



390

2010

Redistributive Taxation, Incentives, and the Intertemporal  
Evolution of Human Capital

Christian Ferrada; Matías Tapia.

PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE  
INSTITUTO DE ECONOMIA

---

Oficina de Publicaciones  
Casilla 76, Correo 17, Santiago  
www.economia.puc.cl

**REDISTRIBUTIVE TAXATION, INCENTIVES, AND THE INTERTEMPORAL  
EVOLUTION OF HUMAN CAPITAL**

**Christian Ferrada  
Matías Tapia\***

**Documento de Trabajo N° 390**

Santiago, Diciembre 2010

## INDEX

### ABSTRACT

1. INTRODUCTION	2
2. THE MODEL	5
2.1 Households	5
2.2 Government	7
2.3 Schools	8
3. SOLVING THE MODEL	9
3.1 Equilibrium in the Education Market	9
3.2 Optimal Choices at the Household Level	13
3.3 Redistributive Taxation and Efficiency of Education Expenditures	16
4. NUMERICAL SIMULATION	17
4.1 Parameter Calibration	18
4.2 Results	20
4.2.1 School Investments	20
4.2.2 Steady-State Human Capital and Income	21
4.2.3 Welfare	24
5. CONCLUSIONS	26
6. REFERENCES	27

# Redistributive Taxation, Incentives, and the Intertemporal Evolution of Human Capital\*

Christian Ferrada

The University of Chicago

Matías Tapia

Instituto de Economía, Universidad Católica de Chile and EH-Clio Lab

December 27, 2010

## Abstract

This paper contributes to the literature on redistributive taxation and human capital dynamics by explicitly analyzing the role of incentives in the education market where human capital is produced. We introduce an explicit education market with heterogeneous private schools in a dynamic stochastic general equilibrium model with overlapping generations and human capital accumulation. We use the model to simulate the effects of taxation on growth, intergenerational mobility, inequality, and welfare. Equalization in education expenditures reduces incentives for differentiation in the education market, with the distribution of education investments shifting towards the least productive schools. This has significant consequences on equilibrium outcomes, and highlights the importance of incorporating the role of intermediation when analyzing redistribution policies.

Keywords: Human capital, school market, redistributive taxation, inequality, efficiency

JEL codes: E24, H21, I21

---

\*We would like to thank comments and suggestions by Felipe Balmaceda, Gary Becker, Francisco Gallego, Oscar Landerretche, Casey Mulligan, José Tessada, and Bernardita Vial, as well as seminar participants at Universidad Católica de Chile, Centro de Economía Aplicada (Universidad de Chile), and Departamento de Economía (Universidad de Chile). Matías Tapia thanks the Esther and T.W. Schultz Endowment and Programa de Inserción Académica 2010 (Vicerrectoría Académica).

# 1 Introduction

Analyzing the role of education funding on growth, inequality, and educational mobility has been an active and important area of research over many years. This paper studies the dynamic effects of different funding schemes in an intergenerational model in which parents optimally choose how much to invest in the education of their children. While the previous literature has not paid close attention to the characteristics of the underlying education market, we incorporate this market explicitly. In particular, we analyze how different redistribution schemes change incentives in a competitive market with heterogeneous producers. In that sense, this paper combines the literature on competition in explicit school markets under different funding regimes (Epple and Romano (1998, 2008), Nechyba (2000), MacMillan (2004), Urquiola (2005), Ferreyra (2007), Urquiola and Veerhogen (2008), Vial (2008), McLeod and Urquiola (2009), Tapia (2010)) with the literature on the dynamic, macroeconomic consequences of different schemes of education (Becker and Tomes (1979), Loury (1981), Glomm and Ravikummar (1992), Benabou (2002), Cunha and Heckman (2007)). This latter type of literature does not explicitly account for the characteristics of the school market where education funding is intermediated.

To do so, we present a DSGE model in the spirit of Benabou (2002), and introduce a competitive education market with heterogeneous producers, in the spirit of Tapia (2010). By doing so, we explicitly consider the precise mechanism under which human capital is accumulated and how the characteristics of the underlying education market affect these outcomes. Changes in the funding of education change the incentives faced by schools, which in turn affect the matching allocations between students and schools, and also the investment decisions of schools. In the model, education outcomes are a joint product of student ability and the school's endogenous quality, which depends on the school's productivity and its investment decision. That way, more productive schools need to invest less to reach a given quality level. The paper studies how different funding regimes (associated to different levels of redistributive taxation) affect the schools' investment and enrollment decisions, which

shape intertemporal aggregate dynamics relative to a world where an explicit market is not considered.

In the context of the model, there are two main margins that determine the overall efficiency of the human capital accumulation process, conditional on an aggregate level of education funding. The first one, extensively discussed in Becker and Tomes (1979), Benabou (2002), and Cunha and Heckman (2007), deals with the distribution of parental investments across households, in particular when capital markets are not complete and families are credit-constrained. The second margin, which is the main contribution of the paper, is related to the equilibrium behavior of schools as incentives change under different funding schemes. This determines school investments and then the associated school-student matches and equilibrium outcomes. There two related aspects that we want to address. First, how much of what schools receive from parents will be actually invested in providing higher education outcomes? Second, how are investments distributed across schools with different productivities? Do the most productive schools make the largest investments?

As mentioned before, most of the papers that have addressed the issue of the institutional setup of education and growth have focused in the production of human capital. Glomm and Ravikumar (1992) present an endogenous growth model in the spirit of Lucas (1988), and discuss the accumulation of human capital and its implications for income growth and distribution. They compare two ways of financing education: private and public. Agents differ in the human capital of their parents, which in turn determines the productivity of educational investments in their human capital. Under this setup, public education (which provides the same educational investment for all students) yields lower per capita income than private education. However, this result is reversed when initial income inequality is sufficiently high. Aguiar-Conraria (2005) shows that in the same model as Glomm and Ravikumar (1992) if parents do not have an altruistic behavior, growth is enhanced under a public education system. Watanabe and Yasuoka (2009) extend also Glomm and Ravikumar (1992)'s model by using a constant relative risk aversion utility function, obtaining that even

under a constant returns to scale production function for human capital, income inequality vanishes in the long run. Zhang (1996) analyzes public subsidies given to private education in a similar context as Glomm and Ravikumar (1992), but he also considers a positive externality of average human capital on the agents' human capital production function. As expected, the subsidy speeds up growth and improves welfare. In a similar study, Kaganovich and Zilcha (1999) analyze the role of education vouchers and growth in a representative-agent model. Cardak (2005) analyzes vouchers in a model where the productivity of human capital accumulation depends on parental human capital. He shows that vouchers can increase economic growth but also under some cases income inequality.

Education and growth has been studied also with an emphasis on the role of heterogeneity in ability, and its implications for the outcomes of private investments in human capital. Han and Mulligan (2001) model private human capital investments with imperfect capital markets when students' abilities and parental altruism can vary. They obtain implications for intergenerational mobility. De Gregorio and Kim (2000) use a similar setup to analyze how the introduction of capital markets allows able students to specialize in education during their youth, while less able students are better-off by working. Chiu (1998) analyzes the role of income inequality in attaining optimal allocations in a world of discrete education investments. De Fraja (2002) and Bohacek and Kapicka (2008) analyze optimal education policies in a world where individual abilities are private information.

The paper provides an analytical setup in which to discuss human capital accumulation, income inequality, and intergenerational mobility in an economy where education is provided by an explicit market of competitive private schools. We show that accounting for the characteristics of the education markets can have significant effects, which are relevant for evaluating the role of redistributive policies. While our model uses a stylized description of schools, and the incentives they face, our results can be interpreted more generally as highlighting the importance of incentives in intermediation. In particular, how the evaluation of redistribution policies, which many times cannot operate directly in terms of the final

target variable (in this case, education outcomes) but only indirectly through an intermediate input (education expenditure), should explicitly analyze the market where resources are intermediated, and the role of incentives within it.

The remainder of the paper is organized as follows. Section 2 introduces the basic model, which combines the models in Benabou (2002) and Tapia (2010). We describe the behavior of households, the government which levies redistributive taxation, and heterogeneous private schools. Section 3 solves for the equilibrium in the school market at every period, and develops the policy function that describes the optimal behavior of households. Section 4 solves the model numerically for levels of redistribution, comparing economies. Finally, Section 5 concludes and gives some discussion about potential extensions.

## 2 The Model

### 2.1 Households

The general setup follows directly from Benabou (2002), which in turn is closely related to previous literature on intergenerational human capital dynamics as Becker and Tomes (1979) and Loury (1981).

We assume the economy is populated by a discrete and sufficiently large number of agents indexed by  $i \in \{0, \dots, N\}$ . These agents are characterized at date  $t$  by the following recursive utility function

$$\ln U_t^i = \max_{c_t^i, l_t^i} \left\{ (1 - \rho) [\ln c_t^i - (l_t^i)^\eta] + \rho \ln \left( (E_t [(U_{t+1}^i)^r])^{1/r} \right) \right\} \quad (1)$$

Agents choose consumption,  $c_t^i$ , and the amount of labor they supply,  $l_t^i$ .  $\rho$  is the intertemporal discount factor,  $\eta$  defines the intertemporal elasticity of substitution as  $1/(1 - \eta)$ , and  $r$  defines the relative risk aversion to lotteries as  $1 - r$ . This formulation is consistent with an overlapping generations model where each agent cares both about her instantaneous utility



and the utility of her only child. Crucially, we assume that there are no financial markets or physical capital, so agents cannot borrow and are only able to save in the form of human capital. There is a unique good, which is used both for consumption and for the production of human capital.

Income is given by

$$y_t^i = (h_t^i)^\lambda (l_t^i)^\mu + \pi_t^i = (h_t^i)^\lambda (l_t^i)^\mu (1 + \pi_t^i) \quad (2)$$

Total income is the sum of two components, labor income and the share of the profits generated in the education market. Labor income depends on  $h_t^i$ , the agent's human capital (which is given when she is an adult) and her choice of labor supply,  $l_t^i$ . As discussed in more detail below, private schools will generate profits in equilibrium; we assume that ownership shares are distributed evenly across households, so that each household receives the same profit flow,  $\pi_t$ .

$$y_t^i = (h_t^i)^\lambda (l_t^i)^\mu (1 + \pi_t^i) = c_t^i (1 + \theta_t) + e_t^i \quad (3)$$

Total income is spent in consumption and in education expenditures,  $e_t^i$

$$y_t^i = c_t^i + e_t^i \quad (4)$$

In the absence of a capital market (or any form of physical or financial capital), agents can only invest and allocate resources across time through human capital. The technology for human capital accumulation depends directly on the child's exogenous and random ability,  $\xi_{t+1}^i$ , the human capital of the parent,  $h_t^i$ , and an education input,  $q_t^i$ . The education input will be bought in an explicit education market served by heterogeneous producers, which is described in Section 2.3. We assume the human capital equation can be expressed as

$$h_{t+1}^i = \kappa \xi_{t+1}^i (h_t^i)^\alpha (q_t^i)^\varepsilon \quad (5)$$

For simplicity, we assume that the ability of the child, an i.i.d. random variable, is not observed by the parent when deciding how much to invest on her education. However, parents know the stochastic process governing ability across generations.

## 2.2 Government

Each period, there is a marginally progressive tax,  $\tau_t$ , on education expenditures,  $e_t^i$ , such that the household's actual education expenditure,  $\hat{e}_t^i$ , is described by

$$\hat{e}_t^i = (\tilde{y}_t/y_t^i)^{\tau_t} e_t^i = (\tilde{y}_t)^{\tau_t} (y_t^i)^{1-\tau_t} s_t^i \quad (6)$$

where  $s_t^i = \frac{e_t^i}{y_t^i}$  is the share of income spent on education and  $\tilde{y}_t$  satisfies the government's budget constraint in period  $t$ :

$$\sum_{i=1}^N e_t^i = \sum_{i=1}^N \hat{e}_t^i \quad (7)$$

$$\sum_{i=1}^N (y_t^i) s_t^i = \sum_{i=1}^N (\tilde{y}_t)^{\tau_t} (y_t^i)^{1-\tau_t} s_t^i \quad (8)$$

As  $\tau_t$  increases, heterogeneity in actual (after tax) education expenditures is reduced. When  $\tau_t = 1$ , there is perfect equalization, and all households spend the same in the education of their child. All households whose income exceeds  $\tilde{y}_t$  pay net positive taxes on education, while households with income below  $\tilde{y}_t$  receive a subsidy. Taxes (subsidies) are strictly increasing (decreasing) in income. We later show that the share of income spent in education is independent of income,  $s_t^i = s_t$ . Using that property, one can easily show that  $\tilde{y}_t$  is a decreasing function of the tax rate and that converges to the economy's mean income when  $\tau_t = 1$ .

## 2.3 Schools

The education market follows closely the model presented in Tapia (2010). Human capital is not produced directly by households, but requires the use of an education input,  $q$ , provided by competitive firms, which we can interpret as schools. From now on, we label this education input as school quality.

All schools have the same production technology and the capacity to serve only one consumer (from now on, a student). Capacity is fixed and cannot be expanded. Schools are run independently, operated as profit-maximizing firms, and have no outside options. There is free entry and no fixed operation costs. Schools differ in their exogenous productivity,  $\gamma_j$ , which is drawn from a common distribution at the beginning of each period. The productivity parameter can be seen as a proxy for the skills of the manager. Each manager runs only one school.<sup>1</sup> All schools disappear exogenously after each period<sup>2</sup> and are replaced by a new set of schools drawn from the same distribution. Thus, all schools solve a purely static problem. In each period, the number of potential schools is  $M$ , with  $M > N$ , so that not all potential schools can operate in equilibrium.

For any given school  $j$ , its observable quality,  $q_j$ , depends on the school's productivity,  $\gamma_j$ , and on the investment it decides to make,  $\theta_j$ . Investments are measured in units of the final good.<sup>3</sup> The production technology for school quality is

$$q_j = g(\gamma_j, \theta_j) \tag{9}$$

where  $g$  is twice-continuously differentiable, with  $g_\theta, g_\gamma > 0$ ,  $g_{\theta\theta}, g_{\gamma\gamma} \leq 0$ , and  $g_{\gamma\theta} \geq 0$ . Also assume that  $g(\gamma_j, 0) = 0$ .

For sake of concreteness, assume that the production of school quality can be described

---

<sup>1</sup>It is very simple to extend the model to one in which managers can open more than one school.

<sup>2</sup>Given that each period in the model is a complete generation, this does not seem to be an unreasonable assumption.

<sup>3</sup>In a more general context, this would include teachers, salaries, etc.

by the following Cobb-Douglas function:

$$q_j = \gamma_j \theta_j \tag{10}$$

Schools perfectly observe the characteristics of each household, as well as the productivity of all other schools. In the equilibrium allocations proposed below, schools will make simultaneous offers to the students they want to enroll.

### 3 Solving the model

#### 3.1 Equilibrium in the Education Market

Equilibrium in the education market can be defined as a sequence of school investments, education qualities, and student-school allocations that solve the static problem in each period.

In this section, we conjecture that optimal education expenditures per households are a constant share,  $s$ , of after tax income, and follow Tapia (2010) to derive equilibrium conditions in the education market under that assumption. In section 3.4, we use those equilibrium conditions to solve the households' intertemporal problem, and verify that the policy function satisfies our initial assumption.

To obtain the equilibrium allocation in any given period  $t$ , rank schools from high to low productivity,  $j$  from 1 to  $M$ , with school  $j = 1$  having the largest productivity ( $\gamma_1$ ). Do the same with students in terms of after-tax income,  $i$  from 1 to  $N$ , with student  $i = 1$  coming from the highest income household ( $y_1$ ).

Let  $\mu : N \rightarrow M$  be a one-to-one matching function. This is, for each student  $i$ ,  $\mu(i)$  corresponds to his associated school, and  $\mu(i) = \mu(k)$  is only true if  $i = k$ . As  $M > N$ ,  $\mu^{-1}(j)$  either corresponds to a student (for the  $N$  schools inside the market) or the empty set (for the  $M - N$  schools that must be inside the market).

Under our conjecture, total education expenditure for any given household does not depend on price (v.g., the demand for education inputs has unitary price elasticity). Thus, all schools will take education expenditures from each household (“tuition”) as given, and will make simultaneous education quality offers to those students they wish to enroll. An offer can be defined as the education input,  $q_{ij}$ , promised by school  $j$  to student  $i$ . School investments are perfectly observable and done simultaneously to tuition payments. As mentioned earlier, school productivities and student abilities are perfectly observable. Define a given set of offers from all schools as an offer profile.

Equilibrium is characterized by an offer profile  $q$  from schools to students and a matching  $\mu$ , with the following properties:

**Definition 1.** *An equilibrium is an offer profile  $q$  from schools to students and a matching  $\mu$  such that there does not exist any student  $i$ , school  $j$ , and an offer  $q'_{ij}$  where:*

$$(i) \ q'_{ij} > q_{i\mu(i)} \text{ and}$$

$$(ii) \ \pi(\gamma_j, q'_{ij}) > \pi(\gamma_j, q_{\mu^{-1}(j)j})$$

$$\text{For the cases in which } \mu^{-1}(j) = \emptyset, \pi(\gamma_j, q_{\mu^{-1}(j)j}) = 0.$$

This is, in any equilibrium allocation there is no student-school pair that will jointly benefit from breaking their equilibrium match and matching to each other.

**Proposition 1.** *The stable market equilibrium can be characterized as follows:*

(a) *Only the  $N$  more productive schools will operate.*

(b) *There is strict assortative matching between more productive schools and higher income households. This is,  $\mu(i) = i$ .*

(c) *For any student  $i$  and matching school  $\mu(i) = i$ , the equilibrium offer will be  $q_{ii}^*$  for all  $i \in \{1..N\}$ . In particular, equilibrium school qualities can be written as*

$$q_{ii}^* = s \left( q_{NN}^* + \sum_{k=1}^{N-i} (y_{N-k} - y_{N+1-k}) \gamma_{N+1-k} \right)$$

for  $i = 1$  to  $N - 1$ , and  $q_{NN}^* = (\kappa m_N) \gamma_{N+1}$ . More productive schools provide strictly higher quality.

(d) School investments are

$$\theta_j^* = \frac{q_{jj}^*}{\gamma_j}$$

for  $j = 1$  to  $N$ .

(e) In equilibrium, more productive schools and more able students are strictly better off.

*Proof.* See Appendix. □

A formal proof is presented in the Appendix. However, we can sketch a (simple) economic argument here. The quality received in equilibrium by any given student  $i$  cannot fall below the maximum quality that could be provided by the marginal school outside the market, which has productivity  $\gamma_{N+1}$ . This is, the quality bought by the student whose willingness to pay is  $sm_i$  must at least exceed the highest feasible offer provided by the fringe school,  $q_{iN+1}^{\max} = sm_i \gamma_{N+1}$ . At  $q_{iN+1}^{\max}$ , the fringe school is actually investing all of its revenue and thus getting zero profits.

It is easy to show, however, that competition between schools must drive implicit prices per unit of quality below the marginal cost of the fringe school, and, except for the lowest income student, equilibrium offers will exceed the minimum level required to prevent entry. More productive schools have an advantage in producing better outcomes at a lower cost and end up serving the students with higher willingness to pay and having a higher actual quality.

For any school  $j < N + 1$ ,  $q_{jj}^*$  is the minimum quality that provides no incentives for its most direct competitor from below,  $j + 1$ , to make a better offer to attract student  $j$ . In equilibrium, school  $j + 1$  makes exactly the same profits on its own equilibrium match,  $j + 1$ , than what it would get by matching  $q_{jj}^*$  with student  $j$ . For school  $j + 1$ , the marginal cost of offering  $q_j^*$  instead of  $q_{j+1}^*$  equals the marginal revenue of enrolling student  $j$  instead of student  $j + 1$ . Thus, the marginal profit of deviating is zero. All other schools make strictly

higher profits in their own equilibrium matches.

Thus, in equilibrium, differences in quality between schools are more than proportional to the (after tax) income differences of the students they enroll. High-income students get higher qualities not only because they can expend more, but because the equilibrium allocation implies that they are attached to high-productivity schools, where they face implicitly lower prices. This is a strong force towards inequality in outcomes, as income differences are magnified by the differences in the quality of the match.

The key insight, however, lies in the way investments are distributed with after-tax income distribution changes. For any given school  $j$ , its equilibrium investment can be written as

$$\begin{aligned}\theta_j^* &= \frac{q_{jj}^*}{\gamma_j} = \frac{s \left( q_{NN}^* + \sum_{k=1}^{N-j} (y_{N-k} - y_{N+1-k}) \gamma_{N+1-k} \right)}{\gamma_j} \\ &= s \left( \frac{y_N \gamma_{N+1}}{\gamma_j} + \sum_{k=1}^{N-j} (y_{N-k} - y_{N+1-k}) \frac{\gamma_{N+1-k}}{\gamma_j} \right)\end{aligned}\quad (11)$$

Thus, in equilibrium, the difference in investments between any two operating schools with different productivities,  $\gamma_j > \gamma_l$ , can be written as:

$$\theta_j^* - \theta_l^* = \frac{s(\gamma_l - \gamma_j)}{\gamma_j \gamma_l} \left( y_N \gamma_{N+1} + \sum_{k=1}^{N-s} (y_{N-k} - y_{N+1-k}) \gamma_{N+1-k} \right) + \sum_{k=N-s}^{N-j} (y_{N-k} - y_{N+1-k}) \frac{\gamma_{N+1-k}}{\gamma_j}\quad (12)$$

The first term is strictly negative and reflects the fact that the quality provided in equilibrium by the school with lower productivity can always be provided at a smaller cost by the more productive school. The second term is the additional investment that the productive school needs to make in order to attain her own equilibrium quality. Thus, whether a more productive school invests more or less is not obvious, and depends both on the distribution of school productivities and the distribution of tuition payments. It is clear, however, that,

for a given distribution of school productivities, larger differences in tuition payments across households will lead to larger differences in equilibrium qualities and, typically, to larger investments by more productive schools. If, on the contrary, tuition payments are more homogeneous, incentives to differentiation are reduced and equilibrium qualities are more similar. As suggested by equation 12, investments in more productive schools will become relatively smaller.

In the extreme, if all households make the same education expenditure, ex-post qualities will be identical and equal to the minimum quality that prevents entry. As there are no incentives to differentiate, more productive schools will make strictly smaller investments, as they need to make a smaller effort to reach the homogeneous equilibrium quality. Thus, schools in the lower end of the productivity distribution will make the bulk of education investments in the economy. Any distribution of education expenditures that has the same mean but a positive variance will necessarily shift equilibrium investments from less productive to more productive schools and unambiguously enhance efficiency in the production of quality.

This is the key insight of the paper and it will shape the results presented in Section 4. In this model, the introduction of an explicit education market implies that redistribution policies affect not only the distribution of tuition payments, but also the distribution of investments across schools of different types. In this context, an evaluation of redistribution policies must not only consider the allocation of resources across different households, but also the effects of such an allocation on the market where those resources are intermediated.

### 3.2 Optimal Choices at the Household Level

Given the discussion on the optimal provision of quality and the government's taxation policy, we now turn our attention to the optimal behavior of households.

The household's problem is to solve the utility maximization problem given by (1) subject to the constraints given by (2), (5), (6), and (10).



**Proposition 2.** *The optimal choices of labor supply and the savings rate are*

$$l_t = \left[ \frac{\mu}{\eta} \left( 1 + \frac{\rho\varepsilon(1-\tau_t)V_{t+1}}{1-\rho} \right) \right]^{\frac{1}{\eta}}$$

$$s_t = \frac{\rho\varepsilon V_{t+1}}{1-\rho + \rho\varepsilon V_{t+1}}$$

Note that the optimal choices depend on taxes and on promised future utilities. However, they do not depend on the human capital of the parent nor on the ability of the child, so  $l_t^i = l_t$  and  $s_t^i = s_t$ . The labor choice decreases with current and expected future taxes, while the savings rate decreases with expected future tax rates. Under a constant profile of taxes  $\tau_t = \tau$  the optimal choices are equivalent to

$$l = \left[ \frac{\mu}{\eta} \left( 1 + \frac{\rho\varepsilon\lambda(1-\tau)}{1-\rho(\alpha + \varepsilon\lambda(1-\tau))} \right) \right]^{\frac{1}{\eta}} \quad (13)$$

$$s = \frac{\rho\varepsilon\lambda}{1-\rho(\alpha + \varepsilon\lambda(1-\tau)) + \rho\varepsilon\lambda} \quad (14)$$

These choices are constant, as conjectured when solving for the education market equilibrium.

Thus, effective parental education spending is a constant share of total income (scaled by  $1 - \tau_t$ ) and it is given by

$$\hat{e}_t^i = \frac{(\tilde{y}_t)^{\tau_t} (y_t^i)^{1-\tau_t} \rho\varepsilon V_{t+1}}{1-\rho + \rho\varepsilon V_{t+1}} \quad (15)$$

Note that this expression is common across households, with  $V_{t+1}$  defining the discounted sum of future utilities and equal to

$$V_{t+1} = (1-\rho)\lambda \sum_{k=0}^{\infty} \rho^k \prod_{j=0}^{k-1} \alpha + \varepsilon\lambda(1-\tau_{t+1+j}) \quad (16)$$

We will use these results to show that under a competitive educational market, the Bellman equation for the household problem is well defined and the education spending of

the parents is effectively a constant share of their income scaled by  $1 - \tau$  as expressed in (15). To that end, we need to make some assumptions about the evolution of the distributions of some key variables. Start by imposing that abilities are drawn every generation from a lognormal distribution, so  $\ln \xi_t^i \sim \mathcal{N}(-\omega^2/2, \omega^2)$ , with  $\omega > 0$ . Assume also that the human capital in the first period is drawn from a lognormal distribution that is independent of the distribution of abilities,  $\ln h_1^i \sim \mathcal{N}(-\omega^2/2, \omega^2)$ . The lognormality of human capital for the next periods is

$$\begin{aligned} \ln h_{t+1}^i &= \ln \kappa + \varepsilon \ln(s_t) + (1 + \varepsilon) \ln \xi_{t+1}^i + (\alpha + \varepsilon \lambda (1 - \tau_t)) \ln h_t^i \\ &\quad + \varepsilon \mu (1 - \tau_t) \ln l_t^i + \varepsilon \ln \delta_t - \varepsilon \ln \hat{\pi}_t + \varepsilon (1 - \tau_t) \ln(1 + \pi_t) + \varepsilon \tau_t \ln \tilde{y}_t \end{aligned} \quad (17)$$

which comes from combining and (2), (5), (6) and (10), and where we defined  $\hat{\pi} = \hat{e}/g = 1 + \pi/g$ . Note that  $g$  comes from the solution of the optimal education supply of schools. We will follow a model where the demand and the supply are determined separately, so possible positive human capital externalities are not internalized.

To show that human capital follows every generation a lognormal distribution, we use the fact that the sum of lognormal and truncated lognormal distributions is also lognormal. The formal proof is omitted, but numerical simulations verify that this is indeed the case. Under the assumptions above and the independence and lognormality of productivities,  $\ln \delta_t^i \sim \mathcal{N}(-\omega_1^2/2, \omega_1^2)$ , with  $\omega$  potentially different from  $\omega_1$ . We can denote the evolution of human capital in time as in  $\ln h_{t+1}^i \sim \mathcal{N}(m_{h_{t+1}}, \Delta_{h_{t+1}}^2)$ .

Combining (1) and (2), the recursive utility function is

$$\ln U_t(h) = \max_{l,s} \left\{ \begin{aligned} &(1 - \rho) [\ln((1 - s)) + \ln(h^\lambda l^\mu (1 + \pi_t)) - l^\eta] \\ &+ \rho \ln \left( (E_t [(U_{t+1}^i(h'))^r])^{1/r} \right) \end{aligned} \right\} \quad (18)$$

subject to

$$h' = \kappa((1 + a_t)s)^\varepsilon (\xi)^{1+\varepsilon} (h)^{\alpha+\varepsilon\lambda(1-\tau_t)} (l)^{\varepsilon\mu(1-\tau_t)} (\delta_t)^\varepsilon (\hat{\pi}_t)^{-\varepsilon} (1 + \pi_t)^{\varepsilon(1-\tau_t)} (\tilde{y}_t)^{\varepsilon\tau_t} \quad (19)$$

Defining  $p(\tau_t) \equiv \alpha + \varepsilon\lambda(1 - \tau_t)$  as the intergenerational persistence of human capital, we can show that

**Proposition 3.** *The value function can be expressed as  $\ln U_t^i = V_t(\ln h_t^i - m_{h_t}) + W_t$ , where  $V_t \equiv (1 - \rho)\lambda \sum_{k=0}^{\infty} \rho^k \prod_{j=0}^{k-1} p(\tau_{t+j})$ ,  $m_{h_t} \equiv \sum_{i=0}^1 \ln h_t^i$ , and the aggregate welfare  $W_t$  is a function of  $\{\tau_{t+k}, \theta_{t+k}, a_{t+k}\}_{k=0}^{\infty}$ .*

**Proposition 4.** *As schools heterogeneity disappears, our model converges exactly to the model in Benabou (2002).*

The households' policy function, together with the equilibrium condition in the education market and the government budget's constraint, fully describe the equilibrium in this economy at any given point in time. In the next section, we characterize the steady-state equilibrium numerically.

Before that, we discuss the potential rationale for redistributive taxation on this context, and the expected channels through which it can affect the intertemporal human capital accumulation.

### 3.3 Redistributive Taxation and Efficiency of Education Expenditures

Redistributive taxation affects the intertemporal evolution of human capital through three channels.

First, it distorts labor and saving decisions, as households reduce their labor effort and invest less out-of-pocket funds on education. Numerically, given the calibration described

below, this channel only has a small effect, as the elasticity of both labor and savings to the tax rate is relatively small.

The second channel, and the one usually emphasized on the previous literature, deals with the absence of a financial market that equalizes the marginal return of investment across households. As redistribution in this model does not change the income distribution within a given generation, but directly changes the distribution of education expenditures, overall efficiency would increase if resources shift towards those households that have the highest rate of marginal return. Given the decreasing marginal return of education inputs in the human capital production function and the fact that ability is uncorrelated with income, this is typically true, as resources move towards those households that are initially investing less. The effect, however, is weakened through the direct effect of parental human capital on human capital production, which, *ceteris paribus*, makes the first cent invested on the average child of a high income (high human capital) household more productive than the average child of a low income (low human capital) household. We show that, in the calibration, the first effect prevails over a significant region, allowing for an important efficiency role for redistributive taxation.

The third channel, associated to the introduction of the education market, deals with the effect of redistribution -or, more correctly, equalization- on school incentives. As discussed in Section 3.1, equalization dampens incentives if schools are heterogeneous, and shifts educational investments towards the least productive schools. For a given level of expenditure, this reduces overall efficiency, and goes against the efficiency-enhancing effect of the second channel.

## 4 Numerical Simulations

As discussed in the previous section, the model cannot be solved analytically once heterogeneous schools are introduced. Thus, we characterize equilibrium allocations by simulating

the model numerically, both in terms of dynamics and the steady state. As our calibration exercise is better suited to provide a qualitative rather than a quantitative description of equilibrium, we use as a benchmark for our results a version of the model with no heterogeneity in schools (this is, a particular case of Benabou (2002)). This allows us to compare the direct impact of incorporating an education market with heterogeneous providers to the baseline overlapping generations' model. We also use the simulations to verify, as assumed earlier in the policy function, that the simulated distribution of human capital over time in fact follows a lognormal distribution.

## 4.1 Parameter Calibration

We start with parameter values close to the ones chosen by Benabou (2002). These values are summarized in the following table:

Table 1. Parameters for the baseline calibration based in Benabou (2002).

Parameter	Value
$\lambda$	0.625
$\mu$	0.375
$\alpha$	0.350
$\varepsilon$	0.400
$\beta$	0.400
$\omega$	1.000
$\rho$	0.360
$r$	0.000
$\omega_1$	0-0.25

Benabou (2002) calibrates his model by targeting specific moments of the steady-state distribution (income inequality, intergenerational persistence). As our model does not have explicit expressions for those moments, we use the calibrated parameters in Benabou (2002), and contrast the results of our model with heterogeneous schools with the model where the education market plays no role (either because education is produced directly at the household level or because schools are perfectly homogeneous)

Parameters for the production function shares,  $\lambda$  and  $\mu$ , are taken from Barro and Sala-i-Martin (1995), excluding physical capital. Intergenerational persistence in the model

with no heterogeneity,  $p(\tau) = \alpha + \beta\lambda(1 - \tau)$ , is used to calibrate (in the baseline scenario)  $\alpha = 0.35$  and  $\beta = 0.4$ . The parameter for idiosyncratic ability shocks,  $\omega$ , which will be crucial for determining the steady state inequality is set at  $\omega = 1$ . The intertemporal elasticity of substitution,  $\varepsilon = \frac{1}{\eta-1}$ , is set at  $\varepsilon = 0.4$ , at the upper end of empirical estimates. The discount factor,  $\rho = 0.36$ , is consistent with an annual discount factor of  $\beta = 0.96$ , compounded over 25 years, a reasonable time span for a given generation. Finally,  $r = 1$ , so that preferences are intertemporally separable (logarithmic).

The last parameters required in order to complete the model are related to the distribution of school productivities. First, we need to specify the number of potential schools,  $M$ , relative to the number of students,  $N$ . We choose  $M = 1.1N$ , so that in any equilibrium the worst 9.1% of schools are left outside of the market. Second, we must choose  $\omega_1$ , which describes the distribution of random school productivities. If  $\omega_1 = 0$ , all schools are identical and have productivity  $\gamma = 1$ . In that case, accounting for the school market is irrelevant for equilibrium outcomes, as all education expenditures are invested in the production of quality with a CRS technology. To introduce heterogeneity, we choose  $\omega_1 = 0.25$ , such that the mean of school productivities is still equal to 1, but we still have a positive (though small) variance. Although not reported, we also simulate the model with different values for  $\omega_1$ ; results are qualitatively similar to the ones presented below.

The model is simulated for economies with 20,000 agents, over 20 periods (this is, 500 “years”). We simulate two types of economies, with and without heterogeneity in school productivities ( $\omega_1 = 0$  and  $\omega_1 = 0.25$ ). For each economy, we run 5 draws of the stochastic distributions and solve the model for values of  $\tau$  ranging from 0% to 100%, at 1% intervals. Finally, for each economy type we summarize results across the 5 draws. We discuss our findings in the next section.

## 4.2 Results

### 4.2.1 School Investments

As mentioned earlier, the model does not have an analytic solution for steady state values. However, given parameters, the model typically converges numerically to a steady state after 5 periods (125 years).

Figures 1 and 2 summarize one of the main insights of the paper by presenting (steady-state) equilibrium school investments under two polar regimes: no redistribution ( $\tau = 0$ ) and complete equalization ( $\tau = 1$ ). Figure 1 presents the economy with heterogeneous schools. Schools are sorted in productivity along the horizontal axis, starting with the most productive school. When there is no redistribution ( $\tau = 0$ ), differences in household expenditures mimic the underlying income distribution. As discussed in Proposition 1, more productive schools match with the households with higher expenditure and, given the calibrated parameters, end up making strictly larger investments, as competitive pressures between schools are strong. In such a world, most productive providers make the bulk of education investments in the economy. When there are no differences in expenditure ( $\tau = 1$ ), competitive pressures between the set of operating schools are at their weakest, as no school has incentive to provide a higher quality than the one that prevents entry. As more productive schools need to make strictly smaller investments to attain any given quality, the least productive schools make the bulk of investments in such a world. Unsurprisingly, such a shift will have first-order effects on outcomes and on the overall efficiency of human capital investments. As way of comparison, Figure 2 replicates the exercise in an economy with homogenous schools. In such a world, expenditures are always fully invested. As all schools are identical, the identity of the investor is irrelevant for the analysis of the effects of that investment. The discussion on efficiency, then, reverts directly to the discussion on the distribution of expenditures and the rates of return across households (this is, whether we have decreasing returns, and to what extent).

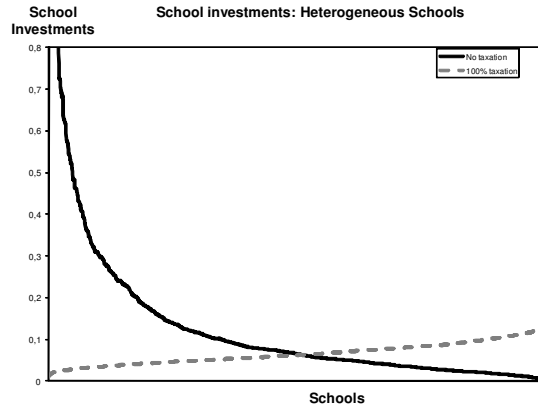


Figure 1: Investments, Heterogeneous Schools

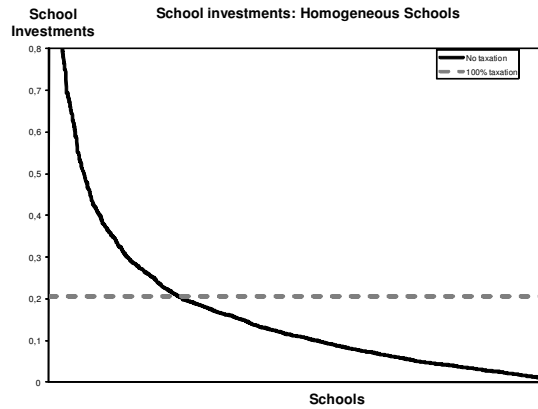


Figure 2: Investments, Homogeneous Schools

The effect of redistribution on competitive pressures is further illustrated in Figure 3, which presents the average markup in the school market (this is, the average revenue over investment ratio) as a function of the tax rate. Higher levels of taxation, which equalize expenditures across households, dampen competitive pressures in the school markets, allowing schools to (on average) charge a larger markup for their services.

#### 4.2.2 Steady-State Human Capital and Income

Figure 4 presents the impact of different redistributive taxes on steady-state average human capital in both economies (homogeneous or heterogeneous producers). Regardless of the



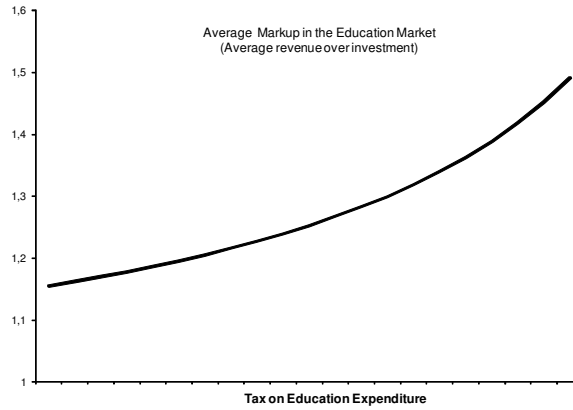


Figure 3: Average School Markups

distribution of school productivities, there is an interior solution for the tax rate that maximizes average human capital, as resources are shifted towards low-income children, where on average marginal investments are more productive due to decreasing returns to scale. However, the human capital maximizing tax rate differs between both economies, and is significantly smaller in the economy with heterogeneous schools. In an economy where all schools are identical, setting a 55% tax on education investments attains the maximum steady-state human capital. When we allow for heterogeneity, the tax rate falls to less than half (20%). As both economies are otherwise identical, this difference is solely explained by the different role played by incentives in both setups. In the first economy, as all expenditures are always fully invested by schools, incentives are not affected by redistribution. Thus, redistribution only impacts efficiency at the household level, as the relative efficiency of the production of market inputs remains unaffected. When schools are not identical, redistribution affects the behavior of schools, changing incentives and, in relative terms, shifting investments towards schools with lower productivity. Thus, there is not only an impact on efficiency at the household level (how expenditures are allocated between households with different characteristics), but also on the production of market inputs. This second effect dampens the efficiency gain associated to redistribution at the household level. Accounting for intermediation has relevant consequences for policy analysis.

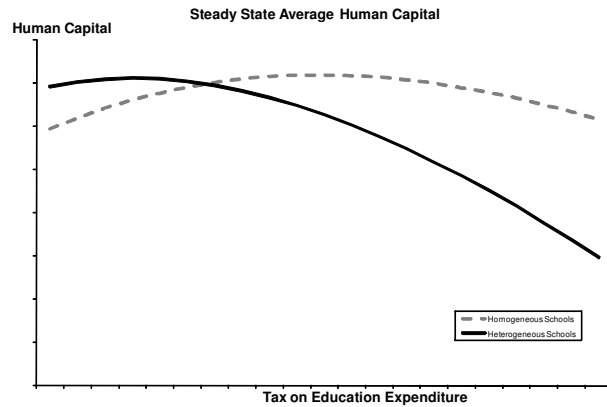


Figure 4: Steady State Human Capital

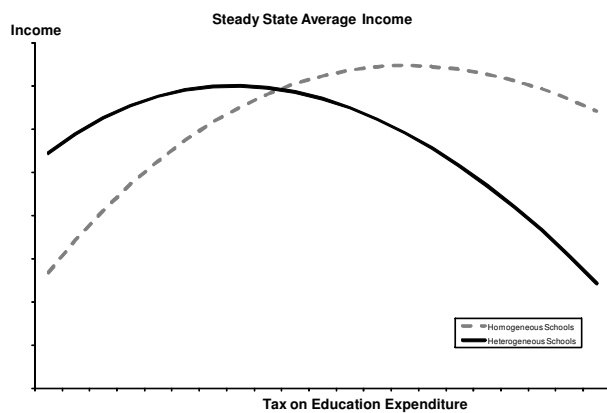


Figure 5: Steady State Income

A similar story applies to steady-state average income, as seen in Figure 5. As the output production function is concave, there is an additional force towards equalization, and accordingly, the tax that maximizes average income is larger in both economies. However, and as with average human capital, the maximizing tax is significantly smaller in the economy with heterogeneous education producers.

We also characterize steady-state income inequality and mobility in both economies. Figure 6 depicts intergenerational mobility, defined as the correlation in income between two successive generations. Even under full equalization in educational expenditures, the correlation is always strictly positive, due to the direct impact of paternal human capital on

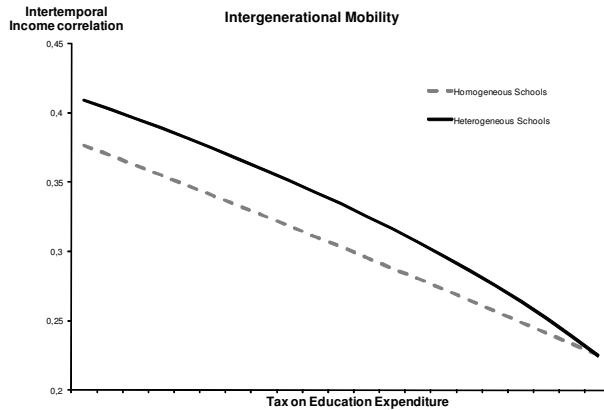


Figure 6: Intergenerational Mobility

the human capital production function. Thus, even if all students receive the same education input, human capital differences will persist in time because part of the capital is transferred directly across generations. Mobility is smaller (the correlation across generations is larger) in an economy with heterogeneity in education production, as differences in education inputs are larger for any given distribution of expenditures. Steady-state income inequality, as measured through the Gini coefficient, is presented in Figure 7. Once again, inequality will be larger in a world with heterogeneous providers, as income differences are amplified in the education market. Inequality falls monotonically with the tax rate, but neither economy attains full equality due to the ability shocks and the legacy of initial distribution of human capital.

### 4.2.3 Welfare

Finally, we look at the welfare implications of different taxation regimes, evaluating the households' steady-state value functions. Figure 8 presents average welfare in both economies, defined as the (unweighted) average across households of the value function evaluated at the steady state. Under this definition, redistributive taxation can increase overall welfare, as there is a region in which the welfare losses of agents that pay positive net taxes in equilibrium are more than compensated with the welfare gains of agents receiving a net subsidy.

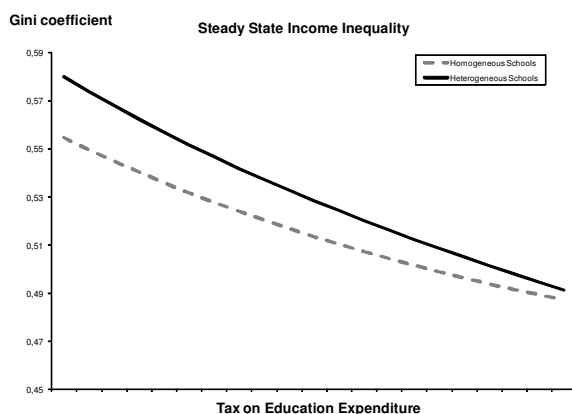


Figure 7: Income Inequality

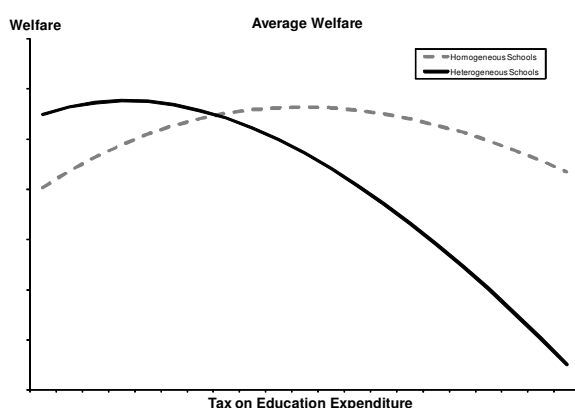


Figure 8: Average Welfare

As in the previous section, the tax rate that maximizes welfare is smaller when we allow for heterogeneity in the production of education.

This is consistent with the fact that, for every income decile (Figure 9), the decile's average welfare is maximized at a smaller tax rate in the economy with heterogeneity in the production function. Even groups at the lower end of the income distribution, which receive net subsidies that are monotonically increasing on the tax rate, would not choose complete redistribution, as the adverse effect on average school effort and the associated deterioration in overall quality offset the benefit of having higher expenditure.

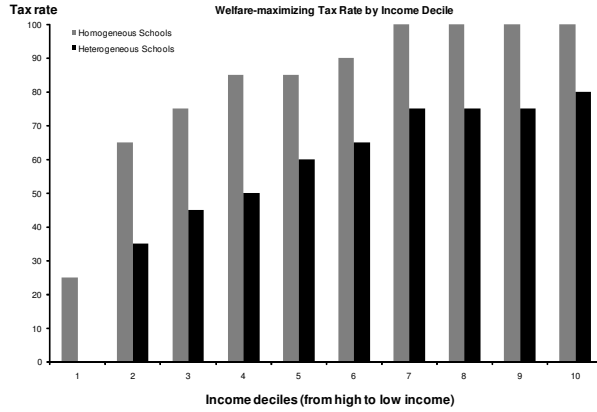


Figure 9: Welfare-maximizing tax rate by income deciles

## 5 Conclusions

This paper has analyzed the intertemporal evolution of human capital under redistributive taxes. To do so, we extended the overlapping generations model in Benabou (2002) to introduce an explicit education market served by private schools with heterogeneous productivity (Tapia, 2010). Schools behave competitively to generate profits, and optimally determine the quality they provide and the tuition they charge. We showed that, as long as education expenditures between parents are heterogeneous, the equilibrium matching function implies that, in every period, the more productive schools serve the students with the higher willingness to pay. Redistributive taxation, which reduces differences in expenditure between households, mutes incentives for differentiation among schools, and shifts investments from more productive to less productive schools. This affects overall efficiency, and reduces the efficiency gains of redistribution associated to the equalization of marginal rates of return across students. We showed numerically that this effect can be important, and income or welfare maximizing taxes in an economy with heterogeneous schools can be significantly smaller than the ones in an economy without an explicit education market, which has been the standard in previous papers.

An obvious objection to our paper is that, in most countries, the education sector is not formed by private, profit-maximizing schools. While this is literally true, we would like to

interpret our results more broadly in terms of the incentives faced by education providers (which can also be teachers within public schools), and how the impact of redistribution on compensation affects equilibrium effort and outcomes. This is, one can interpret our model as a (simple) example of a more general point: namely, the importance of the interaction between redistribution in inputs and incentives on intermediation, and how that affects the outcomes that are the ultimate goal of taxation policy.

## 6 References

[1] Aguiar-Conraria, L. (2005). “Public vs Private Schooling in an Endogenous Growth Model,” *Economics Bulletin*, vol. 9(10), pages 1-6.

[2] Becker, G. and N. Tomes (1979). “An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility.” *Journal of Political Economy*, vol. 87(6), pages 1153-89.

[3] Benabou, R.(2002). “Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?” *Econometrica*, vol. 70(2), pages 33-62.

[4] Bohacek, R. and M., Kapicka (2008). “Optimal Human Capital Policies,” *Journal of Monetary Economics*, vol. 55(1), pages 1-16.

[5] Cardak, B. (2005). “Education Vouchers, Growth and Income Inequality,” *Macroeconomic Dynamics*, vol. 9(1), pages 98-121.

[6] Chiu, W. (1998). “Income Inequality, Human Capital Accumulation and Economic Performance,” *The Economic Journal*, vol. 108(446), pages 44-59.

[7] Cunha, F. and J. Heckman (2007). “The Technology of Skill Formation.” *American Economic Review*, vol. 97(2), pages 31-47.

[8] De Fraja, G. (2002). “The Design of Optimal Education Policies,” *Review of Economic Studies*, vol. 69(2), pages 437-66.

- [9] De Gregorio, J. and S. Kim (2000). “Credit Markets with Differences in Abilities: Education, Distribution, and Growth,” *International Economic Review*, vol. 41(3), pp. 579-607.
- [10] Epple, D. and R. Romano, 1998. “Competition between Private and Public Schools, Vouchers, and Peer-Group Effects,” *American Economic Review*, vol. 88(1), pages 33-62.
- [11] Epple D. & R. Romano, 2008. “Educational Vouchers And Cream Skimming,” *International Economic Review*, vol. 49(4), pp. 1395-1435.
- [12] Ferreyra, M. (2007). “Estimating the Effects of Private School Vouchers in Multi-District Economies.” *American Economic Review*, vol. 97(2), pages 789-817.
- [13] Glomm, G. and B. Ravikumar (1992). “Public versus Private Investment in Human Capital: Endogenous Growth and Income Inequality,” *The Journal of Political Economy*, vol. 100(4), pages 818-834.
- [14] Han, S. and C. Mulligan (2001). “Human Capital, Heterogeneity and Estimated Degrees of Intergenerational Mobility,” *The Economic Journal*, vol. 111(470), pages 207-43.
- [15] Kaganovich, M. and I., Zilcha (1999). “Education, Social Security, and Growth,” *Journal of Public Economics*, vol. 71(2), pages 289-309.
- [16] Loury, G. (1981). “Intergenerational Transfers and the Distribution of Earnings,” *Econometrica*, vol. 49(4), pages 843-867.
- [17] Lucas, R. Jr. (1988). “On the Mechanics of Economic Development,” *Journal of Monetary Economics*, vol. 22(1), pages 3-42.
- [18] MacLeod, B. and M. Urquiola (2009). “Anti-Lemons: School Reputation and Educational Quality.” NBER Working Paper 15112.
- [19] McMillan, R. (2004). “Competition, Incentives, and Public School Productivity.” *Journal of Public Economics*, vol. 88(9-10), pages 1871-1892.
- [20] Nechyba, T. 2000. “Mobility, Targeting, and Private-School Vouchers,” *American Economic Review*, vol. 90(1), pages 130-146.
- [21] Tapia, M. (2010). “Competition, Incentives, and the Distribution of Investments in

PrivateSchool Markets,” Documentos de Trabajo 387, Instituto de Economía. Pontificia Universidad Católica de Chile.

[22] Urquiola, M. (2005). “Does School Choice Lead to Sorting? Evidence from Tiebout Variation.” *American Economic Review*, vol. 95(4), pages 1310-1326.

[23] Urquiola, M. and E. Verhoogen (2009). “Class Size Caps, Sorting, and the Regression Discontinuity Design.” *American Economic Review*, vol. 99(1), pages 179-215.

[24] Vial, B. (2008). “Competitive Equilibrium and Reputation under Imperfect Public Monitoring,” Documentos de Trabajo 327, Instituto de Economía. Pontificia Universidad Católica de Chile.

[25] Watanabe, M. and M. Yasuoka (2009). “Income Growth, Inequality and Preference for Education Investment: a Note,” *Economics Bulletin*, vol. 29(4), pages 3075-3082.

[26] Zhang, Jie (1996). “Optimal Public Investments in Education and Endogenous Growth,” *Scandinavian Journal of Economics*, vol. 98(3), pages 387-404.

## A Proof of Proposition 1

*Proof.* a) Straightforward.

(b) to (e): We can show that under the proposed qualities no school has a profitable deviation.

Take any school with productivity  $\gamma_j > \gamma_{N+1}$ , that in the proposed equilibrium is matched with a student with income  $y_j$ . Equilibrium profits can be written as

$$\pi_j^* = sy_j - \theta_j^* = sy_j - s \left[ \frac{y_N \gamma_{N+1}^{\frac{\zeta}{\varphi}} + \sum_{i=1}^{N-j} (y_{N-i} - y_{N+1-i}) \gamma_{N+1-i}^{\frac{\zeta}{\varphi}}}{\gamma_j^{\frac{\zeta}{\varphi}}} \right]$$

Can the school find a profitable deviation?

Suppose the school tries to enroll the student immediately above in the income distrib-



ution, which has income  $y_{j-1}$ . The school must provide a quality that is at least as good as  $q_{j-1}^*$ , the equilibrium quality of the school in which student  $m - 1$  is enrolled. Profits would then be:

$$\pi_j(y_{j-1}) = sy_{j-1} - \theta_j^{**} = sy_{j-1} - s \left[ \frac{y_N \gamma_{N+1}^{\frac{\zeta}{\varphi}} + \sum_{i=1}^{N-j+1} (y_{N-i} - y_{N+1-i}) \gamma_{N+1-i}^{\frac{\zeta}{\varphi}}}{\gamma_j^{\frac{\zeta}{\varphi}}} \right]$$

Comparing profits, we can see that the school would get exactly the same profits in both cases, and thus has no incentives to deviate:

$$\begin{aligned} \pi_j^* - \pi_j(y_{j-1}) &= \\ &= (sy_j - sy_{j-1}) + \left( \frac{q_{j-1}^*}{\gamma_j^{\frac{\zeta}{\varphi}}} \right)^{\frac{1}{\varphi}} - \left( \frac{q_j^*}{\gamma_j^{\frac{\zeta}{\varphi}}} \right)^{\frac{1}{\varphi}} \\ &= (sy_j - sy_{j-1}) + s \left[ \frac{(y_{j-1} - y_j) \gamma_j^{\frac{\zeta}{\varphi}}}{\gamma_j^{\frac{\zeta}{\varphi}}} \right] \\ &= (sy_j - sy_{j-1}) + s [(y_{j-1} - y_j)] = 0 \end{aligned}$$

At the equilibrium qualities, the marginal revenue that the school would gain from enrolling a richer student is exactly offset by the additional cost of providing higher quality. In fact, if the school tried to go even further upwards the income distribution, and enroll the student with income  $y_{j-2}$ , its profits would actually decrease:

$$\begin{aligned}
\pi_j^* - \pi_j(y_{j-2}) &= (sy_j - sy_{j-2}) + \left(\frac{q_{j-2}^*}{\gamma_j^\zeta}\right)^{\frac{1}{\varphi}} - \left(\frac{q_j^*}{\gamma_j^\zeta}\right)^{\frac{1}{\varphi}} \\
&= (sy_j - sy_{j-2}) + s \left[ \frac{(y_{j-1} - y_j) \gamma_j^{\frac{\zeta}{\varphi}}}{\gamma_j^{\frac{\zeta}{\varphi}}} \right] + s \left[ \frac{(y_{j-2} - y_{j-1}) \gamma_{j-1}^{\frac{\zeta}{\varphi}}}{\gamma_j^{\frac{\zeta}{\varphi}}} \right] \\
&= (sy_j - sy_{j-1}) + s \left[ \frac{(y_{j-1} - y_j) \gamma_j^{\frac{\zeta}{\varphi}}}{\gamma_j^{\frac{\zeta}{\varphi}}} \right] \\
&\quad + (sy_{j-1} - sy_{j-2}) + s \left[ \frac{(y_{j-2} - y_{j-1}) \gamma_{j-1}^{\frac{\zeta}{\varphi}}}{\gamma_j^{\frac{\zeta}{\varphi}}} \right] \\
&= (sy_j - sy_{j-1}) + s(y_{j-1} - y_j) + (sy_{j-1} - sy_{j-2}) \left(1 - \frac{\gamma_{j-1}^{\frac{\zeta}{\varphi}}}{\gamma_j^{\frac{\zeta}{\varphi}}}\right) \\
&= (sy_{j-1} - sy_{j-2}) \left(1 - \frac{\gamma_{j-1}^{\frac{\zeta}{\varphi}}}{\gamma_j^{\frac{\zeta}{\varphi}}}\right) > 0 \text{ as } sy_{j-1} < sy_{j-2} \text{ and } \frac{\gamma_{j-1}^{\frac{\zeta}{\varphi}}}{\gamma_j^{\frac{\zeta}{\varphi}}} > 1
\end{aligned}$$

Profits would be smaller as the quality increase between students  $j - 2$  and  $j - 1$  is implicitly priced at the average cost of school  $j - 1$ , which is strictly smaller than the marginal cost of  $j$ . As the argument can be easily extended for any student  $j - s$ ,  $s > 1$ , there are no profitable deviations in enrolling higher income students. What about enrolling a student with lower income? If school  $j$  tries to get student  $j + 1$ , it has to offer at least  $q_{j+1}^*$  :

$$\begin{aligned}
\pi_j^* - \pi_j(y_{j+1}) &= s(y_j - y_{j+1}) + \left(\frac{q_{j+1}^*}{\gamma_j^\zeta}\right)^{\frac{1}{\varphi}} - \left(\frac{q_j^*}{\gamma_j^\zeta}\right)^{\frac{1}{\varphi}} \\
&= (y_j - y_{j+1}) - \left[ \frac{(y_j - y_{j+1}) \gamma_{j+1}^{\frac{\zeta}{\varphi}}}{\gamma_j^{\frac{\zeta}{\varphi}}} \right] \\
&= (y_j - y_{j+1}) \left(1 - \frac{\gamma_{j+1}^{\frac{\zeta}{\varphi}}}{\gamma_j^{\frac{\zeta}{\varphi}}}\right) > 0 \text{ as } y_j > y_{j+1} \text{ and } \frac{\gamma_{j+1}^{\frac{\zeta}{\varphi}}}{\gamma_j^{\frac{\zeta}{\varphi}}} < 1.
\end{aligned}$$

As the reduction in revenue is larger than cost savings, no school has incentive to move downwards the income distribution.

Thus, at the proposed equilibrium qualities, it is true that, for any  $j < N + 1$ :

$$\begin{aligned}\pi_j^* &= \pi_j(q_{j-1}^*) \\ \pi_j^* &> \pi_j(q_{j-s}^*), \quad s > 1 \\ \pi_j^* &> \pi_j(q_{j+s}^*), \quad s \geq 1\end{aligned}$$

so there are no profitable deviations outside of the equilibrium matches.

The qualities will be sustained in equilibrium by competitive pressures and the incentives the school faces. Given any equilibrium match, an offer below  $h_{i\mu(i)}^* = a_i^\alpha q_i^*$  would allow the school immediately below to make a better offer to student  $i$  and make profit. Therefore, equilibrium quality cannot be below  $q_i^*$ , as the school would lose the student. Any offer above  $h_{i\mu(i)}^* = a_i^\alpha q_i^*$  only increases the cost for school  $\gamma_i$ , without any additional revenue. Therefore, the equilibrium quality will not be above  $q_i^*$ , as the school has no incentive to incur in that cost. □

## B Proof of Propositions 2 and 3

*Proof.* Guessing that the value function satisfies  $\ln U_t^i = V_t \ln h_t^i + B_t$  and replacing into (18) implies

$$\begin{aligned}
V_t \ln h_t^i + B_t &= \rho B_{t+1} + \max_l \{(1 - \rho + \rho \varepsilon (1 - \tau_t) V_{t+1}) \mu \ln l - (1 - \rho) l^n\} \\
&\quad + \max_s \{(1 - \rho) \ln((1 - s)) + \rho \varepsilon V_{t+1} \ln(s)\} \\
&\quad + [(1 - \rho) \lambda + \rho(\alpha + \varepsilon \lambda (1 - \tau_t)) V_{t+1}] \ln h_t^i + (1 - \rho) \ln(1 + \pi_t) \\
&\quad - \rho/r(1 + \varepsilon)rV_{t+1}(1 - (1 + \varepsilon)rV_{t+1})\omega^2/2 \\
&\quad + \rho V_{t+1} (\ln \kappa - \varepsilon \omega_1^2/2 - \varepsilon \ln \hat{\pi}_t + \varepsilon(1 - \tau_t) \ln(1 + \pi_t) + \varepsilon \tau_t \ln \tilde{y}_t) \quad (20)
\end{aligned}$$

Note that the uncertainty comes only from the human capital of the child  $\xi_{t+1}^i$ , which enters into the expected utility of the next period  $U_{t+1}^i$ , but does not affect the investment decisions of the parent, which are shown later to be independent of the uncertainty of the model. We also assume that the realization of the productivity of the schools is known by parents. The maximization problem is strictly concave and then the first order conditions for  $l$  and  $s$  are sufficient for optimality (these conditions are straightforward to derive and are expressed in Proposition 3 in the text). Grouping the terms with  $\ln h_t^i$  in (20)

$$V_t = (1 - \rho) \lambda + \rho(\alpha + \varepsilon \lambda (1 - \tau_t)) V_{t+1} \quad (21)$$

which is recursively equivalent to

$$V_t = (1 - \rho) \lambda \sum_{k=0}^{\infty} \rho^k \prod_{j=0}^{k-1} p(\tau_{t+j}) \quad (22)$$

where

$$p(\tau_t) = \alpha + \varepsilon \lambda (1 - \tau_t)$$

Plugging (22) into (20) results in

$$\begin{aligned}
B_t - \rho B_{t+1} &= (1 - \rho + \rho\varepsilon(1 - \tau_t)V_{t+1})\mu \ln l - (1 - \rho)l^n \\
&+ (1 - \rho) \ln((1 - s)) + \rho\varepsilon V_{t+1} \ln(s) \\
&+ (1 - \rho) \ln(1 + \pi_t) - \rho(1 + \varepsilon)V_{t+1}(1 - (1 + \varepsilon)rV_{t+1})\omega^2/2 \\
&+ \rho V_{t+1} (\ln \kappa - \varepsilon\omega_1^2/2 - \varepsilon \ln \hat{\pi}_t + \varepsilon(1 - \tau_t) \ln(1 + \pi_t) + \varepsilon\tau_t \ln \tilde{y}_t)
\end{aligned}$$

where  $B_t$  satisfies the transversality condition  $\lim_{t \rightarrow \infty} (\rho^t B_t) = 0$ . From (19), it follows that the natural logarithm of the human capital at date  $t + 1$  is

$$\begin{aligned}
\ln h_{t+1}^i &= \ln \kappa + \varepsilon \ln(s_t) + (1 + \varepsilon) \ln \xi_{t+1}^i + (\alpha + \varepsilon\lambda(1 - \tau_t)) \ln h_t^i \\
&+ \varepsilon\mu(1 - \tau_t) \ln l_t^i + \varepsilon \ln \delta_t - \varepsilon \ln \hat{\pi}_t + \varepsilon(1 - \tau_t) \ln(1 + \pi_t) + \varepsilon\tau_t \ln \tilde{y}_t
\end{aligned}$$

Note that as shown in Proposition 3 the labor supply and the savings rate choices do not depend on individual characteristics. As the human capital at date  $t$  follows a lognormal distribution,  $\ln h_t^i \sim \mathcal{N}(m_{h_t}, \Delta_{h_t}^2)$ , the average income,  $\tilde{y}_t$ , given by  $\sum_{i=1}^N (\tilde{y}_t)^{\tau_t} (y_t^i)^{1-\tau_t} = Y_t$  satisfies

$$\tilde{y}_t = \left( \frac{\sum_{i=1}^N y_t^i}{\sum_{i=1}^N (y_t^i)^{1-\tau_t}} \right)^{1/\tau_t} = \left( \frac{E(y_t^i)}{E((y_t^i)^{1-\tau_t})} \right)^{1/\tau_t}$$

which implies that

$$\ln \tilde{y}_t = \lambda m_{h_t} + \lambda^2 (2 - \tau_t) \Delta_{h_t}^2 / 2 + \mu \ln l_t + \ln(1 + \pi_t)$$

This expression is used to obtain the parameters of the distribution of the human capital

at date  $t + 1$

$$\begin{aligned}
m_{h_{t+1}} &= \ln \kappa + \varepsilon \ln(s_t) - (1 + \varepsilon)\omega^2/2 + (\alpha + \varepsilon\lambda)m_{h_t} \\
&\quad + \varepsilon\mu \ln l_t - \varepsilon\omega_1^2/2 - \varepsilon \ln \hat{\pi}_t + \varepsilon \ln(1 + \pi_t) + \varepsilon\tau_t\lambda^2(2 - \tau_t)\Delta_{h_t}^2/2 \\
\Delta h_{t+1}^2 &= (1 + \varepsilon)^2\omega^2 + (\alpha + \varepsilon\lambda(1 - \tau_t))^2\Delta h_t^2 + \varepsilon^2\omega_1^2
\end{aligned}$$

Defining  $B_t \equiv W_t - V_t m_{h_t}$  and using the previous expressions for  $m_{h_{t+1}}$  and  $\ln \tilde{y}_t$  and (21), the difference equation for  $W_t$  satisfies

$$\begin{aligned}
\frac{W_t - \rho W_{t+1}}{1 - \rho} &= \mu \ln l_t - l_t^\eta + \ln((1 - s_t)) + \lambda m_{h_t} \\
&\quad + \ln(1 + \pi_t) + r\rho(1 + \varepsilon)^2(1 - \rho)^{-1}V_{t+1}^2\omega^2/2
\end{aligned}$$

Expanding the difference equation we have that

$$\frac{W_t}{1 - \rho} = \sum_{k=0}^{\infty} \rho^k (\lambda m_{h_{t+k}} + J_{t+k})$$

where  $J_{t+k} = \mu \ln l_{t+k} - l_{t+k}^\eta + \ln((1 - s_{t+k})) + \ln(1 + \pi_{t+k}) + r\rho(1 + \varepsilon)^2(1 - \rho)^{-1}V_{t+k+1}^2\omega^2/2$ .

To develop the expression for  $\sum_{k=0}^{\infty} \rho^k \lambda m_{h_{t+k}}$  note that we can express

$$m_{h_{t+1}} = v m_{h_t} + K_t$$

where  $v = \alpha + \beta\lambda$  and  $K_t = \ln \kappa + \varepsilon \ln(s_t) - (1 + \varepsilon)\omega^2/2 + \varepsilon\mu \ln l_t - \varepsilon\omega_1^2/2 - \varepsilon \ln \hat{\pi}_t + \varepsilon \ln(1 + \pi_t) + \varepsilon\tau_t\lambda^2(2 - \tau_t)\Delta_{h_t}^2/2$ . So

$$m_{h_t} = v^t m_0 + \sum_{i=1}^t v^{i-1} K_{t-i}$$

and then

$$\sum_{t=0}^{\infty} \rho^t (\lambda m_{h_t} + J_t) = \frac{\lambda m_0}{1 - \rho(\alpha + \beta\lambda)} + \frac{1}{1 - \rho(\alpha + \beta\lambda)} \rho\lambda \sum_{t=0}^{\infty} \rho^t K_t + \sum_{t=0}^{\infty} \rho^t J_t$$

So finally

$$\frac{W_t}{1 - \rho} = \frac{\lambda m_t}{1 - \rho(\alpha + \beta\lambda)} + \frac{1}{1 - \rho(\alpha + \beta\lambda)} \rho\lambda \sum_{k=0}^{\infty} \rho^k K_{t+k} + \sum_{k=0}^{\infty} \rho^k J_{t+k}$$

where  $J_{t+k} = \mu \ln l_{t+k} - l_{t+k}^\eta + \ln((1 - s_{t+k})) + \ln(1 + \pi_{t+k}) + r\rho(1 + \varepsilon)^2 (1 - \rho)^{-1} V_{t+k+1}^2 \omega^2/2$  and  $K_{t+k} = \ln \kappa + \varepsilon \ln(s_{t+k}) - (1 + \varepsilon)\omega^2/2 + \varepsilon\mu \ln l_{t+k} - \varepsilon\omega_1^2/2 - \varepsilon \ln \hat{\pi}_{t+k} + \varepsilon \ln(1 + \pi_{t+k}) + \varepsilon\tau_{t+k}\lambda^2 (2 - \tau_{t+k}) \Delta_{h_{t+k}}^2/2$ . The previous difference equation depends on the sequence  $\{\tau_{t+k}, \theta_{t+k}, a_{t+k}\}_{k=0}^{\infty}$ .

Under a constant policy  $\tau_t = \tau$ ,  $s_{t+k} = \bar{s}$ ,  $\delta_{t+k} = \bar{\delta}$ ,  $\hat{\pi}_{t+k} = \bar{\pi}$ , and  $\pi_{t+k} = \bar{\pi}'$  the previous expression reduces to

$$W_0 = \frac{(1 - \rho) \lambda m_0}{1 - \rho(\alpha + \beta\lambda)} + \frac{\rho\lambda K_0}{(1 - \rho(\alpha + \beta\lambda))} + J_0$$

where

$$\begin{aligned} K_0 &= \ln \kappa + \varepsilon \ln(\bar{s}) - (1 + \varepsilon)\omega^2/2 + \varepsilon\mu \ln l(\tau) - \varepsilon\omega_1^2/2 - \\ &\quad \varepsilon \ln \bar{\pi} + \varepsilon \ln(1 + \bar{\pi}') + \varepsilon\tau\lambda^2 (2 - \tau) \omega^2/2, J_0^{-1} \\ &= \mu \ln l(\tau) - l(\tau)^\eta + \ln((1 - \bar{s})/(1 + \bar{\theta})) + \ln(1 + \bar{\pi}') + r\rho(1 + \varepsilon)^2 (1 - \rho)^{-1} \bar{V}^2 \omega^2/2 \end{aligned}$$

and where

$$\bar{V} = (1 - \rho) \lambda / (1 - \rho(\alpha + \varepsilon\lambda(1 - \tau)))$$

□