Higher Educational Standard, Human Capital Accumulation and Intergenerational Mobility in Education: a Theoretical Model*

Stefano Staffolani† and Maria Cristina Recchioni‡

Very preliminary and incomplete DRAFT - not for quotation

English to be revised

September 5, 2015

Abstract

In this paper we assess higher educational settings where universities can be differentiated by their educational standards. Standards are assumed to influence negatively the graduation probability and positively the human capital accumulation during studies. Secondary school graduates choose whether and where to enroll by considering their own talent, their family willingness to finance education and the offered standard(s). We analyze centralization and decentralization in the setting of standards. In the former case, "poor" would prefer a higher standard than "rich". If the standard is set in order to maximize the enrollment rate, human capital accumulation is sub-optimal. Decentralization may perform worst in terms of human capital accumulation if moving costs are not completely financed and it always performs worst in terms of intergenerational mobility in education.

Keywords: education, standards, universities, human capital, inequality.

1 Introduction

There exists a huge literature on wage rate differences among graduates showing that the labour market rewards workers not only according to their educational level, field and grade, but also to the quality of the attended

*We wish to thank Riccardo Lucchetti for his useful comments.

†Department of Economics and Social Sciences, UNIVPM, Ancona - Italy – Corresponding Author, s.staffolani@univpm.it.

‡Department of Management, UNIVPM, Ancona - Italy

university[1]. Recent results highlight that the return to college selectivity is sizeable, but, as underlined by Dale and Krueger (2011), adjusting "for unobserved student ability by controlling for the average SAT score of the colleges that students applied to, our estimates of the return to college selectivity fall substantially and are generally indistinguishable from zero. There were notable exceptions for certain subgroups. ... for students who come from less-educated families... the estimates of the return to college selectivity remain large".

The typical explanation of the wage premium differences among universities lays both on the selection of the enrolled, usually based on previous curricula and entry tests (Hershbein (2013)), and on the actual quality of the attended higher education institution. "Best" universities give their students a better learning environment (class size, tutors, laboratories, better internships opportunities, higher instructional expenditures per student and so on) and therefore allows for a higher human capital accumulation during studies. Nevertheless, a good learning environment is costly and usually requires a higher expenditure per student and higher tuition fees.

Differences among universities can also be due to the commitment required to student to pass the exams, to professors severity, to the whole policies of the university in helping weak students; all these universities characteristics are defined *educational standards* . Briefly, educational standards (also defined learning, or examination, or student workload standards) describe what students are expected to have learned at the end of the course.

In this paper, educational standards are viewed as the determinant of the probability of graduation and of the skills acquired during study in a setting where students are differentiated according to their ability and universities can be vertically differentiated according to the offered standards [2]. As underlined by Bishop and Wobmann (2001), the institutions regulating educational standards that influence students performance considerably are "examination systems, centralized decision-making versus school autonomy.. .and competition in the education system " that are all dimensions of educational standards.

The economic literature on educational standards frequently considers a standard-setter policy maker which maximizes his objective function. Analyzing the "common view" in the literature debate, the seminal paper of

[1]The relationships among workers earnings and the quality of the college they attended have been empirically analyzed by Brewer et al. (1999), Black and Smith (2006) for US, Milla (2012) for Canada, Makiko and Tomohiko (2013) for Japan, Zhang et al. (2013) for China, Lindahl and Regnér (2005) for Sweden, Canaan and Mouganie (2014) for France, Chevalier and Conlon (2003) for UK. All these papers control for college selectivity and some of them take account of individual unobserved heterogeneity among the enrolled.

[2]Educational standards are linked to the "grading standard" literature. For instance, Bagues et al. (2008) argument that the ratio between the number of exams that students passed and the number of exams that students should have taken "might capture both the students true quality and the easiness (or grading standards) of a given institution".

Costrell (1994) highlights that standards could be set at a lower lever than the one parents would like because of egalitarian reasons, because of the lack of accountability and, for U.S., because of the lack of a national curriculum or exams. Difficulties in maintaining high standards arise also on deterioration of parental inputs and on the increased "diversity" of the student population. The Author analyses theoretically the choices of a policy maker who sets the standard in order to maximize his social welfare function, knowing that students choose whether or not to meet the standards and knowing that standards influence students incentives to provide effort. He concludes that an egalitarian policy maker set lower standards than the ones preferred by the median voter. Schwager (2013) deepen the previous analyses by taking into account the direct democratic choice of an examination standard against a utilitarian welfare function. The Author shows that the median preferred standard is inefficiently low if the marginal cost of reaching a higher performance reacts more sensitively to ability for high than for low abilities, and if the right tail of the ability distribution is longer than the left tail. De Paola and Scoppa (2007) analyse how a "badly functioning labor market", where effective skills are scarcely rewarded, affects the behaviour of the policy maker who maximizes a social welfare function. In the case standards are defined taking into account the effective productivity of skills, they will be set at a high level in order to encourage individuals' learning effort. Instead, if standard are set considering exclusively return to school, they will be set at a low level.

Other Authors considered the link between educational standards and inequality. Among them, Betts (1998), by featuring workers with heterogeneous abilities, concludes that higher educational standards increase the earnings of both the most able and the least able workers. Thus an egalitarian social planner may set higher standards than an income-maximizing social planner. The results mitigate the concern that higher standards are necessarily inegalitarian. The paper of Vauteren and Escriche (2006) aims to explain how financial constraints and family background characteristics affect the educational investments of individuals born in low-income families. It shows that an increase in educational standards would help poor individuals with high-ability if it is combined with other non-monetary measures.

The economic literature has also compared the outcomes of a centralized educational standard versus a decentralized one. In the former, standards are equal in all the institutions, whereas, in the latter, different standards exist in different districts or between private and public institutions. One of the conclusions of Costrell (1994) is that decentralization among identical communities lower standards. Conditions under which centralized educational standards raise welfare are deeply analyzed in Costrell (1997) that concludes that high degrees of egalitarianism and cross-district heterogeneity, due to high rates of geographical mobility of graduates (in such a case the benefit of high standards in any district are not fully appropriated by the graduates who meet the standard), favor centralization. Himmler and Schwager (2007) present a model of decentralized graduation standards. They show that a school whose students are disadvantaged on the labor market applies less demanding standards because such students have less incentives to graduate. Brunello and Rocco (2005) compare the public and private educational standards. In their model, private schools can offer a lower educational standard because they attract students with a relatively high cost of effort, who would find the high standards of public schools excessively demanding. Calibrating the model, they conclude that this case is supported for Italy, but not for U.S.

In this paper, we analyze institutional settings where the standards can be either homogeneous (centralization) or heterogeneous (decentralization) and where the standards influence the enrollment choice, the graduation probability and the human capital accumulation. With respect to the previous literature, we explicitly consider that enrollment is a risky choice. Actually, standards act as "prices" of the risky investment in human capital: a higher standard implies a higher risk of dropping-out (investment failure) and a higher wage (investment return)[3]. Furthermore, we do not explicitly consider the policy maker objective function, but we compare human capital accumulation and inequality with respect to different objectives. We take into account that in a centralized setting students do not need to move, whereas in a decentralized one moving costs must be faced in order to enroll in the preferred university.

In the case of decentralization, we assume that students self-select in universities that match their talent at the best, so that more talented students enroll in more difficult universities[4] . This hypothesis implies that we should observe that students who actually choose the university to enroll, the movers, choose the university whose educational standard adapts better

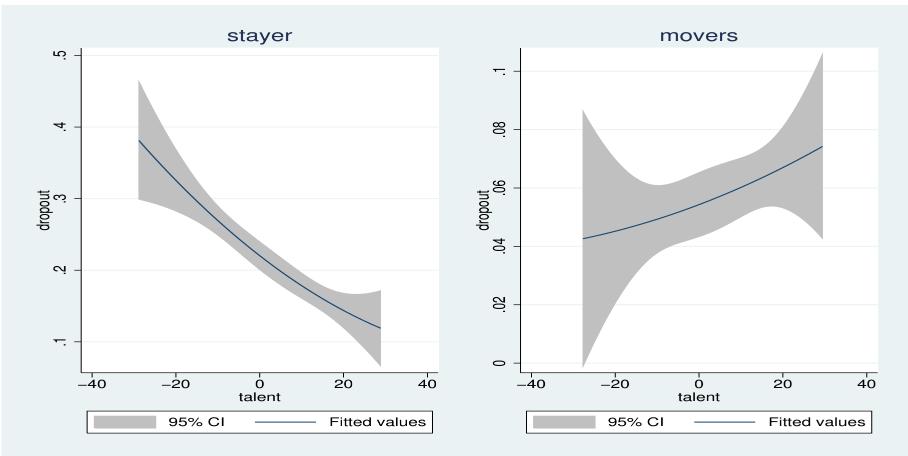
[3]The model can apply to every choice where one among different costly investments can be made and the chosen investment increases long-life utility with a probability depending positively on individual talent and negatively on the return of the investment. For instance, illegal migration decision in the case that the migrant risks to be repatriated if discovered is higher in richer countries. Concerning education, it can apply to all levels of education but is more suitable for college choices. Public compulsory schools usually does not differ too much, at least in many European countries. Even in the secondary school system, once the type of school is taken into account, the quality of education should be similar because frequently all students should reach the same qualification through a uniform State-managed final exam. Furthermore, until secondary school students are supposed to live with their family, and hence the differences in the quality among schools are hardly taken into account, especially for individuals living in small town where the supply of schools can be very limited.

[4]Dillon and Smith (2015) consider the effects of the interaction between student ability and college quality on academic outcomes and future earnings. They find "little evidence to support the "mismatch" hypothesis that college quality and ability interact in substantively important ways".

to their talent whereas the stayers, because of financial constraints, accept the random standard of the closest university. Therefore, the drop-out rate of stayers must depend negatively on their talent, whereas the drop-out rate of movers could be independent of it.

In order to give some "naive" evidence of the above statement and given that Italian higher education is open-access, we analyse the survey "L'inserimento professionale dei diplomati", provided by the Italian Institute of Statistics (ISTAT, 2007). The micro-data collected refer to 25512 individuals who obtained their secondary school degree in 2004, interviewed 3 years later. Information on the chosen university, on family background, on drop-out during studies are available. A proxy for individual talent can be build[5]. Figure 1 shows the relationship between the drop-out rate and the proxy of individual talent for stayers and movers[6] as a fitted quadratic plot with 95% confidence intervals. Figure 1 states that the negative relationship between the talent and the drop-out rate exists only for stayers. Movers are not significantly affected by talent in their drop-out decision.

Figure 1: The relationship between talent and the drop-out probability for stayers and movers, Italy



[5]Secondary school students obtain the final diploma by passing an exam ("esame di maturità") set at Governmental level. In this exam, each student is evaluated by a numerical grade. Following Pigini and Staffolani, 2015, we build a proxy of the individual talent as the residual of an estimation where the grade is regressed on the interaction of the type of school and the province of residence during secondary school.

[6]Movers are those secondary school graduates who moved to other districts even if a university offering the field of study they chose was located in their secondary school district (3556 obs). Among all the students who enroll without moving (11539 obs), stayers are a subsample (3556 obs) chosen considering the one-to-one best match in a propensity score matching model, where the regressors for the moving decision are the student's talent, the type of school (5 classes), the region of residence (21 classes), the mother's education (4 classes) and the father's occupation (9 classes). All the estimates are available from the Authors.

How can this result be explained? The higher commitment once the moving decisions has been made explains the strong difference in the average drop out rate between stayers (22%) and movers (6%), but it cannot explain why talent does not significantly affect movers drop-out. Our explanation is the one proposed above: among movers, more talented students choose universities with a high standard, whereas less talented ones choose universities characterized by low standard.

To summarize, the model is based on the assumptions a) tuition fees and "quality" are equal in all the universities; b) the open access principle holds so that the choice of the university where to enroll is in the students' hands , but subject to the constraint coming from family income because of moving costs; c) different universities can offer different standards.

In this setting, differences in the human capital accumulation during studies (and wage differences) among graduates coming from different universities can therefore exist because of *self-selection*: secondary school graduates select themselves into university according to their talent so that high-standard universities attract more talented students, and because of *educational standards*: for any given talent, high standard universities require stronger effort and commitment from their students.

We analyze theoretically the "centralized standard case", in which each student chooses if enrolling or not at the standard set by the policy maker and the "decentralized standard case" in which each student can either enroll in the university offering his/her preferred standard by sustaining moving costs, or enroll at the closest university (whose standard is random), or not enroll at all.

In the centralized case our theoretical model concludes that the level of the standard that maximizes human capital accumulation is higher than the level of the standard that maximizes the enrollment rate. Therefore, policies oriented to raise the enrollment rate, as the one suggested to European countries from the "Lisbon strategy" , risk to be detrimental to human capital accumulation[7]. Furthermore, we obtain that the level of the standard is negatively correlated with family income: poorer student would prefer a higher standard. The theoretical model suggests that the decentralized system performs better than the centralized one in term of human capital accumulation, but only if the choice of the university where to enroll is achievable for all students, so that only if their moving decisions are completely financed[8]. If this is not the case and the share of students

[7] "the number of 18 to 24 year olds with only lower-secondary level education who are not in further education and training should be halved by 2010"

. `http://www.consilium.europa.eu/en/uedocs/cms_data/docs/pressdata/en/ec/00100-r1.en0.htm`. In our view, the risk is that the goal of broaden access and success rate is mainly pursued by lowering the standard of the courses, as empirically shown in Bratti et al. (2007).

[8] Note that the model considers how financial constraints affect both the enrollment

that cannot afford moving costs is sufficiently high, human capital accumulation can be higher in a centralized system whose standard is efficiently chosen. Given that moving toward the preferred university is more likely for students coming from richer families, a decentralized system would reduce inter-generational mobility in education[9]. Nevertheless, recommendations of the European Commission to the Member States go in the opposite direction of reinforcing decentralization without thoroughly investigating the consequences[10].

The outline of the paper is as follows. Section 2 presents the model. Section 3 illustrates two settings, the centralized and the decentralized and discusses some implications for the university governance. Section 4 draws some conclusions. Appendix A contains the proofs of the main results, Appendix B presents the analyses of the choice of effort during studies made by the enrolled and Appendix C contains some hints on the hypothesis used to graph the results of the model.

2 The Model

We consider a cohort of students, endowed each with a given talent θ which is a real random variable with distribution function:

$$F(\theta) = \int_0^\theta \phi(\theta')d\theta' \qquad \theta \in (-\infty, +\infty) \tag{1}$$

where $\phi(\theta) = 0$, $\theta \in (-\infty, 0)$ and $\phi(\theta) = 0$, $\theta \in (\overline{\theta}, +\infty)$.

They completed a secondary school cycle and should choose whether and where to enroll at university, eventually by moving to other geographical districts[11].

and the moving decisions, but deepen the analysis of the latter.

[9] Actually, richer students seem to not be influenced in their choices by the distance of the university: Denzler and Wolter (2011): "The results also show that distance does not influence study choices among students from the highest socioeconomic group, a finding that further indicates that distance to university is an expression of differences in the cost of a university education."

[10] "Over-regulation and nationally defined courses hinder modernization and the effective management of universities in the EU. To reform their governance, European universities are calling for more autonomy in preparing their courses and in the management of their human resources and facilities. They also want to reinforce public responsibility for the strategic orientation of the whole system. Hence it is not a call for the withdrawal of the State but for a new allocation of tasks. The Commission invites the Member States to relax the regulatory framework so as to allow university leadership to undertake genuine change and pursue strategic priorities."
`http://europa.eu/legislation_summaries/education_training_youth/lifelong_learning/c11078_en.htm`

[11] The relationship between the ex-ante decision to starting university and the university outcomes is theoretically analyzed in the seminal paper of Altonji (1993), that consider that the former decision is made under uncertainty. Oppedisano (2009) present a model of

We assume that their choice depends on the *educational standards* ($s > 0$) offered by the academic institutions that influence both the expected probability of graduation and the human capital accumulation during studies[12].

Assumption 1 *The probability of graduation* $0 < p(s,\theta) < 1$ *is an increasing function of talent and a decreasing function of the standard. It complies with the following:*

$$p'_s(s,\theta) < 0; \quad p'_\theta(s,\theta) > 0; \quad p''_{s\theta}(s,\theta) > 0; \quad p''_{ss}(s,\theta) > 0,$$

$$\lim_{s\to 0} p(s,\theta) = 1 \quad \lim_{s\to\infty} p(s,\theta) = 0 \quad \lim_{\theta\to 0} p(s,\theta) = 0 \qquad \lim_{s\to\infty} \varepsilon_{ps} \leq -1$$

and

$$p(t\,s, t\,\theta) = p(s,\theta), \ \forall t > 0,$$

that is, $p(s,\theta)$ *is homogenous of degree zero in* (s,θ).[13].

The wage rate of undergraduates (both not-enrolled and dropped-out students) is linearly dependent[14] on their human capital:

$$\omega(0,\theta) = A\theta$$

where A is exogenously given. The human capital of students graduated in the university offering the standard s is:

$$\omega(s,\theta) = A\theta[1 + w(s,\theta)]$$

that represents also her/his wage rate.

---

educational choices with uncertainty in countries with open admission policies at university entry.

[12]Both the probability of graduation and the well-being during studies obviously depend on effort. The appendix (B) shows that, given a specific utility function and computing the optimal effort of students, once the optimal effort is substituted in the expected utility function, we obtain the same form of the *premium for enrolling function*

[13]Zero degree homogeneity implies

$$a) \qquad \theta p'_\theta = -sp'_s \qquad \Longrightarrow \varepsilon_{p\theta} = -\varepsilon_{ps}$$

By differenziating with respect to θ:

$$b) \qquad p'_\theta + \theta p''_{\theta\theta} = -sp''_{s\theta}$$

By differenziating with respect to s:

$$c) \qquad p'_s + sp''_{ss} = -\theta p''_{\theta s}$$

[14]Linearity does not affect the results, but strongly simplifies the analysis. For our results, the relation between the wage rate and the talent would not change if $\omega(0,\theta) = A\theta^\alpha$, for any $0 < \alpha \leq 1$.

Assumption 2 *The wage premium from graduation, $w(s,\theta) = \omega(s,\theta) - \omega(0,\theta)$, is an increasing, constant elasticity function of standard and talent, given by:*

$$w(s,\theta) = A\theta s^{\beta} \tag{2}$$

where $\beta < 1$.

Therefore, we are assuming that, coeteris paribus, there should exists a negative relationship between the graduation rate in academic institutions and the average wage earned by graduated[15].

Given the above assumptions and considering risk neutral individuals, the long-life utility associated to the non enrollment choice is given by:

$$V^N(\theta) = \frac{w(0,\theta)}{r} = \frac{A\theta}{r}$$

where r is the discount rate.

We consider an institutional setting where the cost of enrollment is equal in all the universities; consumption during studies depends only on family willingness to finance higher education and, eventually, on the moving decision because of moving costs.

Assumption 3 *The family willingness to finance education (depending on family income and wealth), the exogenous enrollment cost (tuition fees, books and so on) and the student's talent, all define consumption during studies, indicated by $z\theta$. Moving is costly and implies a lower value of z.*

Therefore, the expected intertemporal utility associated with the enrollment choice is given by:

$$V^E(s,\theta,z) = z\theta + \frac{1}{r(1+r)}\left[p(s,\theta)\omega(s,\theta) + (1-p(s,\theta))\omega(0,\theta)\right]$$

where the first addend represents utility during studies, and the second one is the ex-ante long-life utility after studies, given by the discounted sum of the wage rate of skilled and unskilled weighed by the probability of graduation.

The variation in expected utility coming from the enrolling decision is:

$$V^E(s,\theta,z) - V^N(\theta) = z\theta + \frac{1}{r(1+r)}\left[p(s,\theta)w(s,\theta) + A\theta)\right] - \frac{A\theta}{r}$$

[15] Empirical evidence of this assumption is hardly available. For the Italian higher educational setting, Bagues et al. (2008), however, provide some evidence that standard and earnings are positively correlated: ' '... the case of the Italian funds allocation system, which rewards universities according to the number of exams passed by their students. We find that university departments that rank higher according to this indicator actually tend to be signficantly worse in terms of their graduates' performance in the labour market."

We define:

$$Z\theta = r(1+r)\left[\frac{A}{1+r} - z\right]\theta > 0$$

as the difference between utility in the case of non-enrolling discounted for one period ($\frac{A\theta}{1+r}$) and the well-being during studies[16] ($z\theta$). Z proxies the *loss of income while studying*. It is a decreasing function of the family willingness to finance education, z and it is therefore higher for the poor and movers (see Assumption 3)[17]. If $Z < 0$ holds, then the student is always better off by enrolling[18] but, by assuming $Z > 0$ in the above equation, we remove this possibility.

The uni-periodal *expected premium for enrolling function*, defined as $V(s, \theta, Z) = (V^E - V^N)r(1+r)$ becomes:

$$V(s, \theta, Z) \equiv p(s, \theta)w(s, \theta) - Z\theta \tag{3}$$

where the first addend, $p(s, \theta)w(s, \theta)$ indicates the *expected human capital accumulation* during studies because it represents the expected increase in the wage rate due to graduation weighed by the graduation probability. The second addend is the loss of income while studying.

Two peculiarities of the premium for enrolling function should be highlighted: a) higher talented individuals are more willing to enroll than less talented ones if $V(s, \theta, Z)$ is an increasing function of θ; b) each secondary school graduates has a "preferred" standard if the function $V(s, \theta, Z)$ has a maximum at s.

Assume that there exists a positive $\underline{\theta}(s, Z)$ that solves[19]:

$$V(s, \underline{\theta}, Z) = 0 \qquad \Longrightarrow \underline{\theta}(s, Z) > 0 \tag{4}$$

Thereafter, $\underline{\theta}(s, Z)$ defines the talent of the *marginal enrolled* at a given standard s, whose loss of income while studying is Z.

Proposition 1 *More talented and wealthier students are more likely to enroll.*

$$\frac{\partial \underline{\theta}(s, Z)}{\partial s} > 0 \qquad \frac{\partial \underline{\theta}(s, Z)}{\partial Z} > 0$$

■ See Appendix A, Proof 1

For every value of the standard s, the level of the talent $\underline{\theta}(s, Z)$ splits the population of secondary school leavers among the enrolled ($\theta \geq \underline{\theta}(s, Z)$)

---

[16]That, given assumption 3, depends also on the decision of moving, in any.

[17]It also depends positively on the preference for the present, measured by r, if $z < \frac{A}{1+2r}$. Poor families can therefore be defined as the ones that have difficulties in financing the studies of their children (low z) and a high preference for the present (high r).

[18]Actually, some families could finance the enrollment decision with an amount of money higher than the expected earnings of the student as unskilled, so that $Z < 0$.

[19]Otherwise, secondary school leavers would either all enroll or all not-enroll.

and the not enrolled. Note that the talent required to find convenient the enrollment decision, $\underline{\theta}(s, Z)$, is a decreasing function of the family willingness to finance education. Students coming from poor families need, coeteris paribus, a higher talent to choose to enroll.

Figure 2, on the left panel, displays the curve $\underline{\theta}(s, Z)$, that is the level curve $V = V_i$ as a relationship between θ and s for a given value of Z. The same figure, on the right panel, displays the curve $\underline{\theta}(s, Z)$ for two different values of Z, say Z_P for students coming from "poor" families and Z_R for the ones coming from "rich" families. It is straightforward to note that, for every value of the standard s, the talent of the marginal enrolled is higher for the "poor". Thereafter, these level curves will be called *isopremium for enrolling functions*; their properties are analyzed in the following propositions.

Proposition 2 *Each secondary school graduate maximizes the premium for enrolling V by enrolling at the university offering his/her optimal standard $s = s^*(\theta)$ defined as:*

$$\varepsilon_{ps}(s, \theta) + \beta = 0 \qquad \Longrightarrow s^*(\theta) \tag{5}$$

where $\varepsilon_{ps}(s, \theta)$ is the elasticity of the p function to the standard s.
The optimal standard is an increasing function of student's talent.

$$\frac{ds^*(\theta)}{d\theta} > 0$$

■ See Appendix A, Proof 2

For each given talent θ there exists a level of the standard s that maximises the premium for enrolling function and the human capital accumulation, so that $p(s^*(\theta), \theta)w(s^*(\theta), \theta) \geq p(s, \theta)w(s, \theta)\, \forall s$. Students endowed with a higher talent would like to enroll in universities offering higher standards because the higher risk of dropping-out in such universities is compensated by a higher expected wage as graduated. It is worth noting that $s^*(\theta)$ does not depend on the loss of income while studying, Z.

Figure 2 shows the standard that maximizes the premium for enrollment, that is the function $s^*(\theta)$, as a positively sloped straight line (see following propositions). Provided that $V(s^*(\theta), \theta, Z) \geq 0$ each secondary school leaver would like to choose the university offering his/her optimal standard $s^*(\theta)$.

Proposition 3 *The graduation probability does not depend on talent if the student chooses his/her optimal standard.*

$$\frac{\partial p}{\partial \theta}(s^*(\theta), \theta) = 0$$

■ See Appendix A, Proof 3

Intuitively, this result relies of the fact that, by choosing the standard, students self-sort in universities in such a way that the probability of graduation does not depend on individual talent anymore. Assuming that the choice of the optimal standard $s^*(\theta)$ implies mobility, Proposition 3 validates the result illustrated in Figure 1, where the probability of dropping out was not dependent on talent for students that had chosen to enroll in a university outside their district of residence (even if in their district one university offering the same field of study existed).

Furthermore, by virtue of Eq. (41), (i.e. $\varepsilon_{s^*\theta} = 1$) the optimal standard increases linearly with talent. Therefore, in Figure 2 the function $s^*(\theta)$ is depicted as a straight line.

Proposition 4 *There exists a standard, $s_E(Z)$, such that the talent required to enroll is the lowest possible one. That is, $\underline{\underline{\theta}}(Z) = \underline{\theta}(s_E(Z)), (Z)$ is the minimum value of the function $\theta = \underline{\theta}(s, Z)$. The couple $(s_E(Z), \underline{\underline{\theta}}(Z))$ is the solution of the system:*

$$\begin{cases} V(s, \theta, Z) = 0, \\ \varepsilon_{ps}(s, \theta) + \beta = 0. \end{cases} \tag{6}$$

Both $(s_E(Z)$ and $\underline{\underline{\theta}}(Z))$ are positive constant elasticity functions of Z, and the value of the elasticity for both functions is given by $\frac{1}{\beta}$

■ See Appendix A, Proof 4

Proposition 4 and its proof clarify the slope of the isopremium curve shown in Figure 2. In fact, $s_E(Z)$ is the minimizer of the function $\underline{\theta}(s, Z)$, whose minimum value is $\underline{\underline{\theta}}(Z)$. In other words, the standard $s_E(Z)$ is the optimal standard of the marginal enrolled student (i.e. the individual having a lowest talent required to enroll):

$$\varepsilon_{ps}(s, \underline{\theta}(s, Z)) + \beta = 0 \implies s_E(Z), \qquad V(s_E(Z), \theta, Z) = 0 \implies \underline{\underline{\theta}}(Z),$$

and vice versa the talent $\underline{\underline{\theta}}(Z)$ is the lowest talent required to enroll once the student has chosen his/her optimal standard.

$$V(s^*(\theta), \theta, Z) = 0 \implies \underline{\theta}(s, Z), \qquad s^*(\underline{\theta}(s, Z)) \implies s_E(Z),$$

Figure 2 graphically displays the above results by showing, in the left graph, the values of $s_E(Z)$ and $\underline{\underline{\theta}}(Z)$ as the minimum of the isopremium for enrolling function $V = 0$. The same figure, in the right graph, shows how the value of Z affects the equilibrium by considering two different levels, $Z_P > Z_R$. A higher Z implies a higher s_E and a higher $\underline{\underline{\theta}}$ as proved by Proof 4. This means that the marginal enrolled among the poors must have a higher talent than the marginal enrolled among the rich and implies that the standard that minimizes the talent needed to enroll in a poorer population is higher than the one in a richer population:

$$\underline{\underline{\theta}}(Z_P) > \underline{\underline{\theta}}(Z_R) \qquad s_E(Z_P) > s_E(Z_R) \tag{7}$$

Proposition 5 *Human capital accumulation is incresing in the standard for all the enrolled if the standard belongs to a right neighbour of $s_E(Z)$:*

$$\frac{\partial}{\partial s}(pw)(s_E(Z),\theta) > 0, \;\; \theta > \underline{\theta}(s_E(Z), Z), \tag{8}$$

$$\frac{\partial[p(s,\theta)w(s,\theta)]}{\partial s} > 0 \quad \textit{for } s = s_E(Z) \quad \textit{and } \theta > \underline{\theta}(s_E(Z), Z) = \underline{\underline{\theta}}(Z) \tag{9}$$

where $s_E(Z)$ is defined in Proposition 4.

■ See Appendix A, Proof 5

Figure 2 summarizes the results of the previous propositions:

- for every given standard s, the enrolled are those secondary school leavers whose $\theta \geq \underline{\theta}(s, Z)$ (see Proposition 1);
- for every given standard s, a student endowed with a talent θ is more likely to enroll if the family is more willing to finance study (right graph, where $Z_P > Z_R$) (see Proposition 1);
- for every given talent θ, there exists a preferred standard that is linearly increasing in talent, the function $s^*(\theta)$ (see Proposition 2);
- students whose talent is lower than $\underline{\underline{\theta}}(Z)$ in the left graph or lower than $\underline{\underline{\theta}}(Z_P)$ or $\underline{\underline{\theta}}(Z_R)$ in the right graph will never enroll (see Proposition 4);
- the standard that minimizes the talent of the marginal enrolled is decreasing in Z, $s_E(Z_R) < s_E(Z_P)$, right graph (see comments on Proposition 4).

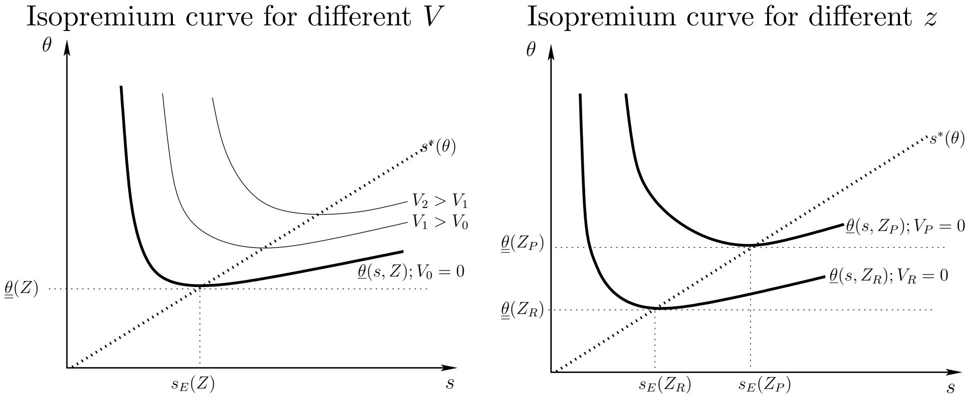
3 The higher education institutional setting

In this section we evaluate different higher educational structures and settings, where the objectives of the policy makers can be the enrollment rate, E, the graduation rate, G, and the human capital accumulation H.

In what follows, we assume that the higher educational setting can be such that:

- all universities offer the some standard s: $\Longrightarrow$ *centralized standard* case, Section 3.1;
- universities offer all the standards $\underline{s} \leq s \leq s^*(\overline{\theta})$: *decentralized standard case*, Section 3.2

Figure 2: Isopremium curve



note: the Figure is based on numerical simulations, with $p(s,\theta) = \frac{\theta}{\theta+s}$

The former case can be found in settings where the policy maker defines strictly criteria that each university must follow and, in its extreme form, can be represented as a setting where the program of each of the courses is strictly defined at national level.

The latter case defines a setting where each university freely choose the curricula of students and the difficulty of the courses and exams. In such a setting, each university could define its own goal and its own strategies[20].

3.1 The centralized standard case

We begin defining the enrollment rate, E^C, the graduation rate, G^C, and the human capital accumulation H^C in the centralized standard case.

Every given standard s defines the minimum talent, $\underline{\theta}(s,Z)$, required to enroll at this standard. The overall enrollment rate is:

$$E^C(s,Z) = \int_{\underline{\theta}(s,Z)}^{\overline{\theta}} \phi(\theta) d\theta \,, \tag{10}$$

the graduation rate[21] is:

$$G^C(s,Z) = \int_{\underline{\theta}(s,Z)}^{\overline{\theta}} p(s,\theta)\phi(\theta) d\theta, \tag{11}$$

[20]Further development of the model should analyze the optimal strategy of the universities in a context of territorial interdependence by using the tools of oligopoly and game theory, by assuming a given objective of each university.

[21]The graduation rate is the ratio between graduated and population of secondary school leavers. To compute the graduation probability in the economy, we should compute $\frac{G^C(s,Z)}{E^C(s,Z)}$.

and the human capital accumulation is:

$$H^C(s,Z) = \int_{\underline{\theta}(s,Z)}^{\overline{\theta}} p(s,\theta)w(s,\theta)\phi(\theta)d\theta\,. \tag{12}$$

Note that the functions E^C, G^C, H^C, all depend on the loss of income while studying Z that in turn depend positively on the cost of enrolling (tuition fees, books and so on) and negatively on the willingness of families in financing higher education.

Proposition 6 *Let $s_E(Z)$, $s_G(Z)$, $s_H(Z)$ be the standards that maximize the enrollment rate E^C, the graduation rate G^C and the human capital accumulation H^C respectively. We have that $s_E(Z)$ is the unique maximizer of E^C and*

$$s_G(Z) < s_E(Z), \qquad s_E(Z) < s_H(Z)\ . \tag{13}$$

■ See Appendix A, Proof 6

From Proposition 6, it emerges that policies aimed to maximize enrollment and graduation are suboptimal for human capital accumulation. In order to push the highest number of secondary school leavers to enroll, the centralized standard, $s_E(Z)$, has to be set at a level that is lower than the one that maximize human capital accumulation. Figure 3 graphically presents the results of proposition 6.

Assume now that, according the the family income, the high school graduates population can be split in two groups, the "poor", whose loss of income while studying is Z_P, and the "rich", whose loss of income is $Z_R < Z_P$.

In this setting, the overall enrollment rate is defined as:

$$E^C(s,q,Z_P,Z_R) = \quad q\int_{\underline{\theta}(\tilde{s},Z_P)}^{\overline{\theta}} \phi(\theta)d\theta + (1-q)\int_{\underline{\theta}(\tilde{s},Z_R)}^{\overline{\theta}} \phi(\theta)d\theta \tag{14}$$

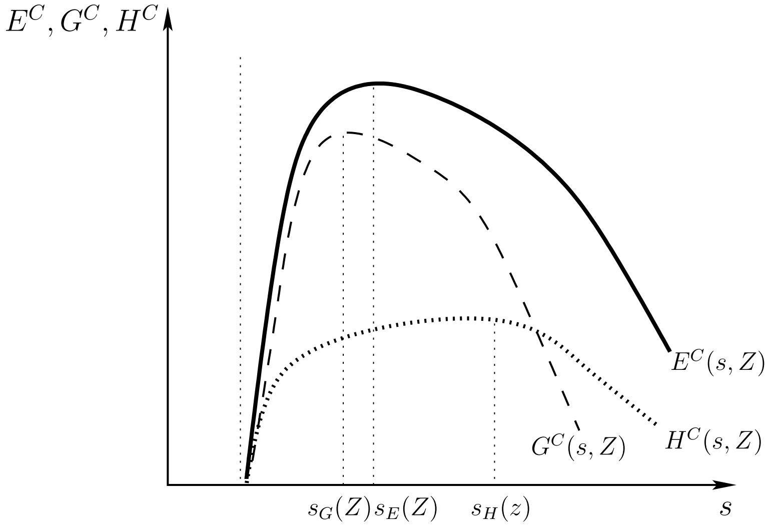
where q is the share of poor in the population and $s_E(Z_R) \leq \tilde{s} \leq s_E(S_P)$ is the standard chosen by the policy maker. The chosen $\tilde{s}$ can be the one that maximizes the overall enrollment rate, or it may depends on the preferences of the policy maker toward the enrollment rate of the poor and the rich, or may be set at the level preferred by the median voter that, in our case, can be $\tilde{s} = s_E(Z_P)$ if $q > 1/2$ or $\tilde{s} = s_E(Z_R)$ in the opposite case.

It is worthwhile to note that the parameter q, instead of representing the share of poor, can also be interpreted as the weight the policy maker assigns to the enrollment rate of the two social classes in his objective function. If, for instance, the policy maker is interested only in the enrollment rate of the poor, then $q = 1$, $\tilde{s} = s_E(Z_P)$.

Eq. (14 can also be written:

$$E^C(\tilde{s}),q,Z_P,Z_R) = \int_{\underline{\theta}(\tilde{s},Z_R)}^{\overline{\theta}} \phi(\theta)d\theta - q\int_{\underline{\theta}(\tilde{s},Z_R)}^{\underline{\theta}(\tilde{s},Z_P)} \phi(\theta)d\theta \tag{15}$$

Figure 3: Enrolment (E), graduation (G) and human capital accumulation (H) rates in the centralized standard case



The figure shows a simulation made using the following functions and parameter values: $p(s, \theta) = \frac{\theta}{\theta + s}$; $w(s\theta) = \theta s^\beta$; uniform ϕ distribution, $\theta = U[0, 2]$; $\beta = 0.2$; $Z = 0.58$

For any given enrollment rate of the rich at the standard $\tilde{s}$, the overall enrollment rate is lower: a) higher is the share of the poor, q; b) higher is inequality, $Z_P - Z_R$.

Proposition 7 *In the centralized standard case, there exists a negative relationship between income inequality (that affects the difference between Z_P and Z_R) and the enrollment rate.*

■ See Appendix A, Proof 7

It is worth noting that Proposition 7 applies also in the case where the policy maker objectives are the graduation rate or the human capital accumulation. In both cases, the homogeneity of the standard requires a unique $\tilde{s}$, that is sub-optimal both for poor and rich students.

A corollary of the previous proposition is that, in the centralized standard case where the policy maker aims to maximize the enrollment rate, or the graduation rate, or human capital accumulation and in presence of income inequality, the maximum level of the objective function is attained only if the loss of income while studying, Z, is equal for all the secondary school leavers.

Proposition 8 *If the policy maker set the homogeneous standard $\tilde{s}$, such that $s_E(Z_R) \leq \tilde{s} \leq s_E(Z_P)$ and if income inequality exists, the human capital accumulation among the poor depends negatively on income inequality* *whereas the human capital accumulation among the rich can depend positively on it.*

■ See Appendix A, Proof 8

Intuitively, the results of Proposition 8 depends on the fact that poor maximizes their human capital accumulation for a level of the standard higher than the one that maximizes enrollment, $s_H(Z_P) \geq s_E(Z_P)$, see Proposition 6, and the standard set by the policy maker is lower than $s_E(Z_P)$. Rich, instead, find the standard set by the policy maker higher than $s_E(Z_R)$ and nearer to the one which maximizes their human capital, but they enroll less.

3.2 The decentralized standard case

As previously mentioned, in this section we assume that all standards are offered by the academic institutions but that enrolling in the preferred universities generates moving costs. Therefore, each secondary school leaver chooses if: a) not enrolling; b) enrolling at the university located in his/her district of residence (at home) without incurring in moving costs, accepting the random standard offered by the local university; c) moving to the university offering his/her optimal standard, facing moving cost. In the case of moving, the loss of income while studying increases from Z to $m\,Z$ with $m \geq 1$. Therefore, a student living in a district where is located a university offering the standard s_0, should compare $V(s_0, \theta, Z)$ with $V(s^*(\theta), \theta, mZ)$ and choosing the higher among the two if positive. In order to simplify the analytics of the model, we assume that $m = 1$ for the rich and $m \to \infty$ for the poor.

Assumption 4 *Poor can not afford moving costs so that they can only choose whether enrolling at the "home" university, whose standard is randomly drawn, or not enrolling. Rich can choose if enrolling at the university offering their preferred standard or not enrolling.*

The previous assumption implies that:

- all secondary school graduates coming from rich families whose talent θ is such that $\theta > \underline{\underline{\theta}}(Z_R)$ will enroll at the university offering their optimal standard. Their graduation probability is $p(s^*(\theta), \theta)$ and their human capital accumulation is $p(s^*(\theta), \theta)w(s^*(\theta), \theta)$.
- all secondary school graduates coming from poor families whose talent θ is such that $\theta > \underline{\underline{\theta}}(s, Z_P)$ will enroll at the university offering the standard s, where s is randomly drawn from a distribution $G(s) = \int_{-\infty}^{s} g(s')ds'$, where $g(s) = 0,\ s < \underline{s}$ and $g(s) = 0,\ s > \overline{s}$[22]. Their

[22] In order to analyze deeply the decentralized case, an accurate analysis on the university management optimal behaviour is necessary.

graduation probability is $p(s,\theta)$ and their human capital accumulation is $p(s,\theta)w(s,\theta)$.

Given q the share of poor in the population of secondary school leavers, the overall enrollment rate in the decentralized setting is:

$$E(q, Z_P, Z_R) = (1-q)\int_{\underline{\theta}(Z_R)}^{\overline{\theta}} \phi(\theta)d\theta + q\int_{\underline{s}}^{\overline{s}}\int_{\underline{\theta}(s,Z_P)}^{\overline{\theta}} \phi(\theta)g(s)d\theta ds \tag{16}$$

The overall graduation rate is:

$$\begin{aligned} G(q, Z_P, Z_R) = (1-q)&\int_{\underline{\theta}(Z_R)}^{\overline{\theta}} p(s^*(\theta),\theta)\phi(\theta)d\theta \\ +q&\int_{\underline{s}}^{\overline{s}}\int_{\underline{\theta}(s,Z_P)}^{\overline{\theta}} p(s,\theta)\phi(\theta)g(s)d\theta ds \end{aligned} \tag{17}$$

The overall human capital accumulation is defined as follows:

$$\begin{aligned} H(q, Z_P, Z_R) = (1-q)&\int_{\underline{\theta}(Z_R)}^{\overline{\theta}} p(s^*(\theta),\theta)w(s^*(\theta),\theta)\phi(\theta)d\theta + \\ +q&\int_{\underline{s}}^{\overline{s}}\int_{\underline{\theta}(s,Z_P)}^{\overline{\theta}} p(s,\theta)w(s,\theta)\phi(\theta)g(s)d\theta ds \end{aligned} \tag{18}$$

Proposition 9 *Given $s_E(Z_R)$ and $s_E(Z_P)$ the unique global maximizers of the enrollment rate of the rich and poor students in the centralized standard case (i.e. $q = 0$ and $q = 1$ respectiv ely), then:*
for $q = 0$,

$$E(0, Z_P, Z_R) = E^C(s_E(Z_R), 1, Z_P, Z_R) \tag{19}$$

for $1-q$ positive and small enough we have:

$$E(q, Z_P, Z_R) < E^C(s_E(Z_P), 1, Z_P, Z_R),\ \ 0 < 1-q << 1, \tag{20}$$

and for any q, $0 < q \leq 1$ we have:

$$E(q, Z_P, Z_R) < E^C(s, q, Z_P, Z_R) + (1-q)\int_{\underline{\theta}(Z_R)}^{\underline{\theta}(s,Z_R)} \phi(\theta)d\theta + q\int_{\underline{\theta}(Z_P)}^{\underline{\theta}(s,Z_P)} \phi(\theta)d\theta \tag{21}$$

■ See Appendix A, Proof 9

Eq. (20) refers to an economic system where the share of rich is low and the policy maker considers only the poor in his objective function. In that case, the centralized case performs better than the decentralized one in term of the overall enrollment rate. From numerical simulations, we obtained that this case always holds, unless the share of poor, q, is very near to 0. In that case, however, it is unlikely that the policy maker set $\tilde{s} = s_E(Z_P)$. It is more likely that, as in the Eq. (19, the standard is near to the one desired by rich, $\tilde{s} \approx s_E(Z_R)$ and, in that case, the enrollment rates in the two systems are equal.

Eq. (21) gives the general form to compare enrollment in the decentralized and the centralized setting. Unfortunately the equations does not gives a clearcut result for intermediate values of the share of the poor in the population, q.

Proposition 10 *Given $s_H(Z_R)$ and $s_H(Z_P)$ the unique global maximizers of the human capital accumulation of the rich and poor students in the centralized standard case (i.e. $q = 0$ and $q = 1$ respectively), for q positive and small enough we have:*

$$H(q, Z_P, Z_R) \geq H^C(s_H(Z_R), Z_R) = H^C(s_H(Z_R), 0, Z_P, Z_R),\ 0 < q << 1 \tag{22}$$

while for $1 - q$ positive and small enough:

$$H(q, Z_P, Z_R) \leq H^C(s_H(Z_P), Z_P) = H^C(s_H(Z_P), 1, Z_P, Z_R),\ 0 < 1{-}q << 1, \tag{23}$$

■ See Appendix A, Proof 10

Proposition 10 states that the decentralized setting give raise to a higher human capital accumulation if the share of poor is low and the standard is set at the level desired by rich and that the decentralized system perform worst if the share of rich is low and the standard is set at the level desired by poor.

Consider a policy maker interested only in the poor. Eq. (23) implies that the preferred institutional setting is the centralized one; in the opposite case, the optimal policy is decentralization in the standard setting process, as clearly implied by Eq. (22).

Putting together Propositions 9 and 10 we can conclude that political parties representing the interests of the poor should prefer a centralized setting with a high educational standard, whereas political parties mainly interested in the rich well-being would prefer a decentralized setting.

4 Conclusions

Higher educational systems are intended to raise human capital of students and to certificate the quality of the students who obtain graduation. These two objectives go together if, as assumed in this paper, both depend on the difficulty faced by each student to obtain the degree. Difficulty depends on commitment required to pass the exams, on granting credit for previous experiences, on the workload required to students, on the whole policy of the academic institution concerning academic progress and , in short, defines the *educational standard* of the academic institution. Standard(s) are focal in the student's decision if enrolling or not and in the choice of the university to attend.

The paper analyses higher educational systems based on the free access principle where universities are (eventually) differentiated only by their standards and where each student is endowed with a given talent and comes from a family whose willingness to finance higher education is given. Students maximize their expected utility choosing if enrolling o not and self-selecting among the offered educational standards.

The theoretical model highlights that the choice of enrolling at university is optimal if the talent of the student is higher than a given threshold, depending on the family income and on the offered standard(s). Furthermore, each level of talent defines uniquely the optimal standard the students would like to choice.

Two higher educational settings are analyzed: the centralized standard one, where the policy maker sets a given standard for all universities, and the decentralized standard one, where all standards are offered from the different universities located in different districts. In that case, each student can decide not to enroll, to enroll in the university located in the district of residence whose standard is random, to enroll in the university offering his/her preferred standard by sustaining additional moving costs.

In the light of the Lisbona strategy, that emphasized the need of a "broaden access"... "and success rate"; and a "more autonomy in preparing their course", we highlighted some controversial concerns about the effectiveness of this suggestions toward two fundamental policy objectives: a) human capital accumulation b) intergenerational mobility in education.

An increase in the number of enrolled and graduated (broaden access) may be reached through many different policies, namely reduction in tuition fees, grants for poor students, or more generally raising public public expenses for the university system in order to raise its efficiency and quality. In the context of our theoretical model, "broaden access" is easily achievable by reducing standards. Nevertheless, this policy can have a cost in term of human capital accumulation because we demonstrated that the homogeneous standard that maximizes enrollment is lower than the one which maximizes human capital accumulation. The standard that maximizes the number of graduates is even lower. Therefore, the objective of increasing the number of enrolled and graduated can be detrimental to human capital accumulation.

Furthermore, from the model emerges that the standard that maximizes enrollment depends negatively on family income. Rich families would prefer a lower standard than poor ones. If the policy maker aims to maximize enrollment by setting the standard at some level among the optimal one for poor and rich, the human capital accumulation among the poor unambiguously reduces and this reduction is more prominent if income inequality is higher. Instead, income inequality can raise human capital accumulation among the rich. In some sense, the higher willingness to enroll among the rich reduces the standard above the one desired by poor, that in turn is lower than the one that maximizes their human capital accumulation. Rich, instead, can benefit from income inequality in human capital accumulation because the standard set by the policy maker is higher than the one that maximizes enrollment, so "nearer" to the one which maximizes their human capital accumulation.

The decentralized standard case probably mimics an economic system where "autonomy" of university exists. The accumulation of human capital is in that case the maximum attainable, but only if all secondary school leavers can choose the preferred standard and can sustain mobility costs. If this is not the case, a share of the secondary school leavers is constrained in the choice among not enrolling and enrolling in the (random) university located at home. So that family income influences not only the enrollment decision, but affects also the human capital accumulation during studies because mobility toward the "preferred" university allows students to accumulate more human capital. Therefore, apart from the well known differences in the enrollment rate among the rich and the poor, another limit to intergenerational mobility in education emerges in higher educational settings based on decentralized standard because the "poor", less geographically mobiles, are more likely to enroll in a university whose standard is sub-optimal and therefore accumulate a lower level of human capital.

We also demonstrated that if the share of "poor" in the population is higher than a given threshold, decentralization in the higher educational setting allows a lower human capital accumulation than centralization, if, in this last case, the standard is chosen properly.

Our results suggest that: a) the "broaden access" principle, by increasing the number of "highly educated" people, can have positive effects on economic growth. Nevertheless, if it is actually get by lowering the standard, it is detrimental to human capital accumulation. b) autonomy of individual universities in setting standards must be linked to policy aimed to completely finance mobility costs of students. Without these policy, intergenerational mobility in education and even human capital accumulation is reduced by 'decentralization.

Further development of the paper could:

- analyze the hypothesis of loans during studies to be refunded during the working life;
- improve the analysis of the "loss of income while studying" including explicitly tuition fees charged by universities;
- analyze the policy maker objective function once the cost of financing ed-

ucation is taken into account;
- analyze the strategic behavior of university in the case of decentralization;
- analyze the demand for graduates and the wage determination;
- empirically verify some of the proposition of the paper.

References

Joseph G Altonji. The Demand for and Return to Education When Education Outcomes Are Uncertain. *Journal of Labor Economics*, 11(1): 48–83, January 1993. URL `http://ideas.repec.org/a/ucp/jlabec/v11y1993i1p48-83.html`.

Manuel Bagues, Mauro Sylos Labini, and Natalia Zinovyeva. Differential Grading Standards and University Funding: Evidence from Italy. Working Papers 2008-07, FEDEA, February 2008. URL `http://ideas.repec.org/p/fda/fdaddt/2008-07.html`.

Julian R Betts. The Impact of Educational Standards on the Level and Distribution of Earnings. *American Economic Review*, 88(1): 266–75, March 1998. URL `http://ideas.repec.org/a/aea/aecrev/v88y1998i1p266-75.html`.

John H. Bishop and Ludger Wobmann. Institutional Effects in a Simple Model of Educational Production. Kiel Working Papers 1085, Kiel Institute for the World Economy, November 2001. URL `http://ideas.repec.org/p/kie/kieliw/1085.html`.

Dan A. Black and Jeffrey A. Smith. Estimating the Returns to College Quality with Multiple Proxies for Quality. *Journal of Labor Economics*, 24 (3):701–728, July 2006. URL `http://ideas.repec.org/a/ucp/jlabec/v24y2006i3p701-728.html`.

Massimiliano Bratti, Chiara Broccolini, and Stefano Staffolani. Mass Tertiary Education, Higher Education Standard and University Reform: A Theoretical Analysis. Working Papers 277, Universita' Politecnica delle Marche (I), Dipartimento di Scienze Economiche e Sociali, February 2007. URL `http://ideas.repec.org/p/anc/wpaper/277.html`.

Dominic J. Brewer, Eric R. Eide, and Ronald G. Ehrenberg. Does It Pay to Attend an Elite Private College? Cross-Cohort Evidence on the Effects of College Type on Earnings. *Journal of Human Resources*, 34(1):104–123, 1999. URL `http://ideas.repec.org/a/uwp/jhriss/v34y1999i1p104-123.html`.

Giorgio Brunello and Lorenzo Rocco. Educational Standards in Private and Public Schools. CESifo Working Paper Series 1418, CESifo Group Munich, 2005. URL `http://ideas.repec.org/p/ces/ceswps/_1418.html`.

Serena Canaan and Pierre Mouganie. Quality of higher education and earnings: Regression discontinuity evidence from the French Baccalaureate. MPRA Paper 62509, University Library of Munich, Germany, October 2014. URL `http://ideas.repec.org/p/pra/mprapa/62509.html`.

Arnaud Chevalier and Gavan Conlon. Does It Pay to Attend a Prestigious University? IZA Discussion Papers 848, Institute for the Study of Labor (IZA), August 2003. URL `http://ideas.repec.org/p/iza/izadps/dp848.html`.

Robert M Costrell. A Simple Model of Educational Standards. *American Economic Review*, 84(4):956–71, September 1994. URL `http://ideas.repec.org/a/aea/aecrev/v84y1994i4p956-71.html`.

Robert M. Costrell. Can centralized educational standards raise welfare? *Journal of Public Economics*, 65(3):271–293, September 1997. URL `http://ideas.repec.org/a/eee/pubeco/v65y1997i3p271-293.html`.

Stacy Dale and Alan B. Krueger. Estimating the Return to College Selectivity over the Career Using Administrative Earning Data. IZA Discussion Papers 5533, Institute for the Study of Labor (IZA), February 2011. URL `http://ideas.repec.org/p/iza/izadps/dp5533.html`.

Stefan Denzler and Stefan C. Wolter. Too Far to Go? Does Distance Determine Study Choices? IZA Discussion Papers 5712, Institute for the Study of Labor (IZA), May 2011.

Eleanor Dillon and Jeffrey A. Smith. The Consequences of Academic Match between Students and Colleges. IZA Discussion Papers 9080, Institute for the Study of Labor (IZA), May 2015. URL `http://ideas.repec.org/p/iza/izadps/dp9080.html`.

Brad J. Hershbein. Worker Signals among New College Graduates: The Role of Selectivity and GPA. Upjohn Working Papers and Journal Articles 13-190, W.E. Upjohn Institute for Employment Research, January 2013. URL `http://ideas.repec.org/p/upj/weupjo/13-190.html`.

Oliver Himmler and Robert Schwager. Double Standards in Educational Standards: Are Disadvantaged Students Being Graded More Leniently? ZEW Discussion Papers 07-016, ZEW - Center for European Economic Research, 2007. URL `http://ideas.repec.org/p/zbw/zewdip/5504.html`.

Lena Lindahl and Hakan Regnér. College Choice and Subsequent Earnings: Results Using Swedish Sibling Data. *Scandinavian Journal of Economics*, 107(3):437–457, 09 2005. URL `http://ideas.repec.org/a/bla/scandj/v107y2005i3p437-457.html`.

Nakamuro Makiko and Inui Tomohiko. The Returns to College Quality in Japan: Does Your College Choice Affect Your Earnings? ESRI Discussion paper series 306, Economic and Social Research Institute (ESRI), November 2013. URL `http://ideas.repec.org/p/esj/esridp/306.html`.

Joniada Milla. University Quality and Labour Market Outcomes of Canadian Youth. Working Papers 1212, University of Guelph, Department of Economics and Finance, 2012. URL http://ideas.repec.org/p/gue/guelph/2012-12.html.

Veruska Oppedisano. Open University Admission Policies and Drop Out Rates in Europe. Working Papers 200944, Geary Institute, University College Dublin, December 2009. URL http://ideas.repec.org/p/ucd/wpaper/200944.html.

M. De Paola and V. Scoppa. Returns to skills, incentives to study and optimal educational standards. *Journal of Economics*, 92(3):229–262, December 2007. URL http://ideas.repec.org/a/kap/jeczfn/v92y2007i3p229-262.html.

Robert Schwager. Majority Vote on Educational Standards. GEMF Working Papers 2013-11, GEMF - Faculdade de Economia, Universidade de Coimbra, March 2013. URL http://ideas.repec.org/p/gmf/wpaper/2013-11..html.

Gonzalo Olcina Vauteren and Luisa Escriche. Education And Family Income: Can Poor Children Signal Their Talent? Working Papers. Serie AD 2006-20, Instituto Valenciano de Investigaciones Económicas, S.A. (Ivie), October 2006. URL http://ideas.repec.org/p/ivi/wpasad/2006-20.html.

Haifeng Zhang, Yang Wu, and Xianguo Yao. Does attending a key school have a quality premium? A regression discontinuity analysis. *Applied Economics Letters*, 20(18):1630–1635, December 2013. URL http://ideas.repec.org/a/taf/apeclt/v20y2013i18p1630-1635.html.

Appendix (A): Proofs.

Very Preliminary version, to be reviewed and synthesized

Proof 1 *Let us begin noting that the function $V(s,\theta,Z)$ has the following properties for $s, Z > 0$:*

$$V(s,0,Z) = 0, \qquad V(s,\overline{\theta},Z) > 0 \tag{24}$$

We show that for any positive s and Z there exists a lowest talent, $\underline{\theta}(s,Z)$, such that the student enrolls. This talent is defined as the smallest positive zero of the function V (i.e. $V(s,\underline{\theta}(s,Z),Z) = 0$). Moreover, we prove that $V(s,\theta,Z) \leq 0$, for all $\theta \leq \underline{\theta}(s,Z)$ and that $V(s,\theta,Z) > 0$ for all $\theta > \underline{\theta}(s,Z)$. In other words, given the standard s and the loss of income while studying Z it is convenient to enroll only when $\theta \geq \underline{\theta}(s,Z)$.

In order to illustrate this, we compute the derivative of Eq. (3) with respect to θ and we rearrange it in term of elasticities (indicated by ε thereafter) as follows:

$$\frac{\partial V}{\partial \theta} = \frac{p(s,\theta)w(s,\theta)\left[\varepsilon_{p\theta}(s,\theta)+1\right] - Z\theta}{\theta}$$

where for $\theta = 0$, $p(s,0) = w(s,0) = 0$ and:

$$\frac{\partial V}{\partial \theta}(s,0,Z) = -Z < 0 \tag{25}$$

Eq. (25) can be written:

$$\frac{\partial V}{\partial \theta} = \frac{V(s,\theta,Z) + p(s,\theta)w(s,\theta)\varepsilon_{p\theta}(s,\theta)}{\theta}, \tag{26}$$

For $\theta = \underline{\theta}(s,Z)$, $V(s,\underline{\theta}(s,\theta),Z) = 0$ and, given $\varepsilon_{p\theta}(s,\theta) > 0$ (see Assumption 1):

$$\frac{\partial V}{\partial \theta}(s,\underline{\theta}(s,Z),Z) > 0 \tag{27}$$

Eqs. (24, 25) and (27) and the fact that V is a smooth function implies that a positive zero of the function V exists. Moreover, Eq. (26) and the definition of $\underline{\theta}(s,Z)$ implies that V is an increasing positive function for $\theta > \underline{\theta}(s,Z)$ and a negative function for $0 < \theta < \underline{\theta}(s,Z)$. This means that there exists a unique positive zero of V. Therefore, we can state that students with talents smaller than $\underline{\theta}$ will not enroll.

The relationship between $\underline{\theta}(s,Z)$ and Z can be derived by considering that, for the marginal individual, $V(s,\underline{\theta}(s,Z),Z) = 0$ and $p(s,\underline{\theta}(s,Z))w(s,\underline{\theta}(s,Z)) = Z\theta$. Equation 26 becomes:

$$\frac{\partial V}{\partial \theta}(s,\underline{\theta}(s,Z),Z) = Z\varepsilon_{p\theta}(s,\underline{\theta}(s,Z)) > 0$$

and, given that $\frac{\partial V}{\partial Z} = -\theta$, *we obtain:*

$$\frac{\partial \underline{\theta}(s,Z)}{\partial Z} = \frac{\underline{\theta}(s,Z)}{Z\varepsilon_{p\theta(s,\underline{\theta}(s,Z)}} > 0 \tag{28}$$

■

Proof 2 *By differentiating Eq. (3) with respect to s, we obtain*[23]*:*

$$\frac{\partial V}{\partial s} = \frac{p(s,\theta)w(s,\theta)}{s}\left(\varepsilon_{ps}(s,\theta) + \beta\right) . \tag{29}$$

For any positive θ *and* Z *the optimal level,* s^**, is the that point of* V *as a function of* s*. Hence, we have to show that there exists a zero of* $\frac{\partial V}{\partial s}$ *and that this is a maximizer of* V*. The existence of a zero of the first order partial derivative follows observing that* $\varepsilon_{ps}(0,\theta) = 0$ *(i.e.* $\frac{\partial V}{\partial s}(0,\theta,Z) > 0$*) and proving that* $\partial\varepsilon_{ps}/\partial s < 0$*, for all positive* s *(i.e.* $\frac{\partial V}{\partial s}(s,\theta,Z) < 0$ *for* $s > s^*$*). This last property guarantees also that* s^* *is a maximizer since we have:*

$$\frac{\partial^2 V}{\partial s^2} = \frac{1}{s}\frac{\partial(p(s,\theta)w(s,\theta))}{\partial s}\left(\varepsilon_{ps}(s,\theta)+\beta\right) - \frac{p(s,\theta)w(s,\theta)}{s^2}\left(\varepsilon_{ps}(s,\theta)+\beta\right)$$
$$+\frac{p(s,\theta)w(s,\theta)}{s}\frac{\partial}{\partial s}\varepsilon_{ps}(s,\theta), \tag{30}$$

so that, by virtue of Eq. (5) we have $\frac{\partial^2 V}{\partial s^2}(s^*,\theta,Z) < 0$ *if* $\partial\varepsilon_{ps}/\partial s < 0$*. Hence, the proof follows showing that the function* $\varepsilon_{ps}(s,\theta)+\beta$ *is a decreasing function of* s *or in other words that* $\partial\varepsilon_{ps}/\partial s < 0$*. Let us demonstrate that* $\frac{\partial\varepsilon_{ps}}{\partial s} < 0$*. The derivative gives* $\frac{\partial\varepsilon_{ps}}{\partial s} = p''_{ss}\frac{s}{p} + p'_s\left(\frac{p - p'_s s}{p^2}\right)$ *and can be written:*

$$\frac{\partial\varepsilon_{ps}}{\partial s} = \frac{1}{p}\left(sp''_{ss} + p'_s(1-\varepsilon_{ps})\right) . \tag{31}$$

Then, since p *is an homogeneous function of degree zero in* (s,θ)*, the equation* $p'_s = -\left(sp''_{ss} + \theta p''_{s\theta}\right)$ *holds (see equation (c) of note 13). By substituting this equation in Eq. (31) we have:*

$$\frac{\partial\varepsilon_{ps}}{\partial s} = \frac{1}{p}\left(sp''_{ss} - \left(sp''_{ss} + \theta p''_{s\theta}\right)(1-\varepsilon_{ps})\right) \Longrightarrow \frac{\partial\varepsilon_{ps}}{\partial s} = -\frac{\theta}{p}p''_{s\theta}(1-\varepsilon_{ps}). \tag{32}$$

By virtue of Assumption 1 we have $p''_{s\theta} > 0$ *and* $\varepsilon_{ps} < 0$ *so that Eq. (32) implies:*

$$\frac{\partial\varepsilon_{ps}}{\partial s} < 0 . \tag{33}$$

[23] $V(s,\theta,Z)$ is concave in s if the second order condition: $2p'_s w'_s + p''_{ss}w + w''_{ss}p < 0$ is satisfied. We assume this to avoid that the optimal standard is defined by a corner solution, $s = 0$ or $s \to \infty$.

Note that Eq. (33) implies that the function ε_{ps} is strictly decreasing. Hence, the existence of a unique solution, s^*, of Equation (5) follows since ε_{ps} is a strictly decreasing continuous function of s with $\varepsilon_{ps}(s, \theta, Z) = 0$ at $s = 0$ and $\varepsilon_{ps}(s, \theta, Z) \leq -1 < -\beta$ as $s \to -\infty$ (see Assumption 1).
As a consequence, the function V has a maximum at $s = s^*(\theta)$ (see Eqs. (33) and (30)).

Finally, keeping in mind that $\varepsilon_{ps} + \beta < 0$ when $s > s^*(\theta)$, Eq. (29) implies that $\partial V/\partial s$ is a strictly decreasing function of s for all $s > s^*(\theta)$ and, as a consequence, $s^*(\theta)$ is the unique positive maximizer of the function V.

By computing:

$$\frac{ds^*(\theta)}{d\theta} = -\frac{\frac{\partial \varepsilon_{ps}}{\partial \theta}}{\frac{\partial \varepsilon_{ps}}{\partial s}} \tag{34}$$

where $\frac{\partial \varepsilon_{ps}}{\partial s}$ is negative as shown in proof 2, so that sign $\left(\frac{\partial s^*(\theta)}{\partial \theta}\right) =$ sign $\left(\frac{\partial \varepsilon_{ps}}{\partial \theta}\right)$ and:

$$\frac{\partial \varepsilon_{ps}}{\partial \theta}(s^*(\theta), \theta) = \left(p''_{s\theta}\frac{s^*(\theta)}{p} - p'_s\frac{s^*(\theta)}{p^2}p'_\theta\right) = \frac{1}{p}\left(s^*(\theta)p''_{s\theta} + \beta p'_\theta\right) \tag{35}$$

where all the derivate in the following equation are computed for $s = s^*(\theta)$ and we substitute $\varepsilon_{ps}(s^*(\theta), \theta) = -\beta$ from Eq. (5). Note that $p'_\theta > 0$ and $p''_{s\theta} > 0$ by assumption 1. Therefore,

$$\frac{ds^*(\theta)}{d\theta} > 0 \tag{36}$$

■

Proof 3 *We should demonstrate that $\frac{dp(s^*(\theta),\theta)}{d\theta} = 0$, for any θ. Let us start computing:*

$$\frac{dp(s^*(\theta), \theta)}{d\theta} = p'_s\frac{ds^*(\theta)}{d\theta} + p'_\theta\,. \tag{37}$$

In order to compute $\frac{ds^*(\theta)}{d\theta}$ given in Eq. (34), we use Eqs. (32) and (35) for $\partial\varepsilon_{ps}/\partial s$ and $\partial\varepsilon_{ps}/\partial\theta$ respectively. Equation 32 when $s = s^*$ (i.e. $\varepsilon_{ps} = -\beta$, see Proposition 3) gives:

$$\frac{\partial \varepsilon_{ps}}{\partial s}(s^*, \theta) = \frac{1}{p}\left(sp''_{ss} + p'_s(1+\beta)\right). \tag{38}$$

Eq. (35) defines:

$$\frac{\partial \varepsilon_{ps}}{\partial \theta} = \frac{1}{p}\left(sp''_{s\theta} + \beta p'_\theta\right), \tag{39}$$

using Equation (a) and (c) of note 13 we obtain:

$$\frac{\partial \varepsilon_{ps}}{\partial \theta} = \frac{1}{p}\left(-\frac{s}{\theta}(p'_s + sp''_{ss}) - \beta\,\frac{s}{\theta}\,p'_s\right) = -\frac{1}{p}\frac{s}{\theta}\left((1+\beta)p'_s + sp''_{ss}\right), \tag{40}$$

so that by substituting Eqs. (38), (40) into Eq. (34) we have:

$$\frac{ds^*(\theta)}{d\theta} = \frac{s^*(\theta)}{\theta}\,. \tag{41}$$

Replacing Eq. (41) into (37) and using equation (a) of note 13 the thesis follows:

$$\frac{\partial p(s^*(\theta),\theta)}{\partial \theta} = p'_s \frac{s^*(\theta)}{\theta} + p'_\theta \qquad \Longrightarrow \quad \frac{dp(s^*(\theta),\theta)}{d\theta} = 0$$

■

Proof 4 *Using Eqs. (25) and (29), we can compute:*

$$\begin{aligned}\frac{\partial \underline{\theta}}{\partial s} = -\frac{\frac{\partial V}{\partial s}}{\frac{\partial V}{\partial \theta}} &= -\frac{\theta}{s}\frac{p(s,\theta)w(s,\theta)\left(\varepsilon_{ps}(s,\theta)+\beta\right)}{p(s,\theta)w(s,\theta)\left(\varepsilon_{p\theta}(s,\theta)+1\right) - Z\theta} \\ &= -\frac{\theta}{s}\;\frac{p(s,\theta)w(s,\theta)\left(\varepsilon_{ps}(s,\theta)+\beta\right)}{V(s,\theta,Z)+p(s,\theta)w(s,\theta)\varepsilon_{p\theta}(s,\theta)}.\end{aligned} \tag{42}$$

Eq. (42) provides the slope of the isopremium for enrolling functions V plotted in Figure 2.
*By computing Eq. (42) for the marginal enrolled student, the one with $\theta = \underline{\theta}(s,Z)$ (see Eq. (**??**)), we obtain:*

$$\frac{\partial \underline{\theta}}{\partial s} = -\frac{\underline{\theta}(s,Z)}{s}\;\frac{\varepsilon_{ps}(s,\underline{\theta}(s,Z))+\beta}{\varepsilon_{p\theta}(s,\underline{\theta}(s,Z))}, \tag{43}$$

so that a stationary point of the function $\underline{\theta}(s,Z)$, $s>0$ is the value $s = s_E$ which satisfies the equation:

$$\varepsilon_{ps}(s,\underline{\theta}(s,Z)) + \beta = 0\,. \tag{44}$$

Finally, we prove that the stationary point $s = s_E$ of the function $\theta = \underline{\theta}(s,Z)$ is a minimizer. In fact, by differentiating Eq. (43) with respect to s and evaluating the result at $s = s_E$ we obtain:

$$\frac{\partial^2 \underline{\theta}}{\partial s^2}(s_E,Z) = -\frac{\underline{\theta}(s_E,Z)}{s_E}\frac{1}{\varepsilon_{p\theta}(s_E,\underline{\theta}(s_E,Z)}\frac{\partial \varepsilon_{ps}}{\partial s}(s_E,\underline{\theta}(s_E,Z)) > 0, \tag{45}$$

so that s_E is a minimizer of $\underline{\theta}(s,Z)$, $s>0$, where $\underline{\underline{\theta}}(Z) = \underline{\theta}(s_E,Z)$.

Let us recall that $s_E(Z)$ e $\underline{\underline{\theta}}(Z)$ satisfy the linear system (6) for any positive value of the cost of failure Z. We can rewrite the two equation of the system as follows:

$$V(s_E(Z),\underline{\underline{\theta}}(Z),Z) = 0, \tag{46}$$

$$Q(s_E(Z),\underline{\underline{\theta}}(Z)) = 0 := \varepsilon_{ps}(s,\theta)+\beta = 0, \tag{47}$$

where Q is given by:

$$Q(s_E(Z), \underline{\underline{\theta}}(Z)) := \varepsilon_{ps}(s_E(Z), \underline{\underline{\theta}}(Z)) + \beta$$

Differentiating Eqs. (46), (47) with respect to Z we obtain that the derivatives of s_E and $\underline{\underline{\theta}}$ with respect to Z are the solution of the following linear system:

$$\begin{bmatrix} V_s(s_E(Z), \underline{\underline{\theta}}(Z), Z) & V_\theta(s_E(Z), \underline{\underline{\theta}}(Z), Z) \\ \\ Q_s(s_E(Z), \underline{\underline{\theta}}(Z)) & Q_\theta(s_E(Z), \underline{\underline{\theta}}(Z)) \end{bmatrix} \begin{bmatrix} \frac{ds_E}{dZ} \\ \\ \frac{d\underline{\underline{\theta}}}{dZ} \end{bmatrix} = \begin{bmatrix} -V_Z(s_E(Z), \underline{\underline{\theta}}(Z), Z) \\ \\ 0 \end{bmatrix} \tag{48}$$

Let us denote with $d(Z)$ the determinant of the matrix appearing in Eq (48). If $\det(Z) \neq 0$ *we have:*

$$\begin{bmatrix} \frac{ds_E}{dZ} \\ \\ \frac{d\underline{\underline{\theta}}}{dZ} \end{bmatrix} = \frac{1}{d(Z)} \begin{bmatrix} Q_\theta(s_E(Z), \underline{\underline{\theta}}(Z)) & -V_\theta(s_E(Z), \underline{\underline{\theta}}(Z), Z) \\ \\ -Q_s(s_E(Z), \underline{\underline{\theta}}(Z)) & V_s(s_E(Z), \underline{\underline{\theta}}(Z), Z) \end{bmatrix} \begin{bmatrix} -V_Z(s_E(Z), \underline{\underline{\theta}}(Z), Z) \\ \\ 0 \end{bmatrix}, \tag{49}$$

that is:

$$\frac{ds_E}{dZ} = -\frac{1}{d(z)} Q_\theta(s_E(Z), \underline{\underline{\theta}}(Z)) \; V_Z(s_E(Z), \underline{\underline{\theta}}(Z), Z) \tag{50}$$

$$\frac{d\underline{\underline{\theta}}}{dZ} = \frac{1}{d(Z)} Q_s(s_E(Z), \underline{\underline{\theta}}(Z)) \; V_Z(s_E(Z), \underline{\underline{\theta}}(Z), Z) \tag{51}$$

*Let us now compute the functions appearing in Eq. (**??**). We have:*

$$V_s = p_s w + p w_s \quad V_\theta = p_\theta w + p w_\theta - Z \quad V_Z = -\theta \tag{52}$$

$$Q_s = \frac{\partial}{\partial s} \epsilon_{ps} \quad Q_\theta = \frac{\partial}{\partial \theta} \epsilon_{ps} \tag{53}$$

Since p is an homogeneous function of degree zero (specifically, $p_\theta = -s p_s/\theta$ ad $\frac{\partial}{\partial s}\epsilon_{ps} = -\frac{\theta}{s}\frac{\partial}{\partial \theta}\epsilon_{ps}$) and the fact that $w_s = \beta w/s$, $w_\theta = w/\theta$, $\epsilon_{ps} = -\beta$, $Z/(pw) = 1/\underline{\underline{\theta}}$ we can rewrite the determinant $d(Z)$ as follows:

$$\begin{aligned} d(Z) &= (p_s w + p w_s)\frac{\partial}{\partial \theta}\epsilon_{ps} - (p_\theta w + p w_\theta - Z)\frac{\partial}{\partial s}\epsilon_{ps} = \\ &\frac{\partial}{\partial \theta}\epsilon_{ps}\left(p_s w + p\beta\frac{w}{s_E}\right) + \frac{pw}{s_E}\left(\frac{p_\theta \underline{\underline{\theta}}}{p} + 1 - \frac{Z\underline{\underline{\theta}}}{pw}\right) \\ &= \frac{\partial}{\partial \theta}\epsilon_{ps}\frac{pw}{s_E}\left(-\frac{p_s s_E}{p}\right) \end{aligned} \tag{54}$$

that is

$$d(Z) = \beta \, \frac{Z \, \underline{\underline{\theta}}}{s_E} \, \frac{\partial}{\partial \theta}\epsilon_{ps} \tag{55}$$

It is easy to see that the determinant $d(Z)$ is positive by virtue of Assumption 1 and Eq. (63). Using (55) in Eqs. (50) and (51) we obtain:

$$\frac{ds_E}{dZ} = \frac{s_E}{\frac{\partial}{\partial\theta}\epsilon_{ps}\,\beta\,Z\,\underline{\underline{\theta}}}\left(\frac{\partial}{\partial\theta}\varepsilon_{ps}\right)\underline{\underline{\theta}} = \frac{1}{\beta}\frac{s_E}{Z} \tag{56}$$

and

$$\begin{aligned}\frac{d\underline{\underline{\theta}}}{dZ} &= \frac{s_E}{\frac{\partial}{\partial\theta}\epsilon_{ps}\,\beta\,Z\,\underline{\underline{\theta}}}\left(\frac{\partial}{\partial s}\varepsilon_{ps}\right)(-\underline{\underline{\theta}}) = \\ &\frac{s_E}{\frac{\partial}{\partial\theta}\epsilon_{ps}\,\beta\,Z\,\underline{\underline{\theta}}}\left(-\frac{\underline{\underline{\theta}}}{s_E}\frac{\partial}{\partial\theta}\varepsilon_{ps}\right)\underline{\underline{\theta}} = \frac{1}{\beta}\frac{\underline{\underline{\theta}}}{Z}\end{aligned} \tag{57}$$

Eqs. (56) and (57) can be rewritten as follows:

$$\begin{aligned}\epsilon_{\underline{\underline{\theta}}Z} &= \frac{Z}{\underline{\underline{\theta}}}\frac{d\underline{\underline{\theta}}}{dZ} = \frac{1}{\beta} \\ \epsilon_{s_E\,Z} &= \frac{Z}{s_E(Z)}\frac{ds_E}{dZ} = \frac{1}{\beta}\end{aligned} \tag{58}$$

This concludes the proof.

■

Proof 5 *First we recall the explicit expression of the derivative of pw with respect to s:*

$$\frac{\partial(pw)}{\partial s}(s,\theta) = \frac{pw}{s}\left(\varepsilon_{ps}(s,\theta)+\beta\right), \tag{59}$$

and then we compute the derivative of $\frac{\partial(pw)}{\partial s}$ with respect to θ in order to show that this derivative is positive when $s = s_E$ and $\theta > \underline{\theta}(s_E, Z)$. We have:

$$\begin{aligned}\frac{\partial^2(pw)}{\partial s\partial\theta}(s,\theta) &= \frac{\partial}{\partial\theta}\left[\frac{pw}{s}\left(\varepsilon_{ps}(s,\theta)+\beta\right)\right] \\ &= \frac{p'_\theta\,w + \frac{p}{\theta}\,w}{s}\left(\varepsilon_{ps}(s,\theta)+\beta\right) + \frac{pw}{s\,\theta}\,\frac{\theta\,sp''_{s\theta}p - sp'_s\,\theta\,p'_\theta}{p^2} \\ &= \frac{pw}{s\,\theta}\left[\left(\frac{\theta\,p'_\theta}{p}+1\right)\left(\varepsilon_{ps}(s,\theta)+\beta\right) + \left(\frac{\theta\,sp''_{s\theta}p - sp'_s\theta\,p'_\theta}{p^2}\right)\right]\end{aligned} \tag{60}$$

A sufficient condition for the function $\frac{\partial^2(p\,w)}{\partial s\partial\theta}(s,\theta)$ being positive is that the function $(\varepsilon_{ps}(s,\theta)+\beta)$ is positive. We highlight that the following two equations holds:

$$\frac{\partial^2(pw)}{\partial s\partial\theta}(s,\theta) = \frac{pw}{s}\,\frac{p''_{s\theta}p - p'_sp'_\theta}{p^2} > 0, \quad (s,\theta) = (s_E, \underline{\theta}(s_E,Z)), \tag{61}$$

and

$$\varepsilon_{ps}(s_E,\theta)+\beta\ >\ 0, \quad \theta>\underline{\theta}(s_E,Z). \tag{62}$$

Eq. (62) holds true since the elasticity ε_{ps} is an increasing function of θ. In fact, the derivative of ε_{ps} with respect to θ given by:

$$\frac{\partial \varepsilon_{ps}}{\partial\theta}(s,\theta)=s\frac{(p''_{s\theta}p-p'_sp'_\theta)}{p^2}\ >\ 0, \tag{63}$$

is positive for any s and θ by virtue of Assumption 1. It is easy to see that Eqs. (60), (61), (62) imply Eq. (8).

■

Proof 6 *We look for a stationary point of E^C. Deriving Eq. (10) we have:*

$$\frac{\partial E^C}{\partial s}(s,Z)=-\phi(\underline{\theta}(s,Z))\frac{\partial}{\partial s}\underline{\theta}(s,Z), \tag{64}$$

and

$$\frac{\partial^2 E^C}{\partial s^2}(s,Z)=-\phi'(\underline{\theta}(s,Z))\left(\frac{\partial}{\partial s}\underline{\theta}(s,Z)\right)^2-\phi(\underline{\theta}(s,Z))\frac{\partial^2}{\partial s^2}\underline{\theta}(s,Z), \tag{65}$$

so that a stationary point is the point $s_E(Z)$ such that

$$\frac{\partial}{\partial s}\underline{\theta}(s_E(Z),Z)=0, \tag{66}$$

that is $s_E(Z)$ is the value of the standard which is the solution of the following equation:

$$\varepsilon_{ps}(s_E(Z),\underline{\theta}(s_E(Z),Z))=-\beta. \tag{67}$$

Using Eq. (67) into Eq. (65) and Eq. (45) we obtain that $s_E(Z)$ is a maximizer of E^C. By virtue of Eq. (43) and Eq. (10) we can conclude that E^C is a decreasing function for $s>s_E(Z)$ and $\theta>\underline{\theta}(s,Z)$ hence $s_E(Z)$ is the unique maximizer.

Now we consider the graduation rate (11). The standard $s=s_G(Z)$ that maximizes the graduation rate is a stationary point of G^C. The partial derivative of G^C with respect to s is given by:

$$\frac{\partial G^C}{\partial s}(s,Z)=-p(s,\underline{\theta}(s,Z))\phi(\underline{\theta}(s,Z))\frac{\partial}{\partial s}\underline{\theta}(s,Z)+\int_{\underline{\theta}(s,Z)}^{\overline{\theta}}p'_s(s,\theta)\phi(\theta)d\theta\,. \tag{68}$$

Computing Eq. (68) at $s=s_E(Z)$ and Assumption 1 ($p'_s(s,\theta)<0,\ \forall\, s\,\theta$) we have:

$$\frac{\partial G^C}{\partial s}(s_E(Z),Z)=+\int_{\underline{\theta}(s_E(Z),Z)}^{\overline{\theta}}p'_s(s_E(Z),\theta)\phi(\theta)d\theta<0, \tag{69}$$

where $p'_s < 0$ by assumption 1. Therefore, $s_G(Z) \leq s_E(Z)$ since G^C is decreasing at $s = s_E(Z)$.

Let us know consider the human capital accumulation Eq. (12). The partial derivative of H^C with respect to s is given by:

$$\frac{\partial H^C}{\partial s}(s,Z) = -p(s,\underline{\theta}(s,Z))As^{\beta}\underline{\theta}(s,Z)\phi(\underline{\theta}(s,Z))\frac{\partial}{\partial s}\underline{\theta}(s,Z) + \int_{\underline{\theta}(s,Z)}^{\overline{\theta}} \frac{\partial}{\partial s}[p(s,\theta)w(s,\theta)]\phi(\theta)d\theta \,. \quad (70)$$

Evaluating (70) at $s = s_E(Z)$ and using Eqs. (66) and (8) in Proposition 5 we obtain:

$$\frac{\partial H^C}{\partial s}(s_E(Z),Z) = \int_{\underline{\theta}(s_E(Z),Z)}^{\overline{\theta}} \frac{\partial\,(pw)}{\partial s}(s_E(Z),\theta)\phi(\theta)d\theta > 0 \,. \quad (71)$$

This concludes the proof.

■

Proof 7 *The overall enrollment rate depends on the chosen standard, $s_E(Z_P) \leq \tilde{s}) \leq s_E(Z_R)$ fixed by the policy maker (see Eq. (67). Given that the enrolment rate of rich is maximized for $s_E(Z_R)$ and the enrollment rate of poor for $s_E(Z_P)$ (see Proposition 4)), both enrollment rates must be sub-optimal. This sub-optimality increases with income inequality (see Eq. (15))*

■

Proof 8 *in order to maximize enrollment, the policy maker sets $\tilde{s}$ such that:*

$$s_E(Z_R) \leq \tilde{s} \leq s_E(Z_P)$$

Proposition 6 states that the standard that maximizes human capital accumulation is higher than the standard that maximizes enrollment, so that:

$$s_H(Z_R) > s_E(Z_R) \qquad s_H(Z_P) > s_E(Z_P)$$

Therefore:

$$s_H(Z_R) \lesseqgtr \tilde{s} \qquad s_H(Z_P) \leq \tilde{s}$$

That is, the standard $\tilde{s}$ set by the policy maker is surely lower that the one, $s_H(Z_P)$, which maximizes the human capital accumulation of the poor.

Furthermore, the larger the difference between $Z_P - Z_R$ is, (i.e the higher the income inequality), the lower the human capital accumulation of the poor students is. On the contrary, an increase in income inequality may both decrease or increase the human capital accumulation of the rich students.

$\blacksquare$

Proof 9 *Let us start setting $q = 1$ in Eq. (16) we have:*

$$E(1, Z_P, Z_R) < \int_{\underline{s}}^{\overline{s}} ds\, g(s) \max_{\underline{s} \leq s \leq \overline{s}} \int_{\underline{\theta}(s,Z_P)}^{\overline{\theta}} \phi(\theta) d\theta\,, \tag{72}$$

which implies (see the proof of Proposition 6):

$$E(1, Z_P, Z_R) < \int_{\underline{\theta}(Z_P)}^{\overline{\theta}} \phi(\theta) d\theta \int_{s_E(Z_P)}^{\overline{s}} ds\, g(s)\,.$$

Since g is probability density function we have:

$$\int_{\underline{s}}^{\overline{s}} ds\, g(s) \leq 1 \tag{73}$$

so that we obtain:

$$E(1, Z_P, Z_R) < \int_{\underline{\theta}(Z_P)}^{\overline{\theta}} \phi(\theta) d\theta = E^C(s_E(Z_P), 1, Z_P, Z_R).$$

By continuity we have that for positive and small enough values of $1 - q$ we have:

$$E(q, Z_P, Z_R) < E^C(s_E(Z_P), 1, Z_P, Z_R), \;\; 1 - q \to 0. \tag{74}$$

Let us consider q, $0 < q \leq 1$. We have:

$$E(q, Z_P, Z_R) < (1 - q) \int_{\underline{\theta}(Z_R)}^{\overline{\theta}} \phi(\theta) d\theta \tag{75}$$

$$+q \int_{\underline{s}}^{\overline{s}} ds\, g(s) \max_{\underline{s} \leq s \leq \overline{s}} \int_{\underline{\theta}(s,Z_P)}^{\overline{\theta}} \phi(\theta) d\theta$$

From Proposition 6 we have:

$$E(q, Z_P, Z_R) < (1 - q) \int_{\underline{\theta}(Z_R)}^{\overline{\theta}} \phi(\theta) d\theta + q \int_{\underline{\theta}(Z_P)}^{\overline{\theta}} \phi(\theta) d\theta \int_{\underline{s}}^{\overline{s}} ds\, g(s)\,, \tag{76}$$

so that from Eq. (76) and Eq. (73) we obtain:

$$E(q, Z_P, Z_R) < (1 - q) \int_{\underline{\theta}(Z_R)}^{\overline{\theta}} \phi(\theta) d\theta + q \int_{\underline{\theta}(Z_P)}^{\overline{\theta}} \phi(\theta) d\theta \tag{77}$$

We rewrite Eq. (77) as follows:

$$E(q, Z_P, Z_R) < (1-q)\int_{\underline{\underline{\theta}}(Z_R)}^{\overline{\theta}} \phi(\theta)d\theta + q\int_{\underline{\underline{\theta}}(Z_P)}^{\overline{\theta}} \phi(\theta)d\theta \pm E^C(s, q, Z_P, Z_R) \tag{78}$$

that is Eq. (21) follows:

$$E(q, Z_P, Z_R) < E^C(s, q, Z_P, Z_R) + (1-q)\int_{\underline{\underline{\theta}}(Z_R)}^{\underline{\theta}(s,Z_R)} \phi(\theta)d\theta + q\int_{\underline{\underline{\theta}}(Z_P)}^{\underline{\theta}(s,Z_P)} \phi(\theta)d\theta \tag{79}$$

Note that the "best standard" s when a uniform standard is preferred is the standard s which minimizes the quantity appearing in the right hand side of Eq. (21) given by:

$$\min_{\underline{s}\leq s\leq \overline{s}} \left[(1-q)\int_{\underline{\underline{\theta}}(Z_R)}^{\underline{\theta}(s,Z_R)} \phi(\theta)d\theta + q\int_{\underline{\underline{\theta}}(Z_P)}^{\underline{\theta}(s,Z_P)} \phi(\theta)d\theta\right] \tag{80}$$

and this value is the value that maximizes the enrollment rate E^C. This concludes the proof.■

Proof 10 *Let us recall the expression of the human accumulation:*

$$H(q, Z_R, Z_P) = (1-q)\int_{\underline{\underline{\theta}}(Z_R)}^{\overline{\theta}} p(s^*(\theta), \theta)w(s^*(\theta), \theta)\phi(\theta)d\theta \tag{81}$$
$$+q\int_{\underline{s}}^{\overline{s}}\int_{\underline{\theta}(s,Z_P)}^{\overline{\theta}} \phi(\theta)p(s, \theta)w(s, \theta)g(s)d\theta ds\,.$$

Letting $q = 0$ since $p(s^*(\theta), \theta)w(s^*(\theta), \theta) > p(s, \theta)w(s, \theta)$ for each $s \geq s_E(Z_R)$ and $\theta \geq \underline{\underline{\theta}}(Z_R)$ we obtain:

$$H(0, Z_R, Z_P) = \int_{\underline{\underline{\theta}}(Z_R)}^{\overline{\theta}} p(s^*(\theta), \theta)w(s^*(\theta), \theta)\phi(\theta)d\theta >$$
$$\int_{\underline{\theta}(s,Z_R)}^{\overline{\theta}} p(s, \theta)w(s, \theta)\phi(\theta)d\theta = H^C(s, Z_R) = H^C(s, 1, Z_P, Z_R)$$
$$\forall s \geq s_E(Z_R). \tag{82}$$

By continuity Eqs. (22) holds for positive and sufficiently small values of q.

Let us consider the case $1-q$ positive and small enough. Remember that the integrand functions appearing in the human accumulation are positive and that the following inequality holds:

$$\int_{\underline{s}}^{\overline{s}} g(s)ds \leq 1$$

Letting $q = 1$ in Eq. (82) we have:

$$H(1, Z_R, Z_P) = \int_{s_E(Z_P)}^{\overline{s}} ds\, g(s) \int_{\underline{\theta}(s,Z_P)}^{\overline{\theta}} p(s,\theta)w(s,\theta)\phi(\theta)d\theta <$$
$$\int_{\underline{s}}^{\overline{s}} ds\, g(s) \max_{\underline{s}\leq s\leq \overline{s}} \int_{\underline{\theta}(s,Z_P)}^{\overline{\theta}} p(s,\theta)w(s,\theta)\phi(\theta)d\theta \leq$$
$$\int_{\underline{\theta}(s_H(Z_P),Z_P)}^{\overline{\theta}} p(s_H(Z_P),\theta)w(s_H(Z_P),\theta)\phi(\theta)d\theta \left(\int_{\underline{s}}^{\overline{s}} ds\, g(s)\right) \leq H^C(s_H(Z_P), Z_P), \tag{83}$$

that is

$$H(1, Z_R, Z_P) \leq H^C(s_H(Z_P), Z_P) = H^C(s_H(Z_P), 0, Z_P, Z_R)\ . \tag{84}$$

By continuity we conclude the proof. ■

Appendix (B): effort during studies as an endogenous variable

Assume that effort during studies, e , with $0 \leq e \leq 1$ affects negatively the utility in the study period and positively the probability of graduation. Assume that the utility during studies is given by: $z\theta - \frac{e^2}{2}$, that the probability of graduation is $e\sqrt{p(s,\theta)}$ and the premium for enrollment wage is $\sqrt{2w(s,\theta)}$. Note that the two last definitions does not change anything in the model because they simply modify the assumptions 1 and 2.

Equation 3 becomes:

$$V = e\sqrt{2p(s,\theta)w(s,\theta)} - Z\theta - \frac{e^2}{2} \tag{85}$$

Each enrolled student choose optimally her effort during studies by maximizing Eq. (85), so that:

$$e^*(s,\theta) = \sqrt{2p(s,\theta)w(s,\theta)}$$

Therefore, effort depends positively on human capital accumulation during studies.

By substituting $e^*(s,\theta($in equation 85, we obtain:

$$V = p(s,\theta)w(s,\theta) - Z\theta$$

that is the equation 3, representing the premium for enrolling function. This result obviously requires a specific form of the function defining the disutility of effort and a specific relationship between the probability of graduation and effort. For these reasons, we preferred to not explicitly model effort in the paper.

Appendix (C): Simulating the model

This section define the endogenous variables of the model using the following specific functional form of the equation $p(h, s)$.

$$p(s, \theta) = \frac{\theta}{s + \theta} \tag{86}$$

that respects all the conditions of assumption 1. Form propostition 2:

$$w(s, \theta) = A\theta s^{\beta}$$

therefore, from equation 3, we obtain

$$V(s, \theta) = A\frac{\theta^2}{s + \theta}s^{\beta} - Z\theta \tag{87}$$

Proposition 1
The marginal enrolled satisfies $V(s, \theta, Z) = 0$ so that:

$$\underline{\theta}(s, Z) = \frac{sZ}{As^{\beta} - Z} \qquad \underline{\theta}(s, Z) > 0 \iff As^{\beta} > Z$$

Furthermore:

$$\frac{\partial \underline{\theta}}{\partial Z} = s\frac{As^{\beta} - Z + sZ}{(As^{\beta} - Z)^2} > 0$$

Propositions 2,3
From equation 87, we obtain:

$$\frac{dV}{ds} = A\left(\frac{\theta}{\theta + s}\right)^2 \left(\beta\frac{\theta + s}{s} - 1\right)$$

so that:

$$\frac{dV}{ds} = 0 \implies -\frac{s}{\theta + s} + \beta = 0$$

therefore:

$$s^*(\theta) = \frac{\beta}{1 - \beta}\theta$$

The same result can be obtained from equation 5, considering that:

$$\varepsilon_{ps} = -\frac{s}{\theta + s}$$

so that:

$$\frac{s}{\theta + s} = \beta \qquad \implies s^*(\theta) = \frac{\beta}{1 - \beta}\theta \tag{88}$$

$s^*(\theta)$ is an increasing function of θ.

By substituting $s^*(\theta)$ in equation 86, we obtain:

$$p(s^*(\theta), \theta) = 1 - \beta$$

so that the probability of graduation does not depend on talent.

Proposition 4
Substituing $s^*(\theta)$ from Eq. (88) into Eq. (87), we obtain:

$$V = \frac{\theta}{\frac{\beta}{1-\beta}\theta + \theta} A\theta \left(\frac{\beta}{1-\beta}\theta\right)^{\beta} - Z\theta \qquad \rightarrow V = \beta^{\beta}(1-\beta)^{1-\beta} A\theta^{\beta+1} - Z\theta$$

Solving for θ:

$$\underline{\underline{\theta}}(Z) = \left(\frac{Z}{\beta^{\beta}(1-\beta)^{1-\beta}}\right)^{\frac{1}{\beta}} \tag{89}$$

Bu sustituting the previous equation in Eq. (**??**):

$$s_E(Z) = \left(\frac{Z}{1-\beta}\right)^{\frac{1}{\beta}} \tag{90}$$

Proposition 5 Computing Eq. (8), with p given in Eq. (86) and w in Eq. (2), we obtain:

$$\frac{\partial pw}{\partial s} = A \left(\frac{\theta}{\theta + s}\right)^2 s^{\beta} \left(\beta\frac{\theta}{s} - (1-\beta)\right)$$

that, computed for $s = s_E(Z)$ as defined in Eq. (90), gives

$$\frac{\partial pw}{\partial s} = A \left(\frac{\theta}{\theta + s_E(Z)}\right)^2 s_(Z)^{\beta} \left(\beta(1-\beta)^{\frac{1}{\beta}} \frac{\theta}{Z^{\frac{1}{\beta}}} - (1-\beta)\right)$$

So that the sign of $\frac{\partial}{\partial s}(pw(s_E(Z))$ depends on:

$$\beta(1-\beta)^{\frac{1}{\beta}-1} \frac{\theta}{Z^{\frac{1}{\beta}}} - 1$$

that depends positively on θ. Computing it for $\theta = \underline{\underline{\theta}}(Z)$ as defined in Eq. (89), we obtain:

$$\beta(1-\beta)^{\frac{1}{\beta}-1} \frac{\left(\frac{Z}{\beta^{\beta}(1-\beta)^{1-\beta}}\right)^{\frac{1}{\beta}}}{Z^{\frac{1}{\beta}}} - 1 = 0$$

Therefore, $\forall \theta > \underline{\underline{\theta}}(Z)$, $\frac{\partial}{\partial s}(pw(s_E(Z)) > 0$