Sorting and Peer Effects

Behrang Kamali-Shahdadi*

March 24, 2016

Abstract

The effect of sorting students based on their academic performances depends not only on direct peer effects but also on indirect peer effects through teachers' efforts. Standard assumptions in the literature are insufficient to determine the effect of sorting on the performances of students and so are silent on the effect of policies such as tracking, implementing school choice, and voucher programs. We show that the effect of such policies depends on the curvature of teachers' marginal utility of effort. We characterize conditions under which sorting increases (decreases) the total effort of teachers and the average performance of students.

Keywords: Indirect Peer Effects, Sorting, Matching, Comparative Statics

JEL classification: I21, D47, C60

1 Introduction

We introduce a model to analyze the effect of student sorting on the total effort of teachers and the average or total performance of students. Our model allows for both direct and indirect peer effects. First, consider the case of homogeneous teachers who choose effort after observing their classes' composition. Each teacher chooses an effort based on the distribution of students' abilities in their classes, i.e., the teacher's choice may depend on the whole distribution, not just the mean of students' abilities. The effect of sorting

*Correspondence: Behrang Kamali Shahdadi: Department of Economics, University of Pennsylvania, (email: behrang@sas.upenn.edu). I am deeply indebted to Rakesh Vohra for his guidance and continuous support. I am grateful to SangMok Lee, Eduardo Azevedo, and Andrew Postlewaite for their valuable advice. I also thank George Mailath, Petra Todd, Hanming Fang, Jere Behrman, Alvaro Sandroni, Steven Matthews, Mallesh Pai, Aislinn Bohren, and the seminar participants at the University of Pennsylvania for their helpful comments and suggestions.

on the teachers' total effort choice is ambiguous. Because the teachers' total effort may increase or decrease as a result of student sorting, the effect of sorting on the average or total performance of students is ambiguous, too. We characterize conditions on the utility function of teachers under which the total effort of teachers strictly increases or strictly decreases by sorting. If the teachers' marginal utility of effort is supermodular (submodular) and convex (concave) in effort, then the total effort of teachers increases (decreases) as a result of sorting of students. Subsequently, in the absence of direct peer effects, if performance is convex (concave) in a teacher's effort, the total performance of an education system increases (decreases) as a result of sorting.[1] We show that this result persists even when we allow for heterogeneous teachers.

Sorting affects students' outcomes because of peer effects. There are two types of peer effects: direct and indirect. Direct peer effects are the result of student-to-student spillovers (see Sacerdote (2000), Sacerdote (2011), and Epple and Romano (2011) for a review of the literature). Indirect peer effects happen through a teacher's effort choice (see Duflo et al. (2011) and Todd and Wolpin (2012)). Duflo et al. (2011) report that both direct and indirect peer effects exist in the data and that the data cannot be explained using only one kind of peer effects.

Two types of sorting are present in an education system: within-school sorting and between-school sorting. Within-school sorting, or tracking, is an explicit policy that sorts students into different classes based on their abilities. Sorting between schools happens in different ways, such as: (i) Sorting between public and private schools.[2] (ii) Sorting as a result of voucher programs.[3] Chakrabarti (2009) states that "There is strong and robust evidence in favor of stratification by ability" as a result of Milwaukee Voucher Program. Hsieh and Urquiola (2006) report that they "find evidence that the voucher

[1]Because most of the results for one set of conditions are parallel in wording to results under the other set of conditions, instead of stating results under each set of conditions separately, we state both results in one statement using parentheses.

[2]see Epple and Romano (1998) and Epple et al. (2002) for more details.

[3]See Barrow and Rouse (2008).

program led to increased sorting, as the 'best' public school students left for the private sector." (iii) Standardized admissions tests. MacLeod and Urquiola (2012) state that "the introduction of standardized admissions tests will lead to stratification by ability." (iv) Public information regarding schools' qualities. Hastings and Weinstein (2007) find that "providing parents with direct information on school test scores resulted in significantly more parents choosing higher-scoring schools for their children.". (v) Different school choice policies.[4] Levin (1998) reports that "evidence is consistent that educational choice leads to greater socioeconomic (SES) and racial segregation of students." We incorporate both types of sorting in our model.

Our paper also contributes to the literature on comparative statics. That literature is focused on the monotonicity of the argmax of a maximization problem;[5] however, to analyze the effect of sorting on the total performance of students, we need to understand the effect of sorting on the total effort of teachers, which depends on supermodularity and submodularity of the argmax of teachers' utility maximization problem. We find conditions under which the argmax of a maximization problem is strictly supermodular or strictly submodular. More concretely, if the marginal utility of a teacher is supermodular (submodular) and convex (concave) in effort, then the argmax of the teacher's maximization problem — the optimal effort of the teacher — is supermodular (submodular); therefore, the total performance of all teachers increases (decreases) as a result of sorting students. In the absence of direct peer effects, the only channel through which sorting changes the total/average performance of students is through indirect peer effects. Hence, if teachers are putting in more effort altogether and performance is a convex function in the teacher's effort, then the total/average students' performance increases. On the other hand, if teachers are putting in less effort altogether and performance is a concave function in the teacher's effort, then the total/average students' performance decreases. We state the results for an education

[4]See Avery and Pathak (2015) for the effect of implementing school choice instead of neighborhood assignment rule on student sorting.

[5]See Topkis (1998), Milgrom and Shannon (1994), and Edlin and Shannon (1998).

system; one can use the same tools to analyze any one-to-many matching with endogenous effort choice and evaluate the effect of different matchings. Moreover, conditions for strict supermodularity (submodularity) of the argmax are derived for a general maximization problem; hence, these results can be used in any maximization problem that has the same structure.

Sorting increases inequality in students' performances; however, if the marginal utility of teachers is supermodular and convex in effort, sorting increases the total/average performance of students. In this situation, sorting is desirable under the Utilitarian welfare function. On the other hand, if the marginal utility of teachers is submodular and concave in effort, sorting decreases the total/average performance of students and increases inequality in students' performances. In this situation, sorting reduces both the Utilitarian welfare function and the Rawlsian welfare function. Affirmative action policies have the opposite effect of sorting, i.e., these policies reduce sorting; therefore, in this situation, affirmative action policies increase both the Utilitarian welfare function and the Rawlsian welfare function.

The effect of sorting on the total/average performance through the channel of indirect peer effects is robust, whether direct peer effects exist, even when teachers are heterogeneous in quality and utility function. Furthermore, if teachers are heterogeneous and classes are sorted, the standard results regarding the benefit of positive assortative matching (PAM) versus negative assortative matching (NAM) may not hold.[6] We show that even if the performance function of classes and the utility of teachers are supermodular, the total performance of students can be higher under negative assortative matching compared with positive assortative matching. Moreover, inequality is lower under negative assortative matching compared to positive assortative matching. To put it differently, the value of the Utilitarian welfare function and the Rawlsian welfare function are higher under NAM compared with PAM.

[6]See Tincani (2014).

The effectiveness of monetary incentives on teachers' effort choices is debated in the literature. There is some evidence of a positive effect of monetary incentives in developing countries (Lavy (2002)). However, in developed countries such as the U.S., the evidence suggests that monetary incentives have an insignificant effect on teachers' effort choices (Fryer (2013)). We consider the U.S. as the main application, i.e., monetary incentives don't affect teachers' effort choices. Under a pay-per-performance system in which a teacher's wage depends on the performance of his/her students, sorting has an impact on the budget of the education system and results in inequality in teachers' salaries. Sorting increases inequality in teachers' salaries when teachers are homogeneous and wage increases are based on students' performances. Sorting increases (decreases) the total payment to teachers if the total students' performance increases (decreases) and payment to teachers is an increasing and a convex (concave) function of class performance. Our model can incorporate monetary incentives, too. We consider a general utility function that can incorporate monetary incentives; hence, we can analyze the effect of sorting under different monetary incentive systems.

In section 2, we set up the model and develop the required mathematical tool to handle the student sorting problem. In section 3, we analyze the effect of student sorting on the total effort of teachers by finding conditions on the utility function of teachers such that the argmax of their utility maximization problem is supermodular (submodular). Subsequently in section 4, we show how sorting affects students' outcomes under three settings: (1) homogeneous teacher with indirect peer effects, (2) homogeneous teacher with direct indirect peer effects, and (3) heterogeneous teacher with direct and indirect peer effects.

2 Model

Let T be a finite set of homogeneous teachers and I a finite set of students, where $|I| = n|T|$. A student $i \in I$ has a type $\theta_i \in \mathcal{R}_+$. The type can represent a student's ability, the prior year's test score, parents' education/income, or any other characteristic that affects the students' performance. We interpret type as ability.

A matching is an assignment of students to teachers, denoted by $\mu : T \to I$, such that $|\mu(t)| = n$, where n is the size of the class. Each student is assigned to only one teacher, i.e., $\mu^{-1}(i)$ is a function. We denote a class by the profile of types $\theta \in \mathcal{R}^n_+$ assigned to it. We denote the class assigned to teacher t by $\theta^t = (\theta_i)_{i \in \mu(t)}$.

There is a measure of performance for each student $i \in \mu(t)$, denoted by $p(e_t, \theta_i, \theta^t)$. We interpret a student's performance as his/her end-of-year test score. There is an aggregate measure of performance for each class t, denoted by $p(e_t, \theta^t)$. We consider the aggregate measure of performance for each class as the average performances of students in that class.

Teacher $t \in T$ chooses an effort $e_t \in [0, 1]$. Each teacher gets a payment — wage plus bonus — based on the aggregate performance of the class, denoted by $w(p(e_t, \theta^t))$.[7] Each teacher t has a utility function $f(e_t, \theta^t)$. The utility function is the same for all teachers.[8] $f(e_t, \theta^t)$ represents the induced utility of a teacher; the utility of a teacher may depend on the performance of his/her class, his/her wage, and the amount of effort he/she exerts.[9] We assume that the wage structure is fixed; hence, if the utility of a teacher depends on the wage, then $f(e, \theta)$ is the induced utility function for a fixed wage structure. The marginal utility of effort at $e = 1$ is strictly negative for any class θ. The marginal utility of effort at $e = 0$ is strictly positive for any class θ.

We assume that a teacher's utility function and performance of his/her class are symmetric functions in students' type, i.e., any permutation of a class θ generates the same

[7] Wage can be a constant function.

[8] We relax this assumption in section 4.3.

[9] For example, the utility function of a teacher can be the non pecuniary utility that he/she gets from his/her class's performance minus the cost of effort, i.e., $f(e, \theta) = u(p(e, \theta)) - c(e)$. The cost of effort may depend on effort and the class composition, i.e., $f(e, \theta) = u(p(e, \theta)) - c(e, \theta)$. The utility function of a teacher may include the utility he/she gets from wages or bonuses plus the non pecuniary utility that he/she gets from his/her class's performance minus the cost of effort, i.e., $f(e, \theta) = u(p(e, \theta)) + v(w(p(e, \theta))) - c(e)$.

performance and utility for a teacher: if θ' is a permutation of θ, then:

$$f(e,\theta) = f(e,\theta'), p(e,\theta_i,\theta) = p(e,\theta_i,\theta'), \text{ and } p(e,\theta) = p(e,\theta').$$

We assume a teacher's utility function and the performance of his/her class are three times continuously differentiable, i.e., $f(e,\theta), p(e,\theta) \in \mathcal{C}^3$.

First, we define sorting of two classes, and then we show the mathematical relationship between sorting and the coordinate-wise maximum and minimum of two classes. We order all the students in the two classes by their types, and then we put the top half of students in one class and bottom half in the other class. This process is called *sorting.*[10]

One-step sorting of two classes is defined as the coordinate-wise maximum and minimum of two classes:

$$\forall \theta, \theta^\dagger : \theta' = \theta \vee \theta^\dagger, \theta'' = \theta \wedge \theta^\dagger,$$

where for any two vectors $\theta = (\theta_1, ..., \theta_n), \theta^\dagger = (\theta_1^\dagger, ..., \theta_n^\dagger)$:

$$\theta \vee \theta^\dagger = (max(\theta_1, \theta_1^\dagger), ..., max(\theta_n, \theta_n^\dagger)), \theta \wedge \theta^\dagger = (min(\theta_1, \theta_1^\dagger), ..., min(\theta_n, \theta_n^\dagger)).$$

Lemma 1 *There exists a reordering of two classes such that sorting is achieved by one-step sorting.*

Proof: In the appendix.

For any class θ, define $\tilde{\theta}$ as the reordering of the vector θ in descending order, i.e., the permutation of the class in which the first element is the greatest type in the class, the second element is the second greatest type in the class, and so on: $\tilde{\theta} = (\tilde{\theta}_1, \ldots, \tilde{\theta}_n)$ such that $\tilde{\theta}_i \geq \tilde{\theta}_{i+1} \quad \forall i = 1, \ldots, n-1$. Given this reordering of the two classes θ^1, θ^2, define a

[10]This process is called sorting only if the two new classes have different student compositions than they did before sorting.

partial ordering of two classes $\succeq^*$ as the vector ordering in $\mathcal{R}^n$:

$$\theta^1 \succeq^* \theta^2 \Leftrightarrow \tilde{\theta}_i^1 \geq \tilde{\theta}_i^2 \quad \forall i = 1, \ldots, n, \tag{1}$$

we call θ^1 *a better class* than θ^2. Note that after sorting of two classes $\theta, \theta^\dagger$, the sorted classes θ', θ'', have the following property:

$$\theta \succeq^* \theta', \theta^\dagger \succeq^* \theta', \theta'' \succeq^* \theta, \theta'' \succeq^* \theta^\dagger.$$

We call θ' *the lower track* and θ'' *the higher track*. Observe that the higher track is a better class than the lower track.

These two classes can be in one school or in two different schools. The former represents within-school sorting; The latter represents between-school sorting. In between-school sorting, every class in one school is a better class than any class in the other school (based on the partial order $\succeq^*$ defined in (1)). However, we may be unable to order two classes in the same school (based on the partial order $\succeq^*$) after between-school sorting. All the following results hold for both between-school sorting and within-school sorting.

A function $h : \mathcal{R}^n \to \mathcal{R}$ is *supermodular* if it is pairwise supermodular in any of its two arguments, i.e., the cross-partial derivatives in any of its two arguments are positive.[11] For example, $p(e, \theta)$ is supermodular if $p_{e\theta_i}(e, \theta) \geq 0, p_{\theta_i\theta_j}(e, \theta) \geq 0 \quad \forall i, j \in I, \forall e, \forall \theta$. If the performance function is supermodular then two types of complementarities exist: complementarity between effort of the teacher and a student's ability and complementarity between students' abilities. If cross-partials are strictly positive, then the function is strictly supermodular. A function is *modular* if the cross-partial derivatives in any of its two arguments are zero. For example, $p(e, \theta)$ is modular if $p_{e\theta_i}(e, \theta) = 0, p_{\theta_i\theta_j}(e, \theta) = 0, \quad \forall e \in [0, 1] \quad \forall i, j, \theta_i \in \mathcal{R}_+, \theta_j \in \mathcal{R}_+$. A function is *submodular* if it is pairwise submodular in any of its two arguments, i.e., the cross-partial derivatives in any of its two arguments

[11]See Topkis (1998).

are negative. For example, $p(e,\theta)$ is submodular if $p_{e\theta_i}(e,\theta) \leq 0, p_{\theta_i\theta_j}(e,\theta) \leq 0 \quad \forall i,j \in I, \forall e, \forall\theta$. If cross-partials are strictly negative, then the function is strictly submodular.

3 Characterizing Teachers' Optimal Efforts

In this section, we characterize the conditions under which the argmax of a maximization problem is strictly supermodular or strictly submodular.[12] Furthermore, these conditions determine the effect of sorting on the total effort of teachers. Each teacher maximizes a utility function; the argmax of a teacher's maximization problem is his/her optimal effort, which is unique under the following assumption (Assumption 1). We show that if the marginal utility of effort is supermodular and convex in effort, then the argmax is supermodular. Subsequently, we show that sorting increases the total effort of teachers in this case. Similarly, we show that if the marginal utility of effort is submodular and concave in effort, then the argmax is submodular. We conclude that sorting decreases the total effort of teachers in this case.

Assumption 1

i) A teacher's utility function is pairwise supermodular in effort and each student's type, i.e., $f_{e\theta_i}(e,\theta) \geq 0 \quad \forall i, e \in [0,1], \theta \in \mathcal{R}^n_+$.

ii) Performance is strictly increasing in the effort of teacher and each student's type, i.e., $p_e(e,\theta) > 0$, $p_{\theta_i}(e,\theta) > 0$.

iii) $f_{ee}(e,\theta) < 0$*, teachers' utility function is strictly concave in effort.*

Assumption 1-i captures a complementarity between a teacher's effort and a student's type in teacher's utility function. To put it differently, the marginal utility of effort is increasing in a student's type. Assumption 1-ii states that the performance of a class increases as

[12] We consider only maximization problems that have a unique argmax.

the teacher puts in more effort or as the ability of a student increases. Assumption 1-iii ensures that a teacher's maximization problem has a unique interior solution given by the first-order condition, i.e., a teacher chooses an effort such that the marginal utility of effort is zero.

Given Assumption 1, increasing a student's type — having a better class — results in a higher effort by the teacher. The following lemma shows this result formally.

Lemma 2 *Given Assumption 1, the optimal effort of a teacher ($e^*(\theta) = \text{argmax}_{e\in[0,1]} f(e,\theta)$) is increasing in any student's type.*

Proof: Increasing any student's type in a class results in a better class: $\theta_i \geq \theta_i' \forall i \Rightarrow \theta \succeq^* \theta'$. Using the Topkis theorem, we have $\theta \succeq^* \theta' \Rightarrow e^*(\theta) \geq e^*(\theta')$. ■

Consider two classes with two teachers. After sorting, the two teachers are assigned to two new sorted classes, the higher track and the lower track. The teacher assigned to the higher track puts in more effort after sorting because the higher track is a better class compared with both initial classes, based on the partial order defined in (1). Sorting increases the effort of the teacher assigned to the higher track and decreases the effort of the teacher assigned to the lower track. Hence, every student in the higher track has a higher performance after sorting, and every student in lower track has a lower performance after sorting. We call this an *increase in inequality* of students' performances.

Proposition 1 *Given Assumption 1, sorting increases inequality in students' performances, i.e., every student in the higher track has a higher performance after sorting, and every student in the lower track has a lower performance after sorting.*

Proof: By Lemma 2, a teacher's effort increases in the higher track. Because performance is increasing in teacher's effort, every student in the higher track has a higher performance after sorting. Similarly for students in the lower track, performance decreases after sorting.
■

COROLLARY 1 *Under the Rawlsian welfare function, i.e., Max-Min of all students' performances, sorting decreases welfare.*

The effect of sorting on the total performance of students, under standard assumptions used in the literature, is ambiguous.

EXAMPLE 1 *(Value-added system)* *This example is inspired by the value-added measure used in Koedel et al. (2015). Suppose* $p(e, \theta_i, \theta) = e\theta_i$ *and*

$$P(e, \theta) = \frac{1}{n} \left(\sum_{i \in \mu(t)} p(e, \theta_i, \theta) - p'(\theta_i) \right),$$

where $p'(\theta_i)$ *is student* i*'s last year's test score. This performance function measures the average of a teacher's contribution to the increase in the students' scores from their scores from last year. Consider the following utility function:*

$$f(e, \theta) = u(P(e, \theta)) - c(e).$$

Suppose $u'(.) > 0$, $u''(.) < 0$, $c'(.) > 0$, $c''(.) \geq 0$, *and* $c'''(.) \geq 0$.

If $u'''(.) \leq 0$, *then sorting decreases the total performance of students; however, if* $u'''(.) \geq \zeta$[13], *then sorting increases the total performance of students. (The proof is in the appendix.)*

To understand the effect of sorting on the average/total performance of students, first we need to analyze another problem: What is the effect of sorting on the total effort of teachers? In the following theorem, we show that if the marginal utility of effort is supermodular and convex in effort, then the argmax of a teacher's utility maximization problem — the optimal effort of a teacher — is strictly supermodular. Under this condition, sorting increases the total effort of teachers. Similarly, if the marginal utility of effort is

[13] $\zeta > max\{-u''(.) \left(\frac{2n}{e}\right), c'''\}$

submodular and concave in effort, then the argmax of the teacher's utility maximization problem — the optimal effort of a teacher — is strictly submodular.

Condition 1 *The marginal utility of effort is supermodular and convex in effort:*

$$f_{e\theta_i\theta_j}(e,\theta) \geq 0, f_{ee\theta_i}(e,\theta) \geq 0, f_{eee}(e,\theta) \geq 0, \forall e \in [0,1], i, j, \theta_i \in \mathcal{R}_+, \theta_j \in \mathcal{R}_+,$$

with at least one strict inequality.

Condition 2 *The marginal utility of effort is submodular and concave in effort:*

$$f_{e\theta_i\theta_j}(e,\theta) \leq 0, f_{ee\theta_i}(e,\theta) \leq 0, f_{eee}(e,\theta) \leq 0, \forall e \in [0,1], i, j, \theta_i \in \mathcal{R}_+, \theta_j \in \mathcal{R}_+,$$

with at least one strict inequality.

Theorem 1

1. *If Condition 1 is satisfied, then the optimal effort is strictly supermodular.*
2. *If Condition 2 is satisfied, then the optimal effort is strictly submodular.*

Proof: In the appendix.

Theorem 1 holds under weaker conditions, which we characterize in the appendix.

The following thought experiment shows the effect of each inequality in Condition 1 on the teacher's total effort. Consider a teacher assigned to two identical classes. The teacher chooses an effort level such that the marginal utility of effort is zero, e^* in Figure 1, for both classes. Hence, the total effort is $2e^*$. After sorting, the teacher's marginal utility changes in both classes. More concretely, the teacher's marginal utility in the higher track shifts upward and in the lower track shifts downward at e^*. Suppose these two shifts are equal, i.e., $s_1 = s_2$ in Figure 1. Because the marginal utility of effort is convex, recall that $f_{eee}(e,\theta) \geq 0$, the increase in the teacher's optimal effort in the higher track is more than

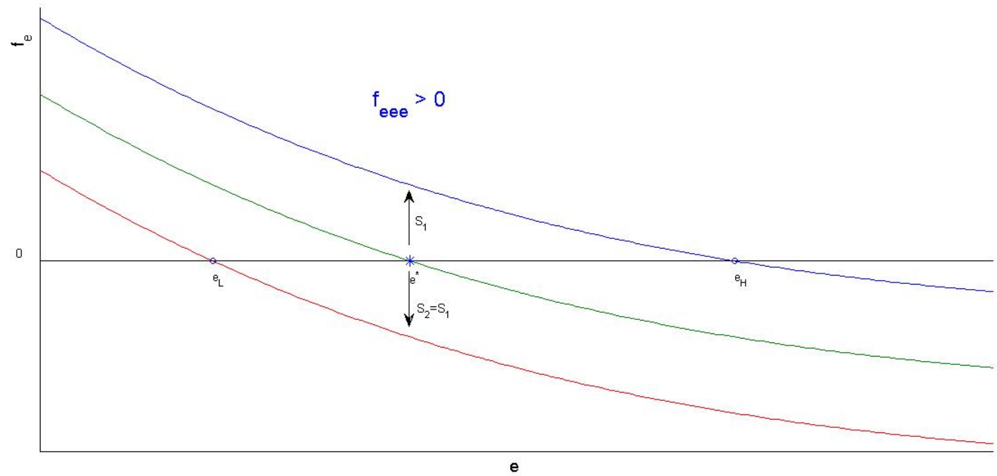

Figure 1: The teacher's marginal utility of effort is convex.

the decrease in the teacher's optimal effort in the lower track, i.e., $e_h - e^* \geq e^* - e_l$ in Figure 1, with strict inequality if the marginal utility of effort is strictly convex.

On top of that, $f_{e\theta_i\theta_j}(e, \theta) \geq 0$ ensures that the marginal utility of effort is supermodular in students' type. Therefore, the upward shift in the marginal utility function in the higher track is greater than or equal to the downward shift in the teacher's marginal utility of effort in the lower track, $s_1 \geq s_2$ in Figure 2. This implies that $e_h - e^* \geq e^* - e_l$ in Figure 2, with strict inequality if the marginal utility of effort is strictly supermodular.

$f_{ee\theta_i}(e, \theta) \geq 0$ ensures that the slope of the marginal utility of effort for the teacher in the higher track is greater than or equal to the slope of the marginal utility of effort for the teacher in the lower track. To put it differently, in Figure 3, the marginal utility of effort for the higher track is flatter than the marginal utility of effort before sorting, which is flatter than the marginal utility of effort for the lower track. Hence, the teacher's optimal effort in the lower track, e_l^*, is to the right of e_l, and the teacher's optimal effort in the higher track, e_h^*, is to the right of e_h. Therefore, $e_h^* - e^* \geq e^* - e_l^*$ with strict inequality if $f_{ee\theta_i}(e, \theta) \geq 0$ holds with strict inequality. To conclude, each of the three inequalities in Condition 1 ensure that the total effort of the teacher increases after sorting. We state the

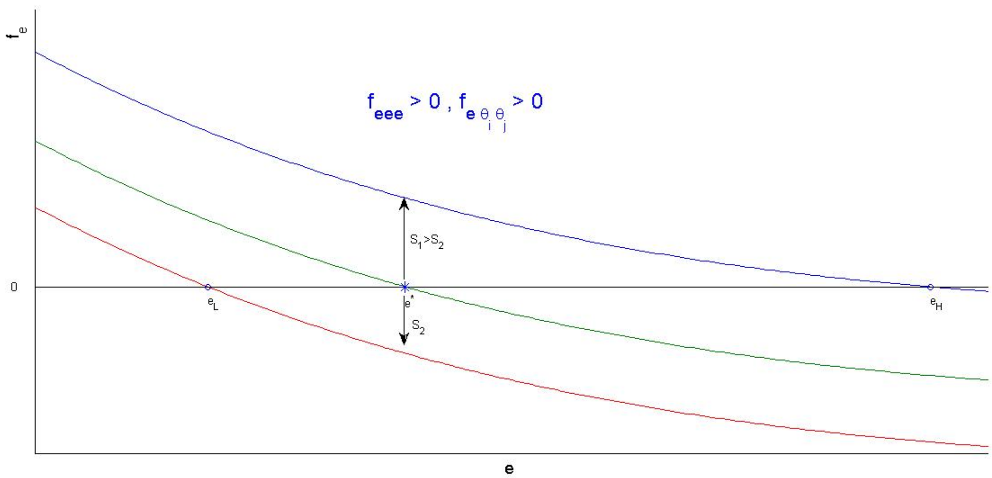

Figure 2: The teacher's marginal utility of effort is pairwise supermodular in students' type.

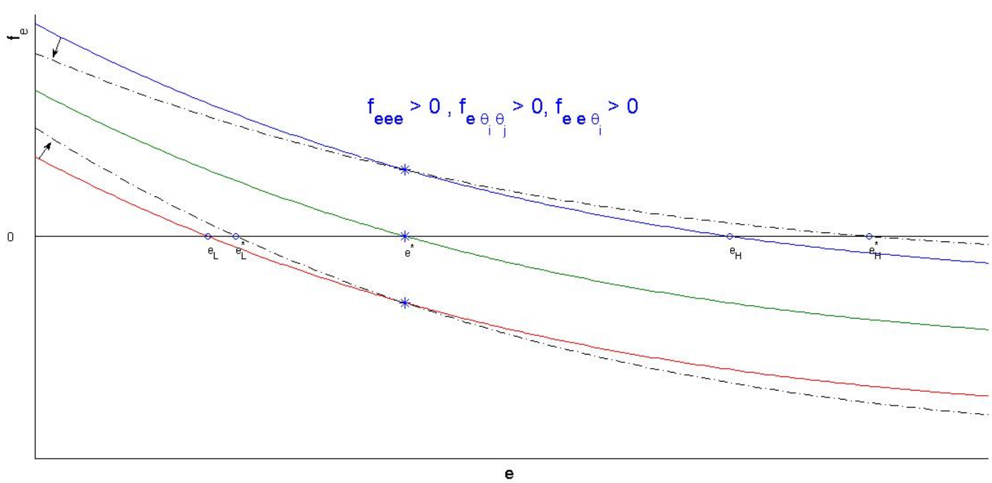

Figure 3: The teacher's marginal utility of effort is pairwise supermodular in the teacher's effort and student's type.

result for any initial class composition in the following proposition.

Proposition 2

1. *Given Condition 1, sorting strictly increases the total effort of teachers.*

2. *Given Condition 2, sorting strictly decreases the total effort of teachers.*

Proof:

1. By Lemma 1, sorting is the result of repeating the one-step sorting process. At each step, by Theorem 1, the total effort strictly increases if a class composition changes. Hence, sorting strictly increases the total effort of teachers.

2. By Lemma 1, sorting is repeating the one-step sorting process. At each step, by Theorem 1, the total effort strictly decreases if a class composition changes. Hence, sorting strictly decreases the total effort of teachers.

■

4 Impact of Sorting on Students' Performances

In this section, we characterize the effect of sorting on the average of students' performances under three different settings: (1) only indirect peer effects exist; (2) both direct and indirect peer effects exist; and (3) teachers are heterogeneous in quality and utility function, and both direct and indirect peer effects exist.

First, we consider a setting with homogeneous teachers and without direct peer effects. The only impact of sorting on students' performances is through indirect peer effects. Using the characterization of the total effort of teachers after sorting, we can characterize the effect of sorting on the average of students' performances. If teachers are putting in more (less) effort in total and the performance function is convex (concave) in effort, then the total/average students' performance increases (decreases). This setting allows us to isolate the effect of teachers' effort and the role of the curvature of the teachers' marginal utility function in answering our main question: Does sorting increase or decrease the average performances of students? Is sorting a desirable outcome based on the utilitarian welfare criterion?

Second, we consider a setting with both direct and indirect peer effects with homogeneous teachers. Duflo et al. (2011) report that both direct and indirect peer effects exist in the data and that excluding either is inconsistent with their data. Sorting has two effects on the average of students' performances. If both effects go in the same direction then we can determine whether sorting increases or decreases the average of students' performances. However, if these two effects go in opposite directions, then the effect of sorting depends on the magnitude of each effect; we provide a general method to evaluate the effect of sorting on average students' performances in this case.

Third, we consider a general environment in which teachers have different qualities and utility functions, and both direct and indirect peer effects exist. We show that — by extending Assumption 1, Condition 1, and Condition 2 to include teachers' type — the previous results are robust. Furthermore, if classes are ordered by the partial ordering $\succ^*$ defined in (1), we can analyze the welfare implications of positive assortative matching (PAM) of teachers and classes compared with negative assortative matching (NAM) of teachers and classes. The curvature of the marginal utility of effort of teachers has an important impact on this welfare comparison. There are simple examples in which the usual results about the benefits of PAM compared with NAM don't hold. More precisely, in these examples, switching from PAM to NAM decreases the inequality and increases the average performances of students, i.e., increases both the utilitarian welfare function and the Rawlsian welfare function.

4.1 Indirect Peer Effects

Consider an environment in which all teachers have the same quality and utility function. Suppose there is no direct peer effect.[14] What is the effect of sorting on the total students' performance in this environment? Can we increase the average performance of students by changing the composition of the classes? In other words, which matching of students to classes maximizes the total performance of students? In this subsection,we show that the answers to these question depend on the curvature of the marginal utility of effort of teachers. More concretely, if Condition 1 is satisfied and performance is convex in effort, sorting increases the total performance of students, i.e., sorting is a desirable outcome based on the utilitarian criterion. On the other hand, if Condition 2 is satisfied and performance is concave in effort, sorting is the least desirable matching of students to classes under both the utilitarian and the Rawlsian criteria.

Suppose there is no direct peer effect, i.e., $p_{e\theta_i}(e,\theta) = 0, p_{\theta_i\theta_j}(e,\theta) = 0 \, \forall e \in [0,1], i, j, \theta_i \in \mathcal{R}_+, \theta_j \in \mathcal{R}_+$. Then, the only effect of sorting on students' performances is through indirect peer effects. The direction of indirect peer effects depends on whether Condition 1 is satisfied and performance is convex in effort or Condition 2 is satisfied and performance is concave in effort. Under the former, teachers put in more effort in total, and the composition of classes for a fixed level effort doesn't affect the total performance of these classes; therefore, sorting increases the total performance. Under the latter, teachers put in less effort in total; hence, sorting decreases total performance.

Theorem 2

1. *If Condition 1 holds and performance is modular and convex in effort, sorting strictly increases the total performance of students.*

2. *If Condition 2 holds and performance is modular and concave in effort, sorting strictly decreases the total performance of students.*

[14] See Foster (2006).

Proof: In the appendix.

The proof specifies a general method for analyzing other situations as well; for example, when Condition 1 is satisfied but performance is concave in effort. Simply put, if equation (5) defined in the appendix is positive (negative), then sorting increases (decreases) the total performance.

Note that even when there is no direct peer effect and the total effort of teachers increases, the total performance may increase or decrease by sorting. For example, consider two teachers who exert the same level of effort for two identical classes; one teacher increases his/her effort by an amount equal to the amount by which the other teacher decreases his/her effort. The total performance of these two classes (strictly) increases if the performance function is (strictly) convex and (strictly) decreases if the performance function is (strictly) concave.

If Condition 1 holds and performance is modular and convex in effort, the effect of sorting depends on the welfare function that we use, i.e., based on different welfare objectives sorting maybe desirable or undesirable. If Condition 2 holds and performance is modular and convex in effort, sorting results in the worst classes' compositions among all other classes' compositions. Therefore, a policymaker needs to consider the curvature of a teacher's marginal utility of effort in order to make a decision that increases or decreases the sorting of students.

Corollary 2

1. *If condition 1 holds and performance is modular and convex in effort, sorting improves the utilitarian welfare function but decreases the Rawlsian welfare function.*

2. *If condition 2 holds and performance is modular and concave in effort, sorting decreases both the utilitarian welfare function and the Rawlsian welfare function.*

Because sorting changes the performance of classes, under a pay-per-performance system, sorting has an effect on the total payment to teachers — budget of an education system — and on inequality in payments to teachers. For example, under the value-added system in the U.S. education system, tracking increases teachers' income inequality; however, it may increase or decrease the average payment to teachers. The following corollary states these effects formally.

Corollary 3

1. *If Condition 1 holds, performance is modular and convex in effort, and payment to teachers is an increasing and convex function of class performance, then sorting strictly increases the total payment to teachers.*

2. *If Condition 2 holds, performance is modular and concave in effort, and payment to teachers is an increasing and concave function of class performance, then sorting strictly decreases the total payment to teachers.*

3. *If payment to teachers is an increasing function of class performance, then sorting strictly increases inequality in teachers' payments.*

4.2 Direct and Indirect Peer Effects

Suppose both direct and indirect peer effects are present. Direct and indirect peer effects may affect the total performances of students after sorting in the same direction or in opposite directions. If both direct and indirect peer effects increases (decrease) the total performances of students, then the effect of sorting is clear, which we state in the following theorem. On the other hand, if they have the opposite effect on the total performances of students, then the effect of sorting on the total performance depends on the magnitude of direct versus indirect peer effects. We provide a general method to evaluate the effect of sorting on the total performances of students when the direct and indirect peer effects have the opposite effect on the total performances of students.

Theorem 3

1. *If Condition 1 holds and performance is increasing, supermodular, and convex in effort, sorting strictly increases the total performance of students.*

2. *If Condition 2 holds and performance is increasing, submodular, and concave in effort, sorting strictly decreases the total performance of students.*

Proof: In the appendix.

Note that Proposition 1 and Corollary 1 hold in this subsection. Hence, sorting increases inequality in students' performance.

Corollary 4

1. *If Condition 1 holds and performance is increasing, supermodular, and convex in effort, sorting improves the utilitarian welfare but decreases the Rawlsian welfare.*

2. *If Condition 2 holds and performance is increasing, submodular, and concave in effort, sorting decreases both the utilitarian welfare and the Rawlsian welfare.*

Consider a situation in which sorting increases the total performances of students because of direct peer effects but decreases the total performances of students because of indirect peer effects. The effect of sorting on the total performance depends on the magnitude of these two forces. A *general method* for finding the effect of the tradeoff between direct and indirect peer effects for any performance function $p(e, \theta)$ and any utility function $f(e, \theta)$ is:

Step 1: Use equation (7) in the appendix to find the sign of the function $\Psi(e, \theta)$, which is defined in equation (7) in the appendix.

Step 2: If $\Psi(e, \theta)$ is positive everywhere, then sorting increases the total performance; if $\Psi(e, \theta)$ is negative everywhere, then sorting decreases the total performance (proof in the appendix).

Step 3: If the sign of $\Psi(e, \theta)$ is unclear for all effort levels and class compositions, we can find the sign of $\Psi(e, \theta)$ for a given level of effort and a given class composition. If $\Psi(e, \theta)$

is positive, then sorting increases the total performance at the given effort level and class composition; if $\Psi(e, \theta)$ is negative, then sorting decreases the total performance at the given effort level and class composition.

In Theorems 2 and 3, we used Assumption 1; specifically, we used the assumption that $p_{\theta_i} \geq 0$ and $f_{e\theta_i} > 0$. Note that we can define the order on θ_i such that $p_{\theta_i} \geq 0$ holds. To put it differently, assumption of $p_{\theta_i} \geq 0$ is without loss of generality; however, the assumption that both p_{θ_i} and $f_{e\theta_i}$ are strictly positive is not without loss of generality. One can relax this assumption: suppose teachers prefer low-ability students, i.e., suppose $f_{e\theta_i} < 0$. Then the optimal effort of a teacher is decreasing in a student's type. Moreover, we can use the general method, i.e., finding the sign of $\Psi(e, \theta)$, to determine the effect of sorting on the total performance of students in this situation.

Using the results from Theorem 3, we can specify the effect of sorting on the teachers' payment when both direct and indirect peer effects exist.

Corollary 5

1. *If Condition 1 holds, performance is increasing and supermodular, and payment to teachers is an increasing and convex function of class performance, then sorting strictly increases the total payment to teachers.*

2. *If Condition 2 holds, performance is increasing and submodular, and payment to teachers is an increasing and concave function of class performance, then sorting strictly decreases the total payment to teachers.*

3. *If payment to teachers is an increasing function of class performance, then sorting strictly increases inequality in teachers' payments.*

4.3 Heterogeneous Teachers and Direct and Indirect Peer Effects

Suppose each teacher has a quality — teacher's type — $q_t \in \mathcal{R}_+$ that enters the performance function and teachers' utility function, i.e., teachers' utility function is $f(e, q, \theta)$ and the performance of a class is $p(e, q, \theta)$. We extend Assumption 1 and Conditions 1 and 2 to include the teachers' type as well.

Assumption 2 *A teacher's utility function is pairwise supermodular in effort and the teacher's type, i.e.,* $f_{eq}(e, q, \theta) \geq 0 \quad \forall e \in [0, 1], q \in \mathcal{R}_+, \theta \in \mathcal{R}^n_+$.

Condition 3 *The marginal utility of effort is supermodular and convex in effort, i.e.,*

$$f_{eq\theta_i} \geq 0, f_{eeq} \geq 0, f_{e\theta_i\theta_j} \geq 0, f_{ee\theta_i} \geq 0, f_{eee} \geq 0, \forall e \in [0, 1], \forall \theta \in \mathcal{R}^n_+, \forall q \in \mathcal{R}_+,$$

with at least one strict inequality.

Condition 4 *The marginal utility of effort is submodular and concave in effort, i.e.,*

$$f_{eq\theta_i} \leq 0, f_{eeq} \leq 0, f_{e\theta_i\theta_j} \leq 0, f_{ee\theta_i} \leq 0, f_{eee} \leq 0, \forall e \in [0, 1], \forall \theta \in \mathcal{R}^n_+, \forall q \in \mathcal{R}_+,$$

with at least one strict inequality.

We maintain the assumption of anonymity for students of a class, i.e., teachers' utility functions and the performance of a class are symmetric functions in students' types. The one-step sorting is defined as before with the assignment of the higher-quality teacher to the upper track. Sorting is defined as: Order all the students in the two classes by their types, then put the top half of the students in one class with the teacher who has the higher quality and put the bottom half in the other class with the teacher who has the lower type. If the two new classes have different student and teacher compositions as did the two classes before sorting, then this process is called *sorting.* After sorting of two classes, we can order the students of these classes by the binary relation $\succ^*$ defined in (1). Define *student sorting with positive assortative matching*(PAM) as the sorting of students of two classes and assigning the higher-quality teacher to the higher-track class and the lower-type teacher to the lower-track class. Sorting is equivalent to student sorting with PAM.

Similarly, define *student sorting with negative assortative matching*(NAM) as the sorting of students of two classes and assigning the lower-quality teacher to the higher-track class and the higher-quality teacher to the lower-track class.

The following theorem states the parallel result of Theorem 2 when teachers are heterogeneous. We drive the result parallel to Proposition 2 — the effect of sorting on the total teachers' effort — in the appendix.

Theorem 4 *Given Assumptions 1 and 2:*

1. *If Condition 3 holds and performance is supermodular and convex, sorting of two classes increases the total performance.*

2. *If Condition 4 holds and performance is submodular and concave, sorting of two classes decreases the total performance.*

Proof: In the appendix.

As we stated in Corollary 3, under a pay-per-performance system — such as the value-added system in the U.S. — sorting changes the inequality in teachers' payment and the total payment to teachers.

Corollary 6 *Suppose Assumptions 1 and 2 are satisfied.*

1. *If Condition 3 holds, performance is supermodular and convex, and payment to teachers is an increasing and convex function of class performance, then sorting strictly increases the total payment to teachers.*

2. *If Condition 4 holds, performance is submodular and concave, and payment to teachers is an increasing and concave function of class performance, then sorting strictly decreases the total payment to teachers.*

3. *If payment to teachers is an increasing function of class performance, then sorting strictly increases inequality in teachers' payments.*

Changing the matching from PAM to NAM when students are sorted — or classes are ordered by the binary relationship $\succ^*$ defined in (1) — has the opposite effect of sorting. The following theorem shows one implication of such a change in the matching of classes and teachers. Note that the standard assumption on supermodularity or submodularity of the performance function is insufficient for analyzing the advantage or disadvantage of PAM versus NAM.

Theorem 5 *Suppose Assumptions 1 and 2 and Condition 4 are satisfied and performance is submodular and concave:*

1. *The total performance under student sorting with NAM is higher than student sorting with PAM.*

2. *Inequality in students' performance under student sorting with NAM is lower than student sorting with PAM.*

Proof: In the appendix.

COROLLARY 7 *Suppose Assumptions 1 and 2 and Condition 4 are satisfied and performance is submodular and concave:*

1. *The total payment to teachers under student sorting with NAM is higher than student sorting with PAM.*

2. *Inequality in teachers' payment under student sorting with NAM is lower than student sorting with PAM.*

Remark 1 ***Affirmative Action****:*

Let I be a set of students and each student belongs to either the minority group or the majority group. These students are assigned to two classes θ^1, θ^2, such that $\theta^1 \succ^* \theta^2$. Suppose the percentage of minority students in class θ^1 in less than ϕ and the percentage of minority students in class θ^2 is more than ϕ.

Affirmative action policies such as implementing a quota, i.e., assigning at least ϕ percent of seats to minority students in each class, have the opposite effect of sorting. Denote the two classes after implementing quotas by θ_a^1,θ_a^2. The following relation between these classes hold: $\theta^1 \succ^* \theta_a^1$, $\theta^2 \succ^* \theta_a^1$, $\theta_a^2 \succ^* \theta^2$, and $\theta_a^2 \succ^* \theta^1$. Affirmative action policies have the opposite effect of sorting on the total/average effort of teachers and the total/average performance of students in Proposition 2 and Theorems 2, 3, and 4.

5 Conclusion

We model an education system in which teachers choose their effort level based on the whole distribution of students, not only the mean of students' abilities. Furthermore, in our model both direct and indirect peer effects exist. The model incorporates both between-school sorting and within-school sorting, i.e., tracking.

We show that the standard assumptions in the literature are insufficient to understand the effect of sorting on the total effort of teachers and the total performance of students. We show that the change in the total performance of students after sorting depends on teachers' utility function. Even in the absence of direct peer effects, when teachers are homogeneous, sorting has an effect on the total effort of teachers and the total performance of students that depends on the curvature of teachers' marginal utility of effort.

We characterize conditions on the utility function of a teacher under which the optimal effort of a teacher is strictly supermodular. Under these conditions, sorting increases the total effort of teachers and the total performance of students, even though sorting increases inequality in students' performances. Therefore, under these conditions, how one evaluates the effect of sorting on students' performances depends on the welfare criteria chosen, i.e., under the Utilitarian criterion, sorting increases welfare; however, under the Rawlsian criterion, sorting decreases welfare. Similarly, we characterize conditions on the utility function of a teacher under which the optimal effort of a teacher is strictly submodular.

Under these conditions, under both welfare criteria, any assignment of students to teachers is strictly preferred to sorting.

Appendix A Proofs of Results

Proof of Lemma 1

Proof: Consider two classes $\theta' = (\theta'_1, ..., \theta'_n)$ and $\theta'' = (\theta''_1, ..., \theta''_n)$. Order all elements of these two vectors in descending order, denoted by $\bar{\theta}_1, \bar{\theta}_2, ..., \bar{\theta}_{2n}$, i.e., $\bar{\theta}_1$ is greater than or equal to all elements of both classes, $\bar{\theta}_2$ is the second greatest element of all elements of both classes, and so on. $\bar{\theta}_1$ is either in class θ' or in class θ''. Consider the permutation of these two classes such that $\bar{\theta}_1$ is the first element of one these two classes. Similarly $\bar{\theta}_2$ is either class θ' or in class θ''. Consider the permutation of these two classes such that $\bar{\theta}_1$ is the first element of one of these two classes and $\bar{\theta}_2$ is the second element of one of these two classes. We can do the same for $\bar{\theta}_1, \bar{\theta}_2, ..., \bar{\theta}_n$. Therefore, we have a permutation of two classes, where $\bar{\theta}_1$ is the first element of one class. Moreover, the first element of the other class is one of the following: $\bar{\theta}_{n+1}, \bar{\theta}_{n+2}, ..., \bar{\theta}_{2n}$. Note that $\bar{\theta}_1 \geq \bar{\theta}_{n+1}, \bar{\theta}_{n+2}, ..., \bar{\theta}_{2n}$. Similarly, $\bar{\theta}_2$ is the second element of one of the two classes, and the second element of the other class is one of $\bar{\theta}_{n+1}, \bar{\theta}_{n+2}, ..., \bar{\theta}_{2n}$. The same is true for any element i between 1 and n. $\bar{\theta}_i$ is the ith element of one these classes, and the ith element of the other class is one of $\bar{\theta}_{n+1}, \bar{\theta}_{n+2}, ..., \bar{\theta}_{2n}$. Furthermore, $\bar{\theta}_i \geq \bar{\theta}_{n+1}, \bar{\theta}_{n+2}, ..., \bar{\theta}_{2n}$. Hence, using coordinate-wise maximum and minimum on these permutations results in having $\bar{\theta}_1, \bar{\theta}_2, ..., \bar{\theta}_n$ in one class and $\bar{\theta}_{n+1}, \bar{\theta}_{n+2}, ..., \bar{\theta}_{2n}$ in the other class. Therefore, by using these permutations for these two classes, one-step sorting is equivalent to sorting. ■

Proof of Example 1

Define $\bar{\theta} = (\frac{1}{n}) \sum_{i \in \mu(t)} \theta_i$. We have:

$$P_{\theta_i} = \frac{e}{n}.$$

$$P_e = \bar{\theta}.$$

$$f_{\theta_i} = \frac{e}{n} u'(.).$$

$$f_{\theta_i,\theta_j} = (\frac{e}{n})^2 u''(.).$$

$$f_{\theta_i,\theta_j,e} = 2(\frac{e}{n}) u''(.) + \bar{\theta}(\frac{e}{n})^2 u'''(.).$$

$$f_e = \bar{\theta} u'(.) - c_e.$$

$$f_{e,e} = (\bar{\theta})^2 u''(.) - c_{e,e}.$$

$$f_{e,e,e} = (\bar{\theta})^3 u'''(.) - c_{e,e,e}.$$

$$f_{\theta_i,e,e} = \frac{2}{n}(\bar{\theta}) u''(.) + \frac{e}{n}(\bar{\theta})^2 u'''(.).$$

Note that $f_{\theta_i,\theta_j,e}$ and $f_{\theta_i,e,e}$ are strictly negative, and $f_{e,e,e}$ is negative. Therefore, Condition 2 is satisfied.

$$P_i^* = P_e e_i + p_i$$

$$P_{i,j}^* = p_{e,e} e_i e_j + p_{e,j} e_i + p_e e_{i,j} + p_{e,i} e_j + P_{i,j}$$

$$= \frac{1}{n}(e_i + e_j) + \bar{\theta} e_{i,j}.$$

Want to show $P_{i,j}^* < 0$:

$$\frac{1}{n}(e_i + e_j) + \bar{\theta} e_{i,j} \Leftrightarrow 2e_i < -n\bar{\theta} e_{i,j}$$

$$\Leftrightarrow \quad -2\frac{f_{e,i}}{f_{e,e}} < -n\bar{\theta}\frac{-1}{f_{e,e}^3}(\Phi)$$

$$\Leftrightarrow \quad 2f_{e,i} < -n\bar{\theta}\frac{1}{f_{e,e}^2}(\Phi)$$

$$\Leftrightarrow \quad \frac{2}{n\bar{\theta}} < \frac{-(\Phi)}{f_{e,e}^2 f_{e,i}},$$

where

$$
\begin{aligned}
-\Phi &= -f_{e,e}^2 f_{e,\theta_i,\theta_j} - f_{e,e,e} f_{e,\theta_i} f_{e,\theta_j} + f_{e,e,\theta_i} f_{e,\theta_j} f_{e,e} + f_{e,e,\theta_j} f_{e,\theta_i} f_{e,e} \\
&= -f_{e,e}^2 f_{e,\theta_i,\theta_j} - f_{e,e,e} f_{e,\theta_i} f_{e,\theta_j} + 2 f_{e,e,\theta_i} f_{e,\theta_j} f_{e,e}.
\end{aligned}
$$

Note that

$$
-f_{e,e}^2 f_{e,\theta_i,\theta_j} \geq 0, -f_{e,e,e} f_{e,\theta_i} f_{e,\theta_j} \geq 0.
$$

Therefore,

$$
-\Phi \geq 2 f_{e,e,\theta_i} f_{e,\theta_j} f_{e,e}.
$$

Because $f_{e,e}^2 f_{e,i} \geq 0$, it is enough to show:

$$
\frac{2}{n\bar{\theta}} < \frac{2 f_{e,e,\theta_i} f_{e,\theta_j} f_{e,e}}{f_{e,e}^2 f_{e,i}} \Leftrightarrow \frac{1}{n\bar{\theta}} < \frac{f_{e,e,\theta_i}}{f_{e,e}}.
$$

Note that $f_{e,e} = (\bar{\theta})^2 u''(.) - c_{e,e}$ and $c_{e,e} > 0$; hence, $f_{e,e} \leq (\bar{\theta})^2 u''(.)$. Therefore, $\frac{f_{e,e,\theta_i}}{f_{e,e}} \geq \frac{f_{e,e,\theta_i}}{(\bar{\theta})^2 u''(.)}$. It is enough to show:

$$
\begin{aligned}
& \tfrac{1}{n\bar{\theta}} < \tfrac{f_{e,e,\theta_i}}{(\bar{\theta})^2 u''(.)} \\
\Leftrightarrow \quad & \tfrac{1}{n} < \tfrac{f_{e,e,\theta_i}}{(\bar{\theta}) u''(.)} \\
& = \tfrac{\frac{2}{n}(\bar{\theta}) u''(.) + \frac{e}{n}(\bar{\theta})^2 u'''(.)}{(\bar{\theta}) u''(.)} \\
& = \tfrac{2}{n} + \tfrac{\frac{e}{n}(\bar{\theta})^2 u'''(.)}{(\bar{\theta}) u''(.)} \\
& = \tfrac{2}{n} + \tfrac{\frac{e}{n}(\bar{\theta})^2}{(\bar{\theta})} \tfrac{u'''(.)}{u''(.)}.
\end{aligned}
$$

Because $\frac{u'''(.)}{u''(.)} \geq 0$, we have:

$$\frac{2}{n} + \frac{\frac{e}{n}(\bar{\theta})^2}{(\bar{\theta})}\frac{u'''(.)}{u''(.)} > \frac{2}{n}.$$

which is what we wanted to show. ■

Proof of Theorem 1

Consider a function $f(e, \theta)$; maximizing with respect to e we have:

$$f_e(e^*, \theta) = 0.$$

By the Implicit Function Theorem, we have:

$$e^* = g(\theta) \Rightarrow \frac{\partial e^*}{\partial \theta_i} = -\frac{\frac{\partial f_e}{\partial \theta_i}}{\frac{\partial f_e}{\partial e}} = -\frac{f_{e\theta_i}}{f_{ee}}. \tag{2}$$

Therefore,

$$\frac{\partial^2 e^*}{\partial \theta_i \partial \theta_j} = -\frac{f_{ee}^2 f_{e\theta_i\theta_j} + f_{eee} f_{e\theta_i} f_{e\theta_j} - f_{ee\theta_i} f_{e\theta_j} f_{ee} - f_{ee\theta_j} f_{e\theta_i} f_{ee}}{f_{ee}^3}. \tag{3}$$

Condition 5

$$f_{ee}^2 f_{e\theta_i\theta_j} + f_{eee} f_{e\theta_i} f_{e\theta_j} - f_{ee\theta_i} f_{e\theta_j} f_{ee} - f_{ee\theta_j} f_{e\theta_i} f_{ee} > 0, \forall e \in [0,1], i, j, \theta_i \in \mathcal{R}_+, \theta_j \in \mathcal{R}_+.$$

Condition 6

$$f_{ee}^2 f_{e\theta_i\theta_j} + f_{eee} f_{e\theta_i} f_{e\theta_j} - f_{ee\theta_i} f_{e\theta_j} f_{ee} - f_{ee\theta_j} f_{e\theta_i} f_{ee} < 0, \forall e \in [0,1], i, j, \theta_i \in \mathcal{R}_+, \theta_j \in \mathcal{R}_+.$$

If Condition 5 is satisfied, then (3) is strictly positive. Moreover, if Condition 1 is satisfied, then Condition 5 is satisfied. Hence, the argmax is strictly supermodular if

Condition 1 is satisfied. If Condition 6 is satisfied, then (3) is strictly negative. Moreover, if condition 2 is satisfied, then Condition 6 is satisfied. Hence, the argmax is strictly submodular if Condition 2 is satisfied.

Note that

$$f_{ee}^2 f_{e\theta_i\theta_j} + f_{eee} f_{e\theta_i} f_{e\theta_j} - f_{ee\theta_i} f_{e\theta_j} f_{ee} - f_{ee\theta_j} f_{e\theta_i} f_{ee},$$

is either zero, strictly positive, or strictly negative locally. Therefore, locally the argmax is either strictly supermodular, strictly submodular, or modular. ■

Proof of Theorem 2

Consider a general performance function $p(e, \theta)$:

$$\frac{\partial(p\big|_{e=e^*})}{\partial\theta_i} = \frac{\partial p}{\partial e}\frac{\partial(e\big|_{e=e^*})}{\partial\theta_i} + \frac{\partial p}{\partial\theta_i}\Big|_{e=e^*} \Rightarrow$$

$$\frac{\partial(p\big|_{e=e^*})}{\partial\theta_i\partial\theta_j} = \left((\frac{\partial^2 p}{\partial e\partial e}\frac{\partial e\big|_{e=e^*}}{\partial\theta_j} + \frac{\partial^2 p}{\partial e\partial\theta_j})\frac{\partial(e\big|_{e=e^*})}{\partial\theta_i} + \frac{\partial p}{\partial e}\frac{\partial^2(e\big|_{e=e^*})}{\partial\theta_i\partial\theta_j} + \frac{\partial p}{\partial e\partial\theta_i}\frac{\partial(e\big|_{e=e^*})}{\partial\theta_j} + \frac{\partial^2 p}{\partial\theta_i\partial\theta_j} \right)\Big|_{e=e^*}. \quad (4)$$

Using (4) when performance is modular, i.e., $p_{e\theta_i}(e, \theta) = 0, p_{\theta_i\theta_j}(e, \theta) = 0 \quad \forall e \in [0, 1], \forall i, j, \theta_i \in \mathcal{R}_+, \theta_j \in \mathcal{R}_+$, we have:

$$\frac{\partial(p\big|_{e=e^*})}{\partial\theta_i\partial\theta_j} = \left((\frac{\partial^2 p}{\partial e\partial e}\frac{\partial e\big|_{e=e^*}}{\partial\theta_j})\frac{\partial(e\big|_{e=e^*})}{\partial\theta_i} + \frac{\partial p}{\partial e}\frac{\partial^2(e\big|_{e=e^*})}{\partial\theta_i\partial\theta_j} \right)\Big|_{e=e^*}. \quad (5)$$

1. By Theorem 1, under Condition 1, we have $\frac{\partial(e\big|_{e=e^*})}{\partial\theta_i\partial\theta_j} > 0$. Performance is convex, hence, we have $\frac{\partial^2 p}{\partial e\partial e} \geq 0$. Moreover, $\frac{\partial e\big|_{e=e^*}}{\partial\theta_i} > 0$ by Lemma 2. Therefore, $\frac{\partial(p\big|_{e=e^*})}{\partial\theta_i\partial\theta_j} > 0$.
2. By Theorem 1, under Condition 2, we have $\frac{\partial(e\big|_{e=e^*})}{\partial\theta_i\partial\theta_j} < 0$. Performance is concave, hence, we have $\frac{\partial^2 p}{\partial e\partial e} \leq 0$. Moreover, $\frac{\partial e\big|_{e=e^*}}{\partial\theta_i} > 0$ by Lemma 2. Therefore, $\frac{\partial(p\big|_{e=e^*})}{\partial\theta_i\partial\theta_j} < 0$.

■

Proof of Theorem 3

Consider a general performance function $p(e,\theta)$, and recall (4):

$$\frac{\partial(p\big|_{e=e^*})}{\partial\theta_i\partial\theta_j} = \left((\frac{\partial^2 p}{\partial e\partial e}\frac{\partial e\big|_{e=e^*}}{\partial\theta_j} + \frac{\partial^2 p}{\partial e\partial\theta_j})\frac{\partial(e\big|_{e=e^*})}{\partial\theta_i} + \frac{\partial p}{\partial e}\frac{\partial(e\big|_{e=e^*})}{\partial\theta_i\partial\theta_j} + \frac{\partial p}{\partial e\partial\theta_i}\frac{\partial(e\big|_{e=e^*})}{\partial\theta_j} + \frac{\partial^2 p}{\partial\theta_i\partial\theta_j}\right)\Bigg|_{e=e^*}. \quad (6)$$

1. $\frac{\partial(p\big|_{e=e^*})}{\partial\theta_i\partial\theta_j} > 0$ because:

   i) By convexity of performance, we have $\frac{\partial^2 p}{\partial e\partial e} \geq 0$.

   ii) By Lemma 2, we have $\frac{\partial e\big|_{e=e^*}}{\partial\theta_i} > 0 \forall i \in I$.

   iii) Performance is supermodular; therefore, we have $\frac{\partial^2 p}{\partial e\partial\theta_i} \geq 0 \forall i \in I$.

   iv) By Assumption 1, performance is increasing in effort $\frac{\partial p}{\partial e} > 0$.

   v) By Theorem 1, under Condition 1, the optimal effort is strictly supermodular $\frac{\partial(e\big|_{e=e^*})}{\partial\theta_i\partial\theta_j} > 0$.

   vi) Performance is supermodular; therefore, we have $\frac{\partial^2 p}{\partial\theta_i\partial\theta_j} \geq 0$.

   Therefore, $\frac{\partial(p\big|_{e=e^*})}{\partial\theta_i\partial\theta_j} > 0$, i.e., sorting increases the total performance of students.

2. $\frac{\partial(p\big|_{e=e^*})}{\partial\theta_i\partial\theta_j} < 0$ because:

   i) By concavity of performance, we have $\frac{\partial^2 p}{\partial e\partial e} \leq 0$.

   ii) By Lemma 2, we have $\frac{\partial e\big|_{e=e^*}}{\partial\theta_i} > 0 \forall i \in I$.

   iii) Performance is submodular; therefore, we have $\frac{\partial^2 p}{\partial e\partial\theta_i} \leq 0 \forall i \in I$.

   iv) By Assumption 1, performance is increasing in effort $\frac{\partial p}{\partial e} > 0$.

   v) By Theorem 1, under Condition 2, the optimal effort is strictly submodular $\frac{\partial(e\big|_{e=e^*})}{\partial\theta_i\partial\theta_j} < 0$.

   vi) Performance is submodular; therefore, we have $\frac{\partial^2 p}{\partial\theta_i\partial\theta_j} \leq 0$.

Therefore, $\frac{\partial(p\big|_{e=e^*})}{\partial\theta_i\partial\theta_j} < 0$, i.e., sorting decreases the total performance of students.

■

General Method

Using (2), (3), and (4), we have:

$$\frac{\partial(p\big|_{e=e^*})}{\partial\theta_i\partial\theta_j} = \left(\left(\left(\frac{\partial^2 p}{\partial e\partial e}\right)\left(-\frac{\frac{\partial^2 f}{\partial e\partial\theta_j}}{\frac{\partial^2 f}{\partial e\partial e}}\right) + \frac{\partial^2 p}{\partial e\partial\theta_j}\right)\left(-\frac{\frac{\partial^2 f}{\partial e\partial\theta_i}}{\frac{\partial^2 f}{\partial e\partial e}}\right)\right.$$
$$+ \frac{\partial p}{\partial e}\left(-\frac{(\frac{\partial^2 f}{\partial e\partial e})^2(\frac{\partial^3 f}{\partial e\partial\theta_i\partial\theta_j})+(\frac{\partial^3 f}{\partial e\partial e\partial e})(\frac{\partial^2 f}{\partial e\partial\theta_i})(\frac{\partial^2 f}{\partial e\partial e\theta_j})-(\frac{\partial^3 f}{\partial e\partial e\partial\theta_i})(\frac{\partial^2 f}{\partial e\partial\theta_j})(\frac{\partial^2 f}{\partial e\partial e})-(\frac{\partial^3 f}{\partial e\partial e\partial\theta_j})(\frac{\partial^2 f}{\partial e\partial\theta_i})(\frac{\partial^2 f}{\partial e\partial e})}{(\frac{\partial^2 f}{\partial e\partial e})^3}\right)$$
$$+ \left.\frac{\partial p}{\partial e\partial\theta_i}\left(-\frac{\frac{\partial^2 f}{\partial e\partial\theta_j}}{\frac{\partial^2 f}{\partial e\partial e}}\right) + \frac{\partial^2 p}{\partial\theta_i\partial\theta_j}\right)\Bigg|_{e=e^*}.$$

Note that $\frac{\partial(p\big|_{e=e^*})}{\partial\theta_i\partial\theta_j}$ depends only on the primitives in this equation. Define:

$$\Psi(e,\theta) = \left(\left(\frac{\partial^2 p}{\partial e\partial e}\right)\left(-\frac{\frac{\partial^2 f}{\partial e\partial\theta_j}}{\frac{\partial^2 f}{\partial e\partial e}}\right) + \frac{\partial^2 p}{\partial e\partial\theta_j}\right)\left(-\frac{\frac{\partial^2 f}{\partial e\partial\theta_i}}{\frac{\partial^2 f}{\partial e\partial e}}\right)$$
$$+ \frac{\partial p}{\partial e}\left(-\frac{(\frac{\partial^2 f}{\partial e\partial e})^2(\frac{\partial^3 f}{\partial e\partial\theta_i\partial\theta_j})+(\frac{\partial^3 f}{\partial e\partial e\partial e})(\frac{\partial^2 f}{\partial e\partial\theta_i})(\frac{\partial^2 f}{\partial e\partial e\theta_j})-(\frac{\partial^3 f}{\partial e\partial e\partial\theta_i})(\frac{\partial^2 f}{\partial e\partial\theta_j})(\frac{\partial^2 f}{\partial e\partial e})-(\frac{\partial^3 f}{\partial e\partial e\partial\theta_j})(\frac{\partial^2 f}{\partial e\partial\theta_i})(\frac{\partial^2 f}{\partial e\partial e})}{(\frac{\partial^2 f}{\partial e\partial e})^3}\right)$$

$$+ \frac{\partial p}{\partial e\partial\theta_i}\left(-\frac{\frac{\partial^2 f}{\partial e\partial\theta_j}}{\frac{\partial^2 f}{\partial e\partial e}}\right) + \frac{\partial^2 p}{\partial\theta_i\partial\theta_j}. \tag{7}$$

If $\Psi(e,\theta) \geq (\leq)0$, then $\frac{\partial(p\big|_{e=e^*})}{\partial\theta_i\partial\theta_j} \geq (\leq)0$, i.e., sorting increases (decreases) the total performance of students. ■

Proof of Theorem 4

First, we extend Theorem 1:

Lemma 3

1. *If Condition 3 holds, sorting strictly increases the total effort of teachers.*

2. *If Condition 4 holds, sorting strictly decreases the total effort of teachers.*

Proof: Consider a function $f(e,q,\theta)$, maximizing with respect to e:

$$f_e(e,q,\theta)=0 \Rightarrow e^* = g(q,\theta)$$
$$\Rightarrow \qquad \frac{\partial e^*}{\partial \theta_i} = -\frac{\frac{\partial f_e}{\partial \theta_i}}{\frac{\partial f_e}{\partial e}} = -\frac{f_{e\theta_i}}{f_{ee}}.$$

Therefore,

$$\frac{\partial^2 e^*}{\partial \theta_i \partial q} = -\frac{f_{ee}^2 f_{e\theta_i q} + f_{eee} f_{e\theta_i} f_{eq} - f_{ee\theta_i} f_{eq} f_{ee} - f_{eeq} f_{e\theta_i} f_{ee}}{f_{ee}^3}. \tag{8}$$

If Condition 3 is satisfied, then (8) is strictly positive. Moreover, Condition 3 implies Condition 1. Therefore, if Condition 3 is satisfied, then (3) is strictly positive, by the same argument as in proof of Theorem 1. Hence, the argmax is strictly supermodular if Condition 3 is satisfied and sorting strictly increases the total effort of teachers. If Condition 4 is satisfied, then (8) is strictly negative. Moreover, Condition 4 implies Condition 2. Therefore, if Condition 4 is satisfied, then (3) is strictly negative, by the same argument as in proof of Theorem 1. Hence, the argmax is strictly submodular if Condition 2 is satisfied and sorting strictly decreases the total effort of teachers. ■

Consider a general performance function $p(e,q,\theta)$, and recall (4):

$$\frac{\partial(p\big|_{e=e^*})}{\partial\theta_i\partial q} = \left(\left(\frac{\partial^2 p}{\partial e\partial e}\frac{\partial e\big|_{e=e^*}}{\partial q} + \frac{\partial^2 p}{\partial e\partial q}\right)\frac{\partial(e\big|_{e=e^*})}{\partial\theta_i} + \frac{\partial p}{\partial e}\frac{\partial(e\big|_{e=e^*})}{\partial\theta_i\partial q} + \frac{\partial p}{\partial e\partial\theta_i}\frac{\partial(e\big|_{e=e^*})}{\partial q} + \frac{\partial^2 p}{\partial\theta_i\partial q}\right)\Bigg|_{e=e^*}.$$

1. To prove the first part of the theorem, we need to show that $p(e,q,\theta)$ is pairwise supermodular in students' types and pairwise supermodular in each student's type and the teacher's type at the optimal effort. In Theorem 3, we established that under these conditions, $p(e,q,\theta)$ is pairwise supermodular in students' types.

   We need to show that $\frac{\partial(p\big|_{e=e^*})}{\partial\theta_i\partial q} > 0$:

i) By convexity of performance, we have $\frac{\partial^2 p}{\partial e \partial e} \geq 0$.

ii) By Lemma 2, we have $\frac{\partial e\big|_{e=e^*}}{\partial \theta_i} > 0 \forall i \in I$.

iii) Because performance is supermodular, we have $\frac{\partial^2 p}{\partial e \partial \theta_i} \geq 0 \forall i \in I$ and $\frac{\partial^2 p}{\partial e \partial q} \geq 0 \forall t$.

iv) By Assumption 1, performance is increasing in effort $\frac{\partial p}{\partial e} > 0$.

v) By Lemma 3, under Condition 3, the optimal effort is strictly supermodular $\frac{\partial (e\big|_{e=e^*})}{\partial \theta_i \partial q} > 0$.

vi) Because performance is supermodular, we have $\frac{\partial^2 p}{\partial \theta_i \partial q} \geq 0$.

Therefore, $\frac{\partial (p\big|_{e=e^*})}{\partial \theta_i \partial q} > 0$. We can conclude that sorting increases the total performance of students.

2. Similarly, to prove the second part of the theorem, we need to show that $p(e, q, \theta)$ is pairwise submodular in students' types and pairwise submodular in each student's type and the teacher's type at the optimal effort. In Theorem 3, we established that under these conditions, $p(e, q, \theta)$ is pairwise submodular in students' types. We need to show that $\frac{\partial (p\big|_{e=e^*})}{\partial \theta_i \partial q} < 0$:

i) By concavity of performance, we have $\frac{\partial^2 p}{\partial e \partial e} \leq 0$.

ii) By Lemma 2, we have $\frac{\partial e\big|_{e=e^*}}{\partial \theta_i} > 0 \forall i \in I$.

iii) Because performance is submodular, we have $\frac{\partial^2 p}{\partial e \partial \theta_i} \leq 0 \forall i \in I$.

iv) By Assumption 1, performance is increasing in effort $\frac{\partial p}{\partial e} > 0$.

v) By Lemma 3, under Condition 4, the optimal effort is strictly submodular $\frac{\partial (e\big|_{e=e^*})}{\partial \theta_i \partial \theta_j} < 0$.

vi) Because performance is submodular, we have $\frac{\partial^2 p}{\partial \theta_i \partial \theta_j} \leq 0$.

Therefore, $\frac{\partial (p\big|_{e=e^*})}{\partial \theta_i \partial \theta_j} < 0$. We can conclude that sorting decreases the total performance of students.

■

Proof of Theorem 5

Consider a general performance function $p(e, q, \theta)$, and recall (4):

$$\frac{\partial(p\big|_{e=e^*})}{\partial\theta_i\partial q} = \left((\frac{\partial^2 p}{\partial e\partial e}\frac{\partial e\big|_{e=e^*}}{\partial q} + \frac{\partial^2 p}{\partial e\partial q}\frac{\partial(e\big|_{e=e^*})}{\partial\theta_i} + \frac{\partial p}{\partial e}\frac{\partial(e\big|_{e=e^*})}{\partial\theta_i\partial q} + \frac{\partial p}{\partial e\partial\theta_i}\frac{\partial(e\big|_{e=e^*})}{\partial q} + \frac{\partial^2 p}{\partial\theta_i\partial q}\right)\Big|_{e=e^*}.$$

By proof of Theorem 4, we know that $\frac{\partial(p\big|_{e=e^*})}{\partial\theta_i\partial q} < 0$.

Consider two classes after student sorting with NAM. By sorting these two classes, we get two new classes with student sorting with PAM. By Theorem 4, sorting decreases the total performance of students, i.e., the total performance of students under NAM is higher than under PAM. Furthermore, by Theorem 4, sorting increases inequality in students' performance, i.e., inequality in students' performance under NAM is higher than under PAM. To put it differently, when two classes are sorted, i.e., student sorting with PAM, changing the matching of teachers and class from PAM to NAM has the opposite effect of sorting. Because when two classes are sorted with NAM, sorting results in two classes that are sorted with PAM. ■

References

Avery, C. and P. A. Pathak (2015): "The Distributional Consequences of Public School Choice," Tech. rep., National Bureau of Economic Research.

Barrow, L. and C. E. Rouse (2008): "School vouchers and student achievement: Recent evidence, remaining questions," *FRB of Chicago Working Paper No. 2008-08.*

Chakrabarti, R. (2009): "Do vouchers lead to sorting under random private-school selection? Evidence from the Milwaukee voucher program," *Evidence from the Milwaukee Voucher Program (July 1, 2009). FRB of New York Staff Report.*

Duflo, E., P. Dupas, and M. Kremer (2011): "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya," *American Economic Review*, 101, 1739–74.

Edlin, A. S. and C. Shannon (1998): "Strict monotonicity in comparative statics," *Journal of Economic Theory*, 81, 201–219.

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

Epple, D. and R. Romano (2011): "Peer effects in education: A survey of the theory and evidence," *Handbook of social economics*, 1, 1053–1163.

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

Foster, G. (2006): "It's not your peers, and it's not your friends: Some progress toward understanding the educational peer effect mechanism," *Journal of public Economics*, 90, 1455–1475.

Fryer, R. (2013): "Teacher Incentives and Student Achievement: Evidence from New York City Public Schools," *Journal of Labor Economics*, 31, 373–427.

Hastings, J. S. and J. M. Weinstein (2007): "Information, school choice, and academic achievement: Evidence from two experiments," Tech. rep., National Bureau of Economic Research.

Hsieh, C.-T. and M. Urquiola (2006): "The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program," *Journal of public Economics*, 90, 1477–1503.

Koedel, C., K. Mihaly, and J. E. Rockoff (2015): "Value-added modeling: A review," *Economics of Education Review.*

Lavy, V. (2002): "Evaluating the effect of teachers group performance incentives on pupil achievement," *Journal of political Economy*, 110, 1286–1317.

Levin, H. M. (1998): "Educational vouchers: Effectiveness, choice, and costs," *Journal of Policy Analysis and management*, 17, 373–392.

MacLeod, W. B. and M. S. Urquiola (2012): "Anti-lemons: school reputation, relative diversity, and educational quality," *IZA discussion paper.*

Milgrom, P. and C. Shannon (1994): "Monotone comparative statics," *Econometrica: Journal of the Econometric Society*, 157–180.

Sacerdote, B. (2000): "Peer effects with random assignment: Results for Dartmouth roommates," Tech. rep., National bureau of economic research.

——— (2011): "Peer effects in education: How might they work, how big are they and how much do we know thus far?" *Handbook of the Economics of Education*, 3, 249–277.

Tincani, M. (2014): "School Vouchers and the Joint Sorting of Students and Teachers," *HCEO Working Paper No. 2014-012.*

Todd, P. E. and K. I. Wolpin (2012): "Estimating a Coordination Game in the Classroom," Tech. rep.

Topkis, D. M. (1998): *Supermodularity and complementarity*, Princeton university press.