University of Leicester

DEPARTMENT OF ECONOMICS

ON THE OPTIMAL ALLOCATION OF STUDENTS WHEN PEER EFFECT WORKS: TRACKING VS MIXING

Marisa Hidalgo-Hidalgo, Universidad Pablo de Olavide, Spain

Working Paper No. 08/18
June 2008

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

Marisa Hidalgo-Hidalgo†

Universidad Pablo de Olavide.

June 3, 2008

Abstract

The belief that both the behavior and outcomes of students are affected by their peers is important in shaping education policy. I analyze two polar education systems -tracking and mixing- and propose several criteria for their comparison. I find that tracking is the system that maximizes average human capital in societies where the distribution of pre-school achievement is not very dispersed. I also find that when peer effects and individuals' pre-school achievement are close substitutes, all risk averse individuals prefer mixing.

Keywords: Human Capital, Efficiency, Peer Effects, Tracking, Mixing

JEL Classification: D63, I28, J24.

---

*Financial support from Spanish Ministry of Education and Science (SEJ 2007-67734), Junta de Andalucía (SEJ-2905) and Instituto Valenciano de Investigaciones Económicas is gratefully acknowledged. **Address for correspondence**: Marisa Hidalgo-Hidalgo, Departamento de Economia (Area Analisis Economico), Universidad Pablo de Olavide, Ctra. Utrera Km.1. E-41013, Sevilla, Spain. Email: mhidalgo@upo.es

†This paper is based on Chapter two of my Ph.D. 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-Ortín. I also acknowledge members of the Department of Economics at University of Leicester, where part of this work was done, for their hospitality.

1 Introduction

Peer effects are at the heart of many recent debates on educational reforms.[1] The critical importance to both parents and policy makers of peer group distribution in school is indisputable. Given the existence of peer effects, defined here as the effect on an individual's academic performance of the ability distribution of her peers, governments should 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 choose whether to stream (track) or mix students of differing abilities within public schools.

This paper analyzes, in a theoretical context, whether tracking or mixing students by ability is optimal. It contributes to this debate by addressing three main questions. First, it asks which system maximizes average human capital at the compulsory 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.

While the influence of peers ability on one's educational achievement is well documented, some relevant issues of this relationship are still being debated.[2] Recently, Hoxby (2000), Ammermueller and Piscke (2006), Ding and Lehrer (2007) and Kang (2007) find evidence of significant peer effects in achievement. While most studies focus on average innate ability within the classroom as the peer-based factor that most strongly impacts on individual achievement, Hoxby and Weingarth (2006) find that students are influenced by students who are similar to them. Finally, there is also evidence regarding the existence of non-linearities. Among others, Ding and Lehrer (2007) suggest that clever students benefit more from having clever mates than weak students do.

The peer group quality affects student achievement positively. However, raising peer quality for every student is an impossible task. From a policy point of view, the more relevant questions are concerned with efficiency issues: for whom does the

[1] See, for example, the 2006 NBER Fall Reporter.

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

peer group matter most? Do clever students or weak students profit more from being confronted with clever peers? To address these issues I introduce a model in which students differ in parental background as well as pre-school achievement. Wealthier individuals have more resources to invest in their kids. In addition wealthier individuals are more educated and care more about education which positively influences their children's achievement upon entering school.[3] Thus, in my model, the two characteristics that define the individual, family background and pre-school achievement, will be positively correlated as well. The production of human capital depends on both students' previous achievement and peer group characteristics. The degree of complementarity between both inputs is shown to be a critical issue in the comparison between tracking and mixing.

When comparing both educational systems, traditional methods focus on mean impacts. However, modern welfare economics emphasizes the importance of accounting for the impact of public policies on distributions of outcomes. My paper advances this literature beyond computing the average achievement under tracking and mixing to derive the distribution of achievement under each educational systems and compare them according to several criteria.

I find that, tracking is the system that maximizes average human capital in societies where the distribution of pre-school achievement is not very dispersed. In this case, the complementarity between peer effects and pre-school achievement drives the result. As some sources of heterogeneity among individuals appear and societies become more dispersed, for example because the pre-school achievement gap between rich and poor students increases, the complementarity effect dilutes: the benefits of the high-achievers do not compensate the losses of the lower-achievers and tracking might not be the system that maximizes average human capital. Then, mixing might maximizes average human capital in this case. I also find that the system that maximizes average human capital 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.

[3] The importance of child investment at early ages has been emphasized by Heckman (2006).

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' pre-school achievement. If they are nearly complementary, then there is no unanimously preferred system in the population. However, if they are close substitutes, it is mixing the system unanimously preferred in the population. In other words, when peer effects matter more for low achievers, then the distribution of human capital under mixing is less "spread" and thus can be considered less risky than the distribution of human capital under tracking.

There are several papers related to this. Epple, Newlon and Romano (2002) study the effects of ability grouping on school competition. They examine the consequences of tracking for the allocation of students of differing abilities and income within and between public schools. De Bartolome (1990) proposes a community model where public-service output depends on input expenditures, on own personal characteristics, and on the peer group effect. He shows that communities may become heterogeneous in composition and (second-best) inefficient and that this equilibrium occurs when the peer group effect is neither "too strong" nor "too weak". Arnott and Rowse (1987) is the paper most related to this one. They analyze the optimal allocation of students and resources when peer effects are present by focusing on the degree of concavity of the peer group effect. However, they fail to consider the existence of a positive dependence between family background and individuals' characteristics and its role in the process of human capital accumulation which is one of the main focus of this paper. They conclude that, when the objective is to maximize mean performance, the optimal allocation of students abilities depends on the properties of the education production function. However they also admit that the (narrow) focus on the degree of concavity of the peer group effects prevents them from seeing the possible dependence of individual welfare on the whole shape of human capital distribution in the population. In this paper, and in line with the most recent empirical evidence, I assume concavity in peer effects and discuss how the complementarity between peer characteristics and individuals' characteristics (a point for which the empirical evidence is still quite mixed) can determine which system maximizes average human capital. In addition to this, my approach contributes to the relevant literature by comparing both systems in terms of the induced distributions of human capital at the end of compulsory school.

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 induced distributions of human capital in these two systems. Section 4 concludes.

2 Model

2.1 Individuals

Population size is 1. Individuals differ in two aspects: their family background and their pre-school achievement, θ_0 where $\theta_0 \in [0, 1]$. To make the model tractable, I assume that family background takes only two values, that is, individuals can have either poor or rich parents with probabilities $1 - \lambda$ and λ, respectively.[4] I denote by $g_b(\theta_0)$ the p.d.f. (probability distribution function) of θ_0 conditional on having family background b, where $b = p, r$ for poor and rich parents respectively. To capture the possibility that some level of positive dependence exists between parental background and pre-school achievement, I assume that $g_p(\theta_0) = \gamma\theta_0^{\gamma-1}$ and $g_r(\theta_0) = 1$, where $\gamma \in (0, 1]$. In Figure 1 I represent in dotted line the C.D.F. of θ_0 for both poor (in black) and rich individuals (in grey) and in solid black line the C.D.F. of θ_0 for the whole population when $\gamma < 1$.

Here Figure 1 (The distribution of pre-school achievement)

Thus, the C.D.F. of pre-school achievement, denoted by $G(\theta_0)$, can be expressed as:

$$G(\theta_0) = \begin{cases} (1-\lambda)\theta_0^{\gamma} + \lambda\theta_0 & if \quad \theta_0 \leq 1 \\ 1 & if \quad \theta_0 > 1. \end{cases} \tag{1}$$

That is, the conditional mean of pre-school achievement depends on parental background. The lower is γ, the higher is the gap in pre-school achievement between poor and rich people. In addition, from (1) we have that pre-school achievement

[4] Alternatively we could interpret the two parent types as black or white, natives or immigrants, etc.

dispersion, measured by the coefficient of variation of θ_0 in the population $cv_{\theta_0}(\gamma, \lambda)$ is:

$$cv_{\theta_0}(\gamma, \lambda) = (1-\lambda)\frac{1}{\sqrt{\gamma(\gamma+2)}} + \lambda\frac{1}{\sqrt{3}}, \tag{2}$$

which is strictly decreasing with γ and λ. That is, as either the pre-school achievement gap or the proportion of poor students decreases the distribution of pre-school achievement becomes less dispersed. Below we analyze the effect of pre-school achievement dispersion on the average human capital under both education systems.

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

2.2 Production of Human Capital

At compulsory level individuals are separated into different groups or classes. To simplify, I consider only two groups. The production of human capital depends on two factors. The first is the individual's pre-school achievement, θ_0. The second is the "peer group" effect that depends on the characteristics of the group in which the individual is placed. These characteristics are summarized by the mean achievement of the group j or "peer" effect, denoted by $\overline{\theta}_0^j$. After attending compulsory education, an individual with pre-school achievement θ_0 ends up with a level of human capital θ_1.[5]

I assume that the production of human capital is a CES of the two inputs, θ_0 and $\overline{\theta}_0^j$. The empirical evidence regarding the relationship between individuals' ability and peer group effect is still mixed. Hoxby and Weingarth (2006) cannot reject the hypothesis according to which high and low-achieving students benefit equally from the presence of high achieving students. However, Ding and Lehrer (2007) find that high-ability students benefit more from having higher-achieving schoolmates than students of lower ability. Thus, I select the following functional form in the analysis in order to study how the complementarity between peers' effect and individuals'

[5] See, among others, Bishop (2006), Epple and Romano (1998), Epple, Romano and Sieg (2003) and Epple, Newlon and Romano (2002) who also assume that peers affect an individual through the mean of their characteristics.

ability affects the comparison between tracking and mixing. In particular:

$$\theta_1(\theta_0, \overline{\theta}_0^j) = A(\rho\theta_0^\beta + (1-\rho)(\overline{\theta}_0^j)^\beta)^{\frac{1}{\beta}}, \tag{3}$$

where $A > 1$, $\rho \in [0,1]$ and $\beta \in (0,1]$. The parameter ρ captures the weight of pre-school achievement on θ_1. The final level of human capital θ_1, is a twice differentiable, increasing and concave function. That is, on the one hand the positive impact of an increase in mean achievement is decreasing, this resulting in an efficiency loss from tracking.[6] On the other hand, Equation (3) allows for the possibility that $\overline{\theta}_0^j$ and θ_0 are either complements or substitutes, since β determines the elasticity of substitution between these two inputs. Note here that, for any $\beta < 1$ we have that $\frac{\partial^2\theta_1}{\partial\theta_0\partial\overline{\theta}_0^j} > 0$, that is, high-achievers benefit most from an increase in mean achievement, which implies that there would be an efficiency gain from tracking.[7] Therefore, Equation (3) clearly sets up a tension between mixing and tracking.[8]

Finally, note that Hoxby and Weingarth (2006) find that the peer effect depends on students' characteristics. In particular they provide some support for a specification in which homogeneity is good, that is, every student learns the most when he or she is with students like him or her(see also Manski and Wise (1983)). In this sense, Equation (3) captures the main features of peer effects on weak students' achievement found by Hoxby and Weingarth (2006). As long as $\beta < 1$, since weak students are closer to the mean within the group than very weak students, the impact of peers is higher for the former than for the latter. Although (3) does not capture the peer impact found by Hoxby and Weingarth (2006) regarding good and very good students, observe that, in general, weak students seem to be the main concern of recent education reforms.[9] Finally note that this pro-homogenity specification would underlie support for tracking. However, even by ruling this possibility out to some

[6] The empirical evidence suggests 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 (see Ding and Lehrer (2007) and Hoxby and Weingarth (2006)).

[7] In particular, for β close to 0, both $\overline{\theta}_0^j$ and θ_0 have some level of complementarity and as β tends to 1 the two factors become perfect substitutes.

[8] See also Benabou (1996) who, in a model of local public finance and community formation, analyzes this general trade-off between complementarity and curvature.

[9] This is clearly the objective that underlies some recent educational policies in the US as for example, the No Child Left Behind Act.

degree, and as we will see below, I get that tracking performs better that mixing in most of cases. Thus, such a specification would only reinforce my main results without adding additional insight.

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 pre-school achievement distribution is the same in both classrooms. The average pre-school achievement within each classroom, denoted $\overline{\theta}_0^m$, coincides with the average pre-school achievement in the population:

$$\overline{\theta}_0^m(\gamma, \lambda) = (1-\lambda)\left(\frac{\gamma}{\gamma+1}\right) + \frac{\lambda}{2}. \tag{4}$$

It can be checked that the average pre-school achievement is increasing with λ the proportion of rich individuals and also with γ, that is, it is decreasing with the pre-school achievement gap. In addition $\overline{\theta_0}^m(\gamma, \lambda) \leq 1/2$ for any γ and λ.

Under mixing, θ_1 will lie in the support $[\underline{m}, \overline{m}]$ where $\underline{m}$ and $\overline{m}$ denote the level of human capital θ_1 acquired under mixing by the "worst" (lowest pre-school achiever) and the "best" (highest pre-school achiever) individual in the population, respectively:

$$\begin{aligned} \underline{m}(\gamma, \lambda) &= A(1-\rho)^{\frac{1}{\beta}}(\overline{\theta}_0^m). & (5) \\ \overline{m}(\gamma, \lambda) &= A(\rho + (1-\rho)(\overline{\theta}_0^m)^\beta)^{\frac{1}{\beta}}. & (6) \end{aligned}$$

Therefore, the C.D.F. of θ_1 under mixing, denoted $F_M(\theta_1)$, is:

$$F_M(\theta_1) = \begin{cases} (1-\lambda)\varphi(\theta_1, \overline{\theta}_0^m)^\gamma + \lambda\varphi(\theta_1, \overline{\theta}_0^m) & if \quad 0 \leq \theta_1 \leq \overline{m} \\ 1 & if \quad \theta_1 > \overline{m}, \end{cases} \tag{7}$$

where $\varphi(\theta_1, \overline{\theta}_0^m) = \left(\frac{1}{\rho}((\frac{\theta_1}{A})^\beta - (1-\rho)(\overline{\theta}_0^m)^\beta)\right)^{\frac{1}{\beta}}$.

It can be checked that, ceteris paribus, if in society A parameter γ is higher than in society B, then the distribution of human capital under mixing $F_M(\theta_1)$ in B will be dominated by the one in A. This is because average pre-school achievement $\overline{\theta}_0^m(\gamma,\lambda)$ (which is higher in A than in B) is the only determinant of the difference in human capital between both societies. In other words, if in A the gap in pre-school achievement is lower than in B, because A implements more effective policies in reducing it at early educational stages than B, then the distribution of human capital under mixing $F_M(\theta_1)$ in A dominates the one in B.

As shown in Figure 2 below, an increase in γ implies an increase in the expected value of θ_1 under mixing. Here, the case $\gamma = 1/4$ is represented in solid line and $\gamma = 3/4$ in dashed line:

Here Figure 2 (The distribution of θ_1 under Mixing)

I denote by $E_M(\theta_1)$ the expected value of θ_1 under mixing, where:

$$\begin{aligned} E_M(\theta_1) &= \int_{\underline{m}}^{\overline{m}} \theta_1 f_M(\theta_1) d\theta_1 \\ &= \int_{\underline{m}}^{\overline{m}} \frac{1}{\rho}\left(\frac{\theta_1}{A\varphi(\theta_1,\overline{\theta}_0^m)}\right)^{\beta}\left((1-\lambda)\gamma\varphi(\theta_1,\overline{\theta}_0^m)^{\gamma}+\lambda\varphi(\theta_1,\overline{\theta}_0^m)\right)d\theta_1, \quad (8)\end{aligned}$$

and $f_M(\theta_1)$ denotes the p.d.f (probability distribution function) of θ_1 under mixing.

From (8) and Figure 2 it can be checked that $E_M(\theta_1)$ is an increasing function of λ, the wealth level in the population, and of γ, as we saw above.

2.3.2 Tracking

Tracking students implies grouping them on the basis of pre-school achievement. For the sake of simplicity I permit only two tracks and I use the median level of pre-school achievement as a threshold for grouping students into one track or the other. Thus, a student is assigned to the high (low) track when his/her pre-school achievement θ_0 is above (below) the median, denoted by $\eta(\gamma,\lambda)$. Thus, $\eta(\gamma,\lambda)$ is such that $G(\eta) = 1/2$.

From Figure 1 we see that the median $\eta(\gamma,\lambda)$ is increasing in λ and γ. In addition, note that the distribution of pre-school achievement θ_0 is right-skewed, that is $\eta(\gamma,\lambda) < \overline{\theta}_0^m(\gamma,\lambda)$ for any $\lambda \in (0,1)$ and $\gamma \in (0,1]$. Thus, $\eta(\gamma,\lambda) \leq 1/2$ for any $\lambda \in (0,1)$ and $\gamma \in (0,1]$ implies that the number of rich students in the high track will be higher than the number of rich students in the population. This captures the empirical evidence found by Brunello and Checchi (2007) among others regarding the socioeconomic composition of the different tracks.

I denote by $\overline{\theta}_0^l$ and $\overline{\theta}_0^h$ the average pre-school achievement in the low and high track, respectively. Thus, given the distributional assumptions on θ_0, I have that:

$$
\begin{aligned}
\overline{\theta}_0^l(\gamma,\lambda) &= (1-\lambda)\frac{\int_0^{\eta(\gamma,\lambda)} \theta_0 g_p(\theta_0)d\theta_0}{\int_0^{\eta(\gamma,\lambda)} g_p(\theta_0)d\theta_0} + \lambda\frac{\eta(\gamma,\lambda)}{2} \\
&= (1-\lambda)\left(\frac{\gamma}{\gamma+1}\right)\eta(\gamma,\lambda) + \lambda\frac{\eta(\gamma,\lambda)}{2} \\
&= \eta(\gamma,\lambda)\overline{\theta}_0^m(\gamma,\lambda), \qquad (9)
\end{aligned}
$$

and that:

$$
\begin{aligned}
\overline{\theta}_0^h(\gamma,\lambda) &= (1-\lambda)\frac{\int_{\eta(\gamma,\lambda)}^1 \theta_0 g_p(\theta_0)d\theta_0}{\int_{\eta(\gamma,\lambda)}^1 g_p(\theta_0)d\theta_0} + \lambda\left(\frac{\eta(\gamma,\lambda)+1}{2}\right) \\
&= (1-\lambda)\left(\frac{\gamma}{\gamma+1}\right)\left(\frac{1-\eta(\gamma,\lambda)^{\gamma+1}}{1-\eta(\gamma,\lambda)^{\gamma}}\right) + \lambda\left(\frac{\eta(\gamma,\lambda)+1}{2}\right). \qquad (10)
\end{aligned}
$$

It can be checked from (9) and (10) that the average pre-school achievement in both the low and the high track is increasing with the median level of pre-school achievement $\eta(\gamma,\lambda)$. Thus, both $\overline{\theta}_0^l(\gamma,\lambda)$ and $\overline{\theta}_0^h(\gamma,\lambda)$ will increase as γ increases. In other words, as the pre-school achievement gap between rich and poor students decreases, the average human capital increases in both the low and the high track.

In the low track, θ_1 lies within the interval $[\underline{l},\overline{l}]$. We denote by $\underline{l}$ and $\overline{l}$ the human capital θ_1 acquired in the low track by the "worst" (lowest pre-school achiever) and the "best" (highest pre-school achiever) individual respectively, that is:

$$
\begin{aligned}
\underline{l}(\gamma,\lambda) &= A(1-\rho)^{1/\beta}\overline{\theta}_0^l. \qquad (11) \\
\overline{l}(\gamma,\lambda) &= A(\rho\eta^\beta + (1-\rho)(\overline{\theta}_0^l)^\beta)^{1/\beta}. \qquad (12)
\end{aligned}
$$

Likewise, in the high track, θ_1 lies within the interval $[\underline{h},\overline{h}]$. We denote by $\underline{h}$ and $\overline{h}$ the human capital θ_1 acquired in the high track by the "worst" (lowest pre-school achiever) and the "best" (highest pre-school achiever) individual, respectively, that is.:

$$\underline{h}(\gamma,\lambda) = A(\rho\eta^{\beta} + (1-\rho)(\overline{\theta}_0^h)^{\beta})^{1/\beta}. \tag{13}$$

$$\overline{h}(\gamma,\lambda) = A(\rho + (1-\rho)(\overline{\theta}_0^h)^{\beta})^{1/\beta}. \tag{14}$$

It can be checked from (12) and (13) above that the support of θ_1 in the low track does not overlap the support of θ_1 in the high track, that is, $\underline{h}(\gamma,\lambda) > \overline{l}(\gamma,\lambda)$ for every $\lambda \in (0,1)$ and $\gamma \in (0,1)$.

The C.D.F. of θ_1 under tracking, denoted by $F_T(\theta_1)$, is:

$$F_T(\theta_1) = \begin{cases} (1-\lambda)\varphi(\theta_1,\overline{\theta}_0^l)^{\gamma} + \lambda\varphi(\theta_1,\overline{\theta}_0^l) & if \quad \underline{l} \le \theta_1 \le \overline{l} \\ 1/2 & if \quad \overline{l} \le \theta_1 \le \underline{h} \\ (1-\lambda)\varphi(\theta_1,\overline{\theta}_0^h)^{\gamma} + \lambda\varphi(\theta_1,\overline{\theta}_0^h) & if \quad \underline{h} \le \theta_1 \le \overline{h} \\ 1 & if \quad \theta_1 > \overline{h}, \end{cases} \tag{15}$$

where $\varphi(\theta_1,\overline{\theta}_0^j) = \left(\frac{1}{\rho}((\frac{\theta_1}{A})^{\beta} - (1-\rho)(\overline{\theta}_0^j)^{\beta})\right)^{\frac{1}{\beta}}$ for $j = l, h$.

As under mixing, and ceteris paribus, if in society A the level of γ is higher than in society B, then the distribution of human capital under tracking $F_T(\theta_1)$ in A will dominate the one in B. Figure 3 represents the case $\gamma = 1/4$ in solid line and $\gamma = 3/4$ in dashed line:

Here Figure 3 (The distribution of θ_1 under Tracking)

The intuition of the previous result is as follows. As the value $\eta(\gamma,\lambda)$ is higher in A than in B, from (9) and (10) we have that the average pre-school achievement in both the low track, $\overline{\theta}_0^l$, and the high track, $\overline{\theta}_0^h$, will also be higher in A than in B. This ensures that the human capital acquired by the students is higher in both tracks. Consequently, as can be checked from (15) and Figure 3, $F_T(\theta_1)$ will be lower in A than in B.

We can conclude that, similar to mixing, if A spends more resources than B on early childhood education then, under tracking, the distribution of human capital in B is dominated by that in A.

The expected value of θ_1 under tracking is:

$$E_T(\theta_1) = \int_{\underline{l}}^{\overline{l}} \theta_1 f_T(\theta_1) d\theta_1 + \int_{\underline{h}}^{\overline{h}} \theta_1 f_T(\theta_1) d\theta_1, \tag{16}$$

where $f_T(\theta_1)$ denotes the p.d.f (probability distribution function) of θ_1 under tracking. Again, the expected value of θ_1 under tracking is increasing in both λ and γ.

3 A comparison of mixing and tracking

First I consider that the educational system is chosen by majority voting and that every individual votes 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 $\theta_0 < \eta$), since under tracking they would be placed into the low track, where they would enjoy a lower peer effect. The other half will prefer tracking (those with $\theta_0 > \eta$), since they would be 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 no system will defeat the other under a majority voting rule.

Second, I consider the government wants to maximize the utility of the worst-off individuals in the society, as might be suggested by some recent education policies in the US, as the No Child Left Behind Act cited above. To do this we have to define first who are the worst-off in our model. If, for example, we take as the worst-off those with pre-school achievement below the median level and with poor parents, the result is quite immediate. Mixing is always better. This comes directly from the properties of the human capital production function (Equation (3)), since maximizing the utility of these individuals will imply to maximize their human capital at compulsory level.[10]

Finally I consider that individuals choose the education system behind the so-called "veil of ignorance". That is, they choose between societies without knowing

[10]Note that this applies to all the individuals with $\theta_0 < \eta(\gamma, \lambda)$, except for that individual with $\theta_0 = 0$.

where they will be placed or what characteristics they will have in each society. The idea of choosing from behind a veil of ignorance, to reflect fairness of societies, has proved very useful in theoretical economics (see the seminal works of Harsanyi (1953 and 1955) and Rawls (1971) and more recently Cremer and Pestieau (1998)) and in empirical (see Johansson-Stenman et al (2002) and Carlsson et al (2003) among others).[11] A possible implication of this approach is that, individuals when choosing from behind the veil of ignorance would choose the alternative that maximizes expected utility or, equivalently they would unanimously agree on the alternative that maximizes a utilitarian welfare function (see Harsanyi (1953)). In my model, to choose an education system from behind the veil of ignorance implies that individuals ignore both the value of θ_0 that they will end up enjoying and their parents' background.[12]

One possibility is just to compare the two systems in terms of average human capital. This is like assuming that all individuals are risk neutral behind the veil of ignorance. I present now and discuss the results regarding this comparison using numerical simulations. The most important result is that the difference between average human capital under the two systems, $E_T(\theta_1) - E_M(\theta_1)$, decreases with β. The following table presents the value of β, for some values of both γ and λ, such that $E_T(\theta_1) - E_M(\theta_1) = 0$, denoted it by $\widehat{\beta}(\gamma, \lambda)$. Thus, for β below (above) $\widehat{\beta}(\gamma, \lambda)$ we have that $E_T(\theta_1) - E_M(\theta_1) > (<) 0$:[13]

[11] Johansson-Stenman et al (2002) and Carlsson et al (2003) analyze the individuals' choice between alternative societies with different income distributions behind a veil of ignorance. In both experiments they instructed the respondents to consider the well-being of their *imaginary grandchild*, that is, to choose the alternative that would be in the interest of their grandchild.

[12] Note that, since $\eta(\gamma, \lambda) < 1/2$, if individuals knew theirs parents' background they would also knew at which track they would be assigned to with higher probability under tracking.

[13] The Coleman Report as well as recent works (see for example Heckman (2006) and references therein) show that families and not schools are the major sources of inequality in school performance, this implying that ρ should be high enough. Finally note that, for example, $\gamma = 0.25$ means that the mean pre-school achievement within the poor represents the 40% of the mean pre-school achievement within the rich.

Table 1. Average Human Capital: $\widehat{\beta}(\gamma,\lambda)$ [a]

	Pre-school achievement gap: γ				
λ	0.2	0.25	0.3	0.35	0.4
1/5	0.930	0.960	0.979	0.992	0.999
1/4	0.931	0.965	0.982	0.996	1
1/3	0.950	0.980	0.997	1	1

(17)

[a] Note that here A=2 and ρ=3/4.

We can conclude that as long as β is not very large, i.e., when $\overline{\theta}_0^j$ and θ_0 have some level of complementarity, then average human capital is always maximized under tracking. When β is close to 1, meaning that the two factors are close 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.

The second lesson we can extract from Table 1 is that as the level of dispersion in pre-school achievement decreases, measured by either an increase in γ or λ (see Equation (2)), tracking maximizes average human capital for a larger interval of values of β. In fact if either γ or λ are sufficiently high then tracking maximizes average human capital for every $\beta \in (0,1]$. Observe that, if either $\gamma = 1$ or $\lambda = 1$ then, from (1) we have that $G(\theta_0) = \theta_0$. That is, the pre-school achievement distribution is uniform and the same for both income groups. In this case, it can be checked that $E_T(\theta_1) - E_M(\theta_1) > 0$ for any ρ and $\beta \in (0,1]$. Clearly, what drives the previous result is the complementarity between the peer group effect and the pre-school achievement level. However, as either γ or λ falls below 1, the dispersion in the pre-school achievement distribution increases and the complementarity effects dilutes. Note that the pre-school achievement distribution of rich students does not change with either γ or λ. As a result, the increased dispersion is a result of either a decrease in the mean pre-school achievement of poor individuals or just an increase in the proportion of poor individuals, whose mean pre-school achievement is lower than that of rich students. Under tracking therefore, those placed in the high track still benefit but their gains might not always compensate the losses of those placed in the low track. As a result, mixing might maximizes average human capital in those cases.

Figure 4 below represents combinations (γ, λ) giving rise to the same value of $\widehat{\beta}$. Recall that $\widehat{\beta}(\gamma, \lambda)$ is the level of complementarity between peer group effects and individual pre-school achievement such that the average human capital under both systems coincides. As it can be checked from Table 1, $\widehat{\beta}(\gamma, \lambda)$ is increasing with both γ and λ. In other words, as society becomes less dispersed in terms of the pre-school achievement distribution, because either the difference in mean pre-school achievement between rich and poor or the proportion of poor individuals decreases, then $\widehat{\beta}(\gamma, \lambda)$ increases. That is, it is required a higher level of sustitutability between peer group effects and individual pre-school achievement, to get mixing as the system that maximizes average human capital.

Here Figure 4 (Average Human Capital)

The general message we can extract from Table 1 and Figure 4 is that tracking is the system that maximizes average human capital in societies where the pre-school achievement is not dispersed. As some sources of dispersion in pre-school achievement appear, because either the gap between poor and rich students or the proportion of poor students increases, then mixing might be better than tracking in maximizing average human capital. This result contrasts to that of Kremer and Maski (1996), where they find that increases in skill-dispersion promote segregation of workers by skill. However, note that they do not consider the existence of peer effects which is a crucial input in explaining the role of skill dispersion on the optimal allocation of individuals with different skill levels. As I have shown above, less dispersion reinforces the complementarity effect between individuals' initial achievement and the peer effect which, in turn, induces segregation among individuals. By contrast, more dispersion dilutes the aforementioned complementarity effect and thus induces integration among individuals.

A last possibility, which has not been previously considered in the literature, is to compare both systems in terms of the whole distribution of human capital. Recall that if the distribution of human capital 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.[14]

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 γ,β and ρ.*

Proof. (i) $F_T(\theta_1) \not\succeq_{FOSD} F_M(\theta_1)$. Using $F_T(\theta_1)$ from (15) and $F_M(\theta_1)$ from (7) we can check that, for any $\theta_1 \in (0, \bar{l}]$, $(F_T(\theta_1) - F_M(\theta_1)) > 0$ for every λ, γ, β and ρ. (ii) $F_M(\theta_1) \not\succeq_{FOSD} F_T(\theta_1)$. Using Equations (15) and (7), we can check that for any $\theta_1 \in [\underline{h}, \overline{h}]$, $(F_T(\theta_1) - F_M(\theta_1)) < 0$ for every λ, γ, β and ρ. ■

Figure 5 illustrates the previous result, where $F_M(\theta_1)$ and $F_T(\theta_1)$ are represented in solid and dashed 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.

Here Figure 5 (No First Order Stochastic Dominance)

Finally, I will consider 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 criterion 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 pre-school achievement, β.[15]

[14] Recall that it is implicitly assumed here that individuals maximize expected utility that depends on human capital.

[15] Note that Proposition 2 holds if and only if F_M and F_T cross only once, which is true from Proposition 1.

Proposition 2 *The preferred system according to second order stochastic dominance depends on β as follows:*
(i) If $\beta < \widehat{\beta}$ then neither system dominates the other.
(ii) If $\beta > \widehat{\beta}$ then mixing dominates tracking.

Proof. First note that $F_T(\theta_1) \not\succeq_{SOSD} F_M(\theta_1)$. Using $F_T(\theta_1)$ from (16) and $F_M(\theta_1)$ from (7) we can check that, $\int_0^{\overline{l}} (F_T(\theta_1) - F_M(\theta_1))d\theta_1 > 0$, for every ρ and γ. Now recall that the expected value of a random variable y can be written as: $E(y) = \overline{y} - \int_0^{\overline{y}} F(y)dy$, 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: $E_T(\theta_1) = \overline{h} - \int_0^{\overline{h}} F_T(\theta_1)d\theta_1$ and, under mixing $E_M(\theta_1) = \overline{m} - \int_0^{\overline{m}} F_T(\theta_1)d\theta_1 = \overline{h} - \int_0^{\overline{h}} F_M(\theta_1)d\theta_1$. Finally note that, if $F_M(\theta_1) \succeq_{SOSD} F_T(\theta_1)$, then the following inequality should hold: $\overline{h} - E_M(\theta_1) \leq \overline{h} - E_T(\theta_1)$. The final result is immediate from Table 1 and Table 2. ∎

This proposition says that when peer effects and pre-school achievement are close substitutes, all averse individuals prefer mixing. That is, when peer effects matter more for low achievers than for high achievers individuals then the distribution of human capital under mixing is less "spread" and thus can be considered less risky than the distribution of human capital under tracking.

Finally note that $\widehat{\beta}(\gamma, \lambda)$ is increasing with γ, ρ and λ. That is, for example, as a society becomes more equal in terms of pre-school achievement between poor and rich individuals, then it is required a higher level of sustitutability between peer effect and pre-school achievement in the production of human capital, to get mixing as the preferred system. In other words, as the government implements policies to reduce the gap in pre-school achievement, mixing will be less and less preferred to tracking.

4 Concluding Remarks

In this paper I have analyzed public intervention in education when the government, taking into account the existence of peer effects, has to decide how to group students. I have considered two different education systems: tracking and mixing.

A number of previous works have studied the optimal education system by focusing on mean achievement. This paper contributes to this line of research by recognizing the existence of a positive dependence between family background and individuals' pre-school achievement and its effect on each of the two educational systems described above. In addition to that, this paper contributes to this literature by comparing the distribution of human capital under each educational system according to several criteria.

The main result of the paper is that tracking is the education system that maximizes average human capital in societies where the distribution of pre-achievement is not very dispersed. As some sources of heterogeneity among individuals appear and societies become more dispersed, for example because the pre-school achievement gap between rich and poor students increases, then mixing becomes the education system that maximizes average human capital.

I have showed that, among risk averse individuals, the preference for mixing versus tracking depends on the degree of complementarity between peer effects and individuals' characteristics. If they are nearly complements then there is no preferred system in the population. However, if they are close substitutes then mixing is the system unanimously preferred.

This paper allows for some extensions. An important one is the introduction of prices which are omitted in this paper under the assumption of free education in both systems. It would be interesting to consider them in the model. 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 to design an optimal educational system. On the other hand, it would be interesting to explore how the education system introduced at compulsory level, either tracking or mixing, can influence students decisions as whether to attend college or not (see Hidalgo-Hidalgo (2005) for a similar analysis with a more stylized model).

Finally I think that the results presented here are relevant for several recent debates in the literature of economics of education. There is increasing evidence that shows the early emergence and persistence of gaps in cognitive and non-cognitive skills (see among others, Carneiro and Heckman (2003)). Studies that highlight the importance of increasing expenditure in early childhood care in pursuing both efficiency and equity provide an interesting illustration. As I have showed in this paper, a government, while reducing the gap in pre-school achievement between rich and poor students, should also choose very carefully the way of grouping them in order to maximize average human capital or to find the preferred education system in the population. Another example is the literature that looks at the heterogeneity in grouping policies across countries and tries to explain it (see Brunello et al (2005) and Ariga et al (2005) among others). As Brunello et al (2005) pointed out, efficiency considerations are not enough in explaining the existing differences that instead might be driven by some distributional concerns of society.

References

[1] Ammermueller, A. and J.S. Pischke (2006): "Peer Effects in European Primary Schools: Evidence from PIRLS," NBER Working Papers No. 12180.

[2] Ariga, K., G. Brunello, R. Iwahashi and L. Rocco (2005): "Why is the timing of school tracking so heterogenous?" IZA Discussion Paper No. 1854.

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

[4] Benabou, R. (1996): "Efficiency and equity in human capital investment: the local connection" *Review of Economic Studies*, 63 (2): 237-264.

[5] Bishop, J. (2006): "Drinking from the fountain of knowledge: student incentive to study and learn externalities, information problems and peer pressure". Handbook of Economics of Education, Vol. 2. Edited by Eric A. Hanushek and Finis Welch. Elsevier.

[6] Brunello, G. M. Giannini and K. Ariga (2005): "The optimal timing of school tracking" in P.Peterson and L. Woessmann, *Schools and the Equal Opportunity Problem*, MIT Press, Cambridge USA.

[7] Brunello, G. and D. Checchi (2007): "Does school tracking affect equality of opportunity? New international evidence" *Economic Policy*, CEPR, CES, MSH, vol. 22, pages 781-861.

[8] Carlsson, F., G. Guptaz and O. Johansson-Stenman (2003): "Choosing from behind a veil of ignorance in India" *Applied Economics Letters*, **10:** 825–827.

[9] Carneiro, P. and J.J. Heckman (2003), in Inequality in America: What Role for Human Capital Policies?" J.J. Heckman, A.B. Krueger, B. Friedman, Eds. MIT Press, Cambridge, MA. Ch. 2, 77-237.

[10] Cremer, H. and P. Pestieau (1998): "Social insurance, majority voting and labor mobility," *Journal of Public Economics*, **68**(3): 397-420

[11] 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.

[12] Ding, W. and S.F. Lehrer (2007): "Do peers affect student achievement in China's secondary schools?" *Review of Economics and Statistics*, 89 (2): 300-312.

[13] Epple, D. and R.E. Romano (1998): "Competition between private and public schools, vouchers and peer group effects" *American Economic Review*, 88(1):33-62.

[14] 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.

[15] Epple, D, R.E. Romano and H. Sieg (2003): "Peer Effects, Financial Aid, and Selection of Students into Colleges and Universities: An Empirical Analysis," *Journal of Applied Econometrics (Special Issue on Empirical Analysis of Social Interactions)*, 18: 501-525.

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

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

[18] Heckman, J. J. (2006): "Skill formation and the economics of investing in disadvantaged children" *Science* **312**(5782): 1900-1902.

[19] Hidalgo-Hidalgo, M. (2005): "Peer group effects and optimal education system" IVIE WP-AD 2005-12.

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

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

[22] Johansson-Stenman, O., Carlsson, F. and Daruvala, D. (2002): "Measuring future grandparents' preferences for equality and relative standing" *Economic Journal,* **112:** 362–83.

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

[24] Kremer, M. and E. Maskin (2008): "Wage inequality and segregation by skill" NBER Working Papers No. 5718, *forthcoming Quarterly Journal of Economics.*

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

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

[27] Rawls, J. (1971) "A Theory of Justice", Harvard University Press, Cambridge MA.

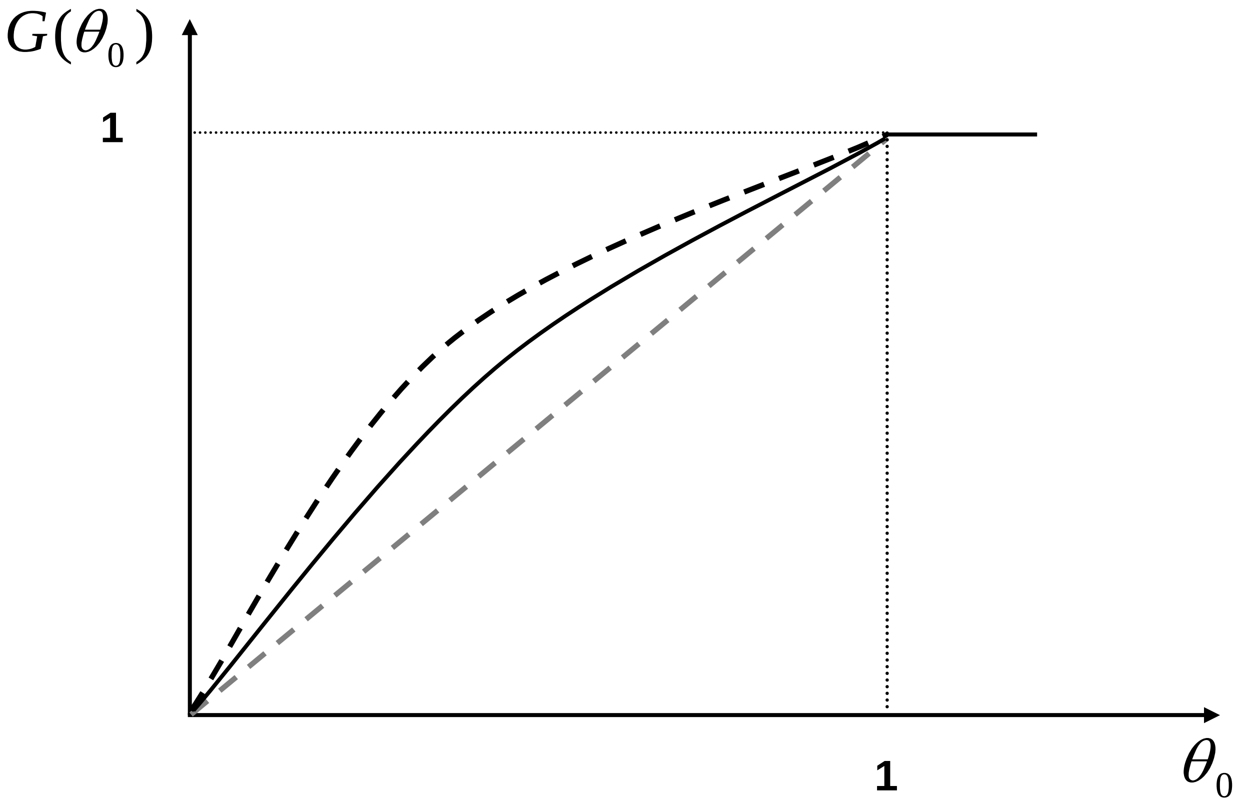

Figure 1: The distribution of pre-school achievement

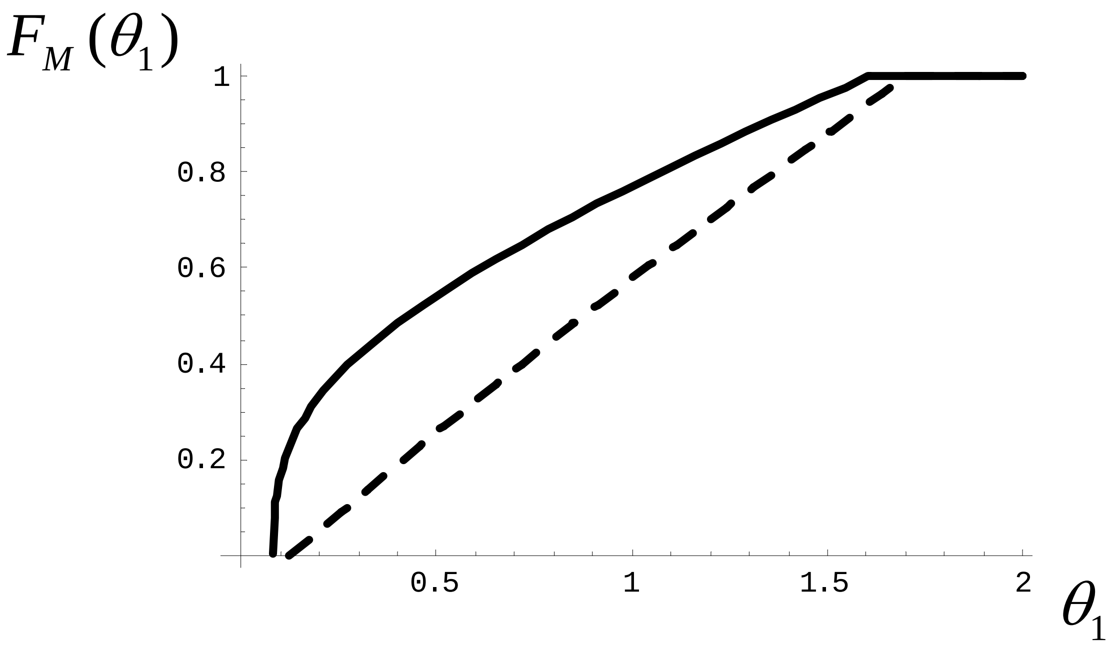

Figure 2: The distribution of θ_1 under Mixing

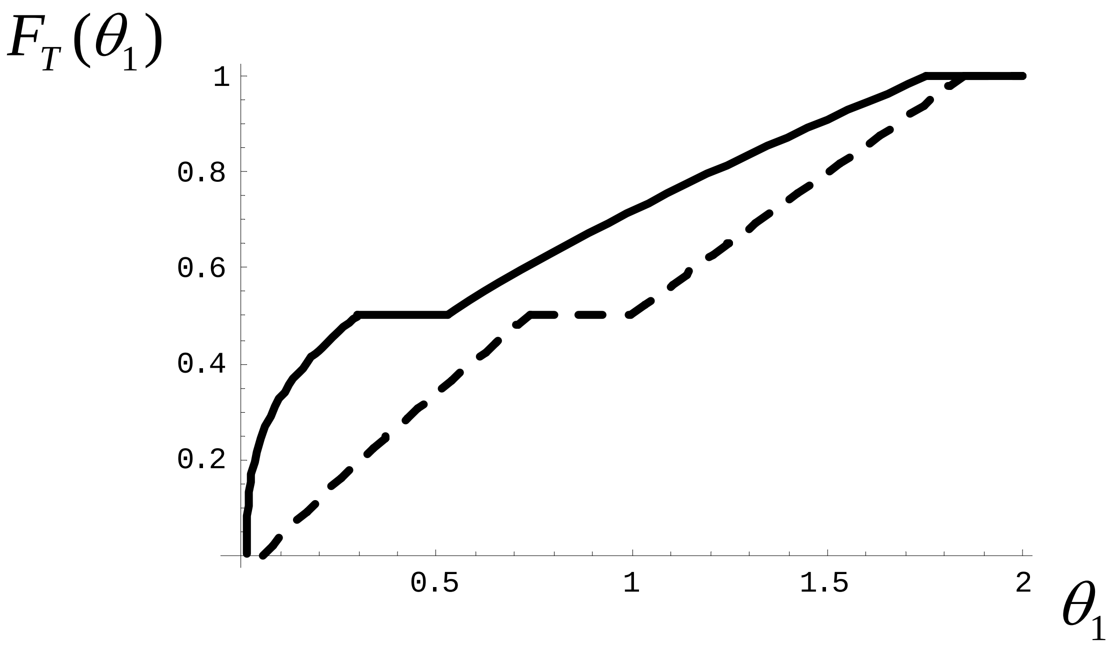

Figure 3: The distribution of θ_1 under Tracking

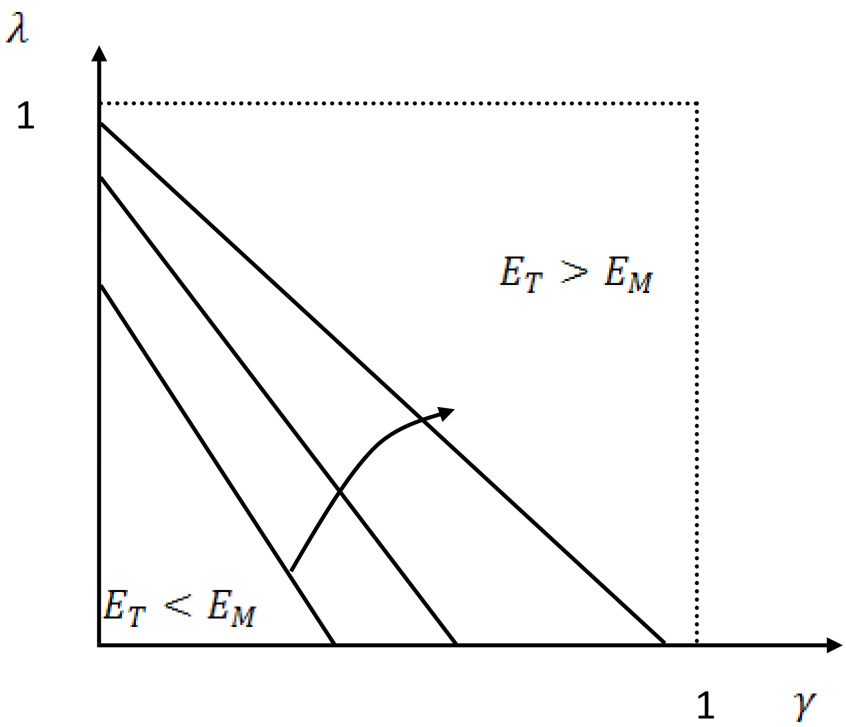

Figure 4: Average Human Capital

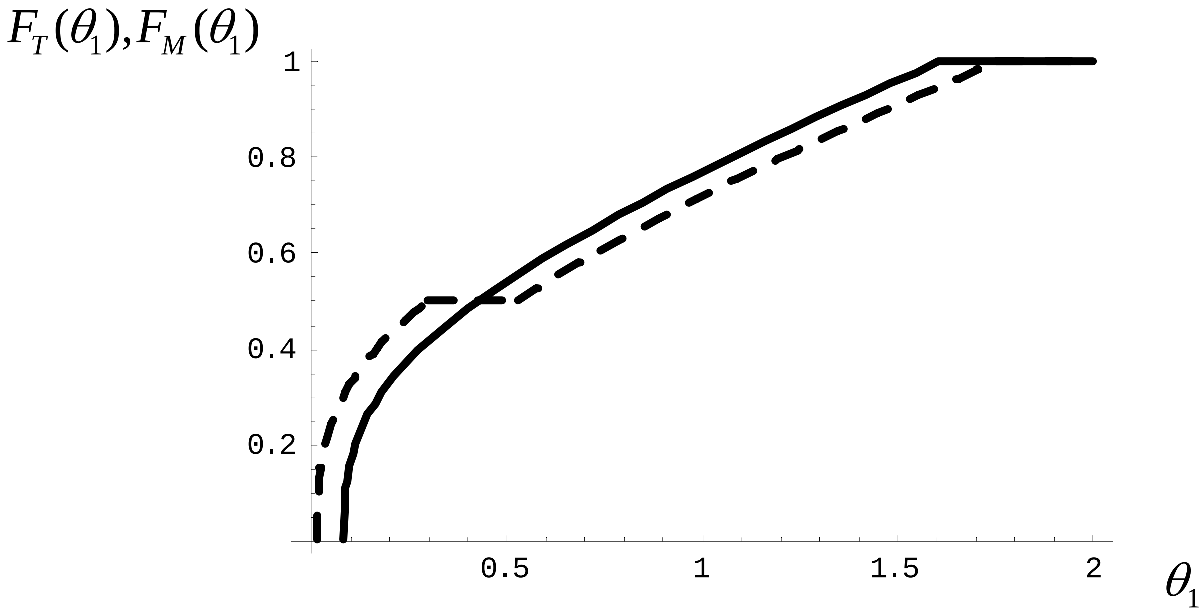

Figure 5: No First Order Stochastic Dominance