A Simple Model of Educational Production

Christian Jaag

2006

A Simple Model of Educational Production

Christian Jaag[*]

January 4, 2006

Abstract

There is a large body of literature on the effect of educational resources on student performance, such as teacher qualification, class size, and physical resources in school. It is dominated by empirical studies which often find ambiguous effects of resource spending on student outcomes. The unique contribution of this paper is the provision of a framework to study educational production with differentiated input factors, which allows for closed-form solutions. We try to interpret the empirical findings on the basis of a simple theoretical model of educational production: Class size, employed school resources and student effort are endogenously determined in order to account for differences in educational achievement. We also discuss the choice of integrated vs. segregated classes. Optimum class size and school quality increase with higher discipline, while in equilibrium overall classroom disruption is equal in all classes.

1 Introduction

This paper lays out a theoretical framework, which allows the analysis of schools' and students' incentives to engage optimally in educational production. The students' motivation in studying is straightforward: They profit from their educational achievement and their accumulated human capital. However, also society profits from the individuals' education via positive externalities and the possibility to tax away some of the gain in wages due to higher productivity.

From an economic point of view, most important in the analysis of education is the elaboration and the solution of the main trade-offs faced by individuals. The attempt of this paper is to model the sequence of decisions in the educational system in order to be able to scrutinize the interaction between the different factors in educational production. We also discuss the case of segregated vs. integrated classes with respect to efficient educational production. For this purpose, the research by Lazear (2001), who develops a theory of optimum class size, and Wössmann (2002), who takes an institutional approach to educational production, are merged to an integrated model of education with closed-form solutions. Decision makers are the government via schools and students. Studying effort and school quality are the choice variables constituting the core of the model. Schools and teachers act on behalf of society as a whole and are not modelled to be independent utility maximizers. It turns out that the empirically missing effect of class size on educational achievement is well explained by our model. Concerning the impact of classroom computers, teacher training and relative teacher pay, our basic model is more optimistic than what the empirical literature suggests. The difference might be due to the unobservability of student types and the resulting misallocation of students to different class types. We will argue against this objection and consider a reconciliation with empirical ecidence in section 6.2.

As already mentioned, effort choices by both students and schools are discussed without considering teacher incentives explicitly. This paper proceeds with section 2 discussing related literature. Section 3 gives an outline of our model with homogeneous students and discusses the process of educational production. In section 4, we present the centralized solution to the school's and the student's optimization problems as a benchmark. Section 5 describes the individual decisions in the order which is implied by the backward induction solution method. In section 6, an extension of the model is given, in which the optimum treatment of differentiated student types is discussed; section 7 concludes.

2 Related Literature

Most of the worldwide policy discussions on school performance concentrate on the factors influencing education at school. Among the schooling inputs generally considered to be important are class size, teacher qualification and expenditure on physical resources (cf. Hanushek, 2002). One of the first influential studies to point out that pure input-based education policies do not make a difference is by Hanushek (1994). He conducts a series of literature reviews, which support the conclusion that increased spending in general and smaller class sizes in particular do not systematically lead to improved student achievement. Hanushek (2001) concludes that educational policy decisions should not focus on school resources because the impact of resources on student achievement is unknown at the time. Addressing the question whether new technologies, such as computer aided instruction, improve learning, Angrist and Lavy (2002) find that there is no such influence.

Rice (2003) describes the importance of various teacher attributes in educational production. Hanushek (1996) argues that there is no consistent evidence linking student achievement to teacher characteristics. However, it could be the case that they are important but that variation in teacher quality is driven by factors that are difficult to measure. Hence, identification of teacher fixed effects requires matched student-teacher data with observations on student achievement and teachers over multiple years. Rockoff (2004) finds that a one standard deviation increase in teacher quality raises students' test scores by approximately 0.1 standard deviations in reading and math. Moreover, also teacher experience adds to student performance: Reading test scores differ by 0.17 standard deviations on average between beginning and experienced teachers. Rivkin, Hanushek, and Kain (2005) find that teachers are a major determinant of student performance, but do not describe teacher quality in terms of specific qualifications and characteristics. Their research identifies teacher quality as the most important school-related factor influencing student achievement. They conclude from their analysis of 400,000 students in 3,000 schools that, while school quality is an important determinant of student achievement, the most important predictor is teacher quality. In comparison, class size, teacher education, and teacher experience play a small role. Ferguson and Ladd (1996) and Sanders and Rivers (1996) also argue that the single most important factor affecting student achievement is teachers. Further, they contend that lower achieving students are the most likely to ben-

efit from increases in teacher effectiveness. However, explicitly testing teachers before hiring them seems not to be able to select high quality teachers: Inquiring the effect of teacher tests on quality, Angrist and Guryan (2003) find that states in the U.S. which introduced teacher tests ended up paying higher teacher salaries with no measurable increase in teacher quality. This suggests that tests are rather a barrier to entry than an effective quality screen.

An important determinant of the per-student cost of schooling is class size. A challenge to the view that resources do not matter has arisen by the STAR (Student Teacher Achievement Ratio) study conducted in the state of Tennessee. There, smaller class sizes with randomly assigned students and teachers resulted in significantly enhanced achievement for children (cf. Mosteller, 1995). Studies based on the STAR experiment find that class size has a significant effect on test scores: Reducing class size from 22 to 15 in the early primary grades increases both math and reading test scores by ca. 0.2 standard deviations (cf. e.g. Krueger, 1999).

Angrist and Lavy (1999) observe that in their data the observed association between class size and student achievement is always perverse. However, using a natural experiment in Israel, they find that reductions in class size induce a significant and substantial increase in the students' math and reading achievement. In Mishel and Rothstein (2002), Hanushek and Krueger lay out their interpretation of basically the same evidence. Notwithstanding their difference of opinion, there is some consensus. The most important is that both agree that smaller classes can matter in some circumstances. There are a number of explanations to account for the inconsistencies among the non-experimental studies on the effect of class size, namely poor measurement of key variables, and model specification issues.
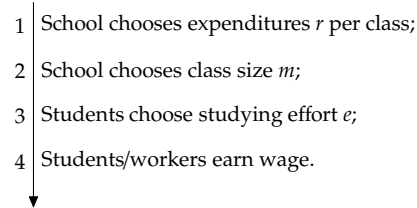
3   Outline of the Model

Our model treats teacher effort in class as a public good subject to negative spillovers from each of the students. It reproduces education decisions by students and schools. For simplicity, we assume that human capital can be traded in for a money wage on a one-to-one basis. Thus, human capital is the numéraire good in the model.

A school and students are the two actors. The basic idea behind the economic model of education in schools is that the involved actors maximize the difference between their benefits and their costs, i.e. their net benefits. Students earn

the market value of their human capital which is taxed at rate τ and bear their individual effort cost. The school maximizes its net surplus from the individuals' human capital minus the cost of its school resources. The exact form of the school's objective and its discussion follow below in the context of optimum class size and schooling resource spending choice.

For the sake of simplicity, we take the income tax rate τ as exogenously given. The decisions in the model are taken in four stages as depicted in figure 1.

1	School chooses expenditures r per class;
2	School chooses class size m;
3	Students choose studying effort e;
4	Students/workers earn wage.

Figure 1 Timeline of decisions in the model.

In the first stage, schools decide on their resources spent. Schooling resources comprise physical resources as well as the quality and education of the hired teachers. In the second stage, schools optimize class size. By separating the expenditure and class size choice of schools, we reproduce the actual decision process in autonomous schools, which typically face exogenous restrictions concerning their infrastructure and physical resources but are free to choose class size in the short run. In the third stage, students maximize their net benefit from schooling by choosing the right amount of effort they are willing to make. In the fourth stage, the acquired human capital in the form of skills is translated into wages on the labor market. This sequence of decisions can be solved by backward induction in order to reach a subgame perfect equilibrium.

A student's direct productivity h from studying during a certain time period is a function of her own effort e and employed school resources per class r. For simplicity, we assume a Cobb-Douglas production function for education:

$$h = e^\alpha r^\beta \qquad \alpha + \beta < 1.$$

Student effort e is controlled by the student herself, reflecting her motivation, time and engagement devoted to learning. r is the amount of educational resources employed in teaching. Contained in this variable are the physical resources in a classroom, such as computers, but also a teacher's education, technical as well as methodical. These resources are the school's choice variable which

it can commit to before class size and student effort are chosen by the respective decision-maker. The assumption of diminishing returns to scale reflects additions to education being incresingly hard to produce.[1]

Formal classroom schooling is a public good subject to congestion. Educational resources spent by a school are basically non-rival. However, there is a congestion effect as in Lazear (2001), since every student has a certain probability of disrupting in class, such that it is not possible to spend all the time at school studying. The ability of a student to learn something in class thus depends on the behavior of others in the same class. If one student disrupts classwork, the entire class suffers; the teacher's and the other students' concentration is diverted from studying. Let π be the probability that a student is not misbehaving at any moment in time, $\pi \in [0, 1[$. Then, the probability that all students in a class of size m are behaving is π^m, which is also equal to the proportion of schooling time during which students are effectively studying. Thus, human capital P is produced according to

$$P = ph, \tag{1}$$

where the productive time in school, $p = \pi^m$. Of course, there are also positive spillovers provided by the students to one another. However, the range of class size where adding students produces net positive externalities is not relevant for an optimizing school. Since an increase in class size reduces cost per student, an optimizing school will increase class size up to the point where additional students have negative effects on each others.

4 Centralized Solution

As a benchmark, we solve the optimum education production problem from a social planner's point of view. There are social benefits from education which consist of (1) an increase in workers' productivity and (2) positive externalities from education.[2] Total costs consist of the students' effort cost and the cost of resources employed in schooling. The social net benefit from education per class writes as

$$V^\# = \max_{e, m, r \in \mathbb{R}_+} \left\{ \rho m \pi^m e^\alpha r^\beta - m c^S e - c^T r \right\}, \tag{2}$$

---

[1]Note that in equations 6 and 8 effort costs are assumed to be linear, such that additional effort coneas at equal cost but is only decreasingly effective.

[2]On the evidence of externalities of education cf. e.g. Acemoglu and Angrist (2001), Lochner and Moretti (2001) and Dee (2003).

where $\rho > 1$ collects all private and external benefits from education and c^S and c^T denote (constant) unit costs of student effort and educational resources, respectively. The first-order conditions with respect to m, r, and e write as

$$m: \quad \rho\pi^m e^\alpha r^\beta + \rho m \pi^m \ln \pi e^\alpha r^\beta \overset{!}{=} c^S e,$$

$$r: \quad \rho\beta m \pi^m e^\alpha r^{\beta-1} \overset{!}{=} c^T,$$

$$e: \quad \rho\alpha m \pi^m e^{\alpha-1} r^\beta \overset{!}{=} m c^S.$$

The optimum values for class size, resource spending, and student effort follow directly:

$$m^\# = -\frac{1-\alpha}{\ln \pi}, \tag{3}$$

$$r^\# = \left[\left(\frac{\alpha}{c^S}\right)^\alpha \left(\frac{\beta m^\#}{c^T}\right)^{1-\alpha} \rho\pi^{m^\#} \right]^{\frac{1}{1-\alpha-\beta}}, \tag{4}$$

$$e^\# = \left[\left(\frac{\alpha}{c^S}\right)^{1-\beta} \left(\frac{\beta m^\#}{c^T}\right)^\beta \rho\pi^{m^\#} \right]^{\frac{1}{1-\alpha-\beta}}. \tag{5}$$

5 Student Effort vs. School Quality

In the decentralized decision process, schools and students optimize their efforts individually. We find the equilibrium values of r, e, m, and P by solving the model backwards.

5.1 Student Effort

As mentioned above, students have their effort as choice variable which they use to maximize their net benefit from education. A student S chooses her level of effort e given the labor income tax rate τ, school resources r and class size m. The student's benefit B^S

$$B^S = (1-\tau) P = (1-\tau) p e^\alpha r^\beta.$$

The costs of a student's effort, C^S are assumed to be linear in effort:

$$C^S = c^S e, \tag{6}$$

where c^S is a cost parameter. These costs represent opportunity costs of the time spent in class and at home and the psychic energy of learning. A student maximizes her net benefit $NB^S = B^S - C^S$ with respect to the level of effort e,

$$e^* = \arg\max_{e \in \mathbb{R}_+} \left\{ (1-\tau)\, pe^{\alpha} r^{\beta} - c^S e \right\}.$$

By increasing her effort, a student raises both educational achievement and total effort cost. The first-order condition yields the optimal level of effort for any given tax rate, class size and level of resource spending by the schools:

$$e^* = \left(\frac{\alpha}{c^S} (1-\tau)\, pr^{\beta} \right)^{\frac{1}{1-\alpha}}.$$

Since school quality and student effort complement each other, student effort increases in the school's resource spending and decreases in class size. An increase in the income tax rate reduces the individual benefit from achievement and therefore leads to a decrease in student effort.

5.2 School Quality

5.2.1 Incentives to Schools

The gross benefit to students is given by their human capital which they acquire at school: $(1-\tau)\, P$. In our model, also teachers benefit from their students' educational success: Be it via performance pay or self motivation. At the moment, we simply assume that a teacher's/school's benefit from education is summarized as σP. In order to analyze the respective incentives to invest, we first have to define school quality: In the literature, school quality is usually measured by class size, teacher pay and -training and technological educational inputs such as classroom computers.[3] In our analysis of a school as an institution, which produces education, we separate the *class size* aspect of school quality from teacher pay and -training and technology which are subsumed under the term *educational resources*.

In the simple model discussed here, we abstract from agency problems between society, the government, schools and teachers and assume that teachers and schools act as one, without own interests involved. Since there are no school-specific

---

[3]For evidence on the effect of school quality on student learning cf. Card and Krueger (1992) and Angrist and Lavy (2001, 2002).

costs, school size does not matter: For simplicity, we may assume that a school consists of one teacher and one class only. The general trade-off for a school with a fixed budget per student is that an increase in class size directly decreases educational attainment but allows for higher resource spending which benefits student learning. This tradeoff is illustrated in section 5.2.4 below. In the next two sections, we first discuss the choice of optimum class size and resource spending separately.

5.2.2 Class Size Optimization

To determine the optimum class size m, we have to look at the benefit B^T and costs C^T of a school, given employed school resources r. The school's benefit per class is given by the number of students per class, m, times each student's contribution to the social benefit perceived by schools σP:

$$B^T = m\sigma P = m\sigma p \left(\frac{\alpha}{c^S} (1 - \tau) \, pr^\beta \right)^{\frac{\alpha}{1-\alpha}} r^\beta. \tag{7}$$

Note that $p = \pi^m$. The school's costs are assumed to be a linear function of its resources it employs per class. They are given by

$$C^T = c^T r, \tag{8}$$

where c^T is a cost parameter. Resource costs consist of the spending for physical inputs, such as books and coputers, as well as teacher training and salary. Abstracting from the natural condition that class size must be integer, the optimum class size m can be determined by maximizing a school's net profit with respect to class size

$$m^* = \arg \max_{m \in \mathbb{R}_+} \left\{ m\sigma p \left(\frac{\alpha}{c^S} (1 - \tau) \, pr^\beta \right)^{\frac{\alpha}{1-\alpha}} r^\beta - c^T r \right\}.$$

An increase in class size reduces the time of effective learning in class as well as the costs per student. From the first-order condition it follows directly that the optimum class size is given by

$$m^* = -\frac{1 - \alpha}{\ln \pi} \quad \forall \pi \neq 1. \tag{9}$$

We are now able to state two results concerning optimum class size:

Result 1 *The optimal class size is larger the better behaved students are.*

Proof.

$$\frac{dm^*}{d\pi} = \frac{1-\alpha}{\pi \left(\ln \pi\right)^2} > 0.$$

∎

Result 1 is intuitive: The optimal class size is larger for groups of students who are well behaved, because these students are less likely to disrupt in class and therefore benefit less from a class size reduction than more disruptive students. West and Wössmann (2003) find that there is indeed sorting of less skilled an possibly more disruptive students into smaller classes between schools as well as within schools.[4]

Result 2 *Overall disruption is constant in π: $\frac{dp}{d\pi} = 0$.*

Proof. Disruption in class is by definition and by equation 9 determined through the individual disruption probability π and the elasticity of human capital with respect to student effort α:

$$p = \pi^{-\frac{1-\alpha}{\ln \pi}}.$$

From this, the result follows directly. ∎

Result 2 also allows for an intuitive interpretation: The school raises class size until the optimal *noise level* in the class room is reached. This is the same for every class because in any case, potential schooling quality[5] is multiplied by the factor that determines actual studying time p. The costs and benefit of changing the class size m do not depend on how productivity is generated (i.e. on the students' effort and the school's spent resources).

5.2.3 School Resources Optimization

The school not only decides on class sizes, but also about how much resources to employ. As already discussed above, schooling resources cover physical inputs to education as well as teacher training and teacher pay. A school's costs C^T and benefits B^T are again given by

$$C^T = c^T r \tag{8a}$$

---

[4]They doubt, however, that the kind of model presented here has any empirical relevance, since it treats schools as unitary actors maximizing educational productivity, thus largely ignoring the interests of individuals working within schools.

[5]By *potential schooling quality* we refer to the schooling quality that would prevail if there were no disturbance in class, as defined above.

and

$$B^T = m\sigma P = mp\sigma \left(\frac{\alpha}{c^S} (1 - \tau) \, pr^\beta \right)^{\frac{\alpha}{1-\alpha}} r^\beta, \qquad (7a)$$

respectively. Concentrating all resource terms, we get

$$B^T = mp\sigma \left(\frac{\alpha}{c^S} (1 - \tau) \, p \right)^{\frac{\alpha}{1-\alpha}} r^{\frac{\beta}{1-\alpha}}.$$

The school finds optimum resource spending by maximizing its net benefit over resources,

$$r^* = \arg\max_{r \in \mathbb{R}_+} \left\{ m\sigma \left(\frac{\alpha}{c^S} (1 - \tau) \right)^{\frac{\alpha}{1-\alpha}} p^{\frac{1}{1-\alpha}} r^{\frac{\beta}{1-\alpha}} - c^T r \right\}.$$

The first-order condition is

$$\frac{d \left(B^T - C^T \right)}{dr} = \frac{\beta}{1 - \alpha} m\sigma \left(\frac{\alpha}{c^S} (1 - \tau) \right)^{\frac{\alpha}{1-\alpha}} p^{\frac{1}{1-\alpha}} r^{\frac{\alpha+\beta-1}{1-\alpha}} - c^T \overset{!}{=} 0.$$

By the concavity of the objective function we can solve for r to get optimum resources as

$$r^* = \left(\frac{1}{c^T} \frac{\beta}{1 - \alpha} m\sigma \left(\frac{\alpha}{c^S} (1 - \tau) \right)^{\frac{\alpha}{1-\alpha}} p^{\frac{1}{1-\alpha}} \right)^{\frac{1-\alpha}{1-\alpha-\beta}}.$$

5.2.4 Illustration of the Optimum School Quality Choice

Figure 2 illustrates the school quality optimization problem with fixed student effort. Class size is displayed on the horizontal axis, school resource spending on the vertical axis. The solid lines are iso-benefit loci per student in the *m-r-*
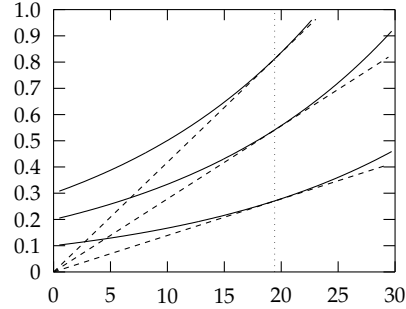
Figure 2

School optimization problem (illustration).

space: In order to keep student achievement constant, an increase in resource

11

spending can be compensated by larger class size. From (7) we directly get $d\frac{B^T}{m} = 0 \Leftrightarrow \frac{dr}{dm} = -\frac{r\ln\pi}{\beta}$. The straight dashed lines through the origin represent iso-cost loci. They result from the per-student cost function $\frac{C^T}{m} = \frac{1}{m}ec^T$, which yields $dC^T = 0 \Leftrightarrow \frac{dr}{dm} = \frac{r}{m}$. Note that the specification of the education production function leads to a similar result as with quasilinearity in the production function: The optimum value of m is independent of any parameter other than student characteristic π and output elasticity with respect to student effort, α.

5.3 Equilibrium

In the decentralized equilibrium, when schools and students optimize their contribution to education on their own behalf, resource spending, student effort, and class size are functions of the cost parameters c^S and c^T, disruption π, the production parameters α and β, the tax rate τ, and the distribution parameter σ.

Result 3 *The equilibrium allocation compares to the centralized solution as follows:*

$$m^* = m^\#,$$

$$r^* = r^\# \Leftrightarrow (1-\tau)^\alpha \left(\frac{\sigma}{1-\alpha}\right)^{(1-\alpha)} = \rho,$$

$$e^* = e^\# \Leftrightarrow (1-\tau)^\alpha \left(\frac{\sigma}{1-\alpha}\right)^{(1-\alpha)} = \rho.$$

Proof. The proof follows directly from a comparison of (10)–(12) in the equilibrium allocation with (3)–(5) from the centralized solution:

$$m^* = -\frac{1-\alpha}{\ln\pi}, \tag{10}$$

$$r^* = \left[\left(\frac{\alpha(1-\tau)}{c^S}\right)^\alpha \left(\frac{\beta m^*\sigma}{c^T(1-\alpha)}\right)^{1-\alpha} \pi^{m^*}\right]^{\frac{1}{1-\alpha-\beta}}, \tag{11}$$

$$e^* = \left[\left(\frac{\alpha(1-\tau)}{c^S}\right)^{1-\beta} \left(\frac{\beta m^*\sigma}{c^T(1-\alpha)}\right)^\beta \pi^{m^*}\right]^{\frac{1}{1-\alpha-\beta}}. \tag{12}$$

∎

Result 3 can be interpreted as follows: The decentralized optimum class size is efficient, whatsoever. However, the optimum choices of resource spending and effort deviate from their first-best values due to three reasons: First, a positive tax rate τ reduces effort; second, the value of of education, which is perceived by schools, σ, may deviate from the actual social value of education ρ. With $\tau = 0$ and $\sigma = \rho$, there remains a third distortion, which results from the students not being able to commit to their effort.

The elasticities of the response of student effort e, resource spending r, and educational achievement P with respect to exogenous variables are summarized in table 1.

Table 1
Comparative static effects of π, c^S, c^T, σ, and τ on optimum e, r, and P.

Effort e	Resources r	Achievement P
$\varepsilon_{e,\pi} = -\frac{1}{\ln \pi}\frac{\beta}{1-\alpha-\beta} > 0$	$\varepsilon_{r,\pi} = -\frac{1}{\ln \pi}\frac{1-\alpha}{1-\alpha-\beta} > 0$	$\varepsilon_{P,\pi} = -\frac{1}{\ln \pi}\frac{\beta}{1-\alpha-\beta} > 0$
$\varepsilon_{e,c^S} = -\frac{1-\beta}{1-\alpha-\beta} < 0$	$\varepsilon_{r,c^S} = -\frac{\alpha}{1-\alpha-\beta} < 0$	$\varepsilon_{P,c^S} = -\frac{\alpha}{1-\alpha-\beta} < 0$
$\varepsilon_{e,c^T} = -\frac{\beta}{1-\alpha-\beta} < 0$	$\varepsilon_{r,c^T} = -\frac{1-\alpha}{1-\alpha-\beta} < 0$	$\varepsilon_{P,c^T} = -\frac{\beta}{1-\alpha-\beta} < 0$
$\varepsilon_{e,\sigma} = \frac{\beta}{1-\alpha-\beta} > 0$	$\varepsilon_{r,\sigma} = \frac{1-\alpha}{1-\alpha-\beta} > 0$	$\varepsilon_{P,\sigma} = \frac{\beta}{1-\alpha-\beta} > 0$
$\varepsilon_{e,1-\tau} = \frac{1-\beta}{1-\alpha-\beta} > 0$	$\varepsilon_{r,1-\tau} = \frac{\alpha}{1-\alpha-\beta} > 0$	$\varepsilon_{P,1-\tau} = \frac{\alpha}{1-\alpha-\beta} > 0$

The more schools are aware of the total benefit of education for society, the higher resource spending, student effort, and educational outcome. On the other hand, the more highly students' private benefits from education are taxed, the smaller their incentives to study hard and – by the complementarity of educational resources and student effort – employed resources and achievement. The positive effect of student behavior on effort, resource spending and achievement cannot be attributed to better overall classroom behavior, since this is constant over student types as shown in result 2. However, since resource spending is a public good within a classroom, resources are more effectively spent in bigger classes. By the complementarity of resources and effort, also effort increases in bigger classes, which results in a better educational achievement.

From the formal considerations presented above, we expect the quality of schooling to be highly correlated with educational spending on schooling quality and not much so with class size because spending directly affects achievement while the class size effect is just offset by differences in the students' disruption: Overall disruption is constant, notwithstanding the students' individual behavior.

While the second expectation is well met by empirical literature[6], the effect of spending on schooling quality, such as teacher quality and physical infrastructure, is ambiguous: On the one hand, Hanushek, Kain, and Rivkin (1998) find that teacher quality alone accounts for at least 7.5 percent of the total variation in student achievement; Angrist and Lavy (2001), too, conclude that teacher training leads to an improvement in test scores. On the other hand, Hanushek (2002a) in his review finds that increased school spending alone contributes little to better student achievement, so do improvements in classroom technology, such as computer-aided instruction, as pointed out by Angrist and Lavy (2002). Hence, our model possibly overstates the effect of resource spending r on educational achievement P.

Hanushek and Wössmann (2005) find strong evidence that student tracking and separation by type increases educational inequality. This is concordant with our finding that employed resources are higher in larger classes with better behaved students. Since overall disruption in class is constant in the student characteristic, educational attainment is the higher the better students behave.

6 Student Heterogeneity

6.1 Observable Student Type

So far, we have allowed for differences in student types, but we have always considered only one type at a time. We proceed by extending the analysis to the question whether it is socially desirable to segregate between students, i.e. whether classes should consist of only one category of student types or whether classes should be mixed. This decision is taken by the government and precedes the optimizations by schools and students.

There are many different approaches to argue for and against segregation by student type. It may be the case that students of a low ability type profit from high ability types if classes are integrated by being transformed into high ability types by being around them. This effect of the peer group on transforming behavior is pointed out and tested in Katz, Kling, and Liebman (2001). Their finding is that children who move from low income areas to higher income areas experienced

---

[6]See e.g. Hoxby (2000a) who estimates that class size does not have a statistically significant effect on student achievement.

lower behavior problems at school. Another argument in favor of integrated classes is that integrated classes reduce educational inequality. If students of different types L, H have different probabilities of disrupting in class (low and high, respectively), $\pi_L > \pi_H$, L-type students receive more in a mixed class than in an all L class. Thus, H-type students benefit from segregation and L-type students may lose by it. However, this argument is controversial: Even L-type students may profit from being in segregated classes if questions asked by H-type students can be considered disruption as far as L-type students are concerned if they can not be understood by L-type students.[7] In general, one has to be cautious in interpreting *disruption* since students who interrupt their teachers more might be more engaged, also benefitting their classmates by raising the value of the educated experience. As Lazear (2001) points out: "The optimal amount of student participation is not zero." Using the narrow meaning of disruption as deterring the teacher from teaching, result 4 holds.

Result 4 *With observable stdent type π, total expected output and welfare are maximized if students are segregated by type.*

Proof. We stick to the two-type case; it is easily extended to the case of a continuous distribution of student types. Suppose that the economy consists of χ H-type students and $(1 - \chi)$ L-type students. First, assume that all classes are of the same size m and that student effort e and resources spent r are the same in each class. Consequentially, also h is uniform across classes. Average achievement per student with segregated classes is

$$\chi \pi_H^m h + (1 - \chi) \pi_L^m h. \tag{13}$$

Since all students in a class share the same behavior, overall probability of disruption in a class with type π students is given by π^m. Achievement in integrated classes is

$$\pi_H^{\chi m} \pi_L^{(1-\chi)m} h. \tag{14}$$

To show that segregation dominates integration, it suffices to show that the difference Δ between 13 and 14 is positive, since the costs are equal by assumtion:

$$\Delta = \chi \pi_H^m h + (1 - \chi) \pi_L^m h - \pi_H^{\chi m} \pi_L^{(1-\chi)m} h. \tag{15}$$

---

[7] A similar argument can be made concerning the segregation between sexes. Betts and Shkolnik (1999) find that the time spent on discipline falls and that on instruction rises as the class becomes more female. By the above logic, boys would want to be in all-girl schools, while girls do not want them there. Also Hoxby (2000b) finds that the gender composition in a class alters classroom conduct.

When $\pi_L = \pi_H$, $\Delta = 0$ because all students are equal. Differentiate 15 with respect to π_H to get

$$\frac{\partial \Delta}{\partial \pi_H} = \chi m \pi_H^{m-1} \left(1 - \frac{\pi_L^{(1-\chi)m}}{\pi_H^{(1-\chi)m}} \right) > 0$$

for $\pi_H > \pi_L$. If we allow class size to differ between segregated classes, we can define m to be the class size that is optimal when student types are mixed. Segregation allows schools to adjust class size for every student type individually, while segregation dominates even when all class sizes are constrained to be at the mixed optimum. Therefore, segregation must maximize expected output if segregated classes can be of different sizes. The same argument is true for the optimal choices of e and r.

Result 2 shows, that overall disruption is independent of student types. This result also holds for mixed classes. Thus, learning effort e and resource spending r depend on class size only. Given any fixed class size, also welfare is maximized with segregation. If, again, we allow the class size to differ, this can only improve welfare. ∎

The more diverse the students' behavior in class, the more different classes should exist. As has been shown in result 1, optimum class size is the larger the less disruptive students are. If students of different types are mixed, the optimum class size can still be met by looking for the optimum overall noise level in class, which would be attained more efficiently though if all students in class are of the same type.[8] Result 4 holds for any peer group effect entering the educational production function multiplicatively: Also if student characteristics add to learning in class, which would amount to assuming that $\pi > 1$, segregated classes dominate over integrated classes by the argument given proposition 4.[9] In this case, of course, the temporal interpretation of the peer group effect cannot be maintained.

In the preceding section dealing with only one student type, we have implicitly assumed that classes consist of one student type only, i.e. that student types are segregated. If the government is maximizing total welfare, it will induce student

---

[8]Hanushek and Wössmann (2005), however, find a tendency for student tracking to reduce mean performance.

[9]Of course, π could then no longer be interpreted as a student's probability of disruption. π is the just a parameter describing individual contributions to learning in class: $\pi < 1$ denotes a negative impact while $\pi > 1$ adds to classroom performance.

segregation indeed, such that all previous results remain valid. One problem remains: Our model is not able to represent the empirical finding that educational resource spending and student achievement are only loosely related, if at all. This shortcoming of the model will be cured partly once we abandon the assumption that student types are observable.

6.2 Unobservable Student Type

With observable student types and therefore free sorting, it is optimal to increase class size with better student behavior. If educational resources are a non-rival within a classroom, the government is willing to spend more resources on bigger classes, thus privileging calmer students. It seems obvious that if student types are not observable, disruptive students want to mimic well behaved students in order to profit from their favored treatment. If effort costs are constant over all student types, $\frac{dP}{d\pi} > 0$ which is implied by $\varepsilon_{P,\pi} = -\frac{1}{\ln \pi} \frac{\beta}{1-\alpha-\beta} > 0$ in table 1. In order to avoid this mimicking behavior, schools can adjust their spending or class size to induce self-selection among students. We assume that students choose the accurate class if their expected net benefit from studying is equal in several class types.

Result 5 *If student types are not observable and differ only in their probability of disrupting in class,*
(a) the level of educational resources spending must be constant over all classes;
(b) even with the constraint of equal resource spending, segregation still dominates integration.

Proof. The student selection constraints write as[10]

$$B_i^S(i) - c^S e_i^*(i) \geq B_i^S(j) - c^S e_i^*(j) \quad \forall (i,j) \in \{L, H\}. \tag{16}$$

In a symmetric equiklibrium and since students are homogeneous with respect to effort costs, $e_i^*(j) = e_i^*(i)$, we can concentrate on human capital production in several classes. In order to meet both inequality constraints, the school system has to assure that

$$B_i^S(i) = B_i^S(j) \quad \forall (i,j) \in \{L, H\}, \tag{17}$$

---

[10] $B_i^S(j)$ denotes the benefit of student i choosing class j.

which, by result 2, is only possible if $r_L = r_H$. From result 4 we know that segregation maximizes educational at any rate. If classes were supposed to be mixed, they would be symmetric in expected terms, such that students were a priori indifferent between any two classes such that we can safely assume that they sort anyway. Hence, segregation is both feasible and superior to integration. The assertion that there exist no asymmetric equilibria is proved by contradiction: Suppose that $r_L \neq r_H$. Then, by 5.1, also $e_L \neq e_H$ and

$$B_L^S(L) - c^S e_L^*(L) \gtrless B_L^S(H) - c^S e_L^*(H) \Leftrightarrow B_H^S(L) - c^S e_H^*(L) \gtrless B_H^S(H) - c^S e_H^*(H).$$

Hence, self-selection is not feasible with $r_L \neq r_H$. ∎

By inducing self-selection through equalization of resource spending, productive efficiency is lost, since the marginal benefit of employed resources is different among different classes. However, since integrated classes are also equal in expected terms, segregated classes are still superior from a welfare point of view. Hence, increased total spending in education has only small effects on student achievement since it must be partly allocated to suboptimal use in order to sustain student selection. These theoretical limitations to efficient resource allocation in education partly explains the findings of Hanushek (1994, 2001) who finds that merely increasing spending on education does not necessarily increase student achievement.

7 Conclusion

In a class, a student who is disruptive or takes up teacher time in ways that are not useful to other students, and affects the learning of others. Thus, class size may have important effects on educational achievement. However, the empirical literature, suggests that class size effects play little or no role. The model in this paper implies that better students are optimally placed in bigger classes and that optimal class size crucially depends on the schools' and students' reaction to higher disruption in class. In the case of symmetric information about students' type, the socially optimal class choice can easily be implemented. However, according to this model, there must be a high correlation between educational resources and student achievement which is not present in the empirical literature. This shortcoming of the model with symmetric information is partly overcome

by relaxing the assumption that student types are observable by schools. In order to induce self-selection among students, educational resource spending is adjusted. This results in a similar educational achievement over all student types – a result which is well matched by empirical observations.

References

ACEMOGLU, D., AND J. ANGRIST (2001): "How Large Are Human Capital Externalities? Evidence from Compulsory Schooling Laws," in NBER *Macroeconomics Annual 2000*, ed. by B. Bernanke, and K. Rogoff. MIT Press, Cambridge.

ANGRIST, J. D., AND J. GURYAN (2003): "Does Teacher Testing Raise Teacher Quality? Evidence from State Certification Requirements," NBER *Working Paper 9545*.

ANGRIST, J. D., AND V. LAVY (1999): "Using Maimonides' Rule to Estimate the Effect of Class Size on Student Achievement," *Quarterly Journal of Economics*, 114(2), 533–575.

——— (2001): "Does Teacher Training Affect Pupil Learning? Evidence from Matched Comparisons in Jerusalem Public Schools," *Journal of Labor Economics*, 19(2), 344–369.

——— (2002): "New Evidence on Classroom Computers and Pupil Learning," *Economic Journal*, 112(482), 735–765.

BETTS, J. R., AND J. L. SHKOLNIK (1999): "The Behavioral Effects of Variations in Class Size: The Case of Math Teachers," *Educational Evaluation and Policy Analysis*, 20, 193–213.

CARD, D., AND A. KRUEGER (1992): "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy*, 100(1), 1–40.

DEE, T. S. (2003): "Are There Civic Returns to Education?," NBER *Working Paper 9588*.

FERGUSON, R. F., AND H. F. LADD (1996): "How and Why Money Matters: An Analysis of Alabama Schools," in *Holding Schools Accountable*, ed. by H. F. Ladd, Washington, D.C. Brookings Press.

HANUSHEK, E. A. (1994): *Making Schools Work: Improving Performance and Controlling Costs*. Brookings Institution Press, Washington, D.C.

——— (1996): "Outcomes, Cost, and Incentives in Schools," in *Improving America's Schools: The Role of Incentives*, ed. by E. A. Hanushek, and D. W. Jorgenson, pp. 29–52, Washington, D.C. National Academy Press.

——— (2001): "Efficiency and Equity in Education," NBER *Reporter*, pp. 15–19.

——— (2002a): "The Importance of School Quality," *Mimeo*.

——— (2002b): "Publicly Provided Education," in *Handbook of Public Economics*, ed. by A. J. Auerbach, and M. Feldstein, vol. 4, chap. 30, pp. 2045–2141. North-Holland, Amsterdam.

HANUSHEK, E. A., J. F. KAIN, AND S. G. RIVKIN (1998): "Teachers, Schools, and Academic Achievement," NBER *Working Paper 6691*.

HANUSHEK, E. A., AND L. WÖSSMANN (2005): "Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence Across Countries," NBER *Working Paper 11124*.

HOXBY, C. M. (2000a): "The Effects of Class Size on Student Achievement: New Evidence from Population Variation," *Quarterly Journal of Economics*, 115(4), 1239–85.

——— (2000b): "Peer Effects in the Classroom: Learning from Gender and Race Variation," NBER *Working Paper 7867*.

KATZ, L. F., J. R. KLING, AND J. B. LIEBMAN (2001): "Moving to Opportunity in Boston: Early Results of a Randomized Housing Mobility Experiment," *Quarterly Journal of Economics*, 116(2), 681–716.

KRUEGER, A. B. (1999): "Experimental Estimates of Educational Production Functions," *Quarterly Journal of Economics*, 114(2), 497–532.

LAZEAR, E. P. (2001): "Educational Production," *Quarterly Journal of Economics*, 66(3), 777–803.

LOCHNER, L., AND E. MORETTI (2001): "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports," NBER *Working Paper 8605*.

MISHEL, L., AND R. ROTHSTEIN (2002): *The Class Size Debate*. Economic Policy Institute, Washington, D.C.

MOSTELLER, F. (1995): "The Tennessee Study of Class Size in the Early School Grades," *The Future of Children*, 5(2), 113–127.

RICE, J. K. (2003): *Teacher Quality: Understanding the Effectiveness of Teacher Attributes*. The Economic Policy Institute, Washington, D.C.

RIVKIN, S. G., E. A. HANUSHEK, AND J. F. KAIN (2005): "Teachers, Schools and Academic Achievement," *Econometrica*, 73(2), 417–458.

ROCKOFF, J. E. (2004): "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data," *American Economic Review Papers and Proceedings*, 94(2), 247–252.

SANDERS, W. L., AND J. C. RIVERS (1996): *Cumulative and Residual Effects of Teachers on Future Student Academic Achievement*. University of Tennessee Value-Added Research and Assessment Center, Knoxville, TN.

WEST, M. R., AND L. WÖSSMANN (2003): "Which School Systems Sort Weaker Students Into Small Classes? International Evidence," *Kiel Working Paper 1145*.

WÖSSMANN, L. (2002): *Schooling and the Quality of Human Capital*. Springer, Berlin.