference curve of low-ability individuals in the (q^i, Y^i) plane is less (more) steep than that of high-ability individuals, the education of low-ability individuals will be distorted downward (upward).

This result is also related to the discussion on production efficiency and redistribution. Diamond and Mirrlees (1971) show that, under a vast range of conditions, there is no scope for distortions in the production "side" of the economy. Seeing education and labor supply as two inputs in the household production problem, the result in this paper can be used to discuss the validity of the Diamond–Mirrlees result. In the full optimal taxation problem analyzed in this paper, the production efficiency result does not hold. The reason is that the Diamond–Mirrlees assumptions grant enough instruments to the government for a fully and independent control of consumers and producers. In the model in this paper, this is not possible. If the incentive constraint is binding and A3 does not hold, any attempt to control the labor-education decisions will have impact in the labor-consumption decision.[8]

The dependence of education policy on the relation between the education-elasticity of wage and ability suggests the importance of knowing at least the sign of this parameter. Recently Ashenfelter and Rouse (1998), Arias, Hallock, and Sosa (1999), Tobias (2003) and Girma and Kedir (2005) have estimated this parameter.[9] All papers reject the invariance of η_ϕ with ϕ (A3). Arias, Hallock, and Sosa (1999) show evidence supporting $\eta_\phi > 0$ (A1). Ashenfelter and Rouse (1998) and Girma and Kedir (2005) show evidence supporting $\eta_\phi < 0$ (A2), whereas Tobias (2003) is not conclusive. The evidence is still too weak to make a strong conclusion since Arias, Hallock, and Sosa (1999) and Ashenfelter and Rouse (1998) use the same database with opposite results, and Girma and Kedir (2005) use data for a different country than the other three studies.[10]

These results differ from those obtained in a model where labor supply is fixed or where income taxes are not allowed. Assuming fixed labor supply or not allowing income taxes, if ability and education are complements, it is desirable not to distort education choice of high-ability individuals, but low-ability individuals will face a

[8] This has been clarified by Naito (1999) and Blackorby and Brett (2004).

[9] I restrict this comment on the empirical literature to papers that control for ability biases. Estimates that do not control for ability biases yield erroneous results, as shown by Arias, Hallock, and Sosa (1999).

[10] Girma and Kedir (2005) use data for Ethiopia, while Arias, Hallock, and Sosa (1999), Ashenfelter and Rouse (1998) and Tobias (2003) use data for the U.S.

downward distortion on education.[11] The results discussed earlier show that, when labor supply is endogenous, and when income taxes are allowed, even if ability and education are complements, for a wide range of sizes of the complementarity the second-best education gap of high-ability and low-ability individuals should be smaller than the first-best one.

3.1 Other dimensions of heterogeneity

Differences in ability is not the only reason lying behind concerns with education policy. Concerns with education policy are also grounded on differences in parameters which are not directly related with the *ability to benefit from education* but with the *ability to pay for education*. In Maldonado (2005), using an extension of the model used in this paper, I explore the cases where individuals differ in exogenous wealth and when they differ in ability and exogenous wealth. In the case where exogenous wealth is the only source of heterogeneity, the main result is that there are no distortions on education levels in the second-best. As in Bovenberg and Jacobs (2005), education taxes are used only to correct the distortion on education resulting from income taxes. This was expected since in that case there are no differences between the education elasticity of wage among different individuals.

When individuals differ in both ability and exogenous wealth, there are additional technical issues that should be resolved. Screening problems when one assumes that individuals differ in more than one dimension are generally difficult to solve. With more than one dimension of heterogeneity, types cannot be ordered easily. Consequently, one cannot know a priori which are the binding incentive constraints, and the traditional no distortion at the top result does not hold (Armstrong and Rochet, 1999). Despite this difficulty, the result in Proposition 1 can be extended to the case where individuals differ in exogenous wealth and ability. Still, the crucial parameter that determines the optimal second-best education-gap between high-ability and low-ability individuals is $\eta(\phi, q)$. If A2 (A1) holds, the education-gap would be shortened (widened) in the second-best with respect to that in the first-best.

4 Concluding comments

In this paper, I have addressed the problem of the design of optimal education policies when income taxation is also designed optimally. To be consistent with the economic literature, this type of problem should be addressed in a setting in which cash transfers are also available and optimally designed, but the level of the transfers is constrained by the information the government has about individuals. I have used nonseparability of education and labor supply to justify the need for intervention in the education market when the government wants to redistribute.

I have shown that the distortion on the education level may not have the same direction as the distortion on labor supply. If the education elasticity of the wage

[11] Under any of these two conditions, the problem is a traditional adverse selection problem (see Laffont and Martimort, 2001, Chap. 2).

function is decreasing in ability, individuals with low ability will face an upward distortion on education together with a downward distortion on labor supply. This result highlights the importance of the estimations of the relation between ability and the education elasticity of the wage function (the return to education). This paper tends a bridge between the empirical and the theoretical literatures on education. As argued earlier, the empirical literature on education has made this type of estimation. However, no other paper in the theoretical literature on education has shown its importance in the formulation of education policy.

The possibility of having marginally progressive or regressive optimal education policies in this model contrasts with the usual finding in the literature. The main reason behind this is that in this paper labor supply is endogenous and the government imposes income taxes. Ignoring income taxation when thinking on education subsidies may yield extremely misleading recommendations. If one ignores the income tax, it may appear that taxing the education of low-ability individuals is optimal if ability and education are complements. However, when one introduces income taxation, this result will be reversed for a wide range of possible wage functions.

Appendix: First-order conditions for the second-best problem

The first-order conditions for the problem in Eq. (11) are given by:

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

$$[R^i]: \quad \pi^i[u'(R^i-q^i)-\lambda]+\sum_{k:k\neq i}[\mu^{ik}u'(R^i-q^i)-\mu^{ki}u'(R^i-q^i)]=0$$

and

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

Acknowledgments I am highly indebted to Helmuth Cremer and Guido Friebel for very useful comments on all the versions of the paper. The anonymous referees and the editor also provided very useful comments. When writing the paper, I had useful discussions with Jean-Jacques Laffont, Pierrre Pestieau, Juan Daniel Oviedo, Michel Le Breton, David Bardey, and Nohora Forero. I wrote this paper at GREMAQ, University of Toulouse and at CORE, University of Louvain. I am grateful to both these institutions for funding and hospitality. Financial support from Banco de la República, Colombia, is gratefully acknowledged.

References

Arias, O., Hallock, K. F., & Sosa, W. (1999). Individual heterogeneity in the returns to schooling: Instrumental variables quantile regression using twins data. *Empirical Economics*, *26*, 7–40.

Armstrong, M., & Rochet, J. C. (1999). Multi-dimensional screening: A user's guide. *European Economic Review*, *43*, 959–979.

Arrow, K. J. (1971). A utilitarian approach to the concept of equality in public expenditures. *Quarterly Journal of Economics*, *85*, 409–415.

Ashenfelter, O., & Rouse, C. (1998). Income, schooling, and ability: Evidence from a new sample of identical twins. *Quarterly Journal of Economics*, *113*, 253–284.

Atkinson, A. B., & Stiglitz, J. E. (1976). The design of tax structure: Direct versus indirect taxation. *Journal of Public Economics*, *6*, 55–75.

Blackorby, C., & Brett, C. (2004). Production efficiency and the direct-indirect tax mix. *Journal of Public Economic Theory*, *5*, 165–180.

Bovenberg, A. L., & Jacobs, B. (2005). Redistribution and education subsidies. *Journal of Public Economics*, *89*, 2005–2035.

Cremer, H., & Gahvari, F. (1997). In-kind transfers, self-selection and optimal tax policy. *European Economic Review*, *41*, 97–114.

Cremer, H., Lozachmeur, J. M., & Pestieau, P. (2004). Social security, retirement age and optimal income taxation. *Journal of Public Economics*, *88*, 2259–2281.

De Fraja, G. (2002). The design of optimal education policies. *Review of Economic Studies*, *69*, 437–466.

Diamond, P. A. (1980). Income taxation with fixed hours of work. *Journal of Public Economics*, *13*, 101–110.

Diamond, P. A., & Mirrlees, J. A. (1971). Optimal taxation and public production. *American Economic Review*, *61*, 8–27.

Diamond, P. A., & Mirrlees, J. A. (1978). A model of social insurance with variable retirement. *Journal of Public Economics*, *10*, 295–336.

Girma, S., & Kedir, A. (2005). Heterogenenity in returns to schooling: Econometric evidence from Ethiopia. *The Journal of Development Studies*, *141*, 1405–1416.

Hare, P. G., & Ulph, D. T. (1979). On education and distribution. *Journal of Political Economy*, *87*, S193–S212.

Krause, A. (2006). Redistributive taxation and public education. *Journal of Public Economic Theory*, *8*(5), 807–819.

Laffont, J. J., & Martimort, D. (2001). *The theory of incentives: The principal-agent model*. Princeton, NJ: Princeton University Press.

Maldonado, D. (2005). *Three essays in the public economics of education*. Ph.D. thesis, University of Toulouse, Toulouse, France.

Mardener, N., & Rochet, J. C. (1995). Is it legitimate to encourage work sharing? *Scandinavian Journal of Economics*, *97*, 621–633.

Mirrlees, J. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies*, *38*, 175–208.

Naito, H. (1999). Re-examination of uniform commodity taxes under a non-linear tax system and its implications for production efficiency. *Journal of Public Economics*, *71*, 165–188.

Stiglitz, J. E. (1982). Self-selection and Pareto efficient taxation. *Journal of Public Economics*, *17*, 213–240.

Tobias, J. L. (2003). Are returns to schooling concentrated among the most able? A semiparametric analysis of the ability–earnings relationships. *Oxford Bulletin of Economics and Statistics*, *65*, 1–29.

Toumala, M. (1986). Optimal income taxation and educational decisions. *Journal of Public Economics*, *30*, 183–198.

Ulph, D. T. (1977). On the optimal distribution of income and educational expenditure. *Journal of Public Economics*, *8*, 341–356.