CENTRO STUDI IN ECONOMIA E FINANZA

CENTRE FOR STUDIES IN ECONOMICS AND FINANCE

WORKING PAPER NO. 59

Optimal educational choice and redistribution when cultural background matters

Elena Del Rey and Marìa del Mar Racionero

May 2001

DIPARTIMENTO DI SCIENZE ECONOMICHE - UNIVERSITÀ DEGLI STUDI DI SALERNO
Via Ponte Don Melillo - 84084 FISCIANO (SA)
Tel. 089-96 3167/3168 - Fax 089-96 3169 – e-mail: csef@unisa.it

CSEF WORKING PAPER NO. 59

Optimal educational choice and redistribution when cultural background matters

Elena Del Rey [*] and Marìa del Mar Racionero[**]

Abstract

Higher education plays an important role in determining lifetime earnings. In turn, the decision to become educated depends to a large extent on family characteristics, such as wealth and cultural background. In this paper, we focus on the interaction between fiscal policies and educational choices when cultural background matters. We derive optimality conditions for a linear income tax and a lump-sum subsidy for education in a dynamic framework in which generations are linked by cultural background. The factors that determine their sign and magnitude include concerns for redistribution, efficiency, and the educational externality on future generations

Keywords: Optimal linear income tax, Subsidies, Higher education, Educational background

JEL classification: H21, H41

Acknowledgements: This research was mainly conducted during our stay at CORE (Center for Operations Research and Econometrics), Université catholique de Louvain, Belgium. Research of the authors was supported by Fundaciòn Ramòn Areces (Spain) and the European Comission through the TMR Program. We are grateful to Robin Boadway, Alejandro Esteller, Maurice Marchand, Pierre Pestieau, Vincent Vandenberghe, and two anonymous referees for their helpful comments. Errors remain ours.

[*] CSEF Università di Salerno.

[**] Corresponding author: School of Economics, Deakin University, VIC 3217, Australia; e-mail: racioner@deakin.edu.au.

Table of contents

1 Introduction

Despite continuing increases in the proportion of the population completing secondary and tertiary education, the educational chances of individuals remain heavily influenced by their family background. Indeed, the decision to undertake higher education seems to rely not only on individual ability to study but also on family characteristics, such as income and level of education.

Several papers have analyzed the effects of family wealth inequality on educational decisions. In Barham *et al.* (1995), for instance, financing for education is obtained from within the family and the unequal distribution of income results in unequal opportunities to acquire higher education. In their model parents are not altruistic but have a better knowledge of their children's abilities to succeed in education and, for this reason, are willing to lend them the funds they need to pursue their studies. Other authors assume that parents are altruistic (Borjas, 1992; Torvik, 1993). Parents leave bequests or invest in the education of their children because they derive a higher utility the higher the welfare of their children. The conclusion is similar to that in Barham *et al.* (1995): differences in income imply differences in opportunity.

However, wealth is not the only family characteristic affecting educational decisions. In fact, the most serious obstacles to attending university in many countries derive from cultural background of the potential student (i.e., the level of education of her family). According to OECD (2000), 'adults with highly educated parents are between two and six times as likely to have gained tertiary qualifications than those with poorly-educated parents. For younger generations this range has narrowed to between two and three times in most countries, but in countries where differences were smallest, they have become larger'.

In this paper, we focus on the effects of differences in cultural background on educational decisions. In doing so we abstract from monetary barriers to education, like liquidity constraints. This approach provides new insights on the dynamics of human capital investment. When the source of the difference in opportunity is wealth inequality, taxation can serve the double objective of equalizing income and opportunity, although at some efficiency cost. We break this ambivalence. In addition, by considering cultural background as an important determinant of children's educational attainment, we are able to identify an intergenerational externality that affects the design of optimal taxes.

We consider a dynamic framework. Individuals differ both in their ability to benefit from education and in the education of their parents (i.e., cultural background). These two variables affect the costs and, thus, the decision to undertake education. Productivity, measured by the wage an individual can earn, is the result of both formal and home education. As a consequence, not only do educated individuals earn higher wages but also, for any given level of attained education, children of educated parents are more productive. Bowles (1972) provides empirical evidence on this point.[1].

In order to understand the issues at stake, we characterize the first best solution. In this benchmark situation, the government can resort to non-costly transfers and achieve any level of desired redistribution. We can then abstract from redistributive issues and identify the levels of labor supply and education demand that maximize the overall utility. We find that more children of educated parents, for whom education is less costly, receive higher education.

We then compare the first best to the laissez-faire solution and show that, due to the fact that children of educated parents are more likely to attain higher education, the decision to study today has an externality effect on the education of future generations. Failing to account for this intergenerational externality leads to underinvestment in human capital at the laissez-faire.

The government can internalize this intergenerational externality by means of a subsidy for education. However, the recipients of these subsidies turn out to be the individuals with higher lifetime earnings (see Hamada (1974), Hare and Ulph (1979) and Creedy (1995), among others). In our framework, we show that the subsidy for education is a regressive policy.

We assume that the government values income distribution and, for this reason, uses an income tax system. For simplicity, we choose a linear income tax. However, redistribution is costly. It negatively affects the educational decisions made by individuals and, hence, the proportions of educated and uneducated parents in the following generation. It also alters labor supply and, therefore, present as well as future tax

[1] We thank an anonymous referee for mentioning this reference to us.

revenues.

Finally, we show that the effect of fiscal policy on the distribution of opportunity is ambiguous in general. However, the conditions under which the educational subsidy can reduce the inequality of opportunity are compatible with it being a regressive policy.

The aim of this paper is to identify the optimal combination of income taxes and subsidies for education that brings together all these considerations. By doing so, we are able to shed some light on the complex relation linking optimal human capital investment with the distribution of income and opportunity.

The paper is organized as follows. In section 2, we describe the model and we provide a characterization of the labor supply and educational choices of individuals. We also analyze the dynamics of the model and compute the first best solution. In section 3, we characterize the behavior of the government. First, we study the effects of fiscal policy on individual decisions. Then, we analyze the conditions for optimality of the labor income tax rate and the subsidy for education. We draw some conclusions in section 4.

2 The Model

2.1 Individual's Labor Supply and Educational Choice

In this section we analyze the behavior of individuals belonging to a given generation. Individuals differ both in their ability to benefit from education and in their family cultural background. Ability to benefit from education, denoted by a, is stochastically determined at birth. For simplicity, we consider that a is uniformly distributed between 0 and 1. An individual's cultural background is represented by the education of his parents, e_{-1}. We assume that the level of education chosen by an individual, e, can take one of two values: either 0, if the individual does not attend university, or 1, if she does. Throughout the paper, let the subindex account for the education level of the parents, e_{-1}, and the superindex for the education level of the individual, e.

Individuals live for one period. First, they decide whether or not to acquire higher education. Studying entails a financial cost that depends on their ability to benefit from education and on the education of their parents. We assume this cost to be $\gamma_{e_{-1}}C(a)$, where the parameter $\gamma_{e_{-1}}$ represents the effect of the parents' education on their children's educational costs. We posit $\gamma_0 > \gamma_1$ to reflect the fact that education is more costly for children of uneducated parents. $C(.)$ is a decreasing and convex function of ability (i.e., $C' < 0, C'' > 0$). The cost of education decreases with ability at a decreasing rate.

Productivities, and thus wages, depend on education and background. We follow Bowles (1972) in making our assumptions about the wage structure. Schooling has a positive effect on wages, so that educated individuals earn higher wages than uneducated individuals of the same cultural background ($w^1_{e_{-1}} > w^0_{e_{-1}}$). But also, due to the effect of home education or inherited human capital, children of educated parents earn higher wages for any given level of education attained ($w^e_1 > w^e_0$).

After the educational choice has been made and wages have been determined, individuals decide how much to work. Let $h(l)$ be the disutility of labor measured in units of consumption at each period, with l being the number of hours spent at work each period. We assume that the disutility of labor is increasing and convex (i.e., $h' > 0, h'' > 0$). The disutility increases with labor at an increasing rate.

Preferences for consumption and leisure are represented by a utility function that is linear in consumption. Individual lifetime utilities are thus given by:

$$\begin{aligned} u^0_{e_{-1}} &= y^0_{e_{-1}} - h(l^0_{e_{-1}}) \qquad && e_{-1} = 0,1 \\ u^1_{e_{-1}} &= y^1_{e_{-1}} - h(l^1_{e_{-1}}) - \gamma_{e_{-1}} C(a) \qquad && e_{-1} = 0,1 \end{aligned} \tag{1}$$

where y stands for labor income (i.e., $y^e_{e_{-1}} = w^e_{e_{-1}} l^{\,e}_{\,e_{-1}}$).

Conditional on e_{-1}, individuals choose the labor supply that maximizes their lifetime utility ($u^e_{e_{-1}}$). At the optimum, they supply the amount of labor that satisfies:

$$w^e_{e_{-1}} = h'\left(l^e_{e_{-1}}\right) \tag{2}$$

The decision to become educated or not is made by comparing lifetime utility with and without education. For each type e_{-1}, it is possible to determine a threshold value of ability above which individuals will acquire higher education. We denote by $\widehat{a}_{e_{-1}}$ the ability level for which individuals of type e_{-1} are indifferent between undertaking higher education or not. From (1):

$$\gamma_{e_{-1}} C(\widehat{a}_{e_{-1}}) = \left(y^1_{e_{-1}} - h\left(l^1_{e_{-1}}\right)\right) - \left(y^0_{e_{-1}} - h\left(l^0_{e_{-1}}\right)\right) \qquad e_{-1} = 0,1 \tag{3}$$

At the threshold ability level $\widehat{a}_{e_{-1}}$, the cost of education equals the gain in terms of net lifetime earnings (i.e., labor income net of the disutility of labor). Children of parents with education e_{-1} whose ability is larger than $\widehat{a}_{e_{-1}}$ (i.e., education costs lower than $\gamma_{e_{-1}} C(\widehat{a}_{e_{-1}})$) will invest in higher education. Individuals of ability $a_{e_{-1}} < \widehat{a}_{e_{-1}}$ will not.

We have assumed that purchasing education is more costly for children of uneducated parents (since $\gamma_0 > \gamma_1$). Unfortunately, this is not enough to drive unambiguous conclusions. Instead, we need that the net benefit of education be higher for children of educated parents. For this reason, we make henceforth the plausible assumption that the wage benefit of investing in education is at least as large for children of educated parents as it is for children of uneducated parents (i.e., $w_1^1 - w_1^0 \geq w_0^1 - w_0^0$). This, together with the other assumptions on the wage structure, yields:[2]

$$\left(y_1^1 - h\left(l_1^1\right)\right) - \left(y_1^0 - h\left(l_1^0\right)\right) \geq \left(y_0^1 - h\left(l_0^1\right)\right) - \left(y_0^0 - h\left(l_0^0\right)\right) \tag{4}$$

Equation (4) together with equation (3, implies that $\widehat{a}_0 > \widehat{a}_1$ (i.e., the threshold ability level is higher if parents are uneducated). Fewer children of uneducated parents attend university and those who do are in average more able than children of educated parents attending university. This result can be summarized in the following proposition.

Proposition 1 *If the wage benefit of investing in education is at least as large for children of educated parents as it is for children of uneducated ones, a higher proportion of the former will undertake higher education.*

2.2 Dynamics of the Model

At the end of the period each individual gives birth to another and dies. Population is thus constant. Given that a is uniformly distributed between 0 and 1, $\widehat{a}_0$ and $\widehat{a}_1$ denote the probabilities of remaining uneducated when having an unfavorable and favorable family background, respectively. Our assumptions allow us to conclude that, as reported by OECD (1998) and discussed in the introduction, children of educated parents are more likely to gain tertiary education than those of non-educated ones.

[2] This is inmediate from the fact that the function $wl - h(l)$ is increasing and convex in w.

The evolution over time of the proportions of educated and uneducated people in this economy can be described by a Markov chain with the following transition matrix:

$$P = \begin{pmatrix} \widehat{a}_0 & 1-\widehat{a}_0 \\ \widehat{a}_1 & 1-\widehat{a}_1 \end{pmatrix}$$

where $(1-\widehat{a}_0)$ is the probability of attending university for a child whose parents have not and $(1-\widehat{a}_1)$ is the probability of attending university when her parents have.

Let π_0 and π_1 denote, respectively, the proportion of uneducated and educated people in each generation. At the steady state:

$$(\pi_0, \pi_1) = (\pi_0, \pi_1)\ P \tag{5}$$

The proportion of educated and uneducated people in the economy replicates itself once the steady state has been reached. We can easily obtain the vector of limiting or steady state probabilities by substituting the matrix P into (5) and using the fact that $\pi_0 + \pi_1 = 1$:

$$\pi_0 = \frac{\widehat{a}_1}{1-\widehat{a}_0+\widehat{a}_1} \tag{6}$$

$$\pi_1 = \frac{1-\widehat{a}_0}{1-\widehat{a}_0+\widehat{a}_1} \tag{7}$$

As expected, there is a close relationship between the proportion of children with educated or uneducated parents that undertake higher education and the proportion of educated individuals in the steady state. If $\widehat{a}_1 = 0$ (i.e., all children of educated parents attain higher education), then $\pi_0 = 0$ and $\pi_1 = 1$ (i.e., all individuals are educated at the steady state). If $\widehat{a}_0 = 1$ (i.e., no child of uneducated parent attains higher education), then $\pi_1 = 0$ and $\pi_0 = 1$ (i.e., no individual is educated at the steady state). If costs associated to education are low enough for the highest ability individuals and high enough for the lowest ability individuals of any background, the steady state will be characterized by positive proportions of both educated and uneducated individuals. We make this assumption henceforth.

2.3 The First Best

In this section we assume that the government is able to achieve, through non-costly transfers, any level of desired redistribution. We can then, at this stage, abstract from redistributive issues and identify the levels of labor supply and education demand that maximize the overall utility.

For given levels of education in this economy, the government would choose the labor supply levels that maximize:

$$\pi_0 Eu_0 + \pi_1 Eu_1$$

where:

$$\begin{aligned} Eu_0 &= \widehat{a}_0\; u_0^0 + \int_{\widehat{a}_0}^{1} u_0^1\; da && (8) \\ Eu_1 &= \widehat{a}_1\; u_1^0 + \int_{\widehat{a}_1}^{1} u_1^1\; da && (9) \end{aligned}$$

stand for the expected utility of children of uneducated parents and educated parents, respectively. The condition for optimality regarding labor supply is:

$$w_{e_{-1}}^e = h'\left(l_{e_{-1}}^e\right) \tag{10}$$

The first best levels of labor supply coincide with the levels that individuals would choose in a decentralized economy without government intervention.

The government also determines the optimal threshold ability levels, $\widehat{a}_0$ and $\widehat{a}_1$, that maximize expected utility. The first best cut-off ability levels will be denoted by $\widehat{a}_0^{FB}$ and $\widehat{a}_1^{FB}$. After some rearrangements, the optimality conditions for interior $\widehat{a}_0$ and $\widehat{a}_1$ are, respectively:

$$\begin{aligned} \frac{Eu_0 - Eu_1}{1-\widehat{a}_0+\widehat{a}_1} + (y_0^0 - h\left(l_0^0\right) - (y_0^1 - h\left(l_0^1\right)) + \gamma_0 C(\widehat{a}_0)) &= 0 && (11) \\ \frac{Eu_0 - Eu_1}{1-\widehat{a}_0+\widehat{a}_1} + (y_1^0 - h\left(l_1^0\right) - (y_1^1 - h\left(l_1^1\right)) + \gamma_1 C(\widehat{a}_1)) &= 0 && (12) \end{aligned}$$

Since $(y_1^1 - h(l_1^1)) - (y_1^0 - h(l_1^0)) \geq (y_0^1 - h(l_0^1)) - (y_0^0 - h(l_0^0))$, $\gamma_0 > \gamma_1$ and $C(.)$ is decreasing, it follows that $\widehat{a}_0^{FB} > \widehat{a}_1^{FB}$.

Proposition 2 *At the first best a higher proportion of children of educated parents attain higher education.*

This result has an intuitive explanation. In this ideal situation, the government controls all the instruments required to accomplish any level of redistribution, including full redistribution. By minimizing the costs associated with the investment in education it manages to generate and redistribute, if desired, more income. This is the reason why fewer children of uneducated than of educated parents enroll in higher education: education is more costly for the former for a given ability level. Equality of opportunity may be defined, in the present context, as university attendance being independent of parental education (i.e., $\widehat{a}_0 = \widehat{a}_1$). The first best is then clearly characterized by the presence of unequal opportunity. However, the equalization of the threshold ability values is not an explicit goal in the objective function of the planner. Therefore, not surprisingly, it does not result.

On the other hand, it is worth noticing that the first best threshold ability values do not coincide with those resulting from the individual decisions in the absence of government intervention, or laissez-faire. Since, at the laissez-faire, the expected lifetime utility is larger for children of educated parents ($Eu_1 > Eu_0$), individuals who make their educational choice in the absence of government intervention end up consuming too little education. In other words, individuals underinvest in education. When cultural background does not matter, the first best threshold ability level and the laissez-faire one coincide. The underinvestment in education, when cultural background matters, is due to the fact that individuals fail to take into account the effect of their educational choice on their children's well-being.

Proposition 3 *Both children of educated and uneducated parents under-invest in education at the laissez-faire.*

At the laissez-faire, individuals maximize their individual utility, neglecting thus the effect of their decisions on future generation's expected income. Due to the individuals' failure to account for this intergenerational externality, there is a role for public policy.

3 The Government Problem

In order to internalize the externality, the government may use a lump-sum subsidy for education. In turn, to redistribute income, the government can use an income tax. For simplicity, we assume the tax schedule to be linear.[3] The linear income tax system is

[3] It can be argued that a linear tax system does not make much sense when we have discrete values for labour income. It would then be equivalent to differentiated lump-sum taxes. However, the assumptions

characterized by a marginal tax rate τ and a demogrant G. While the income tax applies to all individuals, the lump-sum subsidy S is targeted to individuals who undertake higher education. Both labor supply and educational choices of the individual will be affected by these policy variables.

Once taxes and subsidies are introduced, individual lifetime utilities become:[4]

$$\begin{array}{ll} v^0_{e_{-1}} = (1-\tau)\, y^0_{e_{-1}} - h(l^0_{e_{-1}}) + G & e_{-1} = 0, 1 \\ v^1_{e_{-1}} = (1-\tau)\, y^1_{e_{-1}} - h(l^1_{e_{-1}}) - \gamma_{e_{-1}} C(a) + G + S & e_{-1} = 0, 1; \quad a \in (\hat{a}_{e_{-1}}, 1) \end{array} \tag{13}$$

The condition that determines the optimal levels of labor supply is:

$$(1-\tau)\, w^e_{e_{-1}} = h'(l^e_{e_{-1}}) \qquad \forall w, l. \tag{14}$$

whereas the one that determines the optimal educational choice, given cultural background e_{-1}, is:

$$\gamma_{e_{-1}} C(\widehat{a}_{e_{-1}}) = (1-\tau)(y^1_{e_{-1}} - y^0_{e_{-1}}) - \left(h\left(l^1_{e_{-1}}\right) - h\left(l^0_{e_{-1}}\right)\right) + S \tag{15}$$

It is worth analyzing the effects of taxes and subsidies on these optimal choices.

3.1 Comparative Statics

The only policy tool that affects the labor supply decision is the tax rate, since $l = l\left((1-\tau)\, w\right)$. The quasi-linearity assumption imposed on utility implies that labor supply is not subject to income effects and, thus, is not affected by S or G. Given the characteristics of the disutility of labor function, $h\left(.\right)$, it can be easily shown that an increase in the tax rate induces a decrease in labor supply.

The educational choice depends on the tax rate and the lump-sum subsidy for education. By checking the effect of these parameters on the threshold ability values, we are able to determine the induced effect on the number of students of each group, and on the composition of the population at the steady state. Differentiation of (15) yields:

$$\frac{d\widehat{a}_{e_{-1}}}{d\tau} = \frac{y^0_{e_{-1}} - y^1_{e_{-1}}}{\gamma_{e_{-1}} C'(\widehat{a}_{e_{-1}})} > 0 \tag{16}$$

$$\frac{d\widehat{a}_{e_{-1}}}{dS} = \frac{1}{\gamma_{e_{-1}} C'(\widehat{a}_{e_{-1}})} < 0 \tag{17}$$

about the wage structure rule out this possibility.

[4] We employ v once taxes and subsidies are introduced, to differenciate it from u, used before.

A marginal increase in the tax rate involves a disincentive to undertake education. Individuals who become educated will pay higher taxes. On the contrary, the lump-sum subsidy for education provides incentives to study. There exists a relationship between the effects of taxes and subsidies on the threshold ability levels:

$$\frac{d\widehat{a}_{e_{-1}}}{d\tau} = (y^0_{e_{-1}} - y^1_{e_{-1}})\frac{d\widehat{a}_{e_{-1}}}{dS}$$

The term of proportionality in the previous expression, $(y^1_{e_{-1}} - y^0_{e_{-1}})$ (i.e., the labor income difference due to education), determines the quantity by which the subsidy should be increased, when the tax rises, in order to keep $\widehat{a}_{e_{-1}}$ constant. It is worth noticing that, since the differences in earnings induced by education differ across backgrounds, the rule of proportionality differs by types e_{-1}. Therefore, an increase in the subsidy in response to an increase in the tax that seeks to leave $\widehat{a}_0$ unchanged will not be enough to leave $\widehat{a}_1$ unaffected. $\widehat{a}_1$ will increase in response to the tax (i.e., fewer children of educated parents become educated). In order to compensate individuals who gain more from education for an increase in the tax rate we would need to give them higher subsidies.

More generally, it would be interesting to determine how public policy affects the gap between $\widehat{a}_0$ and $\widehat{a}_1$. If we had $\widehat{a}_0 = \widehat{a}_1$, individuals would have the same probability of becoming educated independently of their cultural background. The greater the difference $(\widehat{a}_0 - \widehat{a}_1)$, the greater the inequality of opportunity. It can be shown that:

$$\frac{d\widehat{a}_0}{dS} - \frac{d\widehat{a}_1}{dS} \begin{array}{c} > 0 \\ < 0 \end{array} \Leftrightarrow \frac{\frac{C'(\widehat{a}_1)}{C(\widehat{a}_1)}}{\frac{C'(\widehat{a}_0)}{C(\widehat{a}_0)}} \begin{array}{c} < \\ > \end{array} \frac{(1-\tau)(y^1_0 - y^0_0) - (h(l^1_0) - h(l^0_0)) + S}{(1-\tau)(y^1_1 - y^0_1) - (h(l^1_1) - h(l^0_1)) + S}$$

Since the difference in labor income with and without higher education, net of the disutility of labor, is higher for children of educated parents, the right hand side of the second inequality is smaller than one. Its left hand side is larger than one if $C'(.)/C(.)$ is larger at $\widehat{a}_1$ than it is at $\widehat{a}_0$. In this case, the gap is reduced by the marginal increase in the subsidy for education.

Proposition 4 *If $C'(.)/C(.)$ is non-increasing in ability, a marginal increase in the subsidy for education (financed by a marginal reduction in G) reduces the inequality of opportunity.*

The educational choice depends on the educational costs, which are affected both by the education of the parents and the ability to benefit from education. If $C'(.)/C(.)$ is larger at $\widehat{a}_1$, children of educated parents react less to the subsidy than children of uneducated parents. Consequently, the threshold ability levels come closer to each other (i.e., opportunities equalize).

Even under this last assumption, the effect of the tax rate on equality of opportunity is ambiguous. The relationship between the effects of taxes and subsidies on the gap is given by:

$$\frac{d\widehat{a}_0}{d\tau}-\frac{d\widehat{a}_1}{d\tau}=(y_0^0-y_0^1)\left(\frac{d\widehat{a}_0}{dS}-\frac{d\widehat{a}_1}{dS}\right)+((y_1^1-y_1^0)-(y_0^1-y_0^0))\frac{d\widehat{a}_1}{dS}$$

When the gap is reduced by the subsidy increase, it tends to be increased by the tax increase. However, if the difference in the benefit from education between educated children of educated and uneducated parents is high enough, it may be that the combined increases in taxes and subsidies tend to reduce the gap (i.e., reduce the inequality of opportunity). This is due to the fact that subsidy increases compensate children of uneducated parents more than those of uneducated ones for the rise in the tax. Hence, children of educated parents are relatively more distorted by the tax.

3.2 Optimal Taxes and Subsidies

The objective of the government is to maximize some additive and anonymous Bergson-Samuelson social welfare function. We assume that the government applies a concave transformation $U(.)$ to individual utilities v. The higher the government's aversion to inequality, the more concave $U(.)$ will be. $U(v_{e-1}^e)$ is hence the social valuation of this individual's utility. The government's problem is:

$$\begin{array}{ll}\max\limits_{\tau,G,S} & \pi_0\left(\widehat{a}_0U(v_0^0)+\int_{\widehat{a}_0}^1 U(v_0^1)da\right)+\pi_1\left(\widehat{a}_1U(v_1^0)+\int_{\widehat{a}_1}^1 U(v_1^1)da\right) \\ \text{s.t.} & \tau\left(\pi_0\widehat{a}_0y_0^0+\pi_1\widehat{a}_1y_1^0+\pi_0(1-\widehat{a}_0)y_0^1+\pi_1(1-\widehat{a}_1)y_1^1\right)=G+\pi_1S\end{array} \tag{18}$$

The expression within brackets in the budget constraint represents the tax base. We will denote it by B. In addition, we denote the expected social lifetime utility of a child with uneducated and educated parents, respectively, by:

$$EU(v_0) = \widehat{a}_0U(v_0^0)+\int_{\widehat{a}_0}^1 U(v_0^1)da \tag{19}$$

$$EU(v_1) = \widehat{a}_1U(v_1^0)+\int_{\widehat{a}_1}^1 U(v_1^1)da \tag{20}$$

The resulting Lagrangian is:

$$L = \pi_0 EU(v_0) + \pi_1 EU(v_1) + \lambda(\tau B - G - \pi_1 S)$$

where λ is the Lagrange multiplier associated with the government's budget constraint.

3.2.1 Optimal income tax rate

Following the tradition of the optimal linear taxation literature, we manipulate the first order conditions of problem (18) with respect to G and τ (see Appendix for details) and obtain the equation "à la Sheshinsky":

$$\tau^* = \frac{\dfrac{\dfrac{\partial \pi_0}{\partial \tau}(EU(v_1) - EU(v_0))}{EU'} + cov\left(\dfrac{EU'(v^e_{e_{-1}})}{EU'}, y^e_{e_{-1}}\right) + S\dfrac{\partial \pi_1}{\partial \tau}}{\dfrac{\partial B}{\partial \tau}} \tag{21}$$

This equation provides an implicit solution for the optimal tax rate as a function of the subsidy. It does not give an explicit value for the tax rate, as τ affects as well the right hand side of the expression. However, it allows to isolate the main factors that underlie the sign and magnitude of the tax rate.

The denominator of (21) represents the efficiency term. It accounts for the disincentive effects of income taxation on both labor supply and education. Both effects are negative: when the tax rate increases, the tax base diminishes both due to the decrease in labor supply and to the induced changes in the composition of the population. A higher tax rate implies a decrease in the proportion of educated individuals in the population. Since these individuals earn higher wages, the total effect on the tax revenue is negative.

In the numerator of (21), some terms play in favor of a larger marginal tax rate and others against. The first term measures the difference in expected utility at the steady state as a result of a marginal increase in the tax rate. It thus represents the effect of the tax on the externality. The larger this externality effect, the smaller the marginal tax rate which is compatible with a given efficiency loss.

The covariance term is generally associated with the government's willingness to redistribute income from richer to poorer individuals. If the marginal social valuation of an individual's utility decreases in her labor income, the sign of the covariance is negative. Given that the social welfare function is concave, social marginal utility is decreasing in individual utility. Therefore, if labor incomes and final utilities are equally ordered, the covariance is negative. Due to the presence of the subsidy for education, we cannot be sure that this is the case. In order to shed more light on the sign of the covariance we decompose it into the sum of the inter-group and the expected intra-group covariances:[5]

$$\begin{aligned} cov\left(\frac{EU'(v^e_{e_{-1}})}{EU'}, y^e_{e_{-1}}\right) &= \pi_0\ cov\left(\frac{EU'(v^e_0)}{EU'}, y^e_0\right) + \pi_1\ cov\left(\frac{EU'(v^e_1)}{EU'}, y^e_1\right) \\ &\quad + cov\left(\frac{EU'(v_{e_{-1}})}{EU'}, E(y_{e_{-1}})\right) \end{aligned} \tag{22}$$

Within each cultural background group (children of educated and uneducated parents, respectively), those who study enjoy higher labor incomes and utilities. Since marginal social utility is decreasing in net income, both intra-group covariances are negative. On the other hand, children of educated parents earn higher labor incomes than children of uneducated parents, whether they study or not. Also, the expected lifetime utility is larger for children of educated parents ($EU\,(v_1) > EU\,(v_0)$). The inter-group covariance is also negative. As a consequence, the sign of the covariance term is negative. This, together with the negative sign of the denominator, implies that the marginal tax rate is larger the higher the concern for redistribution.

Finally, the last term in the numerator reflects the fact that the tax induced changes on the proportions of educated individuals affect the subsidy bill. If the tax rate increases, the number of educated individuals, π_1, falls and subsidy payments fall accordingly.

3.2.2 Optimal subsidy for education

By manipulating the first order conditions of problem (18) with respect to G and S, we can also express the optimal level of the subsidy as a function of the tax rate (see

[5] In the inter-group covariance, as we have done before, we drop the superscript to indicate expectation over educational choice.

Appendix):

$$S^* = \frac{\dfrac{\dfrac{\partial \pi_0}{\partial S}(EU(v_0) - EU(v_1))}{EU'} + cov\left(\dfrac{EU'(v^e_{e-1})}{EU'}, \dfrac{\partial v^e_{e-1}}{\partial S}\right) + \tau\dfrac{\partial B}{\partial S}}{\dfrac{\partial \pi_1}{\partial S}} \tag{23}$$

Due to the specification of the utility function in this model, labor supply is not subject to income effects and is hence unresponsive to changes in the subsidy. For this reason, the denominator of equation (23) reflects only the effect of the subsidy on the composition of the population, which is positive: the number of educated individuals increases as a result of a subsidy increase.

The factors that played in favor of a larger tax rate play now against a large lump-sum subsidy for education. In particular, the concern for redistribution implies lower subsidies, the subsidy being a regressive measure. Similarly, the internalization of the externality and the efficiency term, that would induce a lower tax rate, result now in a larger subsidy for education, since individuals will be more willing to both work and undertake education.

Since redistribution, efficiency and the educational externality are simultaneously considered by the government, the optimal policy will be given by the mix of taxes and subsidies that best conciliates these conflicting objectives.

To conclude, recall that we measure the aversion to inequality by the degree of concavity of U. It is illustrative to compare the following two extreme cases. If U is linear, the government maximizes expected utility. This case corresponds to the utilitarian objective: the government is not concerned with redistribution. As a result, the covariance terms disappear from the expressions for optimal taxes and subsidies. Only externality and efficiency considerations matter. If, on the contrary, the degree of concavity is maximal, the government maximizes the utility of the worst-off individual. This case corresponds to the Rawlsian objective. In our framework, the worst-off person is an uneducated individual with uneducated parents. Since the government only cares for the utility of this type of individuals, the effects of the tax and subsidy on the externality on future generations do not affect the design of the optimal policy.

This exercise underlines the fact that, when cultural background matters as suggested by this model, the desirability of subsidizing education relies on efficiency considerations. From a distributive point of view such policy remains regressive. As we have seen, the effect of fiscal policy on the distribution of opportunity is independent from the objective of the government. Thus, the conditions under which subsidies equalize opportunity are compatible with their being regressive.

4 Concluding Remarks

We have analyzed optimal fiscal policies when the lifetime income of an individual depends on her education and, in turn, her decision to become educated is determined by both her ability to benefit from education and her cultural background (i.e., the level of education of her family). To focus on the role of cultural background, we have abstracted from the effect of family wealth on the educational choice, already treated in a number of papers.

We have considered a dynamic framework in which individuals from two different cultural backgrounds decide whether or not to undertake higher education. We have assumed that individuals face different net educational costs according to their innate ability and to the educational level of their parents. Under these assumptions, we have shown that the individual decision to undertake higher education has a positive externality effect on future generations. Failure to account for this intergenerational externality leads to underinvestment in education.

A subsidy for education can serve the purpose of internalizing this externality. However, subsidizing education can be viewed as a regressive policy, since for a given type of parent, the expected lifetime income of an educated individual is higher than that of an uneducated one. On the other hand, due to the fact that the educational choice does not depend on family wealth, fiscal policies aiming at equalizing income will not have an equalizing effect on human capital investment. Furthermore, income taxation has been shown to have disincentive effects on both human capital investment and labor supply. Optimal educational choices and redistribution thus turn out to be opposing objectives.

The optimal policy is then a mix of both instruments, the one that best conciliates the conflicting objectives of the government: redistribution and efficiency, including the educational externality on future generations. Equalizing opportunity is not an explicit objective of the government. For this reason, the optimal policy is independent from and has ambiguous effects on the distribution of opportunity.

Some of our assumptions deserve qualifications. In particular, the discrete nature of the wage structure can be a drawback of our model. Wages are affected directly by education and cultural background but only indirectly by ability, through the cost to acquire education. Although we believe our wage configuration to rather well conciliate the always conflicting objectives of simplicity and realism, we expect to be able to consider a more general framework (e.g., a continuum of wages depending on individual ability) in future research. Similarly, the conclusions we have derived for a linear tax system and lump-sum subsidy, chosen for their tractability, deserve to be challenged by the study of alternative taxing and subsidy schemes.

References

[1] **Barham, V., Boadway, R., Marchand, M. and Pestieau, P.** (1995). 'Education and the poverty trap', *European Economic Review*, **39**, 1257-75.

[2] **Borjas, G. J.** (1992). 'Ethnic capital and intergenerational mobility', *Quaterly Journal of Economics,* **107**, 123-50.

[3] **Bowles, S.** (1972). 'Schooling and Inequality from Generation to Generation', *Journal of Political Economy,* **80**, s219-51.

[4] **Creedy, J.** (1995). *The economics of higher education: An analysis of taxes versus fees*, Edward Elgar Publishing Limited, England.

[5] **Hamada, K.** (1974). 'Income taxation and educational subsidy', *Journal of Public Economics,* **3**, 145-58.

[6] **Hare, P. and Ulph, D.T.** (1979). 'On education and distribution', *Journal of Political Economy,* **87**, S193-212.

[7] **OECD** (2000). *Education at a glance : OECD indicators 2000,* Centre for Educational Research and Innovation, OECD, Paris.

[8] **Torvik, R.** (1993). 'Talent, growth and income distribution', *Scandinavian Journal of Economics,* **95**, 581-96.

1 Derivation of equations *à la Sheshinsky.*

The first-order conditions with respect to the three decision variables, G, τ and S are, respectively:

$$\pi_0\widehat{a}_0U'(v_0^0)+\pi_1\widehat{a}_1U'(v_1^0)+\pi_0\left(1-\widehat{a}_0\right)EU'(v_0^1)+\pi_1\left(1-\widehat{a}_1\right)EU'(v_1^1)\equiv EU'=\lambda \quad (24)$$

$$\frac{\partial\pi_0}{\partial\tau}(EU(v_0)-EU(v_1))+\pi_0\frac{\partial EU(v_0)}{\partial\tau}+\pi_1\frac{\partial EU(v_1)}{\partial\tau}=-\lambda(B+\tau\frac{\partial B}{\partial\tau}-S\frac{\partial\pi_1}{\partial\tau}) \quad (25)$$

$$\frac{\partial\pi_0}{\partial S}(EU(v_0)-EU(v_1))+\pi_0\frac{\partial EU(v_0)}{\partial S}+\pi_1\frac{\partial EU(v_1)}{\partial S}=-\lambda(\tau\frac{\partial B}{\partial S}-S\frac{\partial\pi_1}{\partial S}-\pi_1) \quad (26)$$

Since

$$\frac{\partial EU(v_0)}{\partial\tau}=\frac{\partial\left(\widehat{a}_0U(v_0^0)+\int_{\widehat{a}_0}^1 EU(v_0^1)\,da\right)}{\partial\tau}=\widehat{a}_0U'(v_0^0)\frac{\partial v_0^0}{\partial\tau}+\int_{\widehat{a}_0}^1 EU'(v_0^1)\,da\,\frac{\partial v_0^1}{\partial\tau}$$

$$\frac{\partial EU(v_1)}{\partial\tau}=\frac{\partial\left(\widehat{a}_1U(v_1^0)+\int_{\widehat{a}_1}^1 EU(v_1^1)\,da\right)}{\partial\tau}=\widehat{a}_1U'(v_0^1)\frac{\partial v_0^1}{\partial\tau}+\int_{\widehat{a}_1}^1 EU'(v_1^1)\,da\,\frac{\partial v_1^1}{\partial\tau}$$

and

$$\frac{\partial v_{e_{-1}}^e}{\partial\tau}=-y_{e_{-1}}^e$$

we write (25) as

$$\begin{aligned}
&\frac{\partial\pi_0}{\partial\tau}(EU(v_0)-EU(v_1))-\pi_0(\widehat{a}_0U'(v_0^0)\,y_0^0+\int_{\widehat{a}_0}^1 EU'(v_0^1)\,da\,y_0^1)\\
&-\pi_1(\widehat{a}_1U'(v_0^1)y_1^0+\int_{\widehat{a}_1}^1 EU'(v_1^1)\,da\,y_1^1) \qquad (27)\\
=\;&-\lambda(\pi_0\widehat{a}_0y_0^0+\pi_1\widehat{a}_1y_1^0+\pi_0(1-\widehat{a}_0)y_0^1+\pi_1(1-\widehat{a}_1)y_1^1+\tau\frac{\partial B}{\partial\tau}-S\frac{\partial\pi_1}{\partial\tau})
\end{aligned}$$

where the tax base has been substituted by its full expression in the right hand side.

We know from (24) that $\lambda=EU'$. Then (27) becomes

$$\begin{aligned}
&\frac{\partial\pi_0}{\partial\tau}(EU(v_0)-EU(v_1))-\pi_0(\widehat{a}_0U'(v_0^0)y_0^0+\int_{\widehat{a}_0}^1 EU'(v_0^1)\,da\,y_0^1)\\
&-\pi_1(\widehat{a}_1U'(v_0^1)y_1^0+\int_{\widehat{a}_1}^1 EU'(v_1^1)\,da\,y_1^1) \qquad (28)\\
=\;&-EU'(\pi_0\widehat{a}_0y_0^0+\pi_1\widehat{a}_1y_1^0+\pi_0(1-\widehat{a}_0)y_0^1+\pi_1(1-\widehat{a}_1)y_1^1+\tau\frac{\partial B}{\partial\tau}-S\frac{\partial\pi_1}{\partial\tau})
\end{aligned}$$

Having divided everything by EU' we can group the following terms in (28) as

$$
\begin{aligned}
&-\frac{1}{EU'}(\pi_0(\widehat{a}_0 U'(v_0^0)\, y_0^0 + \int_{\widehat{a}_0}^1 EU'(v_0^1)\, da\, y_0^1) + \pi_1(\widehat{a}_1 U'(v_1^0) y_1^0 + \int_{\widehat{a}_1}^1 EU'(v_1^1)\, da\, y_1^1)) \\
&+(\pi_0 \widehat{a}_0 y_0^0 + \pi_1 \widehat{a}_1 y_1^0 + \pi_0(1-\widehat{a}_0) y_0^1 + \pi_1(1-\widehat{a}_1) y_1^1) \;=\; -\,cov\left(\frac{EU(v_{e_{-1}}^e)}{EU'}, y_{e_{-1}}^e\right)
\end{aligned}
$$

Thus (25) can be written

$$
\frac{\frac{\partial \pi_0}{\partial \tau}(EU(v_0) - EU(v_1))}{EU'} - cov\left(\frac{EU(v_{e_{-1}}^e)}{EU'}, y_{e_{-1}}^e\right) = -\,\tau \frac{\partial B}{\partial \tau} + S\frac{\partial \pi_1}{\partial \tau}
$$

from where we derive the equation à la Sheshinsky for the optimal tax rate:

$$
\tau^* = \frac{\dfrac{\frac{\partial \pi_0}{\partial \tau}(EU(v_1) - EU(v_0))}{EU'} + cov\left(\dfrac{EU(v_{e_{-1}}^e)}{EU'}, y_{e_{-1}}^e\right) + S\dfrac{\partial \pi_1}{\partial \tau}}{\dfrac{\partial B}{\partial \tau}}
$$

Take now (26). Since

$$
\begin{aligned}
\frac{\partial v_0^0}{\partial S} &= \frac{\partial v_1^0}{\partial \tau} = 0 \\
\frac{\partial v_0^1}{\partial \tau} &= \frac{\partial v_1^1}{\partial \tau} = 1
\end{aligned}
$$

and hence

$$
\begin{aligned}
\frac{\partial EU(v_0)}{\partial S} &= \frac{\partial\left(\widehat{a}_0 U(v_0^0) + \int_{\widehat{a}_0}^1 EU(v_0^1)\ da\right)}{\partial S} = \int_{\widehat{a}_0}^1 EU'(v_0^1)\ da \\
\frac{\partial EU(v_1)}{\partial S} &= \frac{\partial\left(\widehat{a}_1 U(v_1^0) + \int_{\widehat{a}_1}^1 EU(v_1^1)\, da\right)}{\partial S} = \int_{\widehat{a}_1}^1 EU(v_1^1)\, da
\end{aligned}
$$

(26) becomes

$$
\begin{aligned}
&\frac{\partial \pi_0}{\partial S}(EU(v_0) - EU(v_1)) + \pi_0 \int_{\widehat{a}_0}^1 EU'(v_0^1)\ da + \pi_1 \int_{\widehat{a}_1}^1 EU(v_1^1)\, da \\
=\;\; & -\lambda\left(\tau\frac{\partial B}{\partial S} - S\frac{\partial \pi_1}{\partial S} - \pi_1\right)
\end{aligned}
$$

that we can write, using $\lambda = EU'$ from (24), as

$$
\begin{aligned}
&\frac{\frac{\partial \pi_0}{\partial S}(EU(v_0) - EU(v_1))}{EU'} + \frac{\pi_0 \int_{\widehat{a}_0}^1 EU'(v_0^1)\ da + \pi_1 \int_{\widehat{a}_1}^1 EU(v_1^1)\, da}{EU'} \\
=\;\; & -\tau\frac{\partial B}{\partial S} + S\frac{\partial \pi_1}{\partial S} + \pi_1
\end{aligned}
$$

and then

$$S^* = \frac{\dfrac{\dfrac{\partial \pi_{\mathbf{0}}}{\partial S}(EU(v_{\mathbf{0}}) - EU(v_1))}{EU'} + cov\left(\dfrac{EU'(v^e_{e-1})}{EU'}, \dfrac{\partial v^e_{e-1}}{\partial S}\right) + \tau \dfrac{\partial B}{\partial S}}{\dfrac{\partial \pi_1}{\partial S}}$$