Center for
European,
Governance and
Economic Development Research

cege

Discussion Papers

Number 289 – August 2016

# ACADEMIC ACHIEVEMENT AND TRACKING – A THEORY BASED ON GRADING STANDARDS

Tim Ehlers
Robert Schwager

Georg-August-Universität Göttingen

ISSN: 1439-2305

# Academic Achievement and Tracking – A Theory Based on Grading Standards

Tim Ehlers*

Robert Schwager†

**Abstract**

We present a theory explaining the impact of ability tracking on academic performance based on grading policies. Our model distinguishes between initial ability, which is mainly determined by parental background, and eagerness to extend knowledge. We show that achievements of low ability students may be higher in a comprehensive school system, even if there are no synergy effects from teaching different students together. This arises because the comprehensive school sets a compromise standard which exceeds the standard from the low ability track. Moreover, if students with lower initial ability have higher eagerness to learn, merging classes will increase average performance.

Keywords: ability tracking, comprehensive school, education, equality of opportunity, peer group effects

JEL: I21, I28, D63

*Department of Economics, University of Goettingen, Platz der Goettinger Sieben 3, 37073 Goettingen (Germany), E-Mail: tehlers@uni-goettingen.de

†Department of Economics, University of Goettingen, Platz der Goettinger Sieben 3, 37073 Goettingen (Germany), E-Mail: rschwag@uni-goettingen.de

# 1 Introduction

A major controversy in education policy concerns whether students should be taught in comprehensive schools or whether classes should be tracked according to ability. For example, in Germany the results of the first PISA tests raised a debate about making the school system more comprehensive. Such a policy is motivated by the idea that mixing good and mediocre students improves the performance of mediocre students without harming the good students too much. This reasoning is largely supported by empirical research which often finds that tracking, or the ability composition of classes, affect the performance of individual students.

While such peer group effects seem to be well documented empirically, the mechanism driving them is rarely discussed. A possible explanation states that good students help mediocre colleagues to pass the exam and in addition learn by explaining the subject. Formally this can be modeled by a learning production function which depends on the average ability in class like in Arnott and Rowse (1987) and in Epple, Newlon and Romano (2002). While we do not question the relevance of this explanation, in this paper we present a complementary theory which is based on the schools' grading policy. We show that the incentives created by grading standards alone can explain many empirical results on tracking without refering to any direct impact of classmates' ability on individual performance.

The idea of the peer goup effect goes back to the 1970's where this effect was first analyzed in the US. In this decade, for example Summers and Wolfe (1977) empirically analyzed the performance of students from the Philadelphia school district. They came to the conclusion that a high share of good students has a positive effect on the mediocre students, while the good students are not harmed. On the other hand the performance of both types is reduced if the share of mediocre students is too high. Summers and Wolfe conclude that in this case both mediocre and good students do not deploy their full potential.

Following up on this early contribution, a large literature has emerged which empirically analyzes peer group effects and the impact of tracking on educational outcomes. Surveys of this literature are provided by Meier and Schütz (2008) and Brunello and Checchi (2007). This research investigates the impact of tracking on average academic performance and on performance of students with different abilities. The latter question is linked to equality of opportunity, in the sense that academic achievement should not depend on the social background. This is especially relevant in the case of early tracking, since family background is likely to strongly influence the ability in the first years of schooling.

In this line of research Argys, Rees and Brewer (1996) find that the abolishment of tracking in the US would result in a large increase in performance of students in low ability classes, but on the other hand would decrease performance of students in high achieving classes. Similarly Woessmann (2010) concludes that less tracking leads to more equality of opportunity. In addition, Hanushek and Wößmann (2006) find a tendency that early tracking reduces average performance, but this cannot be found in all investigated countries. Waldinger (2006) also comes to the conclusion that early tracking of students with non-academic family background results in low academic achievement. He argues that the difference in achievement based on family background is already present before tracking takes place. Tracking does not reinforce this difference, but comprehensive schooling would reduce it.

Several contributions come to somewhat different conclusions, suggesting positive effects of tracking. Figlio and Page (2002) find no evidence that low ability students are harmed in a tracked environment, but that they may, in contrast, gain in a tracked class. The work by Kim, Lee and Lee (2008) uses data from South Korea, where tracking takes place in about half of the existing schools. Their main result states that tracking raises average achievement. It helps students above median ability and does not lower the achievement of students below it. Another study which does not yield negative effects of tracking is provided by Betts and Shkolnik (2000). According to their results, tracking does not significantly change achievement of low ability students. It has a small negative effect on students with moderate ability and a small positive effect on high ability students, which cancel out in average performance.

Summing up, the empirical literature agrees more or less that the family background plays a major role for academic achievement. This is reinforced by tracking, at least when tracking occurs early. Most studies also conclude that equality of opportunity is promoted by comprehensive schooling, in the sense that the achievement gap between students of different abilities or backgrounds is narrowed. In contrast, there does not seem to be clear evidence on whether average achievement rises or falls when students are tracked.

In this paper we provide a simple model which can account for these facts. In the model there are two types of students distinguished by ability. These students are taught either in tracked classes or in a comprehensive school. The instrument of the school is the graduation standard, which is the level of performance required to pass the exam. In setting the standard, the school trades off wages of graduates, which rise in the standard, against effort costs required to meet a more demanding standard.

We consider two different dimensions of ability. The first dimension represents the endowment a student starts with. This endowment is the ability a student has when the tracking decision is taken. It results from previous learning, which is determined by family background and former schooling. Clearly, the earlier tracking occurs, the more relevant is the family background. We assume that students can reach an academic performance corresponding to the endowment without effort cost.

The second dimension of ability is the eagerness to learn or to improve personal achievement. This is the student's potential a teacher can work with. This dimension determines how hard is is for students to raise performance beyond their initial ability. Importantly, we do not exclude the case where students with low initial endowment have high eagerness to learn or vice versa.

We distinguish between the objective function of the student or teacher, and the objective the parents or society may have. While the latter only care for academic performance, the student incurs costs in terms of stress. We assume that the teacher takes these costs into account when setting the standard, implying that he or she chooses a standard below the standard preferred by society. This modeling is motivated by the emphasis which policy makers and researchers put on performance measuring tests like PISA and TIMMS, which disregard effort costs.

In this framework we characterize the standards chosen by tracked schools and by a comprehensive school, and compare the resulting academic achievements of both types of students. We show that it depends on the parameters whether students with lower initial endowment gain from a comprehensive school. In essence this occurs when their eagerness to learn is not too different from the eagerness of students with high initial endowment. In this case also the achievement gap between both types declines when classes are merged. This arises because in a comprehensive school the teacher is forced to set the standard as a compromise. This pushes lower ability students to higher achievement at the cost of stressing them excessively.

In a further result we compare average academic performance in the tracked and comprehensive school system. In line with the empirical literature, both systems may dominate in that respect. The average performance in the comprehensive system is higher if students with lower initial endowment have higher eagerness to learn. In this case the tracked system does not make full use of the learning potential of students with low initial endowment, on which tracking is based. In contrast, if students with low initial endowment also have lower eagerness to learn, average performance goes down when classes are merged.

It is worth noting that these results crucially depend on the two specific elements of our model. First, when all students have the same eagerness to learn and hence differ only in one dimension of ability, then a compromise standard at the comprehensive school necessarily leads to the same average performance as the separate standards of tracked schools. Second, in our model non-tracking will always be dominated by tracking if teachers and society share the same objective function. In the comprehensive school a unique standard must be chosen and one degree of freedom is given up. Therefore maximization of either average performance or welfare is carried out under an additional restriction in the comprehensive school, and a tracked school system is always preferable.

Our paper contributes to the theory of grading, initiated by Costrell (1994) and Betts (1998). We build upon the traditional model of Costrell, where students weigh the advantage of a degree against the disadvantage of exerting sufficient effort to pass the standard which is set by the teacher. Costrell (1997) puts his model into the context of central or decentralized standards. Since a centralized standard is uniform and decentralized standards are specific, this setting is similar to a comparison of untracked and tracked school systems. In addition Costrell's model, similarly to ours, allows for differences in abilities among schools. However, the focus of his analysis is different from ours. His main issue is that individual schools can free-ride on the tough standards of other schools in the case where employers can only observe the average achievement of graduates from all schools. In contrast we focus on tracking according to ability. Moreover, as stressed above, our model extends Costrell's setup by assuming two dimensions of ability and by replacing the ever positive marginal costs of learning by the idea that a certain performance level can be reached without costs.

# 2 The model

There are two types of students $i \in \{l; h\}$ which differ in ability. We consider two dimensions of ability. The first dimension, which we label endowment or initial ability, represents the level of performance a student can achieve without feeling stressed. We assume that students actually like to think, solve problems, and participate in class and that they feel bored if courses do not challenge them enough. Each individual of type $i$ has the same initial endowment $\gamma_{i=\{l;h\}}$. We assume $\gamma_l < \gamma_h$, so that students of type $l$ have lower endowment than students of type $h$. To interpret the nature of differences in the initial endowment, we observe that the level of performance achievable without feeling stressed most likely depends on previous learning. Moreover, it is natural to think that $\gamma_i$ is largely determined by the upbringing and the parental background of students, as mentioned in the introduction.

The second dimension of ability, labeled $a_i$ for type $i$, expresses the ease of learning. This parameter measures how much stress a student feels if he or she pushes performance beyond his or her initial ability $\gamma_i$. We can interpret $a_i$ as the individual intellectual capacity and motivation of the student. Since a student with a low academic background can well be highly motivated or intelligent, we allow for the case $a_l > a_h$. This describes the situation where students of type $l$ have low initial endowment of ability but high eagerness and capacity to learn.

Depending on both dimensions of ability a student has costs $c_i$ to achieve a certain level of education, denoted by $e_i$:

$$c_i(e_i) = \frac{1}{2a_i}(e_i - \gamma_i)^2 \,. \tag{1}$$

Here the inverse of $a_i$ enters the marginal cost of learning. There is a minimum at $e_i = \gamma_i$, where costs are zero. At this point the student's academic performance is just his or her initial ability. The cost function, shown in figure 1, represents the idea that demanding less effort leads to higher costs in terms of being bored.

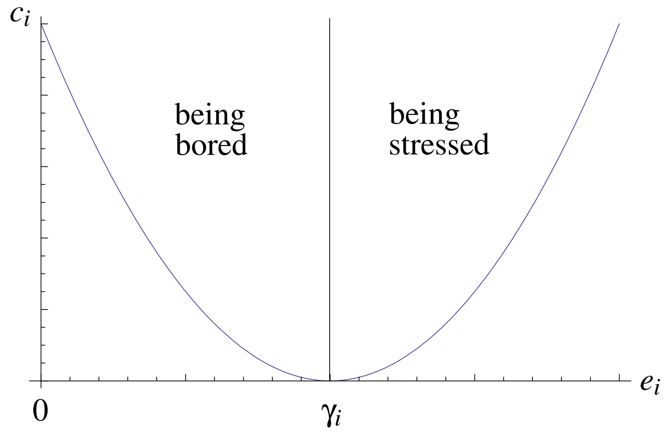



Figure 1: Cost as a function of effort for a student of type $i$. When effort equals initial ability $\gamma_i$, marginal cost of effort is zero.

We now turn to the examination and the labor market. We start in this section by analyzing the case where students are taught in classes tracked according to type. The school or the teacher set a standard $s_i$ which is measured in the same units as the level of education $e_i$. The level of personal education must be at least as high as the standard of the school in order to pass the final exam and graduate. The students decide to become graduates or not and, conditional on this, which academic performance to achieve. This decision is based on the cost of effort and the wages for non-graduates $w_{0i}$ and graduates $w_{1i}$. This formulation assumes that employers observe a student's type $i$ and whether he or she graduated or not. However, the individual performance $e_i$ is unknown to the employer and therefore does not enter the wage.

Conditional on the decision to pass the exam or not, the student chooses performance to maximize utility. We denote the utility achieved in case of passing (not passing) by $u_{1i}$ ($u_{0i}$):

$$\begin{aligned} u_{1i} &= \max_{e_i}\{w_{1i} - c_i(e_i) | e_i \geq s_i\} && \Rightarrow \begin{cases} e_i = s_i & \text{if } s_i \geq \gamma_i \\ e_i = \gamma_i & \text{if } s_i < \gamma_i \end{cases} \\ u_{0i} &= \max_{e_i}\{w_{0i} - c_i(e_i)\} && \Rightarrow e_i = \gamma_i \,. \end{aligned} \tag{2}$$

Anticipating this choice, a student graduates if $u_{1i} \geq u_{0i}$. We assume that students expect $w_{1i} \geq w_{0i}$. One can see from equation (4) below that this expectation is confirmed in equilibrium. The graduation choice is then given by:

$$\max\{u_{0i};u_{1i}\} = \begin{cases} u_{1i} & \text{if } s_i < \gamma_i \\ u_{1i} & \text{if } w_{1i} - c_i(s_i) - w_{0i} \geq 0 \quad \text{and } s_i \geq \gamma_i \\ u_{0i} & \text{if } w_{1i} - c_i(s_i) - w_{0i} < 0 \quad \text{and } s_i \geq \gamma_i\,. \end{cases} \tag{3}$$

In equilibrium, the wage after passing the exam $w_{1i}$ must be equal to the expected productivity of graduates of class $i$. We normalize productivity to be measured in the same units as academic performance. Therefore $w_{1i}$ equals the education level of graduates of class $i$. In the same way the wage $w_{0i}$ is given by the academic performance of non-graduates. From equation (2) we have:

$$\begin{aligned} w_{0i} &= \gamma_i & &\Rightarrow \text{no exam} \\ w_{1i} &= \max\{s_i;\gamma_i\} & &\Rightarrow \text{exam}\,. \end{aligned} \tag{4}$$

We now turn to the choice of standard $s_i$ by the teacher. The teacher maximizes utility of all students. Thus, we assume that the teacher cares about the disutility of learning of his or her students. Inserting (4) and (1) into (3) shows that in the case where $s_i \geq \gamma_i$, the student chooses to graduate if $s_i - \gamma_i \leq 2a_i$. Using this, (2) and (4) in (3) shows that utility of all students of type $i$ is given by:

$$V_i(s_i) = \begin{cases} \gamma_i & \text{if } s_i < \gamma_i \text{ or } s_i - \gamma_i > 2a_i \\ s_i - \frac{1}{2a_i}(s_i - \gamma_i)^2 & \text{if } s_i \geq \gamma_i \text{ and } s_i - \gamma_i \leq 2a_i\,. \end{cases}$$

The optimal standard $s_i^*$ is determined by:

$$\frac{\partial V_i}{\partial s_i} = 1 - \frac{1}{a_i}(s_i - \gamma_i) \stackrel{!}{=} 0 \quad \Rightarrow \quad s_i^* = a_i + \gamma_i\,. \tag{5}$$

The chosen standard reflects both dimensions of ability.

Comparing the two standards, the typical case is given by $s_h^* > s_l^*$, where $h$-students enjoy an advantage compared to $l$-students in terms of total ability:

$$a_h + \gamma_h > a_l + \gamma_l\,. \tag{6}$$

In this case one can clearly label both types of students as $l$-low and $h$-high ability. However we do not rule out the opposite case, where

$$a_h + \gamma_h \leq a_l + \gamma_l\,. \tag{7}$$

Thus, we allow the learning capacity of $l$-students to be so much higher than the one of $h$-students that it overcompensates the disadvantage of initial endowment of the $l$-students.

# 3 Merging classes

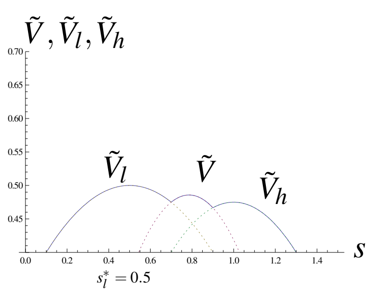
The previous analysis dealt with separated classes $i \in \{l;h\}$. In contrast in this section we consider the case where both classes can be mixed together in one comprehensive school. We denote the share of $h$-students in the comprehensive school by $0 < d < 1$. The teacher sets a common standard $s$ applying to all students in the mixed class. The teacher's objective function is the aggregate utility of all students, denoted by $V(s)$. We continue to assume that employers are able to observe the standard of the school and the type of an applicant $i$. Therefore, for any given standard $s$, individual choices of students are still determined by (2) and (3) and wages are still given by (4), where $s_i$ is replaced by $s$.

Depending on the standard, one of four constellations can occur. First, all students choose a performance equal to their initial ability $\gamma_i$. We denote the value of the school's objective function in this case by $\tilde{V}_0$. Since in this case every student of type $i$ earns a wage equal to $\gamma_i$ and has no cost, it follows $\tilde{V}_0 = \gamma_l(1-d) + \gamma_h d$. This constellation will not occur in equilibrium.

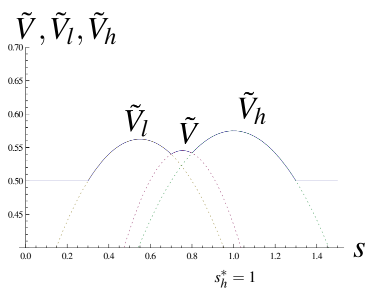
Second, only for the $l$-students the standard is binding, while the $h$-students choose performance $\gamma_h$. Using the wage (4) and the cost function (1), aggregate utility in this case is $\tilde{V}_l(s) = \left[s - \frac{1}{2a_l}(s-\gamma_l)^2\right](1-d) + \gamma_h d$.

Third, both types of students choose to graduate and have to incur effort costs to do so. Then performance of students of both types just meets the standard $s$. Hence, the school's objective is $\tilde{V}(s) = \left[s - \frac{1}{2a_l}(s-\gamma_l)^2\right](1-d) + \left[s - \frac{1}{2a_h}(s-\gamma_h)^2\right] d$.

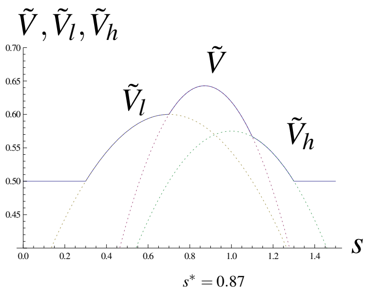
Fourth, $l$-students perform at their initial ability $\gamma_l$, while $h$-students meet the standard. This yields the objective function $\tilde{V}_h(s) = \gamma_l(1-d) + \left[s - \frac{1}{2a_h}(s-\gamma_h)^2\right] d$.




(a) Parameter: $d = 0.5$, $\gamma_h = 0.7$, $\gamma_l = 0.1$, $a_h = 0.3$, $a_l = 0.4$




(b) Parameter: $d = 0.5$, $\gamma_h = 0.7$, $\gamma_l = 0.3$, $a_h = 0.3$, $a_l = 0.25$




(c) Parameter: $d = 0.5$, $\gamma_h = 0.7$, $\gamma_l = 0.3$, $a_h = 0.3$, $a_l = 0.4$

Figure 2: The objective function of the comprehensive school $V(s)$. $\tilde{V}_h$ ($\tilde{V}_l$) describes the value of the schools objective, if one assumes that $h(l)$-students exert effort to pass the exam and $l(h)$-students choose the effort level $\gamma_l$ ($\gamma_h$). $V$ is the upper envelope of $\tilde{V}$, $\tilde{V}_l$ and $\tilde{V}_h$.

Which one of these four cases applies depends on the parameter constellation. This leads to the following definition of $V(s)$ with seven branches defined by parameter restrictions. Some of the restrictions are redundant, but are left for better understanding:

$$V(s) = \begin{cases} \tilde{V}_0 & [1] \quad \text{if } s \leq \gamma_{l,h} \\ \tilde{V}_0 & [2] \quad \text{if } s \leq \gamma_h;\ s > \gamma_l;\ s - \gamma_l > 2a_l \\ \tilde{V}_l(s) & [3] \quad \text{if } s \leq \gamma_h;\ s > \gamma_l;\ s - \gamma_l \leq 2a_l \\ \tilde{V}(s) & [4] \quad \text{if } s > \gamma_{l,h};\ s - \gamma_l \leq 2a_l;\ s - \gamma_h \leq 2a_h \\ \tilde{V}_h(s) & [5] \quad \text{if } s > \gamma_{l,h};\ s - \gamma_l > 2a_l;\ s - \gamma_h \leq 2a_h \\ \tilde{V}_l(s) & [6] \quad \text{if } s > \gamma_{l,h};\ s - \gamma_l \leq 2a_l;\ s - \gamma_h > 2a_h \\ \tilde{V}_0 & [7] \quad \text{if } s > \gamma_{l,h};\ s - \gamma_l > 2a_l;\ s - \gamma_h > 2a_h \end{cases} \tag{8}$$

In the first branch [1] the standard is too low to bind anybody, so all students just perform at the initial endowment. In the next branch [2] outcome is the same, but the standard is above the maximal standard the $l$-students are willing to satisfy and $u_{1l} < u_{0l}$. $h$-students still graduate without effort cost. Branch [3] represents the situation where the $l$-students graduate by just meeting the standard whereas $h$-students still graduate with level $\gamma_h$. On branch [4], the standard is high enough to be binding also for $h$-students. Branches [5] and [6] differ depending on which group first refuses to satisfy the high standard and falls back to initial ability. In branch [5] this is true for the $l$-students and in branch [6] for the $h$-students. The last branch [7] shows a standard higher than anybody will accept to meet.

Notice that the branches [3]-[6] of the expression $V(s)$ are strictly concave. Moreover, observe that $\tilde{V}_l$ and $\tilde{V}_h$ are affine transformations of the objective functions $V_l$ and $V_h$ of the separated classes: $\tilde{V}_l(s) = (1-d)V_l(s) + d\gamma_h$ and $\tilde{V}_h(s) = dV_h(s) + (1-d)\gamma_l$. Consequently in the branches [3], [5], and [6], where one of these functions applies, the optimal standard is $s_l^*$ or $s_h^*$. In branch [4], the optimal standard $s^*$ solves:

$$\frac{\partial \tilde{V}}{\partial s} = \left[1 - \frac{1}{a_l}(s - \gamma_l)\right](1-d) + \left[1 - \frac{1}{a_h}(s - \gamma_h)\right] d \overset{!}{=} 0\,.$$

From this first order condition of $\tilde{V}$, we obtain:

$$\begin{aligned} s^* &= \frac{a_l a_h + d a_l a_h + \gamma_h \gamma_l - d a_h \gamma_l}{a_h(1-d) + a_l d} \\ &= \frac{d a_l}{a_h(1-d) + a_l d} s_h^* + \frac{(1-d)a_h}{a_h(1-d) + a_l d} s_l^*\,. \end{aligned} \tag{9}$$

The standard of a mixed class is a weighted average of the standards chosen in separated classes. The weights combine the population shares $d$ and $(1-d)$ with the ability parameters $a_h$ and $a_l$.

Figure 2 shows the different branches of $V$. The sub-figures 2(a), 2(b) and 2(c) are based on different parameter combinations, where the optimal standard is $s_l^*$, $s_h^*$ and $s^*$ respectively. We will now analyze in which branch of $V$ the optimal standard is located, that is, which of these three standards gives the highest welfare. Pairwise comparison of the local maxima obtained by the three standards shows that

$$\begin{array}{rcccrcl} \tilde{V}(s^*) & \gtreqless & \tilde{V}_l(s_l^*) & \Leftrightarrow & F_l(a_l,a_h,\gamma_l,\gamma_h,d) & \gtreqless & 0 \\ \tilde{V}(s^*) & \gtreqless & \tilde{V}_h(s_h^*) & \Leftrightarrow & F_h(a_l,a_h,\gamma_l,\gamma_h,d) & \gtreqless & 0 \\ \tilde{V}_h(s_h^*) & \gtreqless & \tilde{V}_l(s_l^*) & \Leftrightarrow & F_{lh}(a_l,a_h,d) & \gtreqless & 0, \end{array} \tag{10}$$

where

$$\begin{aligned} F_l(a_l,a_h,\gamma_l,\gamma_h,d) &= a_h\left[a_l(2-d)-2(1-d)(\gamma_h-\gamma_l)\right]-(1-d)(a_l-\gamma_h+\gamma_l)^2 \\ F_h(a_l,a_h,\gamma_l,\gamma_h,d) &= a_h^2 d+d(\gamma_h-\gamma_l)(\gamma_h-\gamma_l-2a_l)-a_h(a_l+a_l d-2d\gamma_h+2d\gamma_l) \\ F_{lh}(a_l,a_h,d) &= a_l(1-d)-a_h d\,. \end{aligned}$$

These functions define boundaries between subsets of the parameter space. Depending on the signs of the three functions $F_l$, $F_h$ and $F_{lh}$, there could be up to eight such subsets. Two of these do not exist, however, since the three functions cannot be all positive or all negative at the same time. For example, if $F_h > 0$ and $F_l > 0$, then equations (10) imply $\tilde{V}_h(s_h^*) > \tilde{V}(s^*) > \tilde{V}_l(s_l^*)$, hence $F_{lh} < 0$.

**Proposition 1** *In the comprehensive school the chosen standard is:*

$$\begin{array}{lll} s^* & = \frac{d a_l}{a_h(1-d)+a_l d}s_h^* + \frac{(1-d)a_h}{a_h(1-d)+a_l d}s_l^* & \text{if } F_l > 0 \text{ and } F_h < 0 \\ s_l^* & = a_l+\gamma_l & \text{if } F_l < 0 \text{ and } F_{lh} > 0 \\ s_h^* & = a_h+\gamma_h & \text{if } F_h > 0 \text{ and } F_{lh} < 0\,. \end{array} \tag{11}$$

**Proof.**

We need to consider all three possible optimalities and check if the needed constrains from (8) are satisfied.

At first we consider $s^*$ **is optimal.**

We need to show that if $F_l > 0$ and $F_h < 0$ hold, $s^*$ satisfies the conditions given in branch [4] of (8).

$s^* > \gamma_l$ is equivalent to $a_l\left[a_h(1-d)+a_l d\right]\left[a_h+d(\gamma_h-\gamma_l)\right] > 0$, which is satisfied in any case.

$s^* > \gamma_h$ is equivalent to $a_h\left[a_h(1-d)+a_l d\right]\left[(\gamma_h-\gamma_l)(1-d)-a_l\right] < 0$, which reduces to $a_l > (\gamma_h-\gamma_l)(1-d)$. The inequality $F_l > 0$ is equivalent to $a_h\left[a_l(2-d)-2(1-d)(\gamma_h-\gamma_l)\right] > (1-$

$d)(a_l-\gamma_h+\gamma_l)^2$. This implies $a_l(2-d)-2(1-d)(\gamma_h-\gamma_l)>0$, which can be transformed into $a_l-(1-d)(\gamma_h-\gamma_l)>\frac{a_l d}{2}$. From this $a_l>(1-d)(\gamma_h-\gamma_l)$ and hence $s^*>\gamma_h$ follows.

$s^*-\gamma_l\leq 2a_l$ is, for $a_l\neq 0$, equivalent to $a_l\geq\frac{d(\gamma_h-\gamma_l)+a_h(2d-1)}{2d}\equiv A_l(a_h)$. $F_h<0$ is equivalent to $a_l>\frac{d(a_h+\gamma_h-\gamma_l)^2}{a_h(1+d)+2d(\gamma_h-\gamma_l)}\equiv B_l(a_h)$. We show that $B_l(a_h)\geq A_l(a_h)$ so that $a_l>B_l(a_h)$ implies $a_l>A_l(a_h)$. To see this, observe that $B_l(a_h)\geq A_l(a_h)$ is equivalent to $-a_h(1-d)d[a_h+d(\gamma_h-\gamma_l)]\,[a_h+a_hd+2d(\gamma_h-\gamma_l)]\leq 0$, which is true in any case.

$s^*-\gamma_h\leq 2a_h$ is, for $a_h\neq 0$, equivalent to $a_h\geq\frac{a_l(1-2d)-(1-d)(\gamma_h-\gamma_l)}{2(1-d)}\equiv A_h(a_l)$. $F_l>0$ implies that $a_l(2-d)-2(1-d)(\gamma_h-\gamma_l)>0$. Hence $F_l>0$ is equivalent to $a_h>\frac{(1-d)(a_l-\gamma_h+\gamma_l)^2}{a_l(2-d)-2(1-d)(\gamma_h-\gamma_l)}\equiv B_h(a_l)$. We show that $B_h(a_l)\geq A_h(a_l)$ so that $a_h>B_h(a_l)$ implies $a_h>A_h(a_l)$. The inequality $B_h(a_l)\geq A_h(a_l)$ is equivalent to $a_l(1-d)d\,[a_l(2-d)-2(1-d)(\gamma_h-\gamma_l)]\,[a_l-(1-d)(\gamma_h-\gamma_l)]>0$. As shown in the proof of $s^*>\gamma_h$, we have $a_l(2-d)-2(1-d)(\gamma_h-\gamma_l)>0$ and $a_l-(1-d)(\gamma_h-\gamma_l)>0$. Hence $B_h(a_l)>A_h(a_l)$ for all $a_l>0$.

**$s_l^*$ is optimal.**

Now we need to show that if $F_l<0$ and $F_{lh}>0$ hold, $s_l^*$ satisfies the conditions given in branch [3] or [6] of (8). We distinguish two cases, depending on whether (6) or (7) holds. In the first case, where $a_h+\gamma_h>a_l+\gamma_l$, we show that branch [3] of (8) applies. In this case $F_l<0$ is equivalent to $a_l<\frac{a_h(2-d)+2(1-d)(\gamma_h-\gamma_l)-\sqrt{a_h}\sqrt{a_h(2-d)^2+4(1-d)d(\gamma_h-\gamma_l)}}{2(1-d)}\equiv C_l(a_h)$. Observe that $C_l(a_h)<\gamma_h-\gamma_l$ is equivalent to $4(1-d)d(\gamma_h-\gamma_l)>0$, which is true. Hence we have $a_l<C_l(a_h)<\gamma_h-\gamma_l$, which proves $s_l^*\leq\gamma_h$. The conditions $s_l^*>\gamma_l$ and $s_l^*-\gamma_l\leq 2a_l$ follow from $s_l^*=a_l+\gamma_l$.

In the case $a_h+\gamma_h\leq a_l+\gamma_l$ the branch [6] of (8) applies. The first two conditions $s_l^*>\gamma_l$ and $s_l^*\leq 2a_l+\gamma_l$ follow directly from $s_l^*=a_l+\gamma_l$. We show that the condition $s_l^*-\gamma_h>2a_h$ follows from $F_l<0$. Inserting $s_l^*=a_l+\gamma_l$, we can rewrite this condition as $a_h<\frac{a_l+\gamma_l-\gamma_h}{2}\equiv C_h(a_l)$. Consider the denominator of $B_h(a_l)$ defined above. From $a_h-a_l<\gamma_l-\gamma_h$ we have $a_l(2-d)-2(1-d)(\gamma_h-\gamma_l)>a_l(2-d)+2(1-d)(a_h-a_l)=a_ld+2(1-d)a_h>0$. Therefore $F_l<0$ is equivalent to $a_h<B_h(a_l)$. We show that $B_h(a_l)<C_h(a_l)$, so that $a_h<B_h(a_l)$ implies $a_h<C_h(a_l)$. The inequality $B_h(a_l)<C_h(a_l)$ is equivalent to $[a_l(-2+d)+2(1-d)(\gamma_h-\gamma_l)]\,(a_l+\gamma_l-\gamma_h)<0$. Since here $a_l+\gamma_l-\gamma_h>a_h>0$, this is equivalent to $a_l(2-d)-2(1-d)(\gamma_h-\gamma_l)>0$, which we just have shown to be true.

**$s_h^*$ is optimal.**

$F_h>0$ is equivalent to $a_l<B_l(a_h)$, where $B_l(a_h)$ is defined above. $s_h^*-\gamma_l>2a_l$ is equivalent to $a_l<\frac{a_h+\gamma_h-\gamma_l}{2}\equiv D_l(a_h)$. We show that $B_l(a_h)<D_l(a_h)$. Knowing $a_h+a_hd+2d(\gamma_h-\gamma_l)>0$ and $d<1$, this is true. Hence $a_l<B_l(a_h)$ implies $a_l<D_l(a_h)$. The conditions $s_h^*>\gamma_h$ and $s_h^*-\gamma_h\leq 2a_h$ follow from $s_h^*=a_h+\gamma_h$. $\square$

From this proposition we can directly read the globally optimal standard. This is $s^*$ if the parameters are such that $F_l > 0$ and $F_h < 0$. If one of these inequalities is not satisfied, the optimal standard is $s_l^*$ or $s_h^*$, depending on the sign of $F_{lh}$.

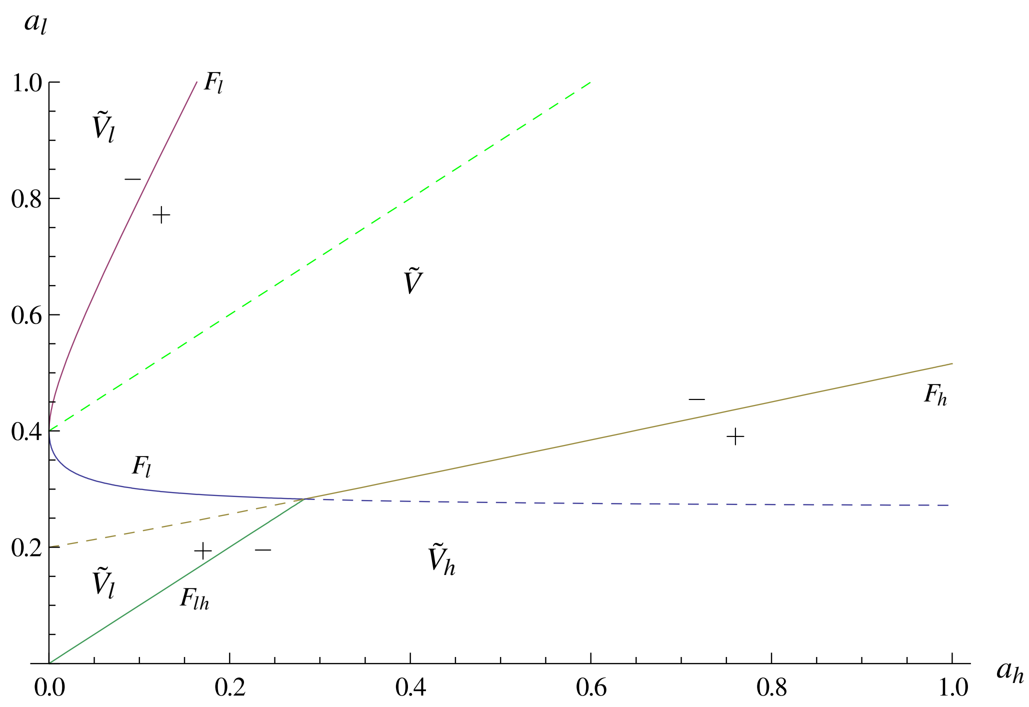



Figure 3: Parameter regions in $a_h - a_l$-space with different optimal standards in the comprehensive school, with $d = 0.5$, $\gamma_l = 0.3$ and $\gamma_h = 0.7$. The labels $\tilde{V}$, $\tilde{V}_l$ and $\tilde{V}_h$ show in which branch of $V$ the global maximum is located. Below (on, above) the dotted line starting at $(a_h = 0; a_l = 0.4)$ one has $a_l + \gamma_l < (=, >) a_h + \gamma_h$.

Figure 3 illustrates in which region of the parameter space each of the three local maxima is the global maximum. This figure is drawn in $a_h - a_l$-space, since the influence of these parameters on the optimal standard is most interesting to study. In this example the other parameters were fixed at $d = 0.5$, $\gamma_l = 0.3$ and $\gamma_h = 0.7$. In the graph we inserted a dotted straight line, starting at $a_h = 0; a_l = 0.4$. Above this line inequality (7) holds. Below this line we have (6), such that labeling the $l$-type as low ability students is appropriate.

In the lower right region of the figure, labeled with $\tilde{V}_h$, the relevant branch of $V$ in equation (8) is [5]. In this region the ability of $h$-students is relatively high in both dimensions compared to the $l$-students. Therefore the teacher sets a standard tailored exactly to $h$-students, accepting that $l$-students will drop out. In the central and upper right region, labeled $\tilde{V}$, branch [4] of $V$ contains the optimum. Abilities of both types do not differ much and hence the teacher sets the compromise standard $s^*$. Finally in the upper left and lower left regions, labeled with $\tilde{V}_l$, the school chooses the standard $s_l^*$ for the $l$-students, whereas the $h$-students perform at $\gamma_h$. They do so for two different reasons. In the upper region with $a_l > 0.4$, where branch [6] of $V$ is relevant and (7) holds, the standard is too high for the $h$-students and they drop out. In the lower part ($a_l < 0.4$, branch [3], and (6) holds), the standard is so low that the $h$-students can meet it without effort cost.

The next proposition provides comparative statics for the case where the optimal standard is $s^*$.

**Proposition 2** *If the optimal standard in a comprehensive school is $s^*$, it increases in $a_l$, $a_h$, $\gamma_l$ and $\gamma_h$. It increases (decreases) in $d$ if $a_h + \gamma_h > (<) a_l + \gamma_l$.*

**Proof.**

Differentiating $s^*$ from (9) we obtain:

$$\frac{\partial s^*}{\partial d} = \frac{a_h a_l (a_h - a_l + \gamma_h - \gamma_l)}{(a_h(1-d) + a_l d)^2} \tag{12}$$

$$\frac{\partial s^*}{\partial \gamma_l} = \frac{a_h(1-d)}{a_h(1-d) + a_l d} > 0$$

$$\frac{\partial s^*}{\partial \gamma_h} = \frac{a_l d}{a_h(1-d) + a_l d} > 0$$

$$\frac{\partial s^*}{\partial a_l} = \frac{a_h(1-d)\,[a_h + d(\gamma_h - \gamma_l)]}{[a_h(1-d) + a_l d]^2} > 0$$

$$\frac{\partial s^*}{\partial a_h} = \frac{a_l d\,[a_l - (\gamma_h - \gamma_l)(1-d)]}{[a_h(1-d) + a_l d]^2} \tag{13}$$

(12) is positive (negative) if $a_h + \gamma_h > (<) a_l + \gamma_l$. (13) is positive if $a_l > (\gamma_h - \gamma_l)(1-d)$. As shown in the proof of Proposition 1, this is true if $s^* > \gamma_h$, which must be the case if $s^*$ is the optimal choice. □

As expected, for all students the standard increases in both dimensions of ability. Moreover, the standard increases in the share of the type of students $i$ whose total ability, measured by $a_i + \gamma_i$, is larger.

We now turn to the question whether two separated classes are preferable to the mixed class. It is important to distinguish between a comparison of utilities and a comparison of academic performances. Regarding utilities, no case is possible where students are better off in the mixed class than in separated classes, because in that case, $s_l^*$ and $s_h^*$ can be optimized separately. Hence, comparison of utilities is a straightforward application of the decentralization theorem by Oates (1972). Notice that the same observation would hold if we assumed that schools maximize academic performance or wages instead of students' utility. In such a model it would be immediate that in the comprehensive school performance of each type can only be worse than in tracked schools.

In contrast, as we will now show, in our model academic performance can also increase by mixing the classes. First, we consider for each type of students separately how their performance changes if classes are merged. For this comparison, observe that in the case where the optimal standard of the comprehensive school is $s_l^*$ ($s_h^*$), the performance of the $h$($l$)-students is $\gamma_h$ ($\gamma_l$), and that the standard $s^*$ is a weighted average of the standards chosen in the tracked schools. With this, Proposition 1 immediately leads to:

**Proposition 3** *If the comprehensive school chooses $s_l^*$ ($s_h^*$), the performance of h(l)-students is lower than in the tracked h(l)-school. If the comprehensive school chooses $s^*$, the performance of l-students is higher than (lower than, equal to) the performance in the tracked l-school if $a_l + \gamma_l < (>, =) a_h + \gamma_h$. If the comprehensive school chooses $s^*$, the performance of h-students is higher than (lower than, equal to) the performance in the tracked h-school if $a_l + \gamma_l > (<, =) a_h + \gamma_h$.*

This proposition shows that our model can generate a positive peer group effect for low ability students, by which we mean the $l$-students, where (6) holds. In a comprehensive school, teachers will need to find a compromise between the standards tailored to individual student types. As long as low ability students are still willing to meet this standard, they will put in more effort than in the separated class. As a consequence one will observe higher test results on their part in the comprehensive school, even if there are no synergy effects from teaching diverse students together. Furthermore, the reduction of the standard for the $h$-students might be quite small when $d$ and/or $a_l$ are relatively large. Then, as (9) shows, the standard of the mixed class is close to the standard of the $h$-class. Given confounding influences, an empirical study might fail to find statistical significance of such a small impact.

The peer group effect obtains only for a subset of the parameter space. As is apparent from figure 3, the learning capacities of both types must not be too different. Otherwise, if one type finds it substantially easier to learn, the school will set optimal incentives for this type and put up with the fact that the other type stops graduating. Furthermore one can show that the $F_h = 0$ curve shifts downwards if $d$ decreases. Hence, a positive peer group effect for the low ability students is more likely when these students are more numerous. In this case the teacher of the comprehensive school puts more weight on their utility and therefore refrains from setting a standard which overburdens them.

Finally it may also happen that the comprehensive school sets the standard $s_l^*$ tailored to the low ability students. This corresponds to the lower left region of Figure 3. In this case mixing classes leaves the performance of $l$-students unchanged and reduces the performance of $h$-students. However, this decline in performance might be very small, since learning capacity of $h$-students is anyway not very large and since they continue to perform at their initial endowment $\gamma_h$. Therefore, although mixing classes obviously does not help in this case, the damage it inflicts is small.

The next proposition deals with the effect of merging classes on aggregate performance.

**Proposition 4** *If $a_l + \gamma_l < a_h + \gamma_h$, the average productivity of students in a comprehensive school exceeds (is equal to, falls short of) the average productivity of students in tracked schools if and only if learning ability of l-students is larger than (is equal to, is smaller than) learning ability of h-students. That is:*

$$s^* \gtreqless ds_h^* + (1-d)s_l^* \Leftrightarrow a_l \gtreqless a_h\,.$$

*If $a_l + \gamma_l > (=) a_h + \gamma_h$, the average productivity of students in a comprehensive school falls short of (is equal to) the average productivity of students in tracked schools.*

**Proof.** From equation (9), we find that $s^* \gtreqless ds_h^* + (1-d)s_l^*$ is equivalent to:

$$(a_l - a_h)(s_h^* - s_l^*) \gtreqless 0\,. \tag{14}$$

In the case $a_l + \gamma_l < a_h + \gamma_h$ we have $s_h^* > s_l^*$, and (14) is equivalent to $a_l \gtreqless a_h$. In the case $a_l + \gamma_l > a_h + \gamma_h$ we have $s_h^* < s_l^*$, and (14) is equivalent to $a_l \lesseqgtr a_h$. Since $\gamma_h > \gamma_l$, in this case $a_l > a_h$ must hold. Hence $s^* < ds_h^* + (1-d)s_l^*$. For $a_l + \gamma_l = a_h + \gamma_h$, we have $s_l^* = s_h^*$, and (14) implies $s^* = ds_h^* + (1-d)s_l^*$. □

This proposition shows that merging classes with heterogeneous students may increase overall academic performance, even when there are no spillover effects between types of students. This occurs when students with low initial ability have higher learning capacity than students with high initial ability. To understand that, consider how the standard is set in the comprehensive school. The teacher will trade off the net-loss incurred by $l$-students when the standard is increased above their optimal standard $s_l^*$ against the net-loss incurred by $h$-students when the standard is decreased below $s_h^*$. Since the learning ability of $l$-students exceeds the learning ability of $h$-students, the net-loss of the latter increases faster than the net-loss of the former. Therefore the optimal standard, where marginal net-losses are equalized, is closer to $s_h^*$ than to $s_l^*$. Hence the optimal standard in the mixed class is higher than the weighted average of the standards of the separated classes.

This kind of result is likely to be relevant in education systems where students are tracked early. It is likely that the allocation to different tracks is mostly determined by the endowment of skills conferred by the family background. At the same time it is well possible that students with low endowment have not yet fully unfolded their potential and correspondingly find it easier to extend their knowledge. In the terminology of our model these students have high learning capacity $a_l$, but low endowment $\gamma_l$. If these students now attend a comprehensive school, average performance of students will increase. Both types of students find a relatively high standard acceptable, but they do so because of different reasons: One group starts with high initial ability and the others are eager to advance.

# 4 Conclusion

In this paper we present a model comparing the choice of examination standards by tracked and untracked schools. The model distinguishes between initial ability and the capacity or willingness to extend ability. When setting the standard, the school or teacher takes the student's disutility of learning into account. Therefore, the resulting choices differ from the standards which maximizes academic performance, which is the focus of PISA and similar studies.

Our findings show that in many cases a comprehensive school will enhance performance of low ability students or even enhance average performance compared to tracked schools with individual standards. In these cases performance of high ability students decreases, but this effect may be so small that it is insignificant in an empirical study. Our model therefore provides a foundation of peer group effects although we abstract from any synergy effect from teaching different student types together.

# References


**Argys, Laura M, Daniel I Rees, and Dominic J Brewer**, "Detracking America's Schools: Equity at Zero Cost?," *Journal of Policy Analysis and Management*, 1996, *15* (4), 623–645.

**Arnott, Richard and John Rowse**, "Peer Group Effects and Educational Attainment," *Journal of Public Economics*, 1987, *32* (3), 287–305.

**Betts, Julian R**, "The Impact of Educational Standards on the Level and Distribution of Earnings," *American Economic Review*, 1998, *88* (1), 266–75.

_ **and Jamie L Shkolnik**, "The Effects of Ability Grouping on Student Achievement and Resource Allocation in Secondary Schools," *Economics of Education Review*, 2000, *19* (1), 1–15.

**Brunello, Giorgio and Daniele Checchi**, "Does School Tracking Affect Equality of Opportunity? New International Evidence," *Economic Policy*, 2007, *22* (52), 782–861.

**Costrell, Robert M.**, "A Simple Model of Educational Standards," *The American Economic Review*, September 1994, *84* (4), 956–971.

_ , "Can Centralized Educational Standards Raise Welfare?," *Journal of Public Economics*, 1997, *65*, 271–293.

**Epple, Dennis, Elizabeth Newlon, and Richard Romano**, "Ability Tracking, School Competition, and the Distribution of Educational Benefits," *Journal of Public Economics*, 2002, *83* (1), 1–48.

**Figlio, David N and Marianne E Page**, "School Choice and the Distributional Effects of Ability Tracking: Does Separation Increase Inequality?," *Journal of Urban Economics*, 2002, *51* (3), 497–514.

**Hanushek, Eric A. and Ludger Wößmann**, "Does Educational Tracking Affect Performance and Inequality? Differences- in-Differences Evidence Across Countries," *The Economic Journal*, 2006, *116* (510), C63–C76.

**Kim, Taejong, Ju-Ho Lee, and Young Lee**, "Mixing versus Sorting in Schooling: Evidence from the Equalization Policy in South Korea," *Economics of Education Review*, 2008, *27* (6), 697–711.

**Meier, Volker and Gabriela Schütz**, "The Economics of Tracking and non-Tracking," *Zeitschrift für Betriebswirtschaft*, 2008, *Special Issue 1*, 23–43.

**Oates, Wallace E**, *Fiscal Federalism*, New York: Harcourt Brace Jovanovich, 1972.

**Summers, Anita A. and Barbara L. Wolfe**, "Do Schools Make a Difference?," *The American Economic Review*, 1977, *67* (4), 639–652.

**Waldinger, Fabian**, "Does Tracking Affect the Importance of Family Background on Students' Test Scores?," *Unpublished manuscript, LSE*, 2006.

**Woessmann, Ludger**, "Institutional Determinants of School Efficiency and Equity: German States as a Microcosm for OECD Countries," *Jahrbücher für Nationalökonomie und Statistik*, 2010, *230* (2), 234–270.