

Ability grouping and incentives

Gerald Eisenkopf
University of Konstanz*[†]

January 31, 2006

Abstract

The paper discusses the impact of ability grouping in secondary education, with a particular consideration of student incentives. Education provides both a signal on unobservable ability and improves productivity after education. Grouping students according to ability sets better incentives in primary education and allows for improved peer group effects in secondary education. Comprehensive education induces higher signalling incentives in secondary education. Optimal education policy allows for some imperfection in the sorting process. Taking social heterogeneity into account, policy makers can influence the sorting process by changing support in primary or secondary education.

Keywords: Educational economics, signalling, human capital formation, tracking, sorting, ability grouping

JEL Classifications: I 20, I 28

*Corresponding Address: University of Konstanz, Box D 144, 78457 Konstanz, Germany, E-Mail: Gerald.Eisenkopf@uni-konstanz.de

[†]The author thanks Oliver Fabel, Ludger Wmann, Gerhard Glomm and seminar participants at the University of Konstanz for helpful discussions. Financial support of the Deutsche Forschungsgemeinschaft (DFG) through the research group “Heterogeneous Labor: Positive and Normative Aspects of the Skill Structure of Labor” is gratefully acknowledged. Comments are very welcome. The usual disclaimer applies.

1 Introduction

The economic literature has devoted great attention to the design of education policy in recent years. Yet, the biggest contribution to educational success derive from student-specific factors, which are very difficult to manipulate for policy makers (e.g. intelligence or family background, see for example in Wößmann (2004)). What is affected both by policy and individual characteristics are the incentives of students to invest effort in education. In this paper, the impact of student incentives on educational policy are discussed, in particular with the problem of tracking students. The analytic focus is predominantly on efficiency, implications for inequality of educational opportunities and inequality of educational outcome are discussed briefly.

A substantial contribution to the research of incentives in education is provided theoretically by de Fraja and Landeras (DFL, 2005) as well as empirically by de Fraja, Oliveira and Zianchi (2005). Otherwise, the discussion about optimal education policy has largely ignored incentives for students (with some notable exceptions, e.g. Angrist and Lavy (2002) or Kremer et al. (2004)). Additionally, the well established literature about education as a signal in the labor market (starting with Spence, 1973) is rarely considered. An exception here is provided by Brunello and Giannini (BG, 2005). A greater emphasis has been put on school choice and school competition (e.g. Hoxby, 2003), on class size (e.g. Lazear, 2001; West and Wößmann, 2005), peer group effects, or the sorting (or synonymously tracking or ability grouping) of students. The peer group effect is captured empirically and theoretically in many different ways, such that sorting implements an efficient solution at the expense of equity (e.g. in Epple et al., 2002), or that sorting raises both equity and efficiency (e.g. Dobbelsteen et al., 2002), or just vice versa (e.g. Argys et al. (1996)). Hence, the evidence for peer group effects and the benefits from sorting is ambiguous, while the impact of sorting on inequality in educational performance and in returns to education is more obvious (Hoxby, 2001, Hanushek and Wößmann, 2005).

The argument in the paper can be summarized as follows. Perfect ability grouping may not be desirable, even if the direct benefits from sorting are bigger for the more able students than the losses for the less able students. Consider two functions of education. Firstly, students acquire skills which increase their post-educational productivity in the labor market. More skills improve the

post-educational wages, which induces students to put some effort into their education. Secondly, schooling provides a signal for the students' "innate" abilities, as more able students are likely to have a higher or better educational degree than less able ones. This ability is important for employers, e.g. as an indicator for the acquisition of job-specific human capital or general problem-solving skills.

However, sorting students according to (expected) educational results into different types of schools already provides a signal about the ability before the education has been completed. This signal reduces the incentives to invest in education as a signal. Such a reduction in incentive strength can explain why the intuitively reasonable gains in peer group effects from tracking are so difficult to observe. DFL (2005) show that increasing teacher incentives via more competition among schools may backfire on student incentives in a similar way.

As a consequence, policy makers face a trade-off between incentives from signalling and higher productivity from ability grouping, if they want to boost educational results. They want to get the best of both effects and distort the ability grouping accordingly.

Taking intertemporal aspects into account, this paper shows that the incentives are not diminished by sorting but transferred to the pre-selection period. If selection takes place in secondary education, most students invest more in primary education to get to the better school.

Considering heterogeneity in private backgrounds, the welfare limits of perfect sorting according to ability (i.e. a perfect meritocracy) become obvious. It is beneficial to have some well-supported students with relatively low ability in a good peer group. These students distort the ability signal from sorting. Policy makers can change the sorting outcome by adjusting per capita support at different schools and in time periods.

Brunello and Giannini also show why neither perfect sorting nor comprehensive schooling provides a strictly dominant, although with an entirely different treatment of the signalling problem. For them "[s]tratifed systems trade the advantages of specialization and signalling against the disadvantages of producing skills with limited flexibility and versatility." (BG, p. 190) For BG, the students cannot influence the admission result and the subsequent signal with more effort. "Their" students have technical and academic abilities and schools provide different types of education. In

contrast to BG, this paper provides a policy instrument to manipulate the sorting process. Finally, other reasons exist why a *sui generis* positive peer group effect can produce negative external effects on the outcome from ability grouping, e.g. higher failure rates for high-ability students (Meier, 2004).

To support the argument, section 2 introduces the analytic framework, identifies a (not achievable) first-best solution and compares second-best outcomes. In section 3 intertemporal effects are analyzed. It becomes obvious that selection in secondary education does not entirely reduce signalling incentives but shifts them to primary education. This paper does not discuss when students should optimally be separated. This problem is analyzed in Brunello et al. (2004). Section 4 then discusses the advantageous effects of imperfect sorting. Section 5 takes heterogeneity regarding the private background into account and shows how a policy maker can manipulate the sorting results. Section 6 concludes.

2 Efficient organization of education

2.1 The model

Suppose one region with two schools ($k \in \{a, b\}$) of equal size and properties. Each school must enrol half of the students in the region. Ability $\theta > 0$ is distributed according to $F(\theta)$ across the students. The ability of student i is known by the student and the policy maker, but not by the future employer of the student. The employers knows $F(\theta)$. Students can invest in effort e . The costs of effort follow the function $C(e)$, with $C'(e) > 0$ and $C''(e) > 0$. Only the student knows about his choice of the effort level. Policy makers can choose between a comprehensive and a selective educational system, i.e. between sorting and integrating students. Under a sorting (or tracking) policy, high ability students go to one school and low ability students to the other one. Otherwise, the ability distribution is identical in both schools. For the moment it is assumed that sorting is perfect.

The identification of efficient solutions requires assumptions on two critical issues. How do input

factors contribute to educational output and who benefits from educational output. Regarding the first issue, the following assumption holds.

Assumption 1 *The post-educational productivity depends on a student's ex-ante ability and the skills he acquired in education.*

For a worker i with ability θ_i and being educated at school k , the productivity is denoted by Π_{ik} , with

$$\Pi_{ik} = \Pi(\theta_i, Q_{ik}(\cdot)) \quad (1)$$

In this equation, $Q_{ik}(\cdot)$ represents the skills the worker has acquired in school k . These skills may represent general human capital. Specific human capital is acquired in the firm, the acquisition depends on the ability of student, θ_i . The skills acquired in school (also called the educational output) depend on the product of the characteristics of the student, the characteristics of his fellow students and additionally the realisation of a random variable. Hence, the function is

$$Q_{ik} = q(\phi_i, \theta_i, e_{ik}, P_k) + \varepsilon_i \quad (2)$$

The random variable $\varepsilon_i \sim N(0, \sigma^2)$ is identically and independently distributed across the students. The function $g(\varepsilon)$ denotes the respective density function. Regarding the function $q(\cdot)$, it increases in all arguments, the second-order derivatives are negative. All cross derivatives are positive. The variable P_k denotes the peer group effect at school k , the variable ϕ_i the private background of student i . Parents and other relatives influence educational output with financial means or intellectual stimulation. More specifically, it is assumed that $\phi > 0$ and that ϕ and θ are statistically independent ($Cov(\phi, \theta) = 0$). The background of each student is private knowledge. Suppose for the moment, that ϕ is identical for all agents. This assumption will be relaxed in section 5.

The peer group effect constitutes an external effect of a student's behavior and characteristics on his fellow students. The literature on peer group effects claims that a high (low) ability of one students supports (harms) the educational performance of his fellow students. In this paper high

(low) effort levels of a student generate a similar effect. Hence, the widely used parameter for peer group effects is modified.

Assumption 2 *The peer group effect is the average product of effort and ability of all students at school k , such that*

$$\frac{\partial^2 P_k}{\partial \theta_i \partial e_{ik}} > 0 \quad (3)$$

and

$$\left(\frac{\partial P_k}{\partial \theta_{ik}} \mid e_{ik} = 0 \right) = 0 \quad (4)$$

hold.

This assumption about the peer group effect implies that ability is worthless to fellow students if a student is not active. The product of average ability and average effort would not have this effect. The peer effect is not directly affected by the private background, as the latter is not directly or at least not entirely observable in the school. The private background influences the peer effect indirectly, as it co-determines the effort level of a student.

Finally, educational output generates the positive effect S on the entire society, with the following properties:

Assumption 3 *The post-educational utility of a student i at school k is denoted with w_{ik} . The utility effect for an individual student i from a marginal increase in social benefits is neglectible ($\frac{\partial w_{ik}}{\partial S} = 0$) The marginal benefit of education for the society is positive ($\frac{\partial S}{\partial q_i} > 0$) and large, such that individually optimizing students always underinvest in education.*

Assumption 3 allows a simpler identification of first best and second best allocations. Since they face a signalling problem, students can overinvest in education. Assumption 3 ensures that more educational output dominates less output and high levels of student effort are generally pareto-superior to low levels. The final look is on post-educational utility.

Assumption 4 *The post-educational utility w_{ik} increases linearly in a student's expected post-educational productivity $E(\Pi_{ik})$, which again depends on the educational output and the employer's*

estimation of a student's ability. The estimation of the ability depends again on the observed educational output:

$$w_{ik} = w(\Pi((E(\theta_i | Q_{ik}), Q_{ik}(\cdot))) \quad (5)$$

Of course, any error in the ability estimation should be observable after some time in a job but any agent will have at least a temporary (negative) benefit from an inaccurate estimation.

2.2 First best solution for a given peer group

The efficient solution maximizes the utility of all agents by choosing the appropriate level of effort:

$$\max_{e_{ik}} S + \int_i \sum_{k=1}^2 w_{ik} - C(e_{ik}) d\theta \quad (6)$$

The optimal effort level for each agent takes the effect of his behaviour on the success of his fellow students and the entire society into account:

$$\frac{\partial S}{\partial q_{ik}} \frac{\partial q_{ik}}{\partial e_{ik}} + \frac{\partial w}{\partial q_{ik}} \frac{\partial q_{ik}}{\partial e_{ik}} + \left(\frac{\partial S}{\partial q_{ik}} \frac{\partial q_{ik}}{\partial P_k} + \frac{\partial w_{ik}}{\partial q_{ik}} \frac{\partial q_{ik}}{\partial P_k} + \sum_{j \neq i} \frac{\partial w_j}{\partial q_j} \frac{\partial q_j}{\partial P_k} \right) \frac{\partial P_k}{\partial e_{ik}} = C'(e_{ik}) \quad (7)$$

The first terms represents the (large) marginal social benefit from educational output. The term $\frac{\partial w_{ik}}{\partial q_{ik}} \frac{\partial q_{ik}}{\partial e_{ik}}$ denotes the direct individual wage impact of increased effort. The term in the brackets bundles all the effects which derive from the increasing peer group effect at school k through increasing effort. The whole society, student i and his fellow students $j \neq i$ all benefit from increased effort of student i .

2.3 No sorting

The ability composition in both schools is identical, I denote the peer group effect with P_k^{ns} , such that $P_a^{ns} = P_b^{ns}$. It was assumed above, that firms only know the distribution of ability and background but not the ability of any single student. Additionally, employers observe the

individual output (Q_{ik}). Without tracking the employing firms estimate ability with the help of Q_{ik} . The wage then depends on expected productivity.

The objective of a student is to maximize the difference between the expected wage and his costs of effort by his choice of effort:

$$\max_{e_{ik}} E [w (E(\theta_i | Q_{ik}), Q_{ik})] - C(e_i) \quad (8)$$

Assumption 3 states that $\frac{\partial w_{ik}}{\partial S} = 0$. Hence, the transformation of the first order condition yields

$$\left(\frac{\partial Q_{ik}}{\partial e_i} + \frac{\partial Q_{ik}}{\partial P_k^{ns}} \frac{\partial P_k^{ns}}{\partial e_i} \right) \left(\frac{\partial w_i}{\partial Q_{ik}(P_k^{ns})} + \frac{\partial w_i}{\partial E(\theta_i)} \frac{\partial E(\theta_i)}{\partial Q_{ik}} \right) = C'(e_i) \quad (9)$$

The first bracket shows the marginal increase in productivity, the second one the marginal wage increase. In this second bracket, the first term ($\frac{\partial w}{\partial Q_{ik}}$) represents the marginal wage increase from increased productivity, while the second represents the wage increase from the improved ability signal ($\frac{\partial w}{\partial E(\theta_i)} \frac{\partial E(\theta_i)}{\partial Q_{ik}}$). The difference from the first best effort (see (7)) is twofold. Firstly, the student does not take into account the external effects of his effort on his peer group and the society. Secondly, the student invests more effort, because better educational results also increase the employers assessment of the student's ability ($\frac{\partial w}{\partial E(\theta_i)} \frac{\partial E(\theta_i)}{\partial Q_{ik}}$). Both differences work in opposite direction. The first one constitutes an underinvestment, as the effect of higher individual effort on his peer group are likely to be positive. The second one constitutes an overinvestment in a *per se* unproductive signal. Note, that this signal is distorted by a student's background ϕ_i , if ϕ_i is not identical for all agents.

2.4 Sorting

Suppose a perfect screening mechanism for θ exists. However, this information is not available for the employer. Then, the students with a high expected output ($E(Q_{i,a}) = q_{ia} > \hat{q}$) go to school $k = a$, the others go to school $k = b$. The threshold between both schools is given by \hat{q} ,

the expected output of an average student. The respective peer group effects are given by P_k^s . The superscript s indicates that students are sorted, ns to a peer group effect in a comprehensive school. Assumption 5 allows a simplification of the analysis:

Assumption 5 *Sorting according to expected productivity improves the peer group effect for the better students and reduces the effect for the worse students, such that*

$$P_a^s > P_k^{ns} > P_b^s \quad (10)$$

and

$$P_a^s + P_b^s \geq 2P_k^{ns} \quad (11)$$

Now, the utility function of a student at school a differs from the one in a non-sorting regime:

$$\max_{e_i, a} E [\Pi((E(\theta_i | Q_{i,a}, q_{i,a} > \hat{q})), Q_{i,a})] - C(e_i) \quad (12)$$

The term $(E(\theta_i | Q_{i,a}, q_{i,a} > \hat{q}))$ denotes the estimation of θ_i from the educational output, given that the individual *expected* output $q_{i,a}$ is greater than \hat{q} . Given a homogeneous private background the sorting outcome is equivalent to a sorting according to ability. The first order condition turns into

$$\left(\frac{\partial Q_{i,a}}{\partial e_i} + \frac{\partial Q_{i,a}}{\partial P_a^s} \frac{\partial P_a^s}{\partial e_i} \right) \left(\frac{\partial w_i}{\partial Q_{i,a}} + \left(\frac{\partial w_i}{\partial E(\theta_i)} \frac{\partial E(\theta_i)}{\partial Q_{i,a}} \mid q_{i,a} > \hat{q} \right) \right) = C'(e_i) \quad (13)$$

For students at school b , the first order condition is like-wise

$$\left(\frac{\partial Q_{i,b}}{\partial e_i} + \frac{\partial Q_{i,b}}{\partial P_b^s} \frac{\partial P_b^s}{\partial e_i} \right) \left(\frac{\partial w_i}{\partial Q_{i,b}} + \left(\frac{\partial w_i}{\partial E(\theta_i)} \frac{\partial E(\theta_i)}{\partial Q_{i,b}} \mid q_{i,a} < \hat{q} \right) \right) = C'(e_i) \quad (14)$$

Comparing (7), (9) and (13) yields proposition 1:

Proposition 1 *Neither a perfect sorting policy nor a non-sorting policy implement an efficient solution. Sorting dominates non-sorting if the peer group effects from sorting ability are sufficiently great and/or if the direct impact of ability on productivity and subsequently wages is sufficiently small.*

Proof. The first statement in the proposition is an immediate consequence from assumption 3. To compare the second-best solutions recall that the overall marginal increase in productivity (i.e. the term in the first bracket: $\left(\frac{\partial Q_{i,a}(P_k)}{\partial e_i} + \frac{\partial Q_{i,a}}{\partial P_k} \frac{\partial P_k}{\partial e_i}\right)$) is greater if students are tracked according to their ability (see assumption 3).

However, the second term in the bracket is smaller for these students. The first summand $\left(\frac{\partial w_i}{\partial Q_{i,k}}\right)$ is identical while the second one is strictly smaller. The term $\left(\frac{\partial E(\theta_i)}{\partial Q_{ik}}\right)$ makes the difference. In case of a non-sorting policy the expected estimated ability increases continuously with increasing educational output such that

$$\lim_{e \rightarrow \infty} E(\theta_i | Q_{ik}) = \infty$$

and

$$\lim_{e \rightarrow 0} E(\theta_i | Q_{ik}) = 0$$

With sorting, the expected output is bound by an upper or, respectively, lower limit. This implies for school a with the better students also

$$\lim_{e \rightarrow \infty} (E(\theta_i | Q_{ia}) | q_{ia} > \hat{q}) = \infty$$

but

$$\lim_{e \rightarrow 0} (E(\theta_i | Q_{ia}) | q_{ia} > \hat{q}) > 0$$

because each student has an expected output $q \geq \hat{q} > 0$. This condition again implies $\theta > 0$, since θ is a factor of q . Hence, the marginal increase of the expected ability for a student at school a is lower than in the case of non-sorting.

For school b with the less able students sorting means

$$\lim_{e \rightarrow \infty} (E(\theta_i | Q_{ib}) | q_{ib} < \hat{q}) < \infty$$

and

$$\lim_{e \rightarrow 0} (E(\theta_i | Q_{ib}) | q_{ib} < \hat{q}) = 0$$

Again, the marginal increase of the expected ability for students at school b is lower than in the case of non-sorting. Even with limitless effort, estimated ability is finite, because employers know that $q_{ib} < \hat{q} < \infty$. If expected output is finite, ability as a factor of expected output is finite, too.

Students with an expected productivity sufficiently higher (lower) than \hat{q} are unaffected by this threshold effect. The lower incentives are less important if $\frac{\partial w_i}{\partial E(\theta_i)}$ is small. Then again, sorting dominates. ■

The lower incentives are caused by the fact that sorting of students already produces a signal about a student's ability. Hence, there is less need to invest in educational output in order to improve the quality signal. A decision on sorting or non-sorting depends on the empirical properties of the peer group effect. If the effects are large, sorting dominates non sorting. As corollary 1 shows sorting this signalling effect is low for students with a very high or low ability.

Corollary 1 *Students with a very high expected output prefer selective schools, students with a very low expected output prefer comprehensive schools. Sorting increases the inequality between these two types of students.*

Proof. For students with a very high expected output,

$$\lim_{q_i \rightarrow \infty} \left(\frac{\partial w_i}{\partial E(\theta_i)} \frac{\partial E(\theta_i)}{\partial Q_{i,a}(P_a^s)} | q_i > \hat{q} \right) = \frac{\partial w_i}{\partial E(\theta_i)} \frac{\partial E(\theta_i)}{\partial Q_{i,a}(P_a^s)} \quad (15)$$

holds, which makes equation (13) similar to equation (9). However, incentives are stronger because the peer group effect is greater. For low expected output students, sorting means

$$\lim_{q_i \rightarrow 0} \left(\frac{\partial w_i}{\partial E(\theta_i)} \frac{\partial E(\theta_i)}{\partial Q_{i,b}(P_b^s)} \mid q_i < \hat{q} \right) = \frac{\partial w_i}{\partial E(\theta_i)} \frac{\partial E(\theta_i)}{\partial Q_{i,b}(P_b^s)} \quad (16)$$

Therefore, the incentives are weaker in a sorting regime because peer group effects are smaller. The diverging preferences imply greater inequality among students under a selective regime. ■

This corollary implies also, that students with a high expected output will leave a comprehensive school, if they can also go to a selective one, e.g. a private school or a selective school in a neighboring jurisdiction. Of course, such an outside option undermines all benefits from comprehensive schooling. Following assumption 5, peer group effects increase with any further sorting, i.e. with more hierarchically stratified schools.

Corollary 2 *Sorting in more than two different levels decreases the incentives to invest in the ability signal.*

Proof. The proof of proposition 1 can be applied. Further sorting decreases the range of possible estimated abilities even more and subsequently the incentives to invest in educational output as an ability signal. ■

3 Intertemporal incentives

Most educational systems do not sort in elementary school but at some later point. In this section, the incentives from a selective school system for pre-selection students are analysed. Consider a two period process. In period $t = 1$ (or in primary education) all students are educated in identical schools. In period $t = 2$ (or secondary education) the students are either sorted or they are kept in the same type of school as in period 1. Now, sorting depends on their observed output in the first period ($Q_{i,k,1}$). This criterion allows for a cheap and simple selection mechanism, I consider

all other alternatives as prohibitively expensive. If $Q_{i,k,1}$ is greater than the output of the average student \hat{Q} , then the student will be sorted into better peer group. Students can choose their respective effort level at the start of each period. The output in period 2 additionally depends on the observed output in period 1, such that

$$Q_{ikt} = q(\phi_i, \theta_i, e_{ikt}, P_{kt}, Q_{i,t-1}) + \varepsilon_{i,t} \quad (17)$$

with $q_{i,0} = 1$ holds. Let Assumption 5 hold for period 2. The separation into two periods changes the incentive structure of education as students in a tracking system have higher incentives in the first period.

The analysis of intertemporal incentives is backwards. Period 2 is analogous to what has been discussed in the previous section, with three notable exceptions. Firstly, the peer group effects are different because sorting is based on the observable output $Q_{i,k,1}$, which includes the random variable $\varepsilon_{i,1}$. Hence, some less able or less engaged students will slip into the better peer group and vice versa. This exception does not imply changes in the qualitative results of the previous section. Secondly, the impact of the output in period 1 has to be taken into account. Therefore, further qualitative differences between a selective and comprehensive system stem from differences in the pre-selection period.

In this period 1, consider first the non-sorting case. The problem of a student is the following:

$$\max_{e_{i,1}} U_{i,1} = w(\Pi_{i,2}) - C(e_{i,1}) \quad (18)$$

which implies

$$\frac{\partial w(\Pi_{i,2})}{\partial Q_{i2}} \frac{\partial Q_{i2}}{\partial Q_{i,1}} \frac{\partial Q_{i,1}}{\partial e_{i,1}} = \frac{\partial w(\Pi_{i,2})}{\partial e_{i,1}} = C'(e_{i,1}) \quad (19)$$

The school indicator k has been ignored in the notation because all schools are identical in both periods. With a comprehensive school system students are only motivated by the effect on their post-educational wages.

For the sorting case, student i has to solve

$$\max_{e_{i,1}} U_{i,1} = \Pr(Q_{i1} > \hat{Q}) w(\Pi_{i,k=a,2}) + \left(1 - \Pr(Q_{i1} > \hat{Q})\right) w(\Pi_{i,k=b,2}) - C(e_{i,1}) \quad (20)$$

From the relevant first order condition follows

$$g(Q_{i1} - \hat{Q}) \Delta w_{i,2} + \Pr(Q_{i1} > \hat{Q}) \frac{\partial w(\Pi_{i,a,2})}{\partial e_1} + \left(1 - \Pr(Q_{i1} > \hat{Q})\right) \frac{\partial w(\Pi_{i,b,2})}{\partial e_1} = C'(e_{i,1}) \quad (21)$$

with

$$\Delta w_{i,2} = w(\Pi_{i,a,2}) - w(\Pi_{i,b,2}) \quad (22)$$

Recall that $g(\cdot)$ is the density function of the random variable. The indicator k has been ignored for all period 1 variables because all schools are identical in the first period. The selective system motivates student in period 1 by increasing its chance to be promoted to the good school $\left(g(Q_{i1} - \hat{Q}) \Delta w_{i,2}\right)$ as well as via the wage increases from overall increased educational output. To judge the efficiency of sorting, one has to compare (19) with (21).

Proposition 2 *Assume that the incentives of an average student in secondary education are identical in the sorting and in the non-sorting regime. Then, the following results hold:*

1. *First period incentives in the sorting regime are lower for the students with low expected output and higher for the average student and for students with very high expected output.*
2. *Increasing uncertainty (increasing σ^2) reduces the incentives for the students with average ability and increases the incentives for the students at the margins of the ability distribution.*

Proof. The term $g(q_{i1} - \hat{Q}) \Delta w_{i,2}$ in equation (21) denotes the increase in probability to get the higher returns from better schooling. Due to the properties of the error term, the following conditions hold for students with very high expected output and for students with very low expected

output:

$$\lim_{q_i \rightarrow 0} g(Q_{i1} - \hat{Q}) = \lim_{q_i \rightarrow \infty} g(Q_{i1} - \hat{Q}) = 0 \quad (23)$$

$$\lim_{q_i \rightarrow 0} \Pr(Q_{i1} > \hat{Q}) = 1 - \lim_{q_i \rightarrow \infty} \Pr(Q_{i1} > \hat{Q}) = 0 \quad (24)$$

The peer group effects (see assumption 3) imply

$$\frac{\partial w(\Pi_{i,a,2})}{\partial e_1} > \frac{\partial w(\Pi_{i,2})}{\partial e_{i,1}}$$

and

$$\frac{\partial w(\Pi_{i,b,2})}{\partial e_1} < \frac{\partial w(\Pi_{i,2})}{\partial e_{i,1}}$$

These implications establish the results for the high ability students and the low ability students.

For the average student,

$$g(Q_{i1} - \hat{Q}) \Delta w_{i,2} + \Pr(Q_{i1} > \hat{Q}) \frac{\partial w(\Pi_{i,a,2})}{\partial e_1} + (1 - \Pr(Q_{i1} > \hat{Q})) \frac{\partial w(\Pi_{i,b,2})}{\partial e_1} > \frac{\partial w(\Pi_{i,2})}{\partial e_{i,1}} \quad (25)$$

has to hold. Since both schools have the same amount of students, the expected Output in period 1 for a student is defined by the threshold value \hat{Q} . such that $g(Q_{i1} - \hat{Q}) = g(0) > 0$. The initial assumption in the proposition implies

$$\frac{1}{2} \frac{\partial w(\Pi_{i,a,2})}{\partial e_1} + \frac{1}{2} \frac{\partial w(\Pi_{i,b,2})}{\partial e_1} = \frac{\partial w(\Pi_{i,2})}{\partial e_{i,1}}$$

which establishes equation 25.

For result 2 notice that $g(Q_{i1} - \hat{Q})$ increases with increasing uncertainty at the margins of the distributions of possible output and increases at $q = \hat{Q}$. ■

The implication of result 1 is crucial. Assume that sorting and non-sorting had been pareto-equivalent in the previous section, in which education was a one-period process. Then take the pre-selection period into account. Now sorting dominates non-sorting because it overall sets better incentives in the first period and thus produces higher outcome in period 2 as well.

4 The positive impact of imperfect sorting

In the previous section, sorting was based on results in period 1. This process is not perfect as it depends in parts on some random effects. Hence, result based sorting can lead to imperfect tracking, as this implies sorting according to marginal productivity in period 1. If marginal productivity changes some of the more productive students are systematically excluded from the better school.

Lemma 1 *Let $\bar{Q}_{k,2}^s$ denote the average output in school k in the selective system in period 2. Then, some level of imperfectness in sorting dominates perfect sorting and comprehensive schooling, if*

$$\frac{\partial^2(\bar{Q}_{a,2}^s + \bar{Q}_{b,2}^s)}{(\partial(P_{a,2} - P_{b,2}))^2} < 0 \quad (26)$$

and

$$\frac{\partial^2 \left(\frac{\partial E(\theta_i)}{\partial Q_{ik}^s} \right)}{(\partial(\Pr(q_{i1} < \hat{q} \mid Q_{i1}^s > \hat{Q})))^2} < 0 \quad (27)$$

are negative and sufficiently small.

Proof. The conditions in the lemma require

- that the marginal benefit from better sorting declines with the quality of sorting and
- that the marginal increase in signalling incentives due to less perfect sorting decreases with lower selection quality.

The first order effects for the these conditions are positive:

$$\frac{\partial(\bar{Q}_{a,2}^s + \bar{Q}_{b,2}^s)}{\partial(P_{a,2} - P_{b,2})} > 0 \quad (28)$$

$$\frac{\partial \left(\frac{\partial E(\theta_i)}{\partial Q_{ik}^s} \right)}{\partial(\Pr(q_{i1} < \hat{q} \mid Q_{i1}^s > \hat{Q}))} > 0 \quad (29)$$

Average productivity increases with the spread in the peer group effects. The observed output has a greater impact on estimated ability, if the probability of wrong sorting increases. Both have

positive, but contradictory, incentive effects. The effects are contradictory, because a better peer group effect means better sorting, but better sorting reduces the probability of being sorted in the wrong type of school. The conditions (26) and (27) mean that marginal increase in post educational productivity decrease with increasing peer group differences and decreasing perfection of sorting. Therefore, the educational policy maker faces a trade-off problem with an interior solution. ■

Lemma 1 says, that imperfect sorting may actually improve welfare. The imperfect sorting encourages the students in both schools to invest more in education as a signal than the overall decrease in productivity due to lower peer group effects discourages them. This result is equivalent to turning the discrete choice problem of (perfect) sorting or integrating into a continuous choice problem. A policy maker will introduce some imperfection in the sorting process to the set the best possible incentives.

5 Imperfect sorting and educational support

The previous sections have shown, that some imperfection in sorting can dominate both comprehensive schooling and perfect sorting. In this section it is shown how the principal can influence the sorting process without introducing a more complicated and more expensive selection process. Suppose ϕ_i is not identical anymore for all students. To simplify the analysis, consider the following restrictions. There are four groups of students of equal size which differ in their respective ability and background level. Students can have either high ability ($\bar{\theta}$) or low ability ($\underline{\theta}$). Their background is either rather supportive ($\bar{\phi}$) or restrictive ($\underline{\phi}$). A less advantageous private background ϕ_i restricts the educational chances of otherwise smart students. For a further simplification assume that all students with the characteristics $(\bar{\theta}, \bar{\phi})$ will be selected to the good school a and all students with the characteristics $(\underline{\theta}, \underline{\phi})$ go to the bad school b . Heterogeneity in the private background introduces imperfect sorting, because it is unclear if the remaining slots in school a will be occupied by students with strong ability or a strong private background. Table 1 shows the outcome of sorting.

Table 1 *Enrolment of students after secondary education*

	Students with ability	
	$\underline{\theta}$	$\bar{\theta}$
and background		
$\underline{\phi}$	b	$xa + (1 - x)b$
$\bar{\phi}$	$xb + (1 - x)a$	a

The variable x denotes the probability of students with the characteristics $(\bar{\theta}, \underline{\phi})$ being enrolled in school a . The principal can modify this probability terms for the two critical groups $(\bar{\theta}, \underline{\phi})$ and $(\underline{\theta}, \bar{\phi})$ by adjusting a support ρ_{kt} to every student at a school k in period t . The educational production function changes into

$$Q_{ik2} = q((\phi_i + \rho_{k2}), \theta_i, e_{ik2}, P_{k2}, Q_{i1}((\phi_i + \rho_{k1}))) + \varepsilon_{i2}$$

To meet participation constraints of the agents, the support has to be greater than the background parameter. ($\rho_{kt} > -\phi$). Since no one can distinguish the students in period 1, ρ_{k1} is identical for all i and k . In period 2, it is possible to distinguish between between students from the different schools in the case of sorting. No more information is important for the identification of the impact of a change in student support on the sorting outcome.

Proposition 3 *Let the conditions (26) and (27) hold. Hence, some degree of imperfect sorting dominates both perfect sorting and comprehensive schooling. Students are sorted according to observed output in period 1. The principal provides some support ρ_{k1} , ρ_{a2} and ρ_{b2} . Then, the principal can modify the signal from sorting in three ways.*

1. *The principal increases support in period 1 to all students (ρ_{k1}) in order to improve the ability signal from the sorting process in period 2.*
2. *The principal increases the relative support for the worse school b in period 2 (ρ_{b2}) in order to get a less clear ability signal from sorting.*

3. *The principal increases the relative support for students at the better school a in period 2 (ρ_{a2}) in order to get a clearer ability signal from sorting.*

Proof. The impact of a change in the support ρ_{k1} is higher for students with $\underline{\phi}$ since the second-order derivatives for Q_{it} are negative. This increase reduces the relative importance of differences in private background for the sorting. Additionally, such a measure enhances the importance of θ_i , since the cross derivatives in $q((\phi_i + \rho_{k2}), \theta_i, e_i, P_k, Q_{i1}((\phi_i + \rho_{ik1})))$ are positive. Hence, the probability of enrolment in school a increases for students with the characteristics $(\bar{\theta}, \underline{\phi})$. (Result 1)

Increasing the support for students at school b in period 2 decreases the incentives in period one because the spread between the outcomes decreases $\left(\frac{\partial g(Q_{i1} - \hat{Q}) \Delta w_i}{\partial \rho_{b2}} < 0 \text{ in equation (21)} \right)$. Due to the positive cross-derivatives and the negative second-order derivatives, this decline in incentive weight is particularly acute for students with high ability $(\bar{\theta})$, because they benefit more strongly from the given support at school b in period 2. (Result 2) Increasing the support at school a boosts incentives of everyone, but, for the same reasons, in particular the incentives of the students with high ability $(\bar{\theta}, \bar{\phi})$. (Result 3)

■

Table 2 offers insight in the relative incentive effects of increasing the support at different schools and stages. The (down-)upward direction indicates an (de-)increase. The Sideward arrows indicate that the overall direction of relative incentives is not clear. The left vs. right indicate that the individual effects go into different directions.

Table 2 *The impact of changes in support on relative incentives for students with heterogenous characteristics.*

	Increasing support at	
	School <i>a</i>	School <i>b</i>
in Period 1	improves ability signal from sorting relative incentive developments: $\bar{\theta}, \underline{\phi} \uparrow$ $\underline{\theta}, \bar{\phi} \downarrow$	
in Period 2	improves ability signal from sorting	distorts ability signal from sorting
	relative incentive developments:	
Effect in period 1	$\bar{\theta} \uparrow$ $\underline{\theta} \downarrow$ $\underline{\phi} \rightarrow$ $\underline{\phi} \leftarrow$	$\bar{\theta} \downarrow$ $\underline{\theta} \uparrow$ $\underline{\phi} \rightarrow$ $\underline{\phi} \leftarrow$
Effect in period 2	$\bar{\theta}, \underline{\phi} \uparrow$ $\underline{\theta}, \bar{\phi} \downarrow$	$\bar{\theta}, \underline{\phi} \uparrow$ $\underline{\theta}, \bar{\phi} \downarrow$

From an equality point of view, increasing support in period 1 promotes the equality of opportunities but reduces equality of outcome. Increasing support for students at the worse school *b* in period 2 supports equality of outcome, but inhibits equality of opportunity. However, increased inequality of opportunity sets better incentives to invest in education as a signal. Therefore, the trade-off between signalling incentives and greater peer-group effects contains an implicit trade-off between more equality of opportunity and more equality of outcomes.

6 Conclusion

The paper has discussed a reason why positive effects from sorting students are not observable, even if standard assumptions about peer group effects are not rejected. Sorting already provides a signal about unobservable ability so that students do not have to work as hard in secondary education any more. Introducing ability grouping transfers incentives from secondary into primary education. The paper has also shown how a policy maker can distort ability grouping by changing student support at the different schools in primary or secondary education. As a result some

inequality in opportunities is necessary to get a maximal educational output.

The implications are easy to summarize. Students incentives matter and can backfire on otherwise sensible education policies. On the other side, the paper introduces an applicable policy mechanism, even if direct relationships between educational resources and results are difficult to observe. In particular, the paper emphasizes the significance of primary education for educational results in a selective system.

For the research agenda, it is important to consider results from primary education when discussing the organization of secondary education. Even though student effort may not be directly observable in the data, the paper provides a simple hypothesis for future empirical research. For example, the results on incentives in primary education indicate that educational performance should increase less strong between primary education and secondary education in a selective schooling system than in a comprehensive one. Support policies in primary and secondary education could be taken into account as controls.

References

Angrist, J. and Lavy, Victor. "The Effect of High School Matriculation Awards: Evidence from Randomized Trials.," NBER, 2003.

Argys, Laura M.; Rees, Daniel I. and Brewer, Dominic J. "Detracking America's Schools: Equity at Zero Cost?" *Journal of Policy Analysis and Management*, 1996, 15(4), pp. 623-45.

Brunello, Giorgio and Giannini, Massimo. "Stratified or Comprehensive? The Economic Efficiency of School Design." *Scottish Journal of Political Economy*, 2005, 51(2), pp. 173-93.

Brunello, Giorgio; Giannini, Massimo and Ariga, Kenn. "The Optimal Timing of School Tracking," IZA, 2004.

De Fraja, Gianni and Landeras, Pedro. "Could Do Better: The Effectiveness of Incentives and Competition in Schools." *Journal of Public Economics*, 2006, 90, pp. 189-213.

De Fraja, Gianni; Oliveira, Tania and Zanchi, Luisa. "Must Try Harder. Evaluating the Role of Effort on Examination Results.," CEPR, 2005.

Dobbelsteen, Simone; Levin, Jesse and Oosterbeek, Hessel. "The Causal Effect of Class Size on Scholastic Achievement: Distinguishing the Pure Class Size Effect from the Effect of Changes in Class Composition." *Oxford Bulletin of Economics and Statistics*, 2002, 64(1), pp. 17-38.

Epple, Dennis; Newlon, Elizabeth and Romano, Richard E. "Ability Tracking, School Competition, and the Distribution of Educational Benefits." *Journal of Public Economics*, 2002, 83, pp. 1-48.

Hanushek, Eric A. and Wößmann, Ludger. "Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence across Countries." *Economic Journal*, 2005, in press.

Hoxby, Caroline. "School Choice and School Competition: Evidence from the United States." *Swedish Economic Policy Review*, 2003, 10, pp. 11-67.

Hoxby, Caroline M. "Peer Effects in the Classroom: Learning from Gender and Race Variation," NBER, 2001.

Kremer, Michael; Miguel, Edward and Thornton, Rebecca. "Incentives to Learn," Cambridge, MA: Harvard University, 2004.

Lazear, Edward P. "Educational Production." *The Quarterly Journal of Economics*, 2001, 116(3), pp. 777-803.

Meier, Volker. "Choosing between School Systems: The Risk of Failure." *Finanzarchiv*, 2004, 60(1), pp. 83-93.

Spence, Michael. "Job Market Signaling." *The Quarterly Journal of Economics*, 1973, 87(3), pp. 355-74.

West, Martin R. and Wößmann, Ludger. "Which School Systems Sort Weaker Students into Smaller Classes? International Evidence." *European Journal of Political Economy*, 2005, in press.

Wößmann, Ludger. "How Equal Are Educational Opportunities? Family Background and Student Achievement in Europe and the Us," CESifo, 2004.