Working papers series

WP ECON 07.14

On the optimal allocation of students when peer effect works: Tracking vs Mixing

Marisa Hidalgo-Hidalgo (U. Pablo de Olavide)

JEL Codes: D63, I28, J24
Keywords: Peer Effects, Tracking, Mixing.

Department of Economics

On the optimal allocation of students when peer effect works: Tracking vs Mixing*

Marisa Hidalgo-Hidalgo†
University Pablo de Olavide

July 2, 2007

Abstract

The belief that the behaviour and outcomes of compulsory school students are affected by their peers has been important in shaping education policy. I analyze two polar education systems -tracking and mixing- and propose several criteria for their comparison. The system that maximizes average human capital, I find, depends crucially on the level of complementarity between peer effects and individuals' ability. I also find that when mean innate ability is much higher among the rich than among the poor, the system that best maximizes average human capital is mixing. However, there is no unanimity in the overall population so as to which system to choose.

Keywords: Peer Effects, Tracking, Mixing

JEL Classification: D63, I28, J24.

*Financial support from the Instituto Valenciano de Investigaciones Económicas and European Social Fund is gratefully acknowledged. **Address for correspondence:** Marisa Hidalgo, Departamento de Economía (Área Análisis Económico), Universidad Pablo de Olavide, Ctra. Utrera, Km.1, E-41013, Sevilla, Spain. E-mail: mhidalgo@upo.es

†This paper is based on Chapter two of my PhD dissertation, completed at Universidad de Alicante under the very helpful supervision of Iñigo Iturbe-Ormaetxe. I am grateful for the comments and suggestions of Gianni De Fraja, Elena Del Rey, Moshe Justman, Francois Maniquet, Francisco Martínez-Mora and Ignacio Ortuño-Ortin. I would like to thank CORE members for their hospitality and Buly Cardak, Hessel Oosterbeek, Laura Romero and Antonio Villar for very fruitful conversations.

1 Introduction

Interest in social interactions, neighborhood effects, and social dynamics has recently undergone a revival. As a consequence of this new trend, a small body of literature has emerged that studies the way neighborhood effects can generate and perpetuate persistent inequality. One of this neighborhood effects is the so-called "peer effect", defined here as the effect on an individual's academic performance of the ability distribution of her peers.[1] The critical importance to both parents and policy makers of peer group distribution in school is indisputable, since peer effects have played an important role in a number of policy debates, including that surrounding the controversial subjects of ability tracking and school desegregation.

Given the existence of peer effects, it is not surprising that governments should want to keep them in mind when planning how best to meet their educational policy objectives. One situation in which peer effects must be carefully considered is when governments must choose whether to stream (track) or mix students of differing abilities within the public school classroom. This paper contributes to this debate by addressing three main questions. First, it asks which system best maximizes average human capital at the compulsory school level. Second, it explores whether the overall population can be said to prefer one of the aforementioned systems-tracking and mixing- over the other. Finally, it considers how the existence of a positive dependence between parental background and individual ability affects the two previous issues.

The practice of grouping students on the basis of ability (tracking), while common in the USA and Europe, remains a controversial one.[2] The main argument in favor of this system is that, by grouping together students of similar abilities levels, teachers can target instruction to a level more closely aligned with other students' needs than

[1] Roemer and Wets (1994) and Streufert (2000) show how economic segregation can lead to inaccurate assessments of the economic payoff to education. The idea is that by depriving children in poor neighborhoods of successful role models (which is an inevitable consequence of economic segregation), they make inferences on the benefits of education that are biased downward.

[2] For the US case, public school teachers reported that only 14.4% and 10.8% of tenth-grade students were in heterogeneous (untracked) math classes in 1988 and 1990 respectively, see Rees et al. (1996).

would be possible in more heterogenous environments. On the other hand critics argue that such segregation deprives disadvantaged students of any positive peer influences that they might have gained by interacting with their more able peers. In keeping with this view, there has been considerable movement in the US towards eliminating ability grouping in public schools.[3]

While the influence of peer ability on one's own educational achievement is well documented, the fine points of this relationship are still being debated. Most studies focus on average innate ability within the classroom as the peer-based factor that most strongly impacts on individual achievement.[4] On the one hand, for example, Evans, Oates and Schwab (1992) find a significant peer group effect that vanishes when they control for endogeneity. On the other hand, Henderson, Mieszkowski, and Sauvageau (1978), Summers and Wolfe (1977) and more recently Robertson and Symons (1996), Hoxby (2000), and Zimmer and Toma (2000) report significant positive influences of higher achieving peers on achievement.

The existence of peer effects and their links with different grouping policies have been studied at the theoretical as well as the empirical level. Most of the literature on the subject tends to analyze the practical effects of grouping students by ability, arriving at the conclusion that ability grouping- unlike mixing - almost inevitably hurts low-track students while favoring high-track ones. However, there is no clear evidence as to whether the losses of the former are offset by the gains of the latter (see, for example, Argys et al.(1996), Betts and Shkolnik (2000), Figlio and Page (2000) and most recently Kang (2007)). Theoretical contributions are fewer in number. Among others we find the works by de Bartolome (1990), Epple, Newlon and Romano (2002) and Arnott and Rowse (1987).

This paper is closely related to that of Arnott and Rowse (1987), who analyze the optimal allocation of students and resources when peer effects are present by focusing on the degree of concavity of peer group effect. They conclude that, when the objective is to maximize mean performance, optimal allocation of students abilities depends on the properties of the education production function. However, they fail to

[3]For example, data from the Schools and Staffing Survey suggest than 20% of schools with programs for gifted children in 1990 had eliminated the programs by 1993 (Figlio and Page (2000)).

[4]See Manski (1993) for details on the difficulties in identifying empirically peer effects.

consider the role of family background in the process of human capital accumulation, despite its importance at the compulsory school level (see Heckman (2000)). They also focus on the role played by the degree of concavity of the peer group effect on the key goal of maximizing mean achievement, even though they admit that this narrow focus prevents them from seeing the possible dependence of individual welfare on the whole shape of human capital distribution in the population.

My approach differs from that of Arnott and Rowse (1987) in two key respects. First, in addition to considering how family background affects student achievement, I also acknowledge the existence of a positive dependence between family background and innate ability, and its effect on the optimal allocation of students. Second, and in light of the most recent empirical evidence, I assume concavity in peer effects and discuss how the complementarity between peer characteristics and individual ability (a point for which the empirical evidence is still quite mixed) can determine which system best maximizes average human capital.[5] In addition, my paper contributes to the relevant literature by comparing both systems in terms of the induced distributions of human capital at the end of compulsory school.

I find that, in societies where the mean ability of rich students greatly exceeds that of their poorer classmates, average human capital is maximized by mixing in most of cases. This is true regardless of the degree to which peer effects and innate ability can be seen as complementary. We find that, under mixing, those societies where mean innate ability is (much) higher among the rich than among the poor will also have (much) higher average human capital. However, this might not be the case if the prevailing education system is tracking. The intuition is that, as the difference between mean ability of rich students versus poor students increases, so too does the threshold separating students in the low ability group from those in the high ability group. As this difference increases, therefore, the average income level of the students remaining in the low ability group tends to diminish while that of the students in the high ability group tends to rise. As a result, average human capital will not always be much higher in those societies where the difference between mean ability of the two

[5] Henderson et al.(1978), Summer and Wolfe (1977) and more recently Zimmer and Toma (2000) and Zimmerman (2003) find a concave relation between average quality of one's peers and own outcome.

income groups is much higher. I also find that the system that best maximizes average human capital at compulsory level depends on the level of complementarity between the peer effect and individuals' innate ability. In particular, when peer effects matter more for low (high) ability students than for high (low) ability students, average human capital is maximized under mixing (tracking), which is the system where low (high) ability students enjoy a stronger peer effect.

Finally my study suggest that, among risk averse individuals the preference for mixing versus tracking depends on the degree of complementarity between the peer effect and individuals' innate ability. If they are nearly complementary, then there is no preferred system in the population. However, in some empirically relevant cases in which the two variables act as substitutes, I find that, it is mixing the system unanimously preferred in the population.

The rest of the paper is organized as follows. Section 2 describes the model and the main features of human capital distribution under the two education systems at compulsory school level. Section 3 compares the induce distributions of human capital in these two systems. Section 4 concludes.

2 Model

2.1 Individuals

Population size is constant at 1. Individuals differ in two aspects: their innate ability, θ_0, and their family background, denoted by z (therefore z could be either the parental income level or the parents' human capital). To make the model tractable, I assume that family background z takes only two values, 1 and $x > 1$ with probabilities $1 - \lambda$ and λ, respectively. I refer to those with income 1 (x) as the poor (rich). To capture the possibility that some level of positive dependence exists between income and innate ability, I assume that innate ability is uniformly distributed on $[0, 1]$ for the poor and on $[0, k]$ for the rich where $k > 1$. Thus, the C.D.F. (cumulative distribution function) of innate ability, denoted by $F(\theta_0)$, can be expressed as:

$$F(\theta_0) = \begin{cases} \left(1 - \lambda + \dfrac{\lambda}{k}\right)\theta_0 & if \quad \theta_0 \leq 1 \\ (1-\lambda) + \dfrac{\lambda}{k}\theta_0 & if \quad 1 \leq \theta_0 \leq k. \\ 1 & if \quad \theta_0 > k. \end{cases} \tag{1}$$

That is, the conditional mean of innate ability depends on the parental income level. Note first that parental income and parental ability are highly correlated and second, according to the empirical evidence found by Plug and Vijverberg (2003), parents' ability and the ability of the child are correlated too.[6] Thus, the two characteristics that define the individual, parental income and innate ability, will be positively correlated as well.

Individuals accumulate human capital by attending compulsory education, which is free of charge, and they are not allowed to work.

Note that mean income is $\lambda x + (1-\lambda)$ and income inequality, measured by the income variance in the population, is $(x-1)^2\lambda(1-\lambda)$. Both are increasing with x. Below we analyze the effect of mean income on the human capital distribution under both education systems.

2.2 Production of Human Capital

At compulsory level individuals are separated into different groups or classes. To simplify matters, I will assume that there are only two groups. The production of human capital depends on three factors. The first is the individual's innate ability, θ_0. The second is the "formal schooling" or "peer group" effect that depends on the characteristics of the group in which the individual is placed. These characteristics are summarized by the mean ability of the group j or "peer" effect, denoted by $\overline{\theta_0}^j$. The third is "informal schooling" and refers to family background effects, captured by z. After attending compulsory education an individual with innate ability θ_0 ends up with a level of human capital θ_1.

[6]In particular Plug and Vijverberg (2003) conclude that about 55-60 percent of the parental ability (measured as IQ) is genetically transmitted.

To conduct the analysis, I must choose a functional form for the production of human capital. I simplify this production function by assuming that it is Cobb-Douglas over "informal schooling" z, and a CES aggregate of the two remaining inputs, θ_0 and $\overline{\theta_0}^j$. This type of production function allows us to evaluate how average human capital under tracking and mixing is affected by the degree of complementarity between the peer group effect and innate ability. Thus, I assume that $\theta_1 = z^{1-\beta_1}\widetilde{\theta_1}^{\beta_1}$ where $\widetilde{\theta_1}$ is a constant returns to scale CES of θ_0 and $\overline{\theta_0}^j$.

These three inputs can be combined in two alternative ways using this Cobb-Douglas specification.[7] Each alternative allows two elasticities of substitution to be equal to 1, and the third one to vary between 0 and infinite. The first alternative is $\theta_1 = \overline{\theta_0}^{j1-\beta_1}\widetilde{\theta_1}^{\beta_1}$, where $\widetilde{\theta_1}$ is a constant returns to scale CES of θ_0 and z, and the second alternative is $\theta_1 = \theta_0^{1-\beta_1}\widetilde{\theta_1}^{\beta_1}$ where $\widetilde{\theta_1}$ is a constant returns to scale CES of $\overline{\theta_0}^j$ and z. These two alternatives restrict the elasticity of substitution between θ_0 and $\overline{\theta_0}^j$ to be equal to 1. This restriction is an important shortcoming since the empirical evidence regarding the relationship between individuals' innate ability and peer group effect is still mixed. Henderson et al. (1978) find no interaction between individual ability and the benefits of an improved peer group, i.e.,$\frac{\partial^2\theta_1}{\partial\overline{\theta_0}^j\partial\theta_0} = 0$, whereas Argys et al. (1996) suggest $\frac{\partial^2\theta_1}{\partial\overline{\theta_0}^j\partial\theta_0} > 0$ and Summers and Wolfe (1977) find some support for higher peer group benefits to lower ability students, that is, $\frac{\partial^2\theta_1}{\partial\overline{\theta_0}^j\partial\theta_0} < 0$. Thus, I use the functional form $\theta_1 = z^{1-\beta_1}\widetilde{\theta_1}^{\beta_1}$ in the analysis in order to study how the complementarity between peers' effect and individuals' innate ability affects the comparison between tracking and mixing. In particular, the specification is given by:

$$\theta_1(\theta_0, \overline{\theta_0}^j, z) = z^{1-\beta_1}(\theta_0^{\beta_2} + (\overline{\theta_0}^j)^{\beta_2})^{\frac{\beta_1}{\beta_2}}, \tag{2}$$

where β_1 and $\beta_2 \in (0,1)$. The final level of human capital θ_1, is a twice differentiable, increasing and concave function. Equation (2) allows for the possibility that $\overline{\theta_0}^j$ and θ_0 are either complements or substitutes, since β_2 determines the elasticity of substitution between these two inputs.[8]

The importance of parental education in the acquisition of human capital at the

[7] See also Krusell et al (2000).

[8] In particular, for β_2 close to 0, both $\overline{\theta_0}^j$ and θ_0 have some level of complementarity and as β_2 tends to 1 the two factors become perfect substitutes.

individual level has been explored theoretically as well as empirically. Feinstein and Symons (1999) find that parental interest is the principal way in which the attainments of each generation are passed to the next. They also suggest the complementarity between parental interest and peer effect. In keeping with these conclusions, I assume that the positive influence of peer effect on the production of human capital rises as parental income increases as can be checked from Equation (2). Finally, empirical evidence establishes that the peer group effect is non-linear: the achievement level of students rises with an improvement in the average quality of their classroom, but this positive effect has decreasing returns.[9]

2.3 Education Systems at Compulsory Level

In this section, I describe the two polar education systems of mixing and tracking and analyze the distribution of human capital at the end of compulsory school under each system.

2.3.1 Mixing

Under mixing the ability distribution is the same in both classrooms. The average ability within each classroom, denoted here as $\overline{\theta_0}^m$ coincides with the average ability in the population:

$$\overline{\theta_0}^m(k,\lambda) = \frac{1-\lambda+k\lambda}{2}. \tag{3}$$

However, as individuals differ in their parents' level of human capital, there will be two income groups within each classroom: the rich and the poor. Among the poor students θ_1 will follow a uniform distribution on the support $[a', c']$, while among the rich students θ_1 will follow a uniform distribution on the support $[b', d']$, where c' and d' denote the level of human capital θ_1 acquired by the "best" (most able) individual in the rich and the poor income group, respectively, and a' and b' denote the level of human capital θ_1 acquired by the "worst" (least able) individual in the rich and the

[9] See Summers and Wolfe (1977), Henderson, Mieszkowski, and Sauvageau (1978), and de Bartolome (1990).

poor income group, respectively:

$$a'(k,\lambda) = (\overline{\theta_0}^m)^{\beta_1} \tag{4}$$

$$b'(k,\lambda) = x^{1-\beta_1}(\overline{\theta_0}^m)^{\beta_1} \tag{5}$$

$$c'(k,\lambda) = (1+(\overline{\theta_0}^m)^{\beta_2})^{\frac{\beta_1}{\beta_2}} \tag{6}$$

$$d'(k,\lambda) = x^{1-\beta_1}(k^{\beta_2}+(\overline{\theta_0}^m)^{\beta_2})^{\frac{\beta_1}{\beta_2}}. \tag{7}$$

Under mixing, therefore, the C.D.F. of human capital at the end of compulsory education, denoted by $F_M(\theta_1)$, is:

$$F_M(\theta_1) = \begin{cases} \left(\theta_1^\beta - (a')^\beta\right)^{\frac{1}{\beta_2}}(1-\lambda) & if \quad 0 \le \theta_1 \le b' \\ \left(\theta_1^\beta - (a')^\beta\right)^{\frac{1}{\beta_2}}(1-\lambda) + \left(\left(\frac{\theta_1}{x^{1-\beta_1}}\right)^\beta - (a')^\beta\right)^{\frac{1}{\beta_2}}\frac{\lambda}{k} & if \quad b' \le \theta_1 \le c' \\ (1-\lambda) + \left(\left(\frac{\theta_1}{x^{1-\beta_1}}\right)^\beta - (a')^\beta\right)^{\frac{1}{\beta_2}}\frac{\lambda}{k} & if \quad c' \le \theta_1 \le d'. \\ 1 & if \quad \theta_1 > d'. \end{cases} \tag{8}$$

where $\beta = \frac{\beta_2}{\beta_1}$.

It can be checked that, ceteris paribus, if in society A the difference in terms of mean ability between the rich and the poor is larger than in society B, then the distribution of human capital under mixing $F_M(\theta_1)$ in society A will dominate the one in society B. This is because mean ability among the rich is the only determinant of the difference in average human capital between the rich and the poor. As shown in Figure 1, below, this implies an increase in the expected value of θ_1 under mixing. Here, the case $k = 1$ is represented in red and $k = 2$ in green:

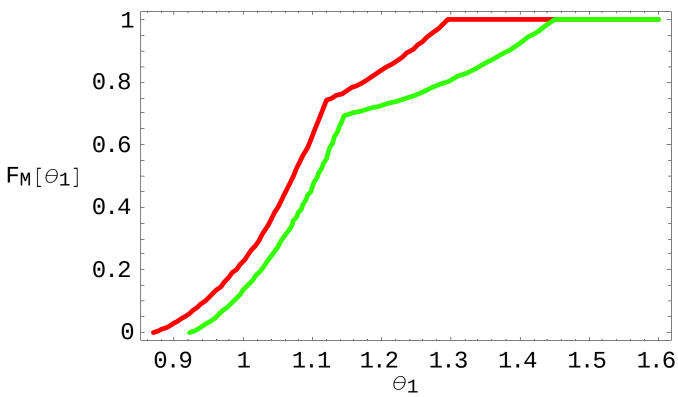

Figure 1: Mixing

Here, $E_M(\theta_1)$ denotes the expected value of θ_1 under mixing, where:

$$E_M(\theta_1) = (1-\lambda)\left(\frac{a'+c'}{2}\right) + \lambda\left(\frac{b'+d'}{2}\right),$$

or, using Equations (4) to (7):

$$\begin{aligned} E_M(\theta_1) \;=\; & \frac{1}{2}\left((\overline{\theta_0}^m)^{\beta_1} + (1+(\overline{\theta_0}^m)^{\beta_2})^{\frac{1}{\beta}}\right) + \\ & \frac{\lambda}{2}\left(x^{1-\beta_1}\left((\overline{\theta_0}^m)^{\beta_1} + (k^{\beta_2}+(\overline{\theta_0}^m)^{\beta_2})^{\frac{1}{\beta}}\right) - \left((\overline{\theta_0}^m)^{\beta_1} + (1+(\overline{\theta_0}^m)^{\beta_2})^{\frac{1}{\beta}}\right)\right). \end{aligned} \tag{9}$$

Thus $E_M(\theta_1)$ is an average of the mean values of θ_1 in the two income groups, with respective weights $(1-\lambda)$ and λ. Finally we also find that $E_M(\theta_1)$ is an increasing function of the wealth level in the population, measured by both x and λ and also increasing with k as we saw above.

2.3.2 Tracking

Tracking students implies grouping them on the basis of innate ability. For the sake of simplicity, I permit only two tracks and use the median level of innate ability as a threshold for grouping students into one track or the other. Thus, a student is assigned to the high (low) track when their ability θ_0 is above (below) the median, denoted by $m(k,\lambda)$:

$$m(k,\lambda) = \frac{k}{2(k(1-\lambda)+\lambda)}. \tag{10}$$

The following assumption ensures that there will be at least one poor student in the high track.

Assumption 1 (A.1): $\lambda < (1/2)$.

The previous assumption ensures that $m(k,\lambda) < 1$ and thereby placing a non-negative proportion of poor students in the high track.[10]

The distribution of human capital within each track is uniform but with different parameters. I denote by $\overline{\theta_0}^h$ and $\overline{\theta_0}^l$ the average ability in the high and low tracks

[10] Note from (3) and (10) that the initial distribution of innate ability, is right-skewed, that is $m(k,\lambda) < \overline{\theta_0}^m(k,\lambda)$ for any $\lambda \in (0,1)$ and $k > 1$.

respectively. Thus, given the distributional assumption on θ_0, I have that:

$$\overline{\theta_0}^{l}(k,\lambda)=\frac{k}{4(k(1-\lambda)+\lambda)} \tag{11}$$

and from (3), (10) and (11) I have that $\overline{\theta_0}^{h}(k,\lambda)=\overline{\theta_0}^{l}(k,\lambda)+\overline{\theta_0}^{m}(k,\lambda)$.

Again, there will be two income groups within each track. In the low track, θ_1 follows a uniform distribution on $[a,c]$ among the poor, while among the rich it follows a uniform distribution on $[b,e]$, where c and e denote the level of human capital θ_1 acquired in the low track by the "best" (most able) individual in the poor and the rich income groups, respectively, and a and b denote the human capital acquired in the low track by the "worst" (least able) individual in the poor and rich groups, respectively, that is:

$$\begin{aligned}
a(k,\lambda) &= (\overline{\theta_0}^{l})^{\beta_1} && (12)\\
b(k,\lambda) &= x^{1-\beta_1}(\overline{\theta_0}^{l})^{\beta_1} && (13)\\
c(k,\lambda) &= (m^{\beta_2}+(\overline{\theta_0}^{l})^{\beta_2})^{\frac{\beta_1}{\beta_2}} && (14)\\
e(k,\lambda) &= x^{1-\beta_1}(m^{\beta_2}+(\overline{\theta_0}^{l})^{\beta_2})^{\frac{\beta_1}{\beta_2}}. && (15)
\end{aligned}$$

Likewise, in the high track, θ_1 follows a uniform distribution on $[d,g]$ among the poor, while among the rich it follows a uniform distribution on $[f,h]$. We denote by d and f the human capital θ_1 acquired in the high track by the "worst" (least able) individual in the poor and rich groups, respectively. We denote by g and h the human capital θ_1 acquired by the "best" (most able) individual in the poor and rich groups, respectively, i.e.:

$$\begin{aligned}
d(k,\lambda) &= (m^{\beta_2}+(\overline{\theta_0}^{h})^{\beta_2})^{\frac{\beta_1}{\beta_2}} && (16)\\
f(k,\lambda) &= x^{1-\beta_1}(m^{\beta_2}+(\overline{\theta_0}^{h})^{\beta_2})^{\frac{\beta_1}{\beta_2}} && (17)\\
g(k,\lambda) &= (1+(\overline{\theta_0}^{h})^{\beta_2})^{\frac{\beta_1}{\beta_2}} && (18)\\
h(k,\lambda) &= x^{1-\beta_1}(k^{\beta_2}+(\overline{\theta_0}^{h})^{\beta_2})^{\frac{\beta_1}{\beta_2}}. && (19)
\end{aligned}$$

From Equation (2) above we have that, given two individuals with the same innate ability level, the one whose parents are rich will always attain a higher level of human capital. Given two individuals with the same parental income level, the most able one will always attain a higher level of human capital. This can also be checked from Equations (12) to (19) above, i.e., first $a < b, c < e, d < f$ and $g < h$, and second $c > a, e > b, g > d$ and $h > f$.

The next assumption ensures that the support of θ_1 in the low track "partially" overlaps the support of θ_1 in the high track.

Assumption 2 (A.2): $g > e > d$.

In other words, the "best" individual in the low track (a rich individual with $\theta_0 = m$) obtains more human capital than the "worst" individual in the high track (a poor individual with $\theta_0 = m$). Moreover, it implies that the "best" individual in the high track among the poor (an individual with $\theta_0 = 1$) obtains more human capital than the "best" individual in the low track among the rich (an individual with $\theta_0 = m$).

This assumption implies a restriction on x, β_1, β_2 and k. For some fixed β_1, β_2 and k, this means that the income level of the rich lies within the following interval:

$$\underline{x}(\beta_1, \beta_2, k, \lambda) < x < \overline{x}(\beta_1, \beta_2, k, \lambda) \tag{20}$$

where $\underline{x}(\beta_1, \beta_2, k, \lambda) = \left(\frac{m^{\beta_2} + (\overline{\theta_0}^h)^{\beta_2}}{m^{\beta_2} + (\overline{\theta_0}^l)^{\beta_2}}\right)^{\frac{\beta_1}{\beta_2(1-\beta_1)}}$ and $\overline{x}(\beta_1, \beta_2, k, \lambda) = \left(\frac{1 + (\overline{\theta_0}^h)^{\beta_2}}{m^{\beta_2} + (\overline{\theta_0}^l)^{\beta_2}}\right)^{\frac{\beta_1}{\beta_2(1-\beta_1)}}$.

That is, x must be high enough to offset the disadvantage of being in the low track, but not too high to offset the disadvantage of having poor parents.

Note from Equations (17) and (18) that Assumption 2 implies that $f(k, \lambda) > g(k, \lambda)$. That is, the "best" poor individual in the high track obtains less human capital than the "worst" rich individual in the high track.

Figure 2 illustrates the different intervals for θ_1 and the relationship between them, for both educational systems.

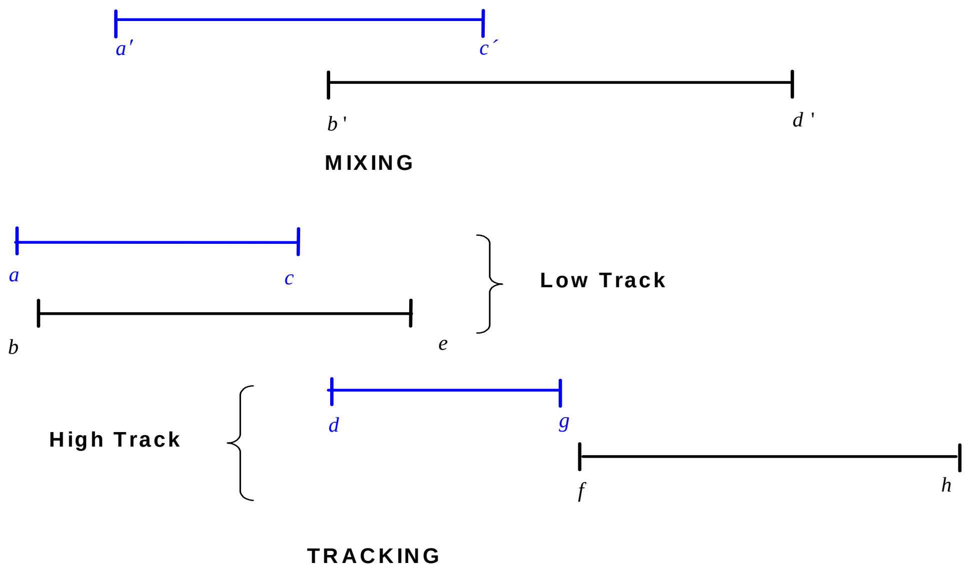

Figure 2: Educational Systems

Under Assumption A.2 the C.D.F. of θ_1 under tracking, denoted by $F_T(\theta_1)$, is as follows:

$$F_T(\theta_1) = \begin{cases} \left(\theta_1^{\beta} - a^{\beta}\right)^{\frac{1}{\beta_2}} (1-\lambda) & if \quad 0 \leq \theta_1 \leq b \\ \left(\theta_1^{\beta} - a^{\beta}\right)^{\frac{1}{\beta_2}} (1-\lambda) + \left(\left(\frac{\theta_1}{x^{1-\beta_1}}\right)^{\beta} - a^{\beta}\right)^{\frac{1}{\beta_2}} \dfrac{\lambda}{k} & if \quad b \leq \theta_1 \leq c \\ m(1-\lambda) + \left(\left(\frac{\theta_1}{x^{1-\beta_1}}\right)^{\beta} - a^{\beta}\right)^{\frac{1}{\beta_2}} \dfrac{\lambda}{k} & if \quad c \leq \theta_1 \leq d \\ \left(\theta_1^{\beta} - \left(\overline{\theta_0}^{h}\right)^{\beta_2}\right)^{\frac{1}{\beta_2}} (1-\lambda) + \left(\left(\frac{\theta_1}{x^{1-\beta_1}}\right)^{\beta} - a^{\beta}\right)^{\frac{1}{\beta_2}} \dfrac{\lambda}{k} & if \quad d \leq \theta_1 \leq e \\ \left(\theta_1^{\beta} - \left(\overline{\theta_0}^{h}\right)^{\beta_2}\right)^{\frac{1}{\beta_2}} (1-\lambda) + m\left(\dfrac{\lambda}{k}\right) & if \quad e \leq \theta_1 \leq g \\ \left(g^{\beta} - \left(\overline{\theta_0}^{h}\right)^{\beta_2}\right)^{\frac{1}{\beta_2}} (1-\lambda) + m\left(\dfrac{\lambda}{k}\right) & if \quad g \leq \theta_1 \leq f \\ (1-\lambda) + \left(\left(\frac{\theta_1}{x^{1-\beta_1}}\right)^{\beta} - \left(\overline{\theta_0}^{h}\right)^{\beta_2}\right)^{\frac{1}{\beta_2}} \dfrac{\lambda}{k} & if \quad f \leq \theta_1 \leq h \\ 1 & if \quad \theta_1 > h. \end{cases} \tag{21}$$

It can be checked that, ceteris paribus, if in society A the difference in terms of mean ability between the rich and the poor is larger than in society B then, and opposite to mixing, the distribution of human capital under tracking $F_T(\theta_1)$ in society A will not dominate the one in society B. Figure 1 represents the case $k = 1$ in blue and $k = 2$ in green:

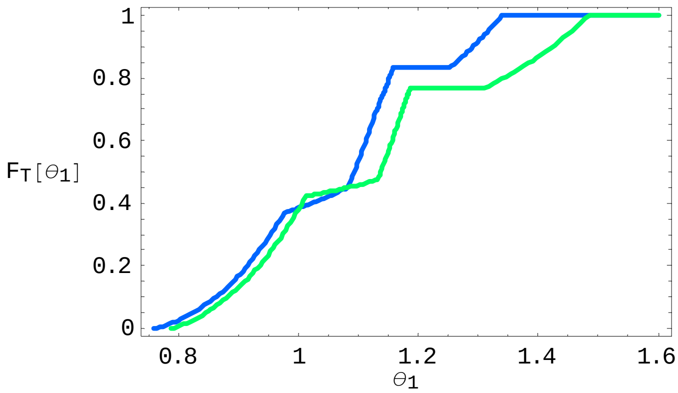

Figure 3: Tracking

The intuition of the previous result is as follows. From (10) and (11) we have that average ability in both the low and the high track, $\overline{\theta_0}^l$ and $\overline{\theta_0}^h$, and the median m will be higher in society A than they are in B. Thus, the proportion of poor (rich) students in the low track will be higher (lower) in society A than in society B, and the reverse occurs in the high track. Consequently, as can be checked from (21) and Figure 3, the decrease in the average human capital of rich students in the low track implies that $F_T(\theta_1)$ will be higher in society A than in society B for intermediate values of θ_1. In addition, the increase in the average human capital in all remaining cases implies that $F_T(\theta_1)$ will be lower in society A than it will be in society B for all remaining values of θ_1.

Therefore, and opposite to mixing, the higher the difference between the mean innate ability of rich and poor kids, the resulting $F_T(\theta_1)$ will not dominate any previous distribution of human capital in the sense of first order stochastic dominance. However, as we see below it implies a higher expected value of θ_1.

The expected value of θ_1 under tracking is:

$$E_T(\theta_1) = m(1-\lambda)\left(\frac{a+c}{2}\right)+(1-m)(1-\lambda)\left(\frac{d+g}{2}\right)+\frac{m}{k}\lambda\left(\frac{b+e}{2}\right)+(1-\frac{m}{k})\lambda\left(\frac{f+g}{2}\right), \tag{22}$$

As with mixing, the expected value of θ_1 is a weighted average of the mean value of θ_1 in the four income groups analyzed above. It is also increasing in both x and λ. Under tracking, the effect of an increase in mean innate ability among rich students (captured by an increase in k) on the average human capital is positive, but not as strong as it is under mixing. The reason for this can be explained as follows. As we saw above, a higher k implies a higher average ability among both high and low track students, and a lower (higher) mean income among low (high) track students. Due to the complementarity between peer effect and family background in the production of human capital (see Equation (2)), the increase in mean income of students in the high track has a clear positive impact on the average human capital in this group. However, the decrease in mean income in the low track implies that average human capital might not always increase in that track.[11]

3 A comparison of mixing and tracking

Let us suppose, first of all, that the educational system is chosen by majority voting and that every individual will vote for the system under which her final level of human capital θ_1 is higher. In this case, exactly half of the population will prefer mixing (those with $\theta_0 < m$), since under tracking they would placed into the low track, where they would enjoy a lower peer effect. The other half will prefer tracking (those with $\theta_0 > m$), since they would placed into the high track, where they would enjoy a higher peer effect. We see that $\frac{1}{2}$ prefers mixing and $\frac{1}{2}$ prefers tracking, which means that choosing one system over the other will always produce winners and losers.

Next I propose to see what happens if individuals must choose the education system without knowing their own characteristics. In particular they ignore the value of θ_0 that they will end up enjoying.[12] Again, I assume that they would like to have

[11] From Equation (22) the impact of k on the average human capital in the low track is:
$\frac{\partial}{\partial k}\left(m(1-\lambda)\left(\frac{a+c}{2}\right)+\frac{m}{k}\lambda\left(\frac{b+e}{2}\right)\right)=$
$(1-\lambda)\left(\frac{\partial m}{\partial k}\left(\frac{a+c}{2}\right)+m\frac{\partial}{\partial k}\left(\frac{a+c}{2}\right)\right)+\lambda\left(\frac{\partial m}{\partial k}\frac{1}{k}\left(\frac{b+e}{2}\right)+\frac{m}{k}\frac{\partial}{\partial k}\left(\frac{b+e}{2}\right)-\frac{m}{k^2}\left(\frac{b+e}{2}\right)\right)$

[12] This approach is known as the "Veil of Ignorance", widely used in modern Welfare Economics (see for example the seminal works of Harsanyi (1953 and 1955) and Rawls (1971)). Under this approach, to evaluate alternative systems, individuals must put themselves behind a hypothetical "veil of ignorance", where they ignore their own characteristics.

as much of θ_1 as possible.

One possibility is just to compare both systems in terms of average human capital.[13] However, due to the complexity of the above human capital production function, I cannot obtain clear analytical results regarding the comparison of average human capital. However we can extract some conclusions using numerical simulations. The most important one is that the difference between average human capital under the two systems, $E_T(\theta_1) - E_M(\theta_1)$, decreases with β_2. The following table presents the value of β_2, for different values of β_1 and k, such that $E_T(\theta_1) - E_M(\theta_1) = 0$, denoted it by $\widehat{\beta_2}$. Thus, for β_2 below (above) $\widehat{\beta_2}$ we have that $E_T(\theta_1) - E_M(\theta_1) > (<) 0$:[14]

Table 1. Average Human Capital : $\widehat{\beta_2}$ (23)

$\beta_1 \setminus k$	1.25	1.5	1.75	2
1/4	0.77	0.74	0.72	0.69
1/2	0.86	0.85	0.82	0.79
3/4	0.95	0.94	0.92	0.910

We can conclude that if β_2 is small, i.e., when $\overline{\theta_0}^j$ and θ_0 have some level of complementarity, then average human capital is always maximized under tracking. As β_2 tends to 1, meaning that as the two factors become closer substitutes, average human capital is maximized under mixing. To put it differently, when peer effects matter more for low (high) ability students than for high (low) ability students, average human capital is maximized under mixing (tracking), which is the system where low (high) ability students enjoy a stronger peer effect.

Note also that, as the role of family background in human capital accumulation diminishes (β_1 increases) tracking maximizes average human capital for most values of β_2. Finally, it can also be checked that in societies where the difference in mean ability of rich versus poor individuals is low (i.e., when k is low), mixing maximizes average human capital for a smaller interval of values of β_2. However, as the difference between the mean ability of individuals in both income groups becomes greater (i.e., as k increases), we have that mixing maximizes average human capital for a larger

[13] This is like assuming that all individuals are risk neutral behind the veil of ignorance.

[14] Note that I additionally assume here that the proportion of rich individuals in the population, λ is equal to 1/4 and $x(\beta_1, k) = \frac{\underline{x}(\beta_1,k)+\overline{x}(\beta_1,k)}{2}$ in each case.

interval of values of β_2. As we saw in the previous section, an increase in k has a clear positive impact on average human capital under mixing, whereas the final impact on average human capital under tracking, although positive, is not as strong as it is under mixing. Thus, ceteris paribus, in societies where the difference between the mean innate ability of rich versus poor students is larger, the difference between the average human capital of each system will be lower.

The following table presents the value of $E_T(\theta_1) - E_M(\theta_1)$ for certain fixed values of x and β_2:[15]

Table 2. $E_T(\theta_1) - E_M(\theta_1)$			
$\beta_2 \setminus x$	$\underline{x}(\beta_2)$	$\widehat{x}(\beta_2)$	$\overline{x}(\beta_2)$
1/4	0.1556	0.1584	0.1611
1/2	0.0304	0.0313	0.0322
3/4	−0.0028	−0.0024	−0.0020
1	−0.0183	−0.0181	−0.0179

(24)

Here, the impact of any increase of the wealth level of the population (measured by x) on the difference in the average level of human capital under one system versus the other depends crucially on the level of complementarity between peer-group characteristics and innate ability. In particular, if these two factors are complements (substitutes), then the difference increases (decreases) in absolute terms, with x. This relationship can be explained as follows. From Equations (9) and (22) we can check that the average level of human capital under both systems is increasing in x. However, the impact of an increase in x on $E_T(\theta_1)$ is much higher than on $E_M(\theta_1)$. This is due to the complementarity between peer effect and family background. Since an increase in x has no impact on peer variables, $\overline{\theta_0}^l$ and $\overline{\theta_0}^h$, or on the composition of the low and the high track m (and as opposed to what happens when k increases), it will have a greater clear positive impact on $E_T(\theta_1)$ than it will on $E_M(\theta_1)$, since the peer variable that rich students enjoy under tracking is higher than the one that they enjoy under mixing. Therefore, as shown in Table 2 , if peers' characteristics

[15] First note that $\widehat{x}(\beta_2) = \frac{\underline{x}(\beta_2)+\overline{x}(\beta_2)}{2}$.In addition we assume here that $\beta_1 = 1/4$, $\lambda = 1/4$ and $k = 1.75$. However, the result are qualitatively the same for different combinations of the previous parameters.

and innate ability are complements (i.e. $\beta_2 < \widehat{\beta_2}$ and tracking maximizes average human capital) then $E_T(\theta_1) - E_M(\theta_1)$ increases with x, whereas if they are substitutes (i.e. $\beta_2 > \widehat{\beta_2}$ and mixing maximizes average human capital) then $E_T(\theta_1) - E_M(\theta_1)$ diminishes, in absolute terms, with x.

Another possibility, which has not been previously considered in the literature, is to compare both systems in terms of human capital distribution. It should be recall that, if the human capital distribution under a given system dominates that of another according to first order stochastic dominance, then all individuals can be said to prefer the former over the latter.

However, it can be checked that neither system dominates the other according to this criterion.

Proposition 1 *$F_r(\theta_1) \not\succeq_{FOSD} F_s(\theta_1)$ for $r, s = M, T$ and $r \neq s$ for any β_1, β_2 and k.*

Proof. (i) $F_T(\theta_1) \not\succeq_{FOSD} F_M(\theta_1)$. Using $F_T(\theta_1)$ from (21) and $F_M(\theta_1)$ from (8) we can check that, for any $\theta_1 \in (0, b]$, $(F_T(\theta_1) - F_M(\theta_1)) > 0$ for every λ, β_1 and β_2 and k. (ii) $F_M(\theta_1) \not\succeq_{FOSD} F_T(\theta_1)$. Using Equations (21) and (8), we can check that for any $\theta_1 \in [f, g]$, $(F_T(\theta_1) - F_M(\theta_1)) < 0$ for every λ, β_1 and β_2 and k. ■

Figure 4 illustrates the previous result, where $F_M(\theta_1)$ and $F_T(\theta_1)$ are represented in red and blue lines respectively. Therefore we can conclude that, regardless of the properties pertaining to the process of human capital accumulation, there is no unanimity in the population so as to which system to choose.

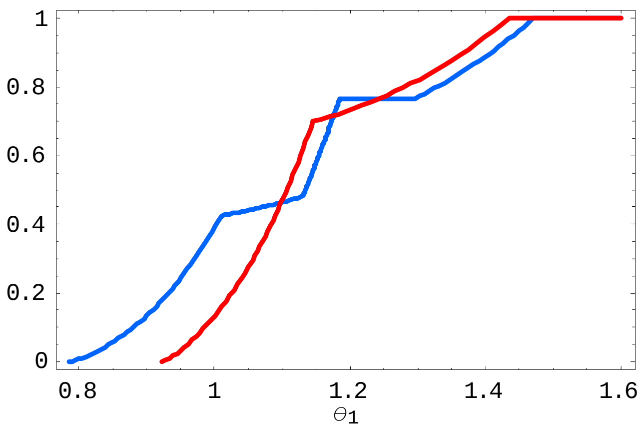

Figure 4: No First Order Stochastic Dominance

Finally, let us suppose that all individuals behind the “veil of ignorance” are risk averse. In this case, they will prefer the less risky distribution of human capital. This criteria leads to the concept of second order stochastic dominance. It can be checked that the preferred system according to this criteria depends on the degree of complementarity between the peer group effect and innate ability, β_2. Below I show first that, when peer effect and innate ability are close complements then, there is no unanimously preferred system in the population.

Proposition 2 *Let* $\beta_2 < \widehat{\beta_2}$ *then* $F_r(\theta_1) \not\succeq_{SOSD} F_s(\theta_1)$ *for* $r, s = M, T$ *and* $r \neq s$ *for any* β_1 *and* k*.*

Proof. (i) $F_T(\theta_1) \not\succeq_{SOSD} F_M(\theta_1)$. Using $F_T(\theta_1)$ from (22) and $F_M(\theta_1)$ from (8) we can check that, $\int\limits_0^b (F_T(\theta_1) - F_M(\theta_1))d\theta_1 > 0$, for every β_1 and k. (ii) $F_M(\theta_1) \not\succeq_{SOSD} F_T(\theta_1)$. Recall that the expected value of a random variable can be written as: $E(y) = \overline{y} - \int\limits_0^{\overline{y}} F(y)dy$, where $\overline{y}$ is the lowest value of y for which $F(y) = 1$. Thus the expected value of θ_1 under tracking can be written as: $E_T(\theta_1) = h - \int\limits_0^h F_T(\theta_1)d\theta_1$ and, under mixing $E_M(\theta_1) = d' - \int\limits_0^{d'} F_T(\theta_1)d\theta_1 = h - \int\limits_0^h F_M(\theta_1)d\theta_1$. Note that, if $F_M(\theta_1) \succeq_{SOSD} F_T(\theta_1)$, then the following inequality should hold: $h - E_M(\theta_1) \leq h - E_T(\theta_1)$. The final result is immediate from Table 1. ■

If the peer group effect and innate ability are close substitutes (i.e., if $\beta_2 > \widehat{\beta_2}$), then there are no clear-cut results as to how mixing and tracking compare under the criteria of second order stochastic dominance. Note that if $\beta_2 > \widehat{\beta_2}$, then we could have that mixing dominates tracking according to this criteria. However this result is only true if F_M and F_T cross only once, in which case, and according to the previous definition of the expected value of a random variable, the following inequality should hold: $h - E_M(\theta_1) \leq h - E_T(\theta_1)$, which is true from Table 1 if

$\beta_2 > \widehat{\beta_2}$. Using Equation (8) for F_M and Equation (21) for F_T we can check that, for any $\theta_1 \in (a, b)$, $F_T(\theta_1) - F_M(\theta_1) > 0$ for every λ, k, β_1 and x, whereas for any $\theta_1 \in (d', h)$, $F_T(\theta_1) - F_M(\theta_1) < 0$ for every λ, k, β_1 and x. Thus, we just can conclude that both cumulative distribution functions cross an odd number of times.

The complexity of the above human capital production function, makes it difficult to obtain clear analytical results for every β_2 and k. However, if we focus on the case where peer effect and individuals' innate ability are perfect substitutes, i.e. where $\beta_2 = 1$, we can find some interesting cases where mixing dominates tracking according to this criteria. In particular, the result will be driven by the level of income inequality in the population. Remember from Section 2 that the income variance is $(x-1)^2\lambda(1-\lambda)$, thus is increasing with x. Therefore, if we take the proportion of rich individuals λ as given, income inequality will be characterized by the value of x. The following proposition shows that if peer effect and innate ability are substitutes and if the income inequality of the population is low enough, then the population will unanimously prefer mixing over tracking.[16]

Proposition 3 *Let $\beta_2 = 1$ and define $\widetilde{x}(\beta_1, k, \lambda) = \left(\frac{1+\overline{\theta_0}^h}{m+\overline{\theta_0}^m}\right)^{\frac{\beta_1}{1-\beta_1}}$. If $x \leq \widetilde{x}(\beta_1, k, \lambda)$ and $\lambda \leq \widetilde{\lambda}(\beta_1, k)$ then $F_M(\theta_1) \succeq_{SOSD} F_T(\theta_1)$ for any β_1 and k.*

Proof. See Appendix. ■

Note first that $x \leq \widetilde{x}(\beta_1, k, \lambda)$ implies that $e(k, \lambda) \leq c'(k, \lambda)$. That is, the "best" poor individual (one with $\theta_0 = 1$) will obtain a higher level of human capital under mixing than will the "best" rich individual from within the low track (an individual with $\theta_0 = 1$). One might consider that mixing represents the public educational system whereas tracking represents a private system comprised of both low and high quality schools. Thus, this condition on x implies that the best public school student can achieve a higher level of human capital than can the best low quality private school student, which seems to be an empirically relevant case in most developed countries (see Martínez-Mora (2006)).

[16] Recall from section 2 that an increase in x can be interpreted both as an increase in the mean income or as increase in the income inequality in the population.

Thus we can conclude that, if peer effect and innate ability are substitutes and societal income level is not too unequal, then individuals prefer mixing. The intuition could be as follows. In Table 2 we saw that if peer effects and individuals' innate ability are close substitutes then the difference between the average human capital under tracking and mixing is lower in societies with lower x. As a result, in poor societies risk averse individuals will prefer mixing.

Finally note that both $\widetilde{x}(\beta_1, k, \lambda)$ and $\widetilde{\lambda}(k)$ are decreasing with k. That is, as society becomes more unequal in terms of the difference in the mean ability between rich and poor, then it will be required a lower proportion of rich individuals in the population and a lower income level of the rich, to get mixing as the preferred system.

4 Concluding Remarks

In this paper I have analyzed public intervention in education when the government has to decide how to group students. I analyze two different education systems: tracking and mixing.

A number of previous works have studied the optimal education system at compulsory level by focusing on mean achievement. This paper contributes to this line of research by introducing critical inputs like family background into the human capital production function, and by recognizing the existence of a positive dependence between family background and individuals' innate ability and its effect on each of the two educational systems described above. In addition to that this paper contributes to this literature by comparing both systems in terms of the induced distributions of human capital at the end of compulsory school.

The paper allows for some extensions. In particular, it might be interesting to check the robustness of my main results against specific features of the model. Here, I assume that individual student achievement rises with increases in the average ability level of their classmates, but at a decreasing rate. Another type of non-linearity with regard to peer group effect arises as a result of the "distance" impact. There is empirical evidence that suggests that peer effects are stronger when the distance between the individual's innate ability and the average innate ability in the classroom is small, and that as this distance increases, peer effects become almost negligible.[17] It might be interesting to model this effect.

It might also be important to relax some of the assumptions presented here. For example, we might consider other distributions of innate ability or introduce the possibility of tracking students only within a certain subset of subjects as in Epple, Newlon and Romano (2002). In addition to adding realism, incorporating this possibility would make it easier for us to design an optimal educational system. On the other hand, it would be interesting to explore how factors introduced by each the two compulsory school systems discussed here can influence students decisions as to whether or not to attend college (see Hidalgo-Hidalgo (2005) for a similar analysis with a more stylized model).

[17] See, for example, Manski and Wise (1983) and more recently, Hoxby and Weingarth (2006).

5 Appendix

Proof. of Proposition 3: I denote by $\widetilde{\lambda}(k)$ the proportion of rich individuals in the population such that $\widetilde{x}(\beta_1, k, \lambda) = \underline{x}(\beta_1, k, \lambda)$ and thus, if $\lambda \leq \widetilde{\lambda}(\beta_1, k)$ then $\widetilde{x}(\beta_1, k) \in [\underline{x}(\beta_1, k), \overline{x}(\beta_1, k)]$ for any β_1 and k. Now define $\widetilde{\widetilde{x}}(\beta_1, k, \lambda) = \left(\frac{1+\overline{\theta_0}^m}{m+\overline{\theta_0}^l}\right)^{\frac{\beta_1}{1-\beta_1}}$ and $x'(\beta_1, k, \lambda) = \left(\frac{m+\overline{\theta_0}^m}{\overline{\theta_0}^m}\right)^{\frac{\beta_1}{1-\beta_1}}$ where both $\widetilde{\widetilde{x}}(\beta_1, k, \lambda) \in [\underline{x}(\beta_1, k, \lambda), \overline{x}(\beta_1, k, \lambda)]$ and $x'(\beta_1, k, \lambda) \in [\underline{x}(\beta_1, k, \lambda), \overline{x}(\beta_1, k, \lambda)]$ for any β_1, k and λ. It can also be checked that $\widetilde{x}(\beta_1, k, \lambda) < \widetilde{\widetilde{x}}(\beta_1, k, \lambda)$ and that $\widetilde{x}(\beta_1, k, \lambda) < \widehat{x}(\beta_1, k, \lambda)$ for any β_1, k and λ. Recall that for every λ, k and β_1, from Equations (8) and (21), we have that for any $\theta_1 \in [a, c]$ $F_T(\theta_1) > F_M(\theta_1)$ whereas for any $\theta_1 \in [d', h]$ $F_T(\theta_1) < F_M(\theta_1)$. In addition note from Equations (6) and (15) that $e(k, \lambda) \leq c'(k, \lambda)$ if and only if $x \leq \widetilde{\widetilde{x}}(\beta_1, k, \lambda)$. It can be checked from Equations (8) and (21) that, if $x \leq \widetilde{\widetilde{x}}(\beta_1, k, \lambda)$ then $F_T(\theta_1) < F_M(\theta_1)$ for all $\theta_1 \in (d, c')$. Now, if we evaluate the two C.D.F. for $\theta_1 = b'$ we can check that $F_T(b') = m(1-\lambda) + (\overline{\theta_0}^m - \overline{\theta_0}^l)\frac{\lambda}{k}$ and $F_M(b') = \overline{\theta_0}^m(x^{\frac{1-\beta_1}{\beta_1}} - 1)(1-\lambda)$. Thus, if $x \leq x'(\beta_1, k, \lambda)$ then $F_T(b') > F_M(b')$. If we evaluate the two C.D.F. for $\theta_1 = g$ we can check that $F_M(g) = (1-\lambda) + \left(\left(\frac{g}{x^{1-\beta_1}}\right)^{\frac{1}{\beta_1}} - \overline{\theta_0}^m\right)\lambda$ and $F_T(g) = (1-\lambda) + m\frac{\lambda}{k}$. Thus, $F_T(g) > (<) F_M(g)$ if and only if $x > (<) \widetilde{x}(\beta_1, k, \lambda)$. Finally, from the definitions of $\widetilde{x}(\beta_1, k, \lambda)$, $\widetilde{\widetilde{x}}(\beta_1, k, \lambda)$ and $x'(\beta_1, k, \lambda)$ it is immediate that if $x < \widetilde{x}(\beta_1, k, \lambda)$ then F_M and F_T cross only once. This completes the proof. ▬

References

[1] Argys, L.M., Rees, D.I. and D.J. Brewer (1996) "Detracking America's schools: equity at zero cost?" *Journal of Policy Analysis and Management*, **5** (4): 623-645.

[2] Arnott, R.J., and J. Rowse (1987) "Peer group effects and educational attainment" *Journal of Public Economics*, **32**: 237-306.

[3] Betts J.R. and J.L. Shkolnik (2000) "The effects of ability grouping on student math achievement and resource allocation in secondary schools" *Economics of Education Review* **19** (1): 1-15.

[4] de Bartolome, C.A.M. (1990) "Equilibrium and inefficiency in a community model with peer group effects" *Journal of Political Economy*, **98** (1): 110-133.

[5] Epple, D. Newlon, E. and R.E. Romano (2002) "Ability Tracking, school competition, and the distribution of educational benefits" *Journal of Public Economics* **83**: 1-48.

[6] Evans, W.N., W. Oates and R.M. Schwab (1992) "Measuring peer groups effects: a study of teenage behavior" *Journal of Political Economy*, **100** (5): 966-991.

[7] Feinstein L. and J. Symons (1999) "Attainment in secondary school" *Oxford Economics Papers* 51 300-321.

[8] Figlio, D.N. and M.E. Page (2000) "School Choice and the distributional Effects of Ability Tracking: Does separation increase equality?" *Journal of Urban Economics* **51** (3):497-514.

[9] Harsanyi, J. (1953) "Cardinal utility in welfare economics and the theory of risk taking" *Journal of Political Economy* **61**, 434-435.

[10] Harsanyi, J. (1955) "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility" *Journal of Political Economy* **63**: 309-321.

[11] Heckman, J. J. (2000) "Policies to foster human capital" Research in Economics **54**(1): 3-56.

[12] Henderson, V., Mieszkowski, P. Y. Sauvageau (1978) “Peer group effects and Educational Production Functions” *Journal of Public Economics* 10, 97:106.

[13] Hidalgo-Hidalgo, M. (2005) “Peer group effects and optimal education system” IVIE WP-AD 2005-12.

[14] Hoxby, C.M. (2000) “Peer effects in the classroom: learning from gender and race variation” NBER WP No. 7867.

[15] Hoxby, C.M. and G. Weingarth (2006) “Taking race out of the equation: school reassignment and the structure of peer effects” Harvard University manuscript.

[16] Kang, C. (2007) “Classroom peer effects and academic achievement: quasi randomization evidence from South Korea” *Journal of Urban Economics* **61**: 458-495

[17] Krusell, P. , L. Ohanian, V. Ríos-Rull and G. Violante (2000) “Capital-Skill complementarity and inequality” *Econometrica* **68** (5): 1029-1053.

[18] Manski, C. F. and D. A. Wise (1983). College Choice in America. Cambridge, Harvard University Press.

[19] Manski, C. F. (1993) “Identification of endogenous social effects: the reflection problem” *Review of Economic Studies* **60:** 531-542.

[20] Martínez-Mora, F. (2006) “The Existence of Non-Elite Private Schools” *Journal of Public Economics* **90**: 1505-1518.

[21] Plug, E. and V. Vijverberg (2003) “Schooling, family background and adoption: is it nature or is it nurture?” *Journal of Political Economy,* **111**: 611-641.

[22] Rawls, J. (1971) “A Theory of Justice”, Harvard University Press, Cambridge MA.

[23] Rees, D.I., L.M. Argys and D.J. Brewer (1996) “Tracking in the United States: Descriptive Statistics from N.E.L.S.” *Economics of Education Review,* **15** (1): 83-89.

[24] Robertson, D. and J. Symons, (2003) "Do peer groups matter? Peer group versus schooling effect on academic achievement" *Economica*, **70**, 31-53.

[25] Roemer, J. E. and R.J.B. Wets (1994) "Neighborhood Effects on Belief Formation and the Distribution of Education and Income" *Working paper of the Department of Economics, University of California.* W.P. series n^o 94-02.

[26] Streufert, P. (2000) "The effect of underclass social isolation on schooling choice" *Journal of Public Economic Theory*, **2**: 461-482.

[27] Summers, A. and B. Wolfe. (1977) "Do schools make a difference?" *American Economic Review* **67** (4): 639-652.

[28] Zimmer, R.W. and E.F. Toma (2000) "Peer effects in private and public schools across countries" *Journal of Policy Analysis and Management* **19** (1): 75-92.

[29] Zimmerman, D. J. (2003) "Peer Effects in Higher Education: Evidence from a Natural Experiment" *Review of Economics and Statistics*, **85**(1): 9-23.