American Economic Review 101 (August 2011): 1739–1774
http://www.aeaweb.org/articles.php?doi=10.1257/aer.101.5.1739

Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya†

By Esther Duflo, Pascaline Dupas, and Michael Kremer*

To the extent that students benefit from high-achieving peers, tracking will help strong students and hurt weak ones. However, all students may benefit if tracking allows teachers to better tailor their instruction level. Lower-achieving pupils are particularly likely to benefit from tracking when teachers have incentives to teach to the top of the distribution. We propose a simple model nesting these effects and test its implications in a randomized tracking experiment conducted with 121 primary schools in Kenya. While the direct effect of high-achieving peers is positive, tracking benefited lower-achieving pupils indirectly by allowing teachers to teach to their level. (*JEL* I21, J45, O15)

To the extent that students benefit from having higher-achieving peers, tracking students into separate classes by prior achievement could disadvantage low-achieving students while benefiting high-achieving students, thereby exacerbating inequality (Denis Epple, Elizabeth Newlon, and Richard Romano 2002). On the other hand, tracking could potentially allow teachers to more closely match instruction to students' needs, benefiting all students. This suggests that the impact of tracking may depend on teachers' incentives. We build a model nesting these effects. In the model, students can potentially generate direct student-to-student spillovers as well as indirectly affect both the overall level of teacher effort and teachers' choice of the level at which to target instruction. Teacher choices depend on the distribution of students' test scores in the class as well as on whether the teacher's reward is a linear, concave, or convex function of test scores. The further away a student's own level is from what the teacher is teaching, the less the student benefits; if this distance is too great, she does not benefit at all.

* Duflo: MIT Economics Department, 50 Memorial Drive, Building E52 room 252G, Cambridge, MA 02142 (e-mail: eduflo@mit.edu); Dupas: UCLA Economics Department, 8283 Bunche Hall, Los Angeles, CA 90095, NBER, CEPR, and BREAD (e-mail: pdupas@econ.ucla.edu); Kremer: Harvard University Department of Economics, Littauer Center, 1805 Cambridge Street, Cambridge, MA 02138 (e-mail: mkremer@fas.harvard.edu). We thank Josh Angrist, Abhijit Banerjee, Michael Greenstone, Caroline Hoxby, Guido Imbens, Brian Jacob, and many seminar participants for helpful comments and discussions. We thank four anonymous referees for their suggestions. We thank the Kenya Ministry of Education, Science and Technology, International Child Support Africa, and Matthew Jukes for their collaboration. We thank Jessica Morgan, Frank Schilbach, Ian Tomb, Paul Wang, Nicolas Studer, and especially Willa Friedman for excellent research assistance. We are grateful to Grace Makana and her field team for collecting all the data. We thank, without implicating, the World Bank and the Government of the Netherlands for the grant that made this study possible.

† To view additional materials, visit the article page at http://www.aeaweb.org/articles.php?doi=10.1257/aer.101.5.1739.

We derive implications of this model, and test them using experimental data on tracking from Kenya. In 2005, 140 primary schools in western Kenya received funds to hire an extra grade one teacher. Of these schools, 121 had a single first-grade class, which they split into two sections, with one section taught by the new teacher. In 60 randomly selected schools, students were assigned to sections based on initial achievement. In the remaining 61 schools, students were randomly assigned to one of the two sections.

We find that tracking students by prior achievement raised scores for all students, even those assigned to lower achieving peers. On average, after 18 months, test scores were 0.14 standard deviations higher in tracking schools than in nontracking schools (0.18 standard deviations higher after controlling for baseline scores and other control variables). After controlling for the baseline scores, students in the top half of the preassignment distribution gained 0.19 standard deviations, and those in the bottom half gained 0.16 standard deviations. Students in all quantiles benefited from tracking. Furthermore, tracking had a persistent impact: one year after tracking ended, students in tracking schools scored 0.16 standard deviations higher (0.18 standard deviations higher with control variables). This first set of findings allows us to reject a special case of the model, in which all students benefit from higher-achieving peers but teacher behavior does not respond to class composition.

Our second finding is that students in the middle of the distribution gained as much from tracking as those at the bottom or the top. Furthermore, when we look within tracking schools using a regression discontinuity analysis, we cannot reject the hypothesis that there is no difference in endline achievement between the lowest scoring student assigned to the higher-achieving section and the highest scoring student assigned to the low-achievement section, despite the much higher achieving peers in the upper section.

These results are inconsistent with another special case of the model, in which teachers are equally rewarded for gains at all levels of the distribution, and so would choose to teach to the median of their classes. If this were the case, instruction would be less well-suited to the median student under tracking. Moreover, students just above the median would perform much better under tracking than those just below the median, for while they would be equally far away from the teacher's target teaching level, they would have the advantage of having higher-achieving peers.

In contrast, the results are consistent with the assumption that teachers' rewards are a convex function of test scores. With tracking, this leads teachers assigned to the lower-achievement section to teach closer to the median student's level than those assigned to the upper section, although teacher effort is higher in the upper section. In such a model, the median student may be better off under tracking and may potentially be better off in either the lower-achievement or higher-achievement section.

The assumption that rewards are a convex function of test scores is a good characterization of the education system in Kenya and in many developing countries. The Kenyan system is centralized, with a single national curriculum and national exams. To the extent that civil-service teachers face incentives, those incentives are based on the scores of their students on the national primary school exit exam given at the end of eighth grade. But since many students drop out before then, the teachers have incentives to focus on the students at the top of the distribution. While these incentives apply more weakly in earlier grades, they likely help maintain a culture in the educational system that is much more focused on the top of the distribution than in the United States. Moreover, teacher training is focused on the curriculum, and many students fall behind it. Indeed, Paul Glewwe, Kremer, and Sylvie Moulin (2009) show that textbooks based on the curriculum benefited only the initially higher-achieving students, suggesting that the exams and associated curriculum are not well suited to the typical student. It may also be the case that teachers find it easier and more personally rewarding to focus their teaching on strong students, in which case, in the absence of specific incentives otherwise, they will tend to do that.

The model also has implications for the effects of the test score distribution in nontracking schools. Specifically, it suggests that an upward shift of the distribution of peer achievement will strongly raise test scores for a student with initial achievement at the top of the distribution, have an ambiguous impact on scores for a student closer to the middle, and raise scores at the bottom. This is so because, while all students benefit from the direct effect of an increase in peer quality, the change in peer composition also generates an upward shift in the teacher's instruction level. The higher instruction level will benefit students at the top; hurt those students in the middle who find themselves further away from the instruction level; and leave the bottom students unaffected, since they are in any case too far from the target instruction level to benefit from instruction. Estimates exploiting the random assignment of students to sections in nontracking schools are consistent with these implications of the model.

While we do not have direct observations on the instruction level and how it varied across schools and across sections in our experiment, we present some corroborative evidence that teacher behavior was affected by tracking. First, teachers were more likely to be in class and teaching in tracking schools, particularly in the high-achievement sections, a finding consistent with the model's predictions. Second, students in the lower half of the initial distribution gained comparatively more from tracking in the most basic skills, while students in the top half of the initial distribution gained more from tracking in the somewhat more advanced skills. This finding is consistent with the hypothesis that teachers are tailoring instruction to class composition, although this could also be mechanically true in any successful intervention.

Rigorous evidence on the effect of tracking on learning of students at various points of the prior achievement distribution is limited, and much of it comes from studies of tracking in the United States, a context that may have limited applicability for education systems in developing countries. Reviewing the early literature, Julian R. Betts and Jamie L. Shkolnik (2000) conclude that while there is an emerging consensus that high-achievement students do better in tracking schools than in nontracking schools and that low-achievement students do worse, the consensus is based largely on invalid comparisons. When they compare similar students in tracking and nontracking high schools, Betts and Shkolnik (2000) conclude that low-achieving students are neither hurt nor helped by tracking; top students are helped; and there is some evidence that middle-scoring students may be hurt.

Another difficulty is that tracking schools may be different from nontracking schools. Jorn-Steffen Pischke and Alan Manning (2006) show that controlling for baseline scores is not sufficient to eliminate the selection bias when comparing students attending comprehensive versus selective schools in the United Kingdom. Three recent studies that tried to address the endogeneity of tracking decisions have found that tracking might be beneficial to students, or at least not detrimental, in the lower-achievement tracks. First, David N. Figlio and Marianne E. Page (2002) compare achievement gains across similar students attending tracking and nontracking schools in the United States. This strategy yields estimates that are very different from those obtained by comparing individuals schooled in different tracks. In particular, Figlio and Page (2002) find no evidence that tracking harms lower-achievement students. Second, Ron Zimmer (2003), also using US data, finds quasi-experimental evidence that the positive effects of achievement-specific instruction associated with tracking overcome the negative peer effects for students in lower-achievement tracks. Finally, Lars Lefgren (2004) finds that, in Chicago public schools, the difference between the achievement of low- and high-achieving students is no greater in schools that track that in schools that do not.

This paper is also related to a large literature that investigates peer effects in the classroom (e.g., Caroline Hoxby 2000; David J. Zimmerman 2003; Joshua D. Angrist and Kevin Lang 2004; see Epple and Romano (2011) for a recent review). While this literature has, mainly for data reasons, focused mostly on the direct effect of peers, there are a few exceptions, and these have results generally consistent with ours. Hoxby and Gretchen Weingarth (2006) use the frequent re-assignment of pupils to schools in Wake County to estimate models of peer effects, and find that students seem to benefit mainly from having homogeneous peers, which they attribute to indirect effects through teaching practices. Victor Lavy, M. Daniele Paserman, and Analia Schlosser (2008) find that the fraction of repeaters in a class has a negative effect on the scores of the other students, in part due to deterioration of the teacher's pedagogical practices. Finally, Damon Clark (2010) finds no impact on test scores of attending selective schools for marginal students who just qualified for the elite school on the basis of their score, suggesting that the level of teaching may be too high for them.

It is impossible to know if the results of this study will generalize until further studies are conducted in different contexts, but it seems likely that the general principle will hold: it will be difficult to assess the impact of tracking based solely on small random variations in peer composition that are unlikely to generate big changes in teacher behavior. Our model suggests that tracking may be particularly beneficial for low-achieving students when teachers' incentives are to focus on students who are above median achievement levels. Education systems are typically complex, having reward functions for schools and teachers that generate various threshold effects at different test score levels. But virtually all developing countries' teachers have incentives to focus on the strongest students. This suggests that our estimate of large positive impacts of tracking would be particularly likely to generalize to those contexts. This situation also seems to often be the norm in developed countries, with a few exceptions, such as the No Child Left Behind program in the United States.

The remainder of this paper proceeds as follows: Section I presents a model nesting various mechanisms through which tracking could affect learning. Section II provides background on the Kenyan education system and describes the study design, data, and estimation strategy. Section III presents the main results on test scores. Section IV presents additional evidence on the impact of tracking on teacher behavior. Section V concludes and discusses policy implications.

I. Model

We consider a model that nests several channels through which tracking students into two streams (a lower track and an upper track) could affect students' outcomes. In particular, the model allows peers to generate both direct student-to-student spillovers as well as to indirectly affect both the overall level of teacher effort and teachers' choice of the level at which to target instruction.[1] However, the model also allows for either of these channels to be shut off. Within the subset of cases in which the teacher behavior matters, we will consider the case in which teachers' payoffs are convex, linear, or concave in student test scores.

Suppose that educational outcomes for student i in class j, y_{ij}, are given by:

$$y_{ij} = x_{ij} + f(\overline{x}_{-ij}) + g(e_j)h(x_j^* - x_{ij}) + u_{ij}, \tag{1}$$

where x_{ij} is the student's pretest score, $\overline{x}_{-ij}$ is the average score of other students in the class, e_j is teacher effort, x_j^* is the target level to which the teacher orients instruction, and u_{ij} represents other i.i.d. stochastic student and class-specific factors that are symmetric and single peaked. In this equation, $f(\overline{x}_{-ij})$ reflects the direct effect of a student's peers on learning, e.g., through peer-to-peer interactions. For simplicity of exposition, in what follows we remove the class indices.

We will focus on the case when h is a decreasing function of the absolute value of the difference between the student's initial score and the target teaching level and is zero when $x_i - x^* > \theta$, although we also consider the possibility that h is a constant, shutting down this part of the model. Furthermore, we assume that $g(\cdot)$ is increasing and concave.

The teacher chooses x^* and e^* to maximize a payoff function P of the distribution of children's endline achievement minus the cost of effort $c(e)$ where $c(\cdot)$ is a convex function. We assume that the marginal cost to teachers of increasing effort eventually becomes arbitrarily high as teacher effort approaches some level $\overline{e}$. We will also consider the case in which the cost of effort is zero below $\overline{e}$, so teachers always choose effort $\overline{e}$, and this part of the model shuts down. There are two kinds of teachers: civil servants, and contract teachers hired to teach the new sections in the Extra-Teacher Program. Contract teachers have higher-powered incentives than civil servants and, as shown in Duflo, Dupas, and Kremer (2010), put in considerably more effort. In particular, we will assume that the reward to contract teachers from any increment in test scores equals λ times the reward to civil service teachers from the same increment in test scores, where λ is considerably greater than 1.

The choice of x^* will depend on the distribution of pretest scores.[2] We assume that within each school the distribution of initial test scores is continuous, strictly

[1] Epple, Newlon, and Romano (2002) consider the equilibrium implications of tracking in public schools in a model where the indirect effect of peer through teacher effort is shut off, but private schools can chose whether or not to track, and students can choose which school to attend.

[2] We rule out the possibility that teachers divide their time between teaching different parts of the class. In this case, tracking could reduce the number of levels at which a teacher would need to teach and thus increase the proportion of time students benefited from instruction. We also rule out fixed costs in adjusting the focus teaching level x^*. If teachers face such fixed costs, they will optimally use some type of Ss adjustment rule for x^*. In this case, teachers will be more likely to change x^* in response to large changes in the composition of student body associated with tracking than in response to small changes associated with random fluctuations in class composition. As

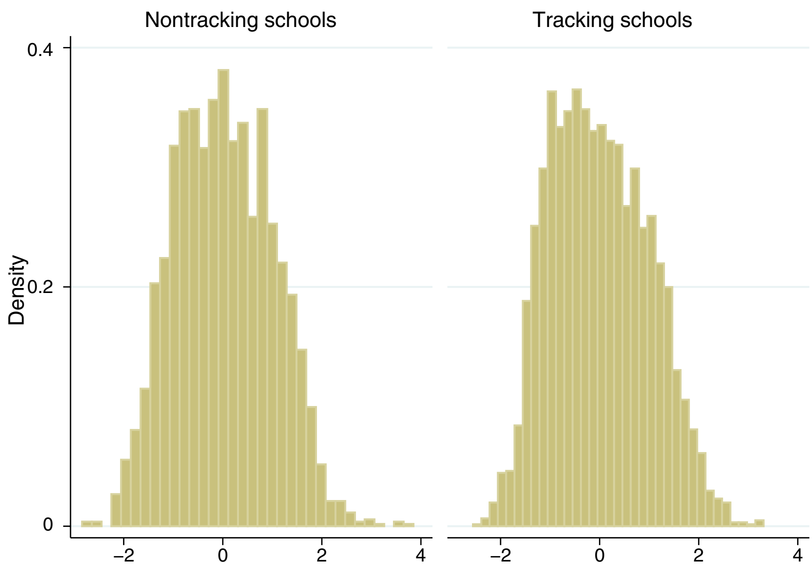


Figure 1. Distribution of Initial Test Scores (*all schools*)

quasi-concave, and symmetric around the median. This appears to be consistent with our data (see Figure 1).

With convexity of teachers' payoffs in both student test scores and teacher effort in general, there could be multiple local maxima for teachers' choice of effort and x^*. Nonetheless, it is possible to characterize the solution, at least under certain conditions. Our first proposition states a testable implication of the special case where peers only affect each other directly.

PROPOSITION 1: *Consider a special case of the model in which teachers do not respond to class composition because $h(\,)$ is a constant and either $g(\,)$ is a constant or the cost of effort is zero below $\overline{e}$. In that case, tracking will reduce test scores for those below the median of the original distribution and increase test scores for those above the median.*

PROOF:

Under tracking, average peer achievement is as high as possible (and, hence, higher than without tracking) for students above the median and as low as possible (and, hence, lower than without tracking) for students below the median.

Note that this proposition would be true even with a more general equation for test scores that allowed for interactions between students' own test scores and those of their peers, as long as students always benefit from higher-achieving peers.

discussed below, we think the evidence is consistent with the hypothesis that some teachers change their teaching techniques even in response to random fluctuations in class composition. Fixed costs of changing x^* may not be that great because this change may simply mean proceeding through the same material more slowly or more quickly.

Proposition 2 describes the optimal choice of x^* for a teacher as a function of the shape of the reward function.

PROPOSITION 2: *If teacher payoffs*, P, *are convex in posttest scores*, *in a non-tracked class the target teaching level*, x^*, *must be above the median of the distribution. If teacher payoffs are linear in posttest scores*, *then* x^* *will be equal to the median of the distribution. If teacher payoffs are concave in posttest scores*, *then* x^* *will be below the median of the distribution.*

PROOF:

In online Appendix.

Building on these results, Proposition 3 considers the overall effect of a uniform shift in the distribution of peer baseline test scores on one's test score, depending on one's place in the initial distribution. Such a shift has two effects: a direct effect (through the quality of the peer group) and an indirect effect (through the teacher effort and choice of x^*). The overall effect depends on $f(\ )$, $h(\ )$, and on the student's initial level.

PROPOSITION 3:

- *If* $f(\ )$ *is increasing in peer test scores, then a uniform marginal increase in peer baseline achievement: will raise test scores for any student with initial score* $x_i > x^*$, *and the effect will be the largest for students with* $x^* < x_i < x^* + \theta$; *will have an ambiguous effect on test scores for students with initial scores* $x^* - \theta \leq x_i \leq x^*$; *and will increase test scores for students with initial scores below* $x^* - \theta$, *although the increase will be smaller than that for students with initial scores greater than* x^*.
- *If* $f(\ )$ *is a constant, so there is no direct effect of peers, then a uniform increase in peer achievement will cause students with* $x_i > x^*$ *to have higher test scores and those with* $x^* - \theta \leq x_i \leq x^*$ *to have lower scores. There will be no change in scores for those with* $x_i < x^* - \theta$.

PROOF:

In online Appendix.

PROPOSITION 4: *Let* x_L^* *denote the target teaching level in the lower section in a tracking school and* x_U^* *denote the target level in the upper section. If payoffs are convex*, x_L^* *will be within distance* θ *of* x_m, *where* x_m *denotes the median of the original distribution. If payoffs are concave*, x_U^* *will be within distance* θ *of* x_m. *If payoffs are linear*, *both* x_U^* *and* x_L^* *will be within distance* θ *of* x_m.

PROOF:

To see this for the convex case, suppose that $x_L^* < x_m - \theta$. Marginally increasing x^* would both increase the number of students at any distance from x^* and the base score x_i of students at any distance from x^*. Thus it would be preferred. Proofs for the other cases are analogous.

PROPOSITION 5: *Denote the distance between x_m and the target teaching level in the upper section x_U^* as D_U and denote the corresponding distance between x_L^* and x_m as D_L. If payoffs are convex and the third derivative is nonnegative, then $D_U > D_L$, so the median student is closer to the target teaching level in the lower track. If payoffs are linear in student scores then $D_U = D_L$. If teacher payoffs are concave in student test scores and the third derivative is nonpositive, then $D_U < D_L$.*

PROOF:

In online Appendix.

This proposition implies that, with a convex reward function, the median student in a tracking school will be closer to the target teaching level if assigned to the lower track than if assigned to the upper track.

PROPOSITION 6: *Teacher effort will be greater in the upper than in the lower section under convexity, equal under linearity, and lesser under concavity. However, for high enough λ, the difference between effort levels of contract teachers assigned to the high- and low-achievement sections will become arbitrarily small.*

PROOF:

In online Appendix.

PROPOSITION 7: *Under a linear teacher payoff function, a student initially at the median of the distribution will score higher if assigned to the upper section than the lower section under tracking.*

PROOF:

Under linear teacher payoffs, a student at the median will experience equal teacher effort in the upper and lower sections, and will be equally far from the target teaching level. However, the student will have stronger peers in the top section.

Note that under convex teacher payoffs, the student at the median will experience higher teacher effort in the top section (compared to the bottom section) and will have stronger peers but will have teaching which is not as good a match for his or her initial achievement. The model therefore offers no definitive prediction on whether the median student performs better in the upper or lower track. The effect of the level of instruction could offset the two positive effects if the teacher payoffs are sufficiently convex, or if the $h(\ )$ function is declining quickly from its peak.

Similarly, if teacher payoffs are concave in student test scores, then the median student would have a more appropriate teaching target level, and better peers, but lower teacher effort in the top section. Once again, there is no clear prediction.

The model, however, has a more definite prediction on the effect of the interaction between the teacher type (contract teacher versus civil-service teacher) and the assignment of the median student.

PROPOSITION 8: *For high enough λ, if teacher payoffs are convex in students' test scores, the gap in endline test scores between a median student assigned to the* *upper track and his counterfactual assigned to the lower track is larger when teachers are civil servants than when teachers are contract teachers. The converse is true if payoffs are concave.*

PROOF:

For high enough λ, the difference in effort levels of contract teachers assigned to the top and bottom tracks becomes arbitrarily small. Hence, with a contract teacher, in the convex case, students assigned to the bottom section do not suffer much from reduced teacher effort, relative to students assigned to the top section. Similarly, with a concave payoff function, students assigned to the bottom class do not benefit much from increased effort with contract teachers.

This model nests, as special cases, models with only a direct effect of peers or only an effect going through teacher behavior. It also nests special cases in which teacher payoffs are linear, concave, or convex in students' test scores. Nevertheless, the model makes some restrictive assumptions. In particular, teacher effort has the same impact on student test score gains anywhere in the distribution. In a richer model, teacher effort might have a different impact on test scores at different places along the distribution. Student effort might also respond endogenously to teacher effort and the target teaching level. In such a model, ultimate outcomes will be a composite function of teacher effort, teacher focus level, and student effort, which in turn would be a function of teacher effort and teaching level. In this case, we conjecture that the results would go through as long as the curvature assumptions on the payoff function were replaced by curvature assumptions on the resulting composite function for payoffs. However, multiplicative separability of e and x^* is important for the results.

Propositions 1, 2, and 4 provide empirical implications that can be used to test whether the data are consistent with the different special cases.

Below we argue that the data are inconsistent with the special case with no teacher response, the special case with no direct effects of peers, and the special case in which teacher payoffs are linear or concave in students' scores. However, our results are consistent with a model in which both direct and indirect effects operate and teachers' payoffs are convex with student test scores, which is consistent with our description of the education system in Kenya.

Note that this model has no clear prediction for the effect of the variance of initial achievement on test scores in an untracked class or for the interaction between the effect of tracking and the initial variance of the distribution.[3]

[3] To see that changes in the distribution of initial scores that increase variance of these scores could reduce average test scores and the effect of tracking, consider an increase in dispersion so no two students are within distance θ of each other. Then teachers can never teach more than one pupil. Average test scores will be low, and tracking will not enable teachers to target the instructional level so as to reach more pupils. To see that changes in the distribution that increase variance could increase the impact of tracking, consider moving from a degenerate distribution, with all the mass concentrated at a single point, to a distribution with some dispersion. Tracking will have no effect on test scores with a degenerate distribution, but will increase average scores with tracking. Increases in dispersion could also increase average test scores in the absence of tracking. To see this, suppose teacher payoffs are very convex, so teachers focus on the strongest student in the class. Suppose also that the highest achieving student's initial score exceeds that of the second highest-scoring pupil by more than θ. Consider a move from the initial distribution to a distribution with the same support, but in which some students were pushed to the boundaries of this support. More students will be within range of the teacher, and, hence, teacher effort and average test scores will rise.

II. The Tracking Experiment: Background, Experimental Design, Data, and Estimation Strategy

A. *Background: Primary Education in Kenya*

Like many other countries, Kenya has a centralized education system with a single national curriculum and national exams. Glewwe, Kremer, and Moulin (2009) show that textbooks based on the curriculum benefited only the initially higher-achieving students, suggesting that the exams and associated curriculum are not well suited to the typical student.

Most primary school teachers are hired centrally through the civil service, and they face weak incentives. As we show in Section V, absence rates among civil-service teachers are high. In addition, some teachers are hired on short-term contracts by local school committees, most of whose members are elected by parents. These contract teachers typically have much stronger incentives, partly because they do not have civil-service and union protection but also because a good track record as a contract teacher can help them obtain a civil-service job.

To the extent that schools and teachers face incentives, the incentives are largely based on their students' scores on the primary school exit exam. Many students repeat grades or drop out before they can take the exam, and so the teachers have limited incentives to focus on students who are not likely to ever take the exam. Extrinsic incentives are thus stronger at the top of the distribution than the bottom. For many teachers, the intrinsic rewards of teaching to the top of the class are also likely to be greater than those of teaching to the bottom of the class, as such students are more similar to themselves and teachers are likely to interact more with their families and with the students themselves in the future.

Until recently, families had to pay for primary school. Students from the poorest families often had trouble attending school and dropped out early. But recently Kenya has, like several other countries, abolished school fees. This led to a large enrollment increase and to greater heterogeneity in student preparation. Many of the new students are first generation learners and have not attended preschools (which are neither free nor compulsory). Students thus differ vastly in age, school preparedness, and support at home.

B. *Experimental Design*

This study was conducted within the context of a primary school class-size reduction experiment in Western Province, Kenya. Under the ETP, with funding from the World Bank, ICS Africa provided 140 schools with funds to hire an additional first grade teacher on a contractual basis starting in May 2005, the beginning of the second term of that school year.[4] The program was designed to allow schools to add an additional section in first grade. Most schools (121) had only one first grade section and split it into two sections. Schools that already had two or more first grade

[4] The school year in Kenya starts in January and ends in November. It is divided into three terms, with month-long breaks in April and August.

sections added one section. Duflo, Dupas, and Kremer (2010) report on the effect of the class size reduction and teacher contracts.

We examine the impact of tracking and peer effects using two different versions of the ETP experiment. In 61 schools randomly selected (using a random number generator) from the 121 schools that originally had only one grade 1 section, grade 1 pupils were randomly assigned to one of two sections. We call these schools the "nontracking schools." In the remaining 60 schools (the "tracking schools"), children were assigned to sections based on scores on exams administered by the school during the first term of the 2005 school year. In the tracking schools, students in the lower half of the distribution of baseline exam scores were assigned to one section, and those in the upper half were assigned to another section. The 19 schools that originally had two or more grade 1 classes were also randomly divided into tracking and nontracking schools, but it proved difficult to organize the tracking consistently in these schools.[5] Thus, in the analysis that follows, we focus on the 121 schools that initially had a single grade 1 section and exclude 19 schools (ten tracking, nine nontracking schools) that initially had two or more.[6]

After students were assigned to sections, the contract teacher and the civil-service teacher were randomly assigned to sections. Parents could request that their children be reassigned, but this occurred in only a handful of cases. The main source of noncompliance with the initial assignment was teacher absenteeism, which sometimes led the two grade 1 sections to be combined. On average across five unannounced school visits to each school, we found the two sections combined 14.4 percent of the time in nontracking schools and 9.7 percent of time in tracking schools (note that the likelihood that sections are combined depends on teacher effort, itself an endogenous outcome, as we show below in Section 5). When sections were not combined, 92 percent of students in nontracking schools and 96 percent of students in tracking schools were found in their assigned section. The analysis below is based on the initial assignment regardless of which section the student eventually joined.

The program lasted for 18 months, which included the last two terms of 2005 and the entire 2006 school year. In the second year of the program, all children not repeating the grade remained assigned to the same group of peers and the same teacher. The fraction of students who repeated grade 1 and thus participated in the program for only the first year was 23 percent in nontracking schools and 21 percent in tracking schools (the p-value of the difference is 0.17).[7]

Table 1 presents summary statistics for the 121 schools in our sample. As would be expected given the random assignment, tracking and nontracking schools look very similar. Since tests administered within schools prior to the program are not

[5] In these schools, the sections that were taught by civil service teachers rather than contract teachers sometimes recombined or exchanged students.

[6] Note that the randomization of schools into tracking and nontracking was stratified according to whether the school originally had one or more grade 1 section.

[7] Students enrolled in grade 2 in 2005 and who repeated grade 2 in 2006 were randomly assigned to either the contract teacher or the civil-service teacher in 2006. All the analysis is based on the initial assignment, so they are excluded from the study and excluded from the measures of peer composition at endline. Students who repeated grade 1 in 2006 remain in the dataset and are included in the measures of peer composition at endline. New pupils who joined the school after the introduction of the program were assigned to a class on a random basis. However, since the decision for these children to enroll in a treatment or control school might be endogenous, they are excluded from the analysis. The number of newcomers was balanced across school types (tracking and nontracking) at six per school on average.

Table 1—School and Class Characteristics, by Treatment Group, Pre– and Post–Program Start

	All ETP schools				
	Nontracking schools		Tracking schools		*p*-value tracking = nontracking
	Mean	SD	Mean	SD	
Panel A. Baseline school characteristics					
Total enrollment in 2004	589	*232*	549	*198*	0.316
Number of government teachers in 2004	11.6	*3.3*	11.9	*2.8*	0.622
School pupil/teacher ratio	37.1	*12.2*	35.9	*10.1*	0.557
Performance in national exam in 2004 (out of 400)	255.6	*23.6*	258.1	*23.4*	0.569
Panel B. Class size prior to program inception (March 2005)					
Average class size in first grade	91	*37*	89	*33*	0.764
Proportion of female first grade students	0.49	*0.06*	0.49	*0.05*	0.539
Average class size in second grade	96	*41*	91	*35*	0.402
Panel C. Class size six months after program inception (October 2005)					
Average class size in first grade	44	*18*	42	*15*	0.503
Range of class sizes in sample (first grade)	19–98		20–97		
Panel D. Class size in year two of program (March 2006)					
Average class size in second grade	42	*17*	42	*20*	0.866
Range of class sizes in sample (second grade)	18–93		21–95		
Number of schools	61		60		121
	Within tracking schools				
	Assigned to bottom section		Assigned to top section		*p*-value top = bottom
	Mean	SD	Mean	SD	
Panel E. Comparability of two sections within tracking schools					
Proportion female	0.49	*0.09*	0.50	*0.08*	0.38
Average age at endline	9.04	*0.59*	9.41	*0.60*	0.00
Average standardized baseline score (mean 0, SD 1 at school level)	−0.81	*0.04*	0.81	*0.04*	0.00
Average SD within section in standardized baseline scores	0.49	*0.13*	0.65	*0.13*	0.00
Average standardized endline score (mean 0, SD 1 in nontracking group)	−0.15	*0.44*	0.69	*0.58*	0.00
Average SD within section in standardized endline scores	0.77	*0.23*	0.88	*0.20*	0.00
Assigned to contract teacher	0.53	*0.49*	0.46	*0.47*	0.44
Respected assignment	0.99	*0.02*	0.99	*0.02*	0.67
	Within nontracking schools				
	Section A (assigned to civil-service teacher)		Section B (assigned to contract teacher)		*p*-value *A* = *B*
Panel F. Comparability of two sections within nontracking schools					
Proportion female	0.49	*0.06*	0.49	*0.06*	0.89
Average age at endline	9.07	*0.53*	9.00	*0.45*	0.45
Average standardized baseline score (mean 0, SD 1 at school level)	0.003	*0.10*	0.002	*0.11*	0.94
Average SD within section in standardized baseline scores	1.005	*0.08*	0.993	*0.08*	0.43
Average standardized endline score (mean 0, SD 1 in nontracking group)	0.188	*0.46*	0.047	*0.48*	0.10
Average SD within section in standardized endline scores	0.937	*0.24*	0.877	*0.24*	0.16

Note: School averages. *p*-values are for tests of equality of the means across groups.

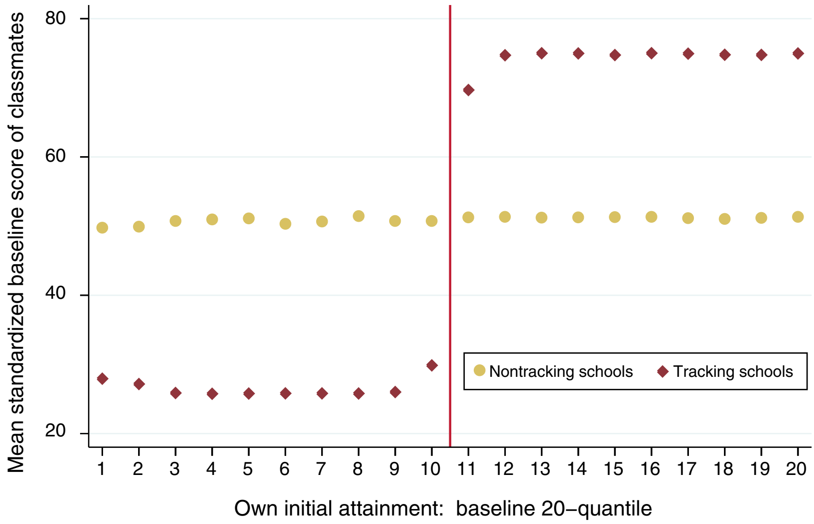


FIGURE 2. EXPERIMENTAL VARIATION IN PEER COMPETITION

Note: Each dot corresponds to the average peer quality across all students in a given 20-quantile, for a given treatment group.

comparable across schools, they are normalized such that the mean score in each school is zero and the standard deviation is one. Figure 2 shows the average baseline score of a student's classmates as a function of the student's own baseline score in tracking and nontracking schools. Average nonnormalized peer test scores are not correlated with the student's own test score in nontracking schools but, consistent with the discontinuous assignment at the fiftieth percentile for most schools, there is sharp discontinuity at the fiftieth percentile in tracking schools.[8] The baseline exams are a good measure of academic achievement, in that they are strongly predictive of the endline test we administered, with a correlation of 0.47 in the nontracking schools and 0.49 in tracking schools. In tracking schools, the top section has somewhat more girls, and students are 0.4 years older.

C. *Data*

The sample frame consists of approximately 10,000 students enrolled in first grade in March 2005. The key outcome of interest is student academic achievement, as measured by scores on a standardized math and language test first administered in all schools 18 months after the start of the program. Trained enumerators administered the test, which was then graded blindly by enumerators. In each school, 60 students (30 per section) were drawn from the initial sample to participate in the

[8] Peer quality is slightly more similar for children below and above the fiftieth percentile than for students at other percentiles because the assignment procedure used a manually computed ranking variable that was very strongly correlated with the ranking based on the actual school grades but had a few discrepancies (due to clerical errors). Thus, some children close to the median who should have been assigned to one section wound up in the other one. We are using the rank based on the actual school grade as our control variable in what follows, in case the ranking variable that was used for assignment was in fact manipulated.

tests. If a section had more than 30 students, students were randomly sampled (using a random number generated before enumerators visited the school) after stratifying by their position in the initial distribution. Part of the test was designed by a cognitive psychologist to measure a range of skills students might have mastered at the end of grade 2. Part of the test was written, and part was orally administered one-to-one by trained enumerators. Students answered math and literacy questions ranging from identifying letters and counting to subtracting three-digit numbers and reading and understanding sentences.

To limit attrition, enumerators were instructed to go to the homes of sampled students who had dropped out or were absent on the day of the test, and to bring them to school for the test. It was not always possible to find those children, however, and the attrition rate on the test was 18 percent. There was no difference between tracking and nontracking schools in overall attrition rates. The characteristics of those who attrited are similar across groups, except that girls in tracking schools were less likely to attrit in the endline test (see Appendix Table A1). Transfer rates to other schools were similar in tracking and nontracking schools. In total, we have endline test score data for 5,795 students.

To measure whether program effects persisted, children sampled for the endline were tested again in November 2007, one year after the program ended. During the 2007 school year, students were overwhelmingly enrolled in grades for which their school had a single section, so tracking was no longer an option. Most students had reached grade 3, but repeaters were also tested. The attrition for this longer-term follow-up was 22 percent, only 4 points higher than attrition at the endline test. The proportion of attritors and their characteristics do not differ between the two treatment arms (Appendix Table A1).

We also collected data on grade progression and dropout rates, and student and teacher absence. Overall, the dropout rate among grade 1 students in our sample was low (below 0.5 percent). Several times during the course of the study, enumerators went to the schools unannounced and checked, upon arrival, whether teachers were present in school and whether they were in class and teaching. On those visits, enumerators also took a roll call of the students.

D. *Empirical Strategy*

Measuring the Impact of Tracking.—To measure the overall impact of tracking on test scores, we run regressions of the form:

$$(E1) \qquad y_{ij} = \alpha T_j + X_{ij}\beta + \epsilon_{ij},$$

where y_{ij} is the endline test score of student i in school j (expressed in standard deviations of the distribution of scores in the nontracking schools),[9] T_j is a dummy equal to 1 if school j was tracking, and X_{ij} is a vector including a constant and child and

[9] We have also experimented with an alternative specification of the endline test score for math, which uses item response theory to give different weights to questions of different levels of difficulty (the format of the language test was not appropriate for this exercise). The results were extremely similar (results available from the authors), so we focus on the standardized test scores in this version.

school control variables (we estimate a specification without control variables and a specification that controls for baseline score, whether the child was in the bottom half of the distribution in the school, gender, age, and whether the section is taught by a contract or civil-service teacher).

To identify potential differential effects for children assigned to the lower and upper section, we also run:

$$(E2) \qquad y_{ij} = \alpha T_j + \gamma T_j \times \beta_{ij} + X_{ij}\beta + \epsilon_{ij},$$

where B_{ij} is a dummy variable that indicates whether the child was in the bottom half of the baseline score distribution in her school (B_{ij} is also included in X_{ij}). We also estimate a specification where treatment is interacted with the initial quartile of the child in the baseline distribution. Finally, to investigate flexibly whether the effects of tracking are different at different levels of the initial test score distribution, we run two separate nonparametric regressions of endline test scores on baseline test scores in tracking and nontracking schools and plot the results.

To understand better how tracking works, we also run similar regressions using as dependent variable a more disaggregated version of the test scores: the test scores in math and language, and the scores on specific skills. Finally, we also run regressions of a similar form, using as outcome variable teacher presence in school, whether the teacher is in class teaching, and student presence in school.

Nontracking Schools.—Since children were randomly assigned to a section in these schools, their peer group is randomly assigned, and there is some naturally occurring variation in the composition of the groups.[10] In the sample of nontracking schools, we start by estimating the effect of a student's peer average baseline test scores by OLS (this is the average of the section excluding the student him or herself):

$$(E3) \qquad y_{ij} = \kappa \overline{x}_{-ij} + X_{ij}\beta + \nu_j + \epsilon_{ij},$$

where $\overline{x}_{-ij}$ is the average peer baseline test score in the section to which a student was assigned.[11] The vector of control variables X_{ij} includes the student's own baseline score x_{ij}. Since students were randomly assigned within schools, our estimate of the coefficient of $\overline{x}_{-ij}$ in a specification including school fixed effects will reflect the causal effect of peers' prior achievement (both direct through peer-to-peer learning, and indirect through adjustment in teacher behavior to the extent to which teachers change behavior in response to small random variations in class composition). Although our model has no specific prediction on the impact of the variance, we also include the variance of the peers' test scores, as an independent variable in one specification.

[10] On average across schools, the difference in baseline scores between the two sections is 0.17 standard deviations, with a standard deviation of 0.13. The 25th–75th percentiles interval for the difference is [0.07 – 0.24].

[11] There were very few reassignments, but we always focus on the initial random assignment: that is, we consider the test scores of the other students *initially assigned* to the class to which a student was *initially assigned* (regardless of whether they eventually attended that class).

The baseline grades are not comparable across schools (they are the grades assigned by the teachers in each school). However, baseline grades are strongly correlated with endline test scores, which are comparable across schools. Thus, to facilitate comparison with the literature and with the regression discontinuity estimates for the tracking schools, we estimate the impact of average endline peer test scores on a child's test score:

$$(E4) \qquad y_{ij} = \kappa \overline{y}_{-ij} + X_{ij}\beta + \nu_j + \epsilon_{ij}.$$

This equation is estimated by instrumental variables, using $\overline{x}_{-ij}$ as an instrument for $\overline{y}_{-ij}$.

Measuring the Impact of Assignment to Lower or Upper Section.—Tracking schools provide a natural setup for a regression discontinuity (RD) design to test whether students at the median are better off being assigned to the top section, as would be true in the special case of the model in which teacher payoffs were linear in test scores.

As shown in Figure 2, students on either side of the median were assigned to classes with very different average prior achievement of their classmates: the lower-scoring member was assigned to the bottom section, and the higher-scoring member was assigned to the top section. (When the class had an odd number of students, the median student was randomly assigned to one of the sections.)

Thus, we first estimate the following reduced form regression in tracking schools:

$$(E5) \qquad y_{ij} = \delta B_{ij} + \lambda_1 P_{ij} + \lambda_2 P_{ij}^2 + \lambda_3 P_{ij}^3 + X_{ij}\beta + \epsilon_{ij},$$

where P_{ij} is the percentile of the child on the baseline distribution in her school.

Since assignment was based on scores within each school, we also run the same specification, including school fixed effects:

$$(E6) \qquad y_{ij} = \delta B_{ij} + \lambda_1 P_{ij} + \lambda_2 P_{ij}^2 + \lambda_3 P_{ij}^3 + X_{ij}\beta + \epsilon_{ij} + \nu_j.$$

To test the robustness of our estimates to various specifications of the control function, we also run specifications similar to equations ($E5$) and ($E6$), estimating the polynomial separately on each side of the discontinuity, and report the difference in test scores across the discontinuity. Finally, we follow Guido W. Imbens and Thomas Lemieux (2008) and use a Fan locally weighted regression of the relationship between endline test scores and baseline percentile on both sides of the discontinuity.

Note that this is an unusually favorable setup for a regression discontinuity design. There are 60 different discontinuities in our dataset, rather than just one, as in most regression discontinuity applications, and the number of different discontinuities in principle grows with the number of schools.[12] We can therefore run a specification including only the pair of students straddling the median.

$$(E7) \qquad y_{ij} = \delta B_{ij} + X_{ij}\beta + \epsilon_{ij} + \nu_j.$$

[12] Dan A. Black, Jose Galdo, and Jeffrey A. Smith (2007) also exploit a series of sharp discontinuities in their estimation of a re-employment program across various sites in Kentucky.

Since the median will be at different achievement levels in different schools, results will be robust to sharp nonlinearities in the function linking pre- and post-test achievement.

These reduced form results are of independent interest, and they can also be combined with the impact of tracking on average peer test scores for instrumental variable estimation of the impact of average peer achievement for the median child in a tracking environment. Specifically, the first stage of this regression is:

$$\overline{y}_{-ij} = \pi B_{ij} + \phi_1 P_{ij} + \phi_2 P_{ij}^2 + \phi_3 P_{ij}^3 + X_{ij}\beta + \epsilon_{ij} + \nu_j,$$

where $\overline{y}_{-ij}$ is the average endline test scores of the classmates of student i in school j. The structural equation:

$$(E8) \qquad y_{ij} = \kappa \overline{y}_{-ij} + \lambda_1 P_{ij} + \lambda_2 P_{ij}^2 + \lambda_3 P_{ij}^3 + X_{ij}\beta + \nu_j + \epsilon_{ij}$$

is estimated using B_{ij} (whether a child was assigned to the bottom track) as an instrument for $\overline{y}_{-ij}$.

Note that this strategy will give an estimate of the effect of peer quality for the median child in a tracking environment, where having high-achieving peers on average also means that the child is the lowest-achieving child of his section (at least at baseline), and having low-achieving peers means that the child is the highest-achieving child of his track.

III. Results

In Section IIIA, we present reduced form estimates of the impact of tracking, showing that tracking increased test scores throughout the distribution and thus rejecting the special case of the model in which higher-achieving peers raise test scores directly but there is no indirect effect through changing teacher behavior. In Section IIIB, we use random variation in peer composition in nontracked schools to assess the implications of Proposition 3, and to argue that the data is not consistent with the special case of the model in which there are no direct effects of peers. In Section IIIC, we argue that the data are inconsistent with the special case of the model in which teacher incentives are linear in student test scores, because the median student in tracking schools scores similarly whether assigned to the upper or lower section. We conclude that the data is most consistent with a model in which peer composition affects students both directly and indirectly, through teacher behavior, and in which teachers face convex incentives. In this model, teachers teach to the top of the distribution in the absence of tracking, and teaching can improve learning for all children.

A. *The Impact of Tracking by Prior Achievement and the Indirect Impact of Peers on Teacher Behavior*

A striking result of this experiment is that tracking by initial achievement significantly increased test scores throughout the distribution.

Table 2 presents the main results on the impacts of tracking. At the endline test, after 18 months of treatment, students in tracking schools scored 0.139 standard deviations (with a standard error of 0.078 standard deviations) higher than students in nontracking schools overall (column 1, panel A, Table 2). The estimated effect is somewhat larger (0.176 standard deviations, with a standard error of 0.077 standard deviations) when controlling for individual-level covariates (column 2). Both sets of students, those assigned to the upper track and those assigned to the lower track, benefited from tracking (in row 2, column 3, panel A, the interaction between being in the bottom half and in a tracking school cannot be distinguished from zero, and the total effect for the bottom half is 0.156 standard deviations, with a p-value of 0.04). When we look at each quartile of the initial distribution separately, we find positive point estimates for all quartiles (column 4).

Figure 3 provides graphical evidence suggesting that all students benefited from tracking. As in David S. Lee (2008), it plots a student's endline test score as a function of the baseline test score using a second-order polynomial estimated separately on either side of the cutoff in both the tracking and nontracking schools. The fitted values in tracking schools are systematically above those for nontracking schools, suggesting that tracking increases test scores regardless of the child's initial test score in the distribution of test scores.

Overall, the estimated effect of tracking is relatively large. It is similar in magnitude to the effect of being assigned to a contract teacher (shown in row 6 of Table 2), who, as we will show in Table 6, exerted much higher levels of effort than civil-service teachers. It is also interesting to contrast the effect of tracking with that of a more commonly proposed reform, class size reduction. In other contexts, studies have found a positive and significant effect of class size reduction on test scores (Angrist and Lavy 1999; Alan Krueger and Diane Whitmore 2002). In Duflo, Dupas, and Kremer (2010), however, we find that in the same exact context, class size reduction per se (without a change in teachers' incentives) generates an increase in test scores of 0.09 standard deviation after 18 months (though insignificant), but the effect completely disappears within one year after the class size reduction stops.

The effect of tracking persisted beyond the duration of the program. When the program ended after 18 months, three quarters of students had then reached grade 3, and in all schools except five, there was only one class for grade 3. The remaining students had repeated and were in grade 2 where, once again, most schools had only one section (since after the end of the program they did not have funds for additional teachers). Thus, after the program ended, students in our sample were not tracked any more (and they were in larger classes than those both tracked and nontracked students had experienced in grades 1 and 2). Yet, one year later, test scores of students in tracking schools were still 0.163 standard deviations greater (with a standard error of 0.069 standard deviations) than those of students in nontracking schools overall (column 1, panel B, Table 2). The effect is slightly larger (0.178 standard deviations) and more significant with control variables (column 2, panel B), and the gains persist both for initially high and low achieving children. A year after the end of the program, the effect for the bottom half is still large (0.135 standard deviations, with a p-value of 0.09), although the effect for students in the bottom quartile is insignificant (panel B, column 4).

Table 2—Overall Effect of Tracking

	Total score				Math score		Literacy score	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Short-run effects (after 18 months in program)								
(1) Tracking school	0.139	0.176	0.192	0.182	0.139	0.156	0.198	0.166
	(0.078)*	(0.077)**	(0.093)**	(0.093)*	(0.073)*	(0.083)*	(0.108)*	(0.098)*
(2) In bottom half of initial distribution × tracking school			−0.036		0.04		−0.091	
			(0.07)		(0.07)		(0.08)	
(3) In bottom quarter × tracking school				−0.045		0.012		−0.083
				(0.08)		(0.09)		(0.08)
(4) In second-to-bottom quarter × tracking school				−0.013		0.026		−0.042
				(0.07)		(0.08)		(0.07)
(5) In top quarter × tracking school				0.027		−0.026		0.065
				(0.08)		(0.07)		(0.08)
(6) Assigned to contract teacher		0.181	0.18	0.18	0.16	0.161	0.16	0.16
		(0.038)***	(0.038)***	(0.038)***	(0.038)***	(0.037)***	(0.038)***	(0.038)***
Individual controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,795	5,279	5,279	5,279	5,280	5,280	5,280	5,280
Total effects on bottom half and bottom quarter								
Coeff (Row 1) + Coeff (Row 2)			0.156		0.179		0.107	
Coeff (Row 1) + Coeff (Row 3)				0.137		0.168		0.083
F-test: total effect = 0			4.40	2.843	5.97	3.949	2.37	1.411
p-value (total effect for bottom = 0)			0.038	0.095	0.016	0.049	0.127	0.237
p-value (effect for top quarter = effect for bottom quarter)				0.507		0.701		0.209
Panel B. Longer-run effects (a year after program ended)								
(1) Tracking school	0.163	0.178	0.216	0.235	0.143	0.168	0.231	0.241
	(0.069)**	(0.073)**	(0.079)***	(0.088)***	(0.064)**	(0.075)**	(0.089)**	(0.096)**
(2) In bottom half of initial distribution × tracking school			−0.081		−0.027		−0.106	
			(0.06)		(0.06)		(0.06)	
(3) In bottom quarter × tracking school				−0.117		−0.042		−0.152
				(0.09)		(0.10)		(0.085)*
(4) In second-to-bottom quarter × tracking school				−0.096		−0.073		−0.091
				(0.07)		(0.07)		(0.07)
(5) In top quarter × tracking school				−0.028		−0.04		−0.011
				(0.07)		(0.06)		(0.08)
(6) Assigned to contract teacher		0.094	0.094	0.094	0.061	0.061	0.102	0.103
		(0.032)***	(0.032)***	(0.032)***	(0.031)**	(0.031)**	(0.031)***	(0.031)***
Individual controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,490	5,001	5,001	5,001	5,001	5,001	5,007	5,007
Total effects on bottom half and bottom quarter								
Coeff (Row 1) + Coeff (Row 2)			0.135		0.116		0.125	
Coeff (Row 1) + Coeff (Row 3)				0.118		0.126		0.089
p-value (total effect for bottom = 0)			0.091	0.229	0.122	0.216	0.117	0.319
p-value (effect for top quarter = effect for bottom quarter)				0.365		0.985		0.141

Notes: The sample includes 60 tracking and 61 nontracking schools. The dependent variables are normalized test scores, with mean 0 and standard deviation 1 in the nontracking schools. Robust standard errors clustered at the school level are presented in parentheses. Individual controls included: age, gender, being assigned to the contract teacher, dummies for initial half/quarter, and initial attainment percentile. We lose observations when adding individual controls because information on the initial attainment could not be collected in some of the nontracking schools.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

This overall persistence is striking, since in many evaluations, the test score effects of even successful interventions tend to fade over time (e.g., Abhijit V. Banerjee et al. 2007; Tahir Andrabi et al. 2008). This indicates that tracking may have helped

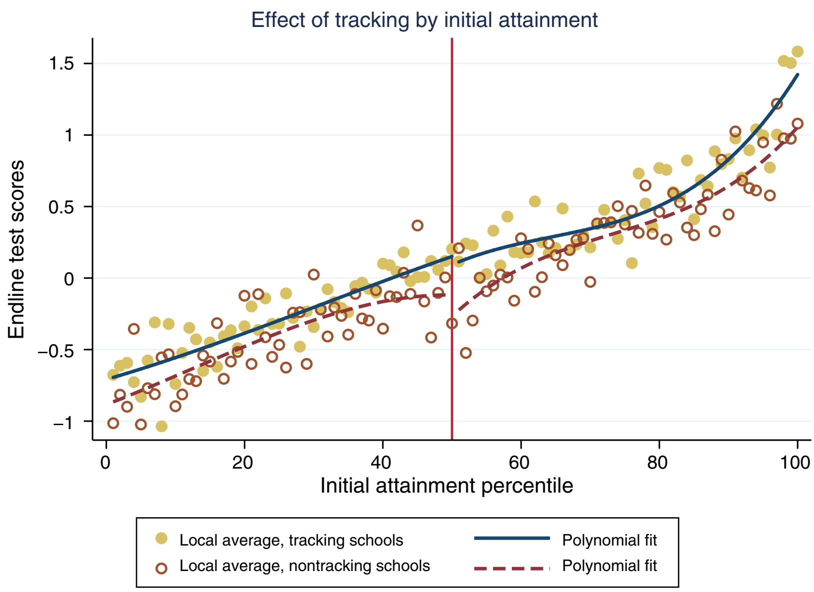


Figure 3. Local Polynomial Fits of Endline Score by Initial Attainment

Notes: Dots represent local averages. The fitted values are from regressions that include a second order polynomial estimated separately on each side of the percentile = 50 threshold.

students master core skills in grades 1 and 2 and that this may have helped them learn more later on.[13]

Under Proposition 1, this evidence of gains throughout the distribution is inconsistent with the special case of the model in which pupils do not affect each other indirectly through teacher behavior but only directly, with all pupils benefiting from higher scoring classmates.

Table 3 tests for heterogeneity in the effect of tracking. We present the estimated effect of tracking separately for boys and girls in panel A. Although the coefficients are not significantly different from each other, point estimates suggest that the effects are larger for girls in math (panel A). For both boys and girls, initially weaker students benefit as much as initially stronger students.

Panel B presents differential effects for students taught by civil-service teachers and contract teachers. This distinction is important, since the impact of tracking could be affected by teacher response, and contract and civil-service teachers have different experience and incentives.

While tracking increases test scores for students at all levels of the pretest distribution assigned to be taught by contract teachers (indeed, initially low-scoring students

[13] We also find (in results not reported here to save space) that initially low-achieving girls in tracking schools are 4 percentage points less likely to repeat grade 1. Since the program continued in grade 2, students who repeated lost the advantage of being in a small class, and of being more likely to be taught by a contract teacher. Part of the effect of tracking after the end of grade one may be due to this. In the companion paper, we estimate the effect of the class size reduction program in nontracking schools to be 0.16 standard deviations on average. At most, the repetition effect would therefore explain an increase in $0.04 \times 0.16 = 0.0064$ standard deviations in test scores. Furthermore, it is present only for girls, while tracking affects both boys and girls.

Table 3—Testing for Heterogeneity in Effect of Tracking on Total Score

	Short run: after 18 months in program			Longer run: a year after program ended		
	Effect of tracking on total score for		Test (top = bottom)	Effect of tracking on total score for		Test (top = bottom)
	Bottom half (1)	Top half (2)	*p*-value (3)	Bottom half (4)	Top half (5)	*p*-value (6)
Panel A. By gender						
Boys	0.130	0.162	0.731	0.084	0.206	0.168
	(0.076)*	(0.100)		(0.083)	(0.084)**	
Girls	0.188	0.222	0.661	0.190	0.227	0.638
	(0.089)**	(0.104)**		(0.098)*	(0.089)**	
Test (boys = girls): *p*-value	0.417	0.470		0.239	0.765	
Panel B. By teacher type						
Regular teacher	0.048	0.225	0.155	0.086	0.198	0.329
	(0.088)	(0.120)*		(0.099)	(0.098)**	
Contract teacher	0.255	0.164	0.518	0.181	0.246	0.605
	(0.099)**	(0.118)		(0.094)*	(0.103)**	
Test (regular = contract): *p*-value	0.076	0.683		0.395	0.702	
Panel C. By age						
Younger than average	0.151	0.287	0.135	0.146	0.309	0.062
	(0.088)*	(0.107)***		(0.093)	(0.088)***	
Older than average	0.154	0.047	0.274	0.169	0.111	0.593
	(0.095)	(0.098)		(0.109)	(0.092)	
Test (younger = older): *p*-value	0.976	0.002		0.818	0.008	

Notes: The sample includes 60 tracking and 61 nontracking schools. The dependent variables are normalized test scores, with mean 0 and standard deviation 1 in the nontracking schools. Robust standard errors clustered at the school level are presented in parentheses. Individual controls included: age, gender, being assigned to the contract teacher, dummies for initial half, and initial attainment percentile.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

assigned to a contract teacher benefited even more from tracking than initially high-scoring students), initially low-scoring students did not benefit from tracking if assigned to a civil-service teacher. In contrast, tracking substantially increased scores for initially high-scoring students assigned to a civil-service teacher. Below, we will present evidence that this may be because tracking led civil-service teachers to increase effort when they were assigned to the upper track, but not when assigned to the lower track, while contract teachers exert high effort in all situations. This is consistent with the idea that the cost of effort rises very steeply as a certain effort level is approached. Contract teachers are close to this level of effort in any case and therefore have little scope to increase their effort, while civil-service teachers have more such scope.

B. *Random Variation in Peer Composition and the Direct Effect of Peers*

The local random variation in peer quality in nontracking schools helps us test whether the opposite special case in which peers affect each other only indirectly, through their impact on teacher behavior, but not directly, can also be rejected.

Recall that Proposition 3 implies that the impact of a uniform increase in peer achievement on students at different levels of the distribution depends on whether or not there are direct peer effects. Namely, a uniform increase in peer achievement increases test scores at the top of the distribution in all cases, but effects on students in the middle and at the bottom of the distribution depend on whether there are also direct, positive effects of high-achieving peers. In the presence of such effects, the impact on students in the middle of the distribution is ambiguous, while for those at the bottom it is positive, albeit weaker than the effects at the top of the distribution. In the absence of such direct effects, there is a negative impact on students in the middle of the distribution and no impact at the bottom.

The random allocation of students between the two sections in nontracking schools generated substantial random variation which allows us to test those implications (see footnote 10).[14] We can thus implement methods to evaluate the impact of class composition similar to those introduced by Hoxby (2000), with the difference that we use actual random variation in peer group composition but have a lower sample size. The results are presented in Table 4. Similar approaches are proposed by Michael A. Boozer and Stephen E. Cacciola (2001) in the context of the STAR experiment and David S. Lyle (2007) for West Point cadets, who are randomly assigned to a group of peers.

On average students benefit from stronger peers: the coefficient on the average baseline test score is 0.35 with a standard error of 0.15 (column 1, Table 4, panel A). This coefficient is not comparable with other estimates in the literature since we are using the school grade sheets, which are not comparable across schools, and so we are standardizing the baseline scores in each school. Thus, in panel B, we use the average baseline scores of peers to instrument for their average endline score (the first stage is presented in panel C). If effects were linear, column 1 would imply that a one standard deviation increase in average peer endline test score would increase the test score of a student by 0.445 standard deviations, an effect comparable to those found in previous work, with the exception of Lyle (2007), which finds insignificant peer effects with a similar strategy.[15]

More interestingly, as shown in columns 4 to 6, the data are consistent with Proposition 3 in the presence of direct peer effects—the estimated effect is 0.9 standard deviations in the top quartile, insignificant and negative in the middle two quartiles, and 0.5 standard deviations in the bottom quartile. The data thus suggest that peers affect each other both directly and indirectly.[16]

[14] We used only the initial assignment (which was random) in all specifications, not the section the student eventually attended.

[15] Of course, these estimates come from variations in peer test scores that are smaller than one standard deviation, and the extrapolation to one standard deviation may not actually be legitimate: the linear approximation is valid only locally. However, presenting the results in terms of the impact of a one standard deviation change in peers' test scores allows us to compare our results to that of the literature, which also uses local variation in average test scores and generally expresses the results in terms of the impact of a one standard deviation increase in average test scores. Note that even with this normalization, the results are not quite comparable to those of papers which estimate the effect of a standard deviation in average baseline test scores on endline test scores: those results would be scaled down, relative to the ones we present here, by the size of the relationship between baseline and endline scores.

[16] Controlling for the standard deviation of the test scores (column 2) does not change the estimated effect of the mean, and we cannot reject the hypothesis that the standard deviation of scores itself has no effect, though the standard errors are large.

TABLE 4—PEER QUALITY: EXOGENOUS VARIATION IN PEER QUALITY (NONTRACKING SCHOOLS ONLY)

	All			25th–75th percentiles only	Bottom 25th percentiles	Top 25th percentiles only
	Total score (1)	Math score (2)	Lit score (3)	Total score (4)	Total score (5)	Total score (6)
Panel A. Reduced form						
Average baseline score of classmates[a]	0.346	0.323	0.293	−0.052	0.505	0.893
	(0.150)**	(0.160)**	(0.131)**	(0.227)	(0.199)**	(0.330)***
Observations	2,188	2,188	2,188	2,188	2,188	2,188
School fixed effects	x	x	x	x	x	x
Panel B. IV						
Average endline score of classmates (predicted)	0.445	0.470	0.423	−0.063	0.855	1.052
	(0.117)***	(0.124)***	(0.120)***	(0.306)	(0.278)***	(0.368)***
Observations	2,188	2,188	2,189	1,091	524	573
School fixed effects	x	x	x	x	x	x
Panel C. First-Stage for IV: average endline score of classmates						
	Average total score	Average math score	Average lit score	Average total score	Average total Score	Average total score
Average (standardized) baseline score of classmates‡	0.768	0.680	0.691	0.795	0.757	0.794
	(0.033)***	(0.033)***	(0.030)***	(0.056)***	(0.066)***	(0.070)***

Notes: Sample restricted to the 61 nontracking schools (where students were randomly assigned to a section). Individual controls included but not shown: gender, age, being assigned to the contract teacher, and own baseline score. Robust standard errors clustered at the school level in parentheses.

[a] This variable has a mean of 0.0009 and a standard deviation of 0.1056. We define classmates as follows: two students in the same section are classmates; two students in the same grade but different sections are not classmates.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

C. *Are Teacher Incentives Linear? The Impact of Assignment to Lower versus Upper Section: Regression Discontinuity Estimates for Students near the Median*

Recall from Proposition 7 that under a linear payoff schedule for teachers, the median student will be equidistant from the target teaching level in the upper and lower sections but will have higher-achieving peers and therefore perform better in the upper section. Under a concave payoff schedule, teacher effort will be greater in the lower section but the median student will be better matched to the target teaching level in the upper section, potentially creating offsetting effects. Finally, if teacher payoffs are convex in student test scores, the median student will be closer to the target teaching level in the lower section but, on the other hand, will have lower-achieving peers and experience lower teacher effort. These effects go in opposite directions, so that the resulting impact of the section in which the median child is assigned is ambiguous. In this section, we present regression discontinuity estimates of the impact of assignment to the lower or upper section for students near the median in tracking schools. We argue that the test score data are inconsistent with linear payoffs but consistent with the possibility that teachers face a convex payoff function and focus on students at the top of the distribution. (Later, we rule out the concave case.)

The main thrust of the regression discontinuity estimates of peer effects are shown in Figure 3, discussed above. As is apparent from the figure, there is no discontinuity in test scores at the fiftieth percentile cutoff in the tracking schools, despite the strong discontinuity in peer baseline scores observed in Figure 2 (a difference of

Table 5—Peer Quality: Regression Discontinuity Approach (Tracking Schools Only)

	Total						
	Specification 1: With third order polynomial in baseline attainment		Specification 2: With second order polynomial in baseline attainment estimated separately on either side		Specification 3: With local linear regressions	Specfication 4: Pair around the median	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Reduced form							
Estimated effect of bottom section at 50th percentile	0.010	0.001	−0.045	−0.051	−0.016	0.034	0.027
	(0.093)	(0.079)	(0.106)	(0.089)	(0.204)	(0.136)	(0.145)
Observations (students)	2,959	2,959	2,959	2,959	2,959	149	149
School fixed effects	No	Yes	No	Yes	No	No	Yes
Panel B. IV							
Mean total score of peers	−0.012	−0.002				−0.068	−0.004
	(0.117)	(0.106)				(0.205)	(0.277)
Observations (students)	2,959	2,959				149	149
School fixed effects	No	Yes				No	Yes
Panel C. First stage for IV							
In bottom half of initial distribution	−0.731	−0.743				−0.612	−0.607
	(0.047)***	(0.021)***				(0.090)***	(0.058)***
Observations (students)	2,959	2,959				149	149
R^2	0.42	0.78				0.25	0.57
School fixed effects	No	Yes				No	Yes

Notes: Sample restricted to the 60 tracking schools (where students were tracked into two sections by initial attainmnent). Students in the bottom half of the initial distribution were assigned to the "bottom section" where the average peer quality was much lower than in the top section (see Figure 2). Panel A, columns 1–2 and 6–7: the score was regressed on a dummy "assigned to bottom section" and individual controls (age, gender, dummy for being assigned to contract teacher and, for columns 1 and 2, a polynomial in initial percentile). We present the estimated coefficient of the dummy "assigned to bottom section." Standard errors clustered at school level. Panel A, columns 3–5: The estimated effect of being assigned to the bottom section is the difference between the estimates of the expectation function estimated separately on either side of the 50th percentile. In columns 3–4, the score was regressed on a second order polynomial in initial percentile fully interacted with a dummy for "bottom section." In column 5, the score was estimated through local linear regression (bandwidth = 2). Bootstrapped standard errors clustered at the school level. Regressions in columns 6–7 include 1 pair of students per school: The top student in the bottom section and the bottom student in the top section. The number of observations is greater than 120 due to ties in some schools. In panel B, the mean score of class peers is instrumented by the dummy "In bottom half of initial distribution" and controls.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

1.6 standard deviations in the baseline scores). The relationship is continuous and smooth throughout the distribution.[17]

A variety of regression specifications show no significant effect of students near the median of the distribution being assigned to the bottom half of the class in tracking schools (Table 5, panel A). Columns 1 and 2 present estimates of equations ($E5$) and ($E6$), respectively: the endline test score is regressed on a cubic of original percentile of a child in the distribution in his school, and a dummy for whether he is in the bottom half of the class. Column 6 presents estimates of equation ($E7$), and column 7 adds a school fixed effect. To assess the robustness of these results,

[17] This result is robust to a series of specifications. When we use a linear fit, rather than a polynomial, we again do not see an effect of the section in which the students were placed for students in the middle of the distribution (figure not shown). In Appendix Figure A1, we reproduce Figure 3 (for tracking schools) with a quadratic fit for total score in panel A and also find no discontinuity. In panel B, we use a Fan locally weighted regression with a biweight kernel and a bandwidth of 2.0, and we again see no discontinuity at the threshold for being assigned to the bottom track.

columns 3 through 5 specify the control function in the regression discontinuity design estimates in two other ways: column 5 follows Imbens and Lemieux (2008) and shows a Fan locally weighted regression on each side of the discontinuity.[18] The specifications in columns 3 and 4 are similar to equations ($E5$) and ($E6$), but the cubic is replaced by a quadratic allowed to be different on both sides of the discontinuity. The results confirm what the graphs show: despite the big gap in average peer achievement, the marginal students' final test scores do not seem to be significantly affected by assignment to the bottom section.

Panel B shows instrumental variable estimates of the impact of classmates' average test score. We use the average endline score of classmates (because the baseline scores are school specific) and instrument it using the dummy for being in the "bottom half" of the initial distribution. The first stage is shown in panel C and shows that the average endline test scores of a child's classmates are about 0.76 standard deviations lower if she was assigned to the bottom section in a tracking school. The IV estimates in panel B are all small and insignificant. For example the specification in column 2, which has school fixed effects and uses all the data, suggests that an increase in one standard deviation in the classmates' average test score *reduces* a child's test score by 0.002 standard deviations, a point estimate extremely close to zero. The 95 percent confidence interval in this specification is $[-0.21; 0.21]$, which excludes large effects.[19]

Overall, these regression discontinuity results allow us to reject the third special case, in which teachers have linear incentives and consequently target the median child in the distribution of the class.

Taken together, the test score results are consistent with a model in which students influence each other both directly and indirectly through teacher behavior, and teachers face convex payoffs in pupils' test scores, and thus tend to target their teaching to the top of the class. This model can help us interpret our main finding that tracking benefits all students: for higher-achieving students, tracking implies stronger peers and higher teacher effort, while for lower-achieving students, tracking implies a level of instruction that better matches their needs.

Nevertheless, the result that the median student benefits just as much from tracking if assigned to the top or the bottom track of the classroom is striking. Under convexity of the teacher payoff function, it suggests that the gain experienced by the median student assigned to the lower track in terms of level of instruction compensates exactly the loss (relative to the median student assigned to the upper track) in terms of teacher effort and direct effect of peers. This is a somewhat surprising coincidence. Furthermore, we have not yet rejected the possibility that teacher payoffs are concave in student test scores. Recall that under concavity, students in the bottom half of the distribution may gain from greater teacher effort under tracking (proposition 6). Fortunately, convexity and concavity have different implications for how the impact of being assigned to the top or the bottom track should differ, depending on whether the median student is assigned to the civil-service teacher or

[18] Since the result is completely insensitive to the choice of bandwidth, we do not implement the cross-validation strategy they recommend.

[19] With the caveat, mentioned above, that there may be a direct impact of one's rank in the class, which would violate the identification assumption that the only channel of impact of being assigned to the lower section is through classmates' average score.

the regular teacher (proposition 8), which gives us a sharper prediction to test. Our model predicts a greater response of teacher effort to tracking among civil teachers than among contract teacher (we will test this directly below). Under convexity, this implies that the median student should benefit more from being assigned to the top section (rather than the bottom section) with a civil-service teacher than with a contract teacher. Appendix Figure A2 shows the plot of the regression discontinuity estimated separately for contract and for civil service teachers.[20] The pattern supports the prediction of the convex case. For students assigned to contract teachers, we see no discontinuity at the median in the difference between tracking and nontracking schools. For students assigned to civil-service teachers, we see a graphical illustration of the result above: there is no effect of tracking for children at the bottom, but there is a jump at the median in the effect of tracking (the difference between the two lines). The point estimate of the difference in the limit from above and from below of the tracking effect is 0.11 standard deviations (although it is not significant).

Additional functional form assumptions would allow us to make more predictions on how the shape of the relationship between baseline quantile and endline achievement varies with tracking status and the shape of teachers' payoffs. Assume, for example, that there is a linear relationship between own baseline quantile and endline test score, that the effect of peers is also linear, and that the effect of teacher instruction is as described above. Then, if the teacher teaches to the top, the gains to tracking should be higher at the top of both tracks than at the bottom: the relationship between baseline quantile and endline test scores should be convex within each track. In Figure 3, we observe this pattern in the upper track, but not in the lower track. Second, the relationship between baseline quantile and endline test score in nontracking schools should be linear at the bottom of the distribution, and steeper at the top (since the top students are closer to the instruction level), which is consistent with what we see in Figure 3.[21]

The next section examines data on teacher behavior, arguing that it is also inconsistent with the hypothesis that teacher payoffs are concave in student test scores, but consistent with the hypothesis that payoffs are convex in student scores.

IV. Teacher Response to Tracking

This section reports on tests of implications on the model related to teacher behavior. Subsection A argues that the evidence on teacher behavior is consistent with the idea that teachers face convex payoff incentives in pupil test scores and inconsistent with the hypothesis of concavity. Subsection B presents some evidence that the patterns of changes in test scores are consistent with the hypothesis that teachers change their focus teaching level x^*, in response to tracking.

[20] Note that the data are from different schools on the left and on the right of the median in each of these plots, since in a given school, the contract teacher is assigned one class, and the civil teacher is assigned the other one.

[21] In future work, it would be interesting to take this line of argument further and estimate a parametric version of the model, using both cross-sectional variation and the variation generated by the experiment. The model would likely be overidentified, which would allow it to be formally tested.

TABLE 6—TEACHER EFFORT AND STUDENT PRESENCE

	All teachers		Government teachers		ETP teachers		Students
	Teacher found in school on random school day (1)	Teacher found in class teaching (unconditional on presence) (2)	Teacher found in school on random school day (3)	Teacher found in class teaching (unconditional on presence) (4)	Teacher found in school on random school day (5)	Teacher found in class teaching (unconditional on presence) (6)	Student found in school on random school day (7)
Tracking school	0.041	0.096	0.054	0.112	−0.009	0.007	−0.015
	(0.021)**	(0.038)**	(0.025)**	(0.044)**	(0.034)	(0.045)	(0.014)
Bottom half × tracking school	−0.049	−0.062	−0.073	−0.076	0.036	−0.004	0.003
	(0.029)*	(0.040)	(0.034)**	(0.053)	(0.046)	(0.057)	(0.007)
Years of experience teaching	0.000	−0.005	0.002	0.002	−0.002	−0.008	
	(0.001)	(0.001)***	(0.001)*	(0.001)	(0.003)	(0.008)	
Female	−0.023	0.012	−0.004	0.101	−0.034	−0.061	−0.005
	(0.018)	(0.026)	(0.020)	(0.031)***	(0.032)	(0.043)	(0.004)
Assigned to contract teacher							0.011
							(0.005)**
Assigned to contract teacher × tracking school							0.004
							(0.008)
Observations	2,098	2,098	1,633	1,633	465	465	44,059
Mean in non-tracking schools	0.837	0.510	0.825	0.450	0.888	0.748	0.865
F (test of joint significance)	2.718	9.408	2.079	5.470	2.426	3.674	5.465
p-value	0.011	0.000	0.050	0.000	0.023	0.001	0.000

Notes: The sample includes 60 tracking and 61 nontracking schools. Linear probability model regressions. Multiple observations per teacher and per student. Standard errors clustered at school level. Region and date of test dummies were included in all regressions but are not shown.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

A. *Teacher Effort and the Curvature of the Teacher Payoff Function*

Estimates of the impact of tracking on teacher's effort are presented in Table 6. Our measure of teacher effort is whether the teacher was present in school during unannounced visits, and whether she was found in class and teaching.

Recall that the model does not yield a clear prediction for whether tracking should increase or decrease teacher effort overall. However, the model predicts that the effort level might vary across tracks (upper or lower). Namely, proposition 6 implies that if teacher payoffs are convex in student test scores, then teachers assigned to the top section in tracking schools should exert more effort than those assigned to the bottom section. On the other hand, if payoffs are concave in student test scores, teachers should put in more effort in the lower section than the upper section.

We find that teachers in tracking schools are significantly more likely both to be in school and to be in class teaching than those in nontracking schools (columns 1 and 2, Table 6).[22] Overall, teachers in tracking schools are 9.6 percentage points (19 percent) more likely to be found in school and teaching during a random spot check than their counterparts in nontracking schools. However, the negative coefficient on

[22] The specification is similar to equation ($E2$), though the set of control variables includes teacher age and experience teaching.

the interaction term between "tracking" and "bottom half" shows that teacher effort in tracking schools is higher in the upper section than the lower sections, consistent with the hypothesis that teacher payoffs are convex in student test scores.

Recall that the model also suggests that if teachers face strong enough incentives (high enough λ) then the impact of tracking on their effort will be smaller because they have less scope to increase effort. To test this, we explore the impact of tracking on teacher effort separately for civil-service teachers and new contract teachers, who face very different incentives. Contract teachers are on short-term (one year) contracts and have incentives to work hard to increase their chances both of having their short-term contracts renewed, and of eventually being hired as civil-service teachers. In contrast, the civil service teachers have high job security and promotion depends only weakly on performance. Civil service teachers thus may have more scope to increase effort.

We find that the contract teachers attend more than the civil-service teachers, are more likely to be found in class and teaching (74 percent versus 45 percent for the civil-service teacher), and their absence rate is unaffected by tracking. In contrast, the civil-service teachers are 5.4 percentage points more likely to be in school in tracking schools than in nontracking schools when they were assigned to the top section, and the difference is significant (recall that teacher assignment to each section was random, so this is indeed the causal effect of being assigned to a group of strong students, rather than a nontracked group). However, the difference disappears entirely for civil-service teachers assigned to the bottom section: the interaction between tracking and bottom section is minus 7.3 percentage points, and is also significant. The effect is even stronger for finding teachers in their classrooms: overall, these civil-service teachers are 11 percentage points more likely to be in class and teaching when they are assigned to the top section in tracking schools than when they are assigned to nontracking schools. This represents a 25 percent increase in teaching time. When civil-service teachers are assigned to the bottom section, they are about as likely to be teaching as their counterparts in nontracking schools.[23] Students' attendance is not affected by tracking or by the section they were assigned to (column 7).

These results on teacher effort also shed light on the differential impact of tracking across students observed in Table 3. Recall that among students who were assigned to civil service teachers, tracking created a larger test score increase in the top section than in the bottom section, but this was not the case for students of contract teachers. The effort data suggest that, for students of civil service teachers, the tracking effect is larger for the upper stream because they benefit not only from (potentially) more appropriate teaching and better peers, but also from higher effort. For students of contract teachers, the "higher effort" margin is absent.

[23] Note that these results on the efforts of civil-service teachers are obtained in an environment where there is also a high-effort teacher present. In our companion paper (Duflo, Dupas, and Kremer 2010) we find that civil-service teachers in both tracking and nontracking schools reduced their effort relative to schools which did not receive any extra teacher. Being assigned to the upper track only mitigated this effect but did not improve teacher's effort beyond the effort level observed in the absence of any contract teacher program. As such, our results may not generalize to a context where tracking is achieved by hiring extra civil-service teachers.

Table 7—Effect of Tracking by Level of Complexity and Initial Attainment

	Mathematics			Test	Literacy			
	Difficulty level 1 (1)	Difficulty level 2 (2)	Difficulty level 3 (3)	Coeff (col 3) = coeff (col 1) (4)	Reading letters (5)	Spelling words (6)	Reading words (7)	Reading sentences (8)
(1) In bottom half of initial distribution	−1.43 (0.09)***	−1.21 (0.08)***	−0.49 (0.05)***		−3.86 (0.33)***	−4.05 (0.42)***	−4.15 (0.40)***	−1.15 (0.21)***
(2) Tracking school	0.15 (0.10)	0.16 (0.12)	0.21 (0.10)**	$\chi^2 = 0.66$ p-value = 0.417	1.63 (0.65)**	1.00 (0.78)	1.08 (0.75)	0.38 (0.34)
(3) In bottom half of initial distribution × tracking school	0.18 (0.14)	0.08 (0.12)	−0.10 (0.08)	$\chi^2 = 3.97$ p-value = 0.046	−0.42 (0.46)	−0.61 (0.61)	−0.39 (0.56)	−0.44 (0.30)
Constant	4.93 (0.23)***	1.82 (0.22)***	0.57 (0.16)***		11.64 (1.00)***	10.06 (1.20)***	10.12 (1.12)***	3.94 (0.56)***
Observations	5,284	5,284	5,284		5,283	5,279	5,284	5,284
Maximum possible score	6	6	6		24	24	24	24
Mean in nontracking schools	4.16	1.61	0.67		6.99	5.52	5.00	2.53
SD in nontracking schools	2.02	1.62	0.94		6.56	7.61	7.30	3.94
Total effect of tracking on bottom half:								
Coeff (row 2) + coeff (row 3)	0.33	0.24	0.11	$\chi^2 = 2.34$ p-value = 0.126	1.21	0.39	0.69	−0.06
F Test: coeff (row 2) + coeff (row 3) = 0	3.63	6.39	4.42		4.74	0.70	1.82	0.09
p-value	0.06	0.01	0.04		0.03	0.40	0.18	0.76

Notes: The sample includes 60 tracking and 61 nontracking schools. Robust standard errors clustered at the school level are presented in parentheses. Difficulty level 1: addition or subtraction of 1 digit numbers. Difficulty level 2: addition or subtraction of 2 digit numbers, and multiplication of 1 digit numbers. Difficulty level 3: addition or subtraction of 3 digit numbers.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

B. *Adjustment in the Level of Teaching and Effects on Different Skills*

The model suggests teachers may adjust the level at which they teach in response to changes in class composition. For example, a teacher assigned to students with low initial achievement might begin with more basic material and instruct at a slower pace, providing more repetition and reinforcement. With a group of initially higher-achieving students, the teacher can increase the complexity of the tasks and pupils can learn at a faster pace. Teachers with a heterogeneous class may teach at a relatively high level that is inappropriate for most students, especially those at the bottom.

While we unfortunately do not have direct evidence on the material teachers covered, Table 7 reports specifications similar to equation ($E2$), but with test scores disaggregated by specific skill for math and language. The differential impact of tracking on strong and weak students' mastery of easy and hard material is consistent with the hypothesis that teachers adjusted their teaching to fit their classroom's composition. The equations are estimated jointly in a simultaneous equation framework (allowing for correlation between the error terms). There is no clear pattern for language, but the estimates for math suggest that, while the total effect of tracking on children initially in the bottom half of the distribution (thus assigned to the bottom section in the tracking schools) is significantly positive for all levels of difficulty, these children gained from tracking more than other students on the easiest questions and less on the more difficult questions. The interaction "tracking times bottom half" is positive for the easiest skills, and negative for the hardest skills. A chi-square test allows us to reject equality of the coefficients of the interaction in the "easy skills" regression and the "difficult skills" regression at the 5 percent level. Conversely, students assigned to the upper section benefited less on the easiest questions, and more on the difficult questions (in fact, they did not significantly benefit from tracking for the easiest questions, but they did significantly benefit from it for the hardest questions).

Overall, this table provides suggestive evidence that tracking allowed teachers the opportunity to focus on the skills that children had not yet mastered, although the estimates are not very precise.[24] An alternative explanation for these results, however, is that weak students stood to gain from any program on the easiest skills (since they had not mastered them yet, and in 18 months they did not have time to master both easy and strong skills), while strong students had already mastered them and would have benefited from any program on the skills they had not already mastered. The ordinal nature of test score data makes regression interaction terms difficult to interpret conclusively, which further weakens the evidence.

V. Conclusion

This paper provides experimental evidence that students at all levels of the initial achievement spectrum benefited from being tracked into classes by initial achievement. Despite the critical importance of this issue for the educational policy in both developed and developing countries, there is surprisingly little rigorous evidence addressing it, and to our knowledge this paper provides the first experimental evaluation of the impact of tracking in any context, and the only rigorous evidence in a developing country context.

After 18 months, the point estimates suggest that the average score of a student in a tracking school is 0.14 standard deviations higher than that of a student in a nontracking school. These effects are persistent. One year after the program ended, students in tracking schools performed 0.16 standard deviations higher than those in nontracking schools.

Moreover, tracking raised scores for students throughout the initial distribution of student achievement. A regression discontinuity design approach reveals that students who were very close to the fiftieth percentile of the initial distribution within their school scored similarly on the endline exam whether they were assigned to the top or bottom section. In each case, they did much better than their counterparts in nontracked schools.

We also find that students in nontracking schools scored higher if they were randomly assigned to peers with higher initial scores. This effect was very strong for students at the top of the distribution, absent for students in the middle of the

[24] We estimated a version of equation ($E2$) allowing the effect to vary by quarter of the distribution for each skill, and the patterns are very similar, with progressively weaker students benefiting the most from tracking for the easiest skills, and progressively stronger students benefiting the most for the hardest skills. We also estimated a version of equation ($E2$) separately by teacher type. We find that the effects observed in Table 7 are much stronger for students assigned to contract teachers than for those assigned to civil-service teachers. This is because lower section students assigned to the civil-service teachers did not benefit from tracking, as seen in Table 3.

distribution and positive but not as strong at the bottom of the distribution. Together, these results suggest that peers affect students both directly and indirectly by influencing teacher behavior, in particular teacher effort and choice of target teaching level. Under the model, the impact of tracking will depend on teachers' incentives, but in a context in which teachers have convex payoffs in student test scores, tracking can lead them to refocus attention closer to the median student.

These conclusions echo those reached by Geoffrey D. Borman and Gina M. Hewes (2002), who find positive short- and long-term impacts of "Success for All." One of the components of this program, first piloted in the United States by elementary schools in Baltimore, Maryland, is to regroup students across grades for reading lessons targeted to specific performance levels for a few hours a day. Likewise, Banerjee et al. (2007), who study a remedial education and a computer-assisted learning program in India, found that both programs were very effective, mainly because they allowed students to learn at their own levels of achievement. Finally, our results match those of Zimmer (2003), who finds that, in the United States, tracking has overall a positive effect on lower-achieving students, for whom the benefit of having more tailored instruction under tracking offsets the reduction in peer quality. On the other hand, Tuomas Pekkarinen, Roope Uusitalo, and Sari Kerr (2009) find positive effects of ending school tracking in Finland on the performance of students from lower ability background.

A central challenge of educational systems in developing countries is that students are extremely diverse, and the curriculum is largely not adapted to new learners. These results show that grouping students by preparedness or prior achievement and focusing the teaching material at a level pertinent to them could potentially have large positive effects with little or no additional resource cost.

Our results may have implications for debates over school choice and voucher systems. A central criticism of such programs is that they may wind up hurting some students if they lead to increased sorting of students by initial academic achievement and if all students benefit from having peers with higher initial achievement. Furthermore, tracking in public school would affect the equilibrium under these programs. Epple, Newlon, and Romano (2002) study theoretically how tracking in public schools would affect the decision of private schools to track students, and the welfare of high- and low-achieving students. They find that, if the only effect of tracking was through the direct effects of the peer group, tracking in public schools would increase enrollment and raise average achievement in public schools, but that high-achieving students would benefit at the expense of low-achieving students. Our results suggest that, at least in some circumstances, tracking can potentially benefit all students, which would have implications for the school choice equilibrium in contexts with school choices.

Note that some aspects of our setup are specific to the experimental nature of the work and would be different in a scaled-up version. The effects may be muted in this setting because the experiment was known to be relatively short-lived, and teachers may not have fully adjusted their lesson plans. On the other hand, since teachers were randomly assigned to each section and class size was also constant, resources were similar for nontracked classes and the lower and upper sections under tracking. However, in other contexts, policymakers or school officials could target more resources to either the weaker or stronger students. Thomas Piketty

(2004) notes that tracking could allow more resources to be devoted to weaker students, promoting catch-up of weaker students. Compensatory policies of this type are not unusual in developed countries, but in some developed countries and almost all developing countries, more resources are devoted to stronger students, consistent with the assumption of convex payoffs to test scores in the theoretical framework above. Indeed, even in developed countries, the best teachers are often assigned to the stronger students.

If the best teachers are assigned to the highest-achieving students, the initially lower achieving students could be hurt by tracking, so caution is needed in generalizing from these results in which teacher ability was held constant between tracking and nontracking schools.[25] Of course tendencies for strong teachers to seek high-achieving students could perhaps be mitigated if evaluations of a teacher's performance were on a value-added basis, rather than based on endline scores.

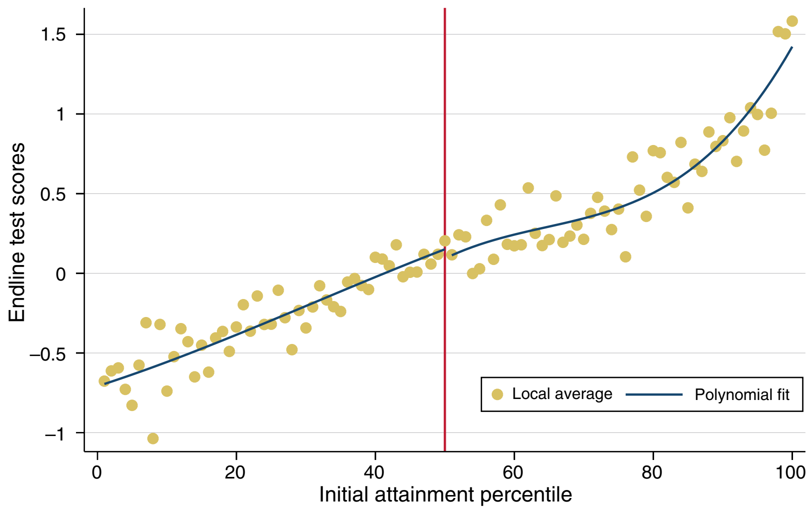
It is an open question whether similar results would be obtained in different contexts. The model provides some evidence on features of the context that are likely to affect the impact of tracking: initial heterogeneity, high scope to increase teacher effort (at least through increased presence), and the relative incentives teachers face to teach low- and high-achieving students. For example, in a system where the incentive is to focus on the weakest students, and there is not much scope to adjust teacher effort, tracking could have a very strong positive effect on high-achieving students, and a weak or even negative effect on weak students, who would lose strong peers without the benefit of getting more appropriately focused instruction. Going beyond the model, it seems reasonable to think that the impact of tracking might also depend on availability of extra resources to help teachers deal with different types of students (such as remedial education, teacher aides, lower pupil to teacher ratio, computer-assisted learning, and special education programs).

We believe that tracking might be reasonably likely to have a similar impact in other low income countries in sub-Saharan Africa and South Asia, where the student population is often heterogeneous, and the educational system rewards teachers for progress at the top of the distribution. Our reduced form results may not apply to the United States or other developed countries where teachers' incentives may differ, student and teacher effort levels are typically higher, and the distribution of student ability at baseline may be much narrower. However, we hope that our analysis may still provide useful insights to predict the situations in which tracking may or may not be beneficial in these countries, and on the type of experiments that would shed light on this question.

[25] Note, however, that in our setting it seems likely that if choice had been allowed, the more powerful teachers would have been assigned to the stronger group, and since the more powerful teachers were the civil-service teachers, who also happen to be the worst teachers, this would have benefited the weak students.

APPENDIX

Panel A. Quadratic fit



Note: The dots are the average score. The fitted values are from regressions that include a second order polynomial estimated separately on each side of the percentile = 50 threshold.

Panel B. Fan locally weighted regression

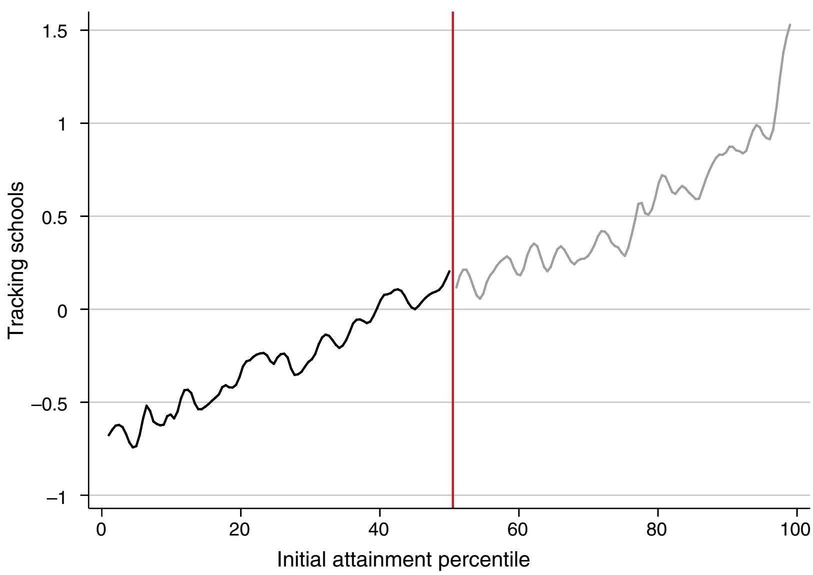


FIGURE A1. PEER QUALITY AND ENDLINE SCORES IN TRACKING SCHOOLS

Note: Fitted values from Fan's locally weighted regressions with quartic (biweight) kernels and a bandwidth of 2.0.

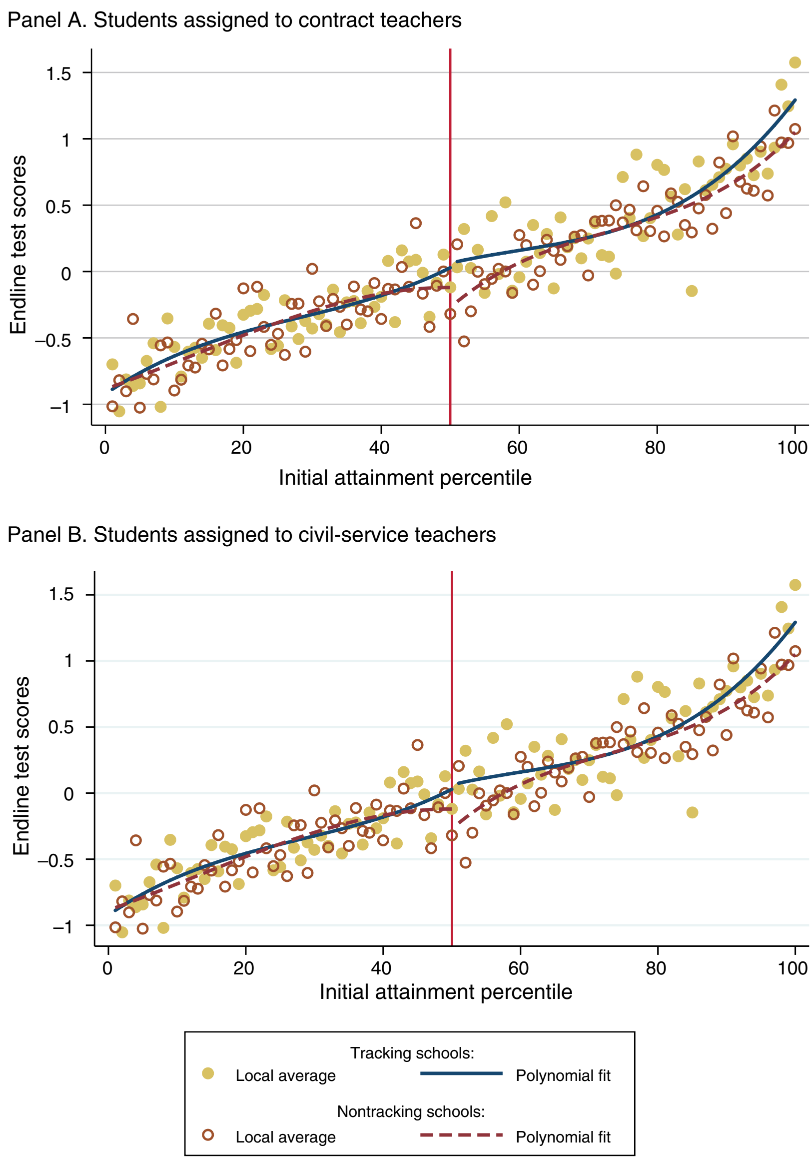


FIGURE A2. DIFFERENTIAL EFFECTS FOR MEDIAN STUDENT, BY TEACHER TYPE

Note: The data are from different schools on the left and on the right of the median in each of these plots, since in a given school, the contract teacher is assigned one class, and the civil teacher is assigned the other one.

Appendix Table A1—Does Attrition Vary Across Tracking and Nontracking Schools?

	At endline test (after 18 months in program)				At long-run follow-up test (a year after program ended)	
	Transferred to other school (1)	If not transferred: missed test (2)	Total attrition (3)	Total attrition, subsample around median (4)	Total attrition (5)	Subsample around median (6)
Tracking school	0.003	0.011	0.014	−0.015	0.000	−0.087
	(0.012)	(0.019)	(0.022)	(0.066)	(0.028)	(0.067)
In bottom half of initial distribution	0.027	−0.013	0.014	−0.006	0.050	0.042
	(0.016)*	(0.020)	(0.025)	(0.075)	(0.028)*	(0.073)
In bottom half of initial distribution × tracking school	−0.012	0.02	0.007	0.043	0.006	0.041
	(0.013)	(0.019)	(0.023)	(0.083)	(0.026)	(0.080)
Girl	0.012	0.029	0.041	0.020	0.024	0.055
	(0.009)	(0.014)**	(0.016)**	(0.049)	(0.016)	(0.060)
Girl × tracking school	0.004	−0.048	−0.044	−0.026	−0.022	0.039
	(0.013)	(0.016)***	(0.019)**	(0.067)	(0.021)	(0.079)
Assigned to contract teacher	−0.006	−0.019	−0.025	−0.078	−0.014	−0.086
	(0.010)	(0.011)*	(0.014)*	(0.046)*	(0.015)	(0.055)
Assigned to contract teacher × tracking school	0.018	0.009	0.027	0.056	0.038	0.142
	(0.015)	(0.019)	(0.025)	(0.076)	(0.026)	(0.092)
In bottom half of initial distribution × assigned to contract teacher × tracking school	0.01	−0.035	−0.025	−0.01	−0.011	0.031
	(0.019)	(0.030)	(0.036)	(0.093)	(0.037)	(0.104)
Constant	0.034	0.172	0.205	−0.737	0.23	2.11
	(0.018)*	(0.026)***	(0.032)***	(1.664)	(0.034)***	(1.747)
Observations	7,345	7,345	7,345	517	7,340	515
Mean	0.057	0.119	0.175	0.175	0.224	0.224

Notes: OLS regressions; standard errors clustered at school level. Additional controls not shown: a third degree polynomial in the student's percentile in the initial attainment distribution. Columns 4 and 6: restricted to students within 0.1 standard deviation of median at baseline.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

REFERENCES

Andrabi, Tahir, Jishnu Das, Asim Khwaja, and Tristan Zajonc. 2008. "Do Value-Added Estimates Add Value? Accounting for Learning Dynamics." Unpublished.

Angrist, Joshua D., and Kevin Lang. 2004. "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program." *American Economic Review*, 94(5): 1613–34.

Angrist, Joshua D., and Victor Lavy. 1999. "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *Quarterly Journal of Economics*, 114(2): 533–75.

Banerjee, Abhijit V., Shawn Cole, Esther Duflo, and Leigh Linden. 2007. "Remedying Education: Evidence from Two Randomized Experiments in India." *Quarterly Journal of Economics*, 122(3): 1235–64.

Betts, Julian R., and Jamie L. Shkolnik. 2000. "Key Difficulties in Identifying the Effects of Ability Grouping on Student Achievement." *Economics of Education Review*, 19(1): 21–26.

Black, Dan A., Jose Galdo, and Jeffrey A. Smith. 2007. "Evaluating the Worker Profiling and Reemployment Services System Using a Regression Discontinuity Approach." *American Economic Review*, 97(2): 104–07.

Boozer, Michael A., and Stephen E. Cacciola. 2001. "Inside the 'Black Box' of Project Star: Estimation of Peer Effects Using Experimental Data." Yale University Economic Growth Center Discussion Paper 2001.

Borman, Geoffrey D., and Gina M. Hewes. 2002. "The Long-Term Effects and Cost-Effectiveness of Success for All." *Educational Evaluation and Policy Analysis*, 24(4): 243–66.

Clark, Damon. 2010. "Selective Schools and Academic Achievement." *The B.E. Journal of Economic Analysis & Policy*, 10(1): Article 9.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2010. "School Governance, Pupil-Teacher-Ratios, and Teacher Incentives: Experimental Evidence from Kenyan Primary Schools." Unpublished.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2011. "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya: Dataset." *American Economic Review*, http://www.aeaweb.org/articles.php?doi=10.1257/aer.101.5.95.

Epple, Dennis, and Richard Romano. 2011. "Peer Effects in Education: A Survey of the Theory and Evidence." In *Handbook of Social Economics*, Vol. 1B, ed. Jess Benhabib, Alberto Bisin, and Matthew Jackson, 1053–1163. Amsterdam: North-Holland, Elsevier.

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

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

Glewwe, Paul, Michael Kremer, and Sylvie Moulin. 2009. "Many Children Left Behind? Textbooks and Test Scores in Kenya." *American Economic Journal: Applied Economics*, 1(1): 112–35.

Hoxby, Caroline. 2000. "Peer Effects in the Classroom: Learning from Gender and Race Variation." National Bureau of Economic Research Working Paper 7867.

Hoxby, Caroline, and Gretchen Weingarth. 2006. "Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects." Unpublished.

Imbens, Guido W., and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics*, 142(2): 615–35.

Krueger, Alan, and Diane Whitmore. 2002. "Would Smaller Classes Help Close the Black-White Achievement Gap?" In *Bridging the Achievement Gap*, ed. John E. Chubb and Tom Loveless, 11–46.Washington: Brookings Institution Press.

Lavy, Victor, M. Daniele Paserman, and Analia Schlosser. 2008. "Inside the Black of Box of Ability Peer Effects: Evidence from Variation in Low Achievers in the Classroom." National Bureau of Economic Research Working Paper 14415.

Lee, David S. 2008. "Randomized Experiments from Non-Random Selection in U.S. House Elections." *Journal of Econometrics*, 142(2): 675–97.

Lefgren, Lars. 2004. "Educational Peer Effects and the Chicago Public Schools." *Journal of Urban Economics*, 56(2): 169–91.

Lyle, David S. 2007. "Estimating and Interpreting Peer and Role Model Effects from Randomly Assigned Social Groups at West Point." *Review of Economics and Statistics*, 89(2): 289–99.

Pekkarinen, Tuomas, Roope Uusitalo, and Sari Kerr. 2009. "School Tracking and Development of Cognitive Skills." Institute for the Study of Labor (IZA) Discussion Paper 4058.

Piketty, Thomas. 2004. "L'Impact de la Taille des Classes et de la Ségrégation Sociale sur la Réussite Scolaire dans les Ecoles Françaises: Une Estimation a Partir du Panel Primaire 1997. " Unpublished.

Pischke, Jorn-Steffen, and Alan Manning. 2006. "Comprehensive Versus Selective Schooling in England in Wales: What Do We Know?" National Bureau of Economic Research Working Paper 12176.

Zimmer, Ron. 2003. "A New Twist in the Educational Tracking Debate." *Economics of Education Review*, 22(3): 307–15.

Zimmerman, David J. 2003. "Peer Effects in Academic Outcomes: Evidence from a Natural Experiment." *Review of Economics and Statistics*, 85(1): 9–23.

This article has been cited by:

1. Jazla Fadda. Climate Change: An Overview of Potential Health Impacts Associated with Climate Change Environmental Driving Forces 77-119. [Crossref]
2. Jane Cooley Fruehwirth, Jessica Gagete-Miranda. 2019. Your peers' parents: Spillovers from parental education. *Economics of Education Review* **73**, 101910. [Crossref]
3. Juan E. Saavedra, Emma Näslund-Hadley, Mariana Alfonso. 2019. Remedial Inquiry-Based Science Education: Experimental Evidence From Peru. *Educational Evaluation and Policy Analysis* **41**:4, 483-509. [Crossref]
4. Antonia Grohmann, Sahra Sakha. 2019. The effect of peer observation on consumption choices: evidence from a lab-in-field experiment. *Applied Economics* **51**:55, 5937-5951. [Crossref]
5. Matias Berthelon, Eric Bettinger, Diana I. Kruger, Alejandro Montecinos-Pearce. 2019. The Structure of Peers: The Impact of Peer Networks on Academic Achievement. *Research in Higher Education* **60**:7, 931-959. [Crossref]
6. Norman Kerle, Saman Ghaffarian, Raphael Nawrotzki, Gerald Leppert, Malte Lech. 2019. Evaluating Resilience-Centered Development Interventions with Remote Sensing. *Remote Sensing* **11**:21, 2511. [Crossref]
7. Lex Borghans, Ron Diris, Wendy Smits, Jannes de Vries. 2019. The long-run effects of secondary school track assignment. *PLOS ONE* **14**:10, e0215493. [Crossref]
8. Philip Hallinger. 2019. Science mapping the knowledge base on educational leadership and management in Africa, 1960–2018. *School Leadership & Management* **39**:5, 537-560. [Crossref]
9. Michael Gilraine. 2019. A Method for Disentangling Multiple Treatments from a Regression Discontinuity Design. *Journal of Labor Economics* . [Crossref]
10. Nikolaj Harmon, Raymond Fisman, Emir Kamenica. 2019. Peer Effects in Legislative Voting. *American Economic Journal: Applied Economics* **11**:4, 156-180. [Abstract] [View PDF article] [PDF with links]
11. Yu Zhang, Fei Qin, Juanjuan Liu. 2019. Improving education equality and quality: Evidence from a natural experiment in China. *International Journal of Educational Development* **70**, 102078. [Crossref]
12. Erica K. Chuang, Barbara S. Mensch, Stephanie R. Psaki, Nicole A. Haberland, Meredith L. Kozak. 2019. PROTOCOL: Policies and interventions to remove gender-related barriers to girls' school participation and learning in low- and middle-income countries: A systematic review of the evidence. *Campbell Systematic Reviews* **15**:3. . [Crossref]
13. Markus Baldauf, Joshua Mollner. 2019. Pedaling peers: The effect of targets on performance. *Journal of Economic Behavior & Organization* . [Crossref]
14. Malte Reichelt, Matthias Collischon, Andreas Eberl. 2019. School tracking and its role in social reproduction: reinforcing educational inheritance and the direct effects of social origin. *The British Journal of Sociology* **70**:4, 1323-1348. [Crossref]
15. Rebecca Dizon-Ross. 2019. Parents' Beliefs about Their Children's Academic Ability: Implications for Educational Investments. *American Economic Review* **109**:8, 2728-2765. [Abstract] [View PDF article] [PDF with links]
16. Eugen Dimant. 2019. Contagion of pro- and anti-social behavior among peers and the role of social proximity. *Journal of Economic Psychology* **73**, 66-88. [Crossref]
17. Isaac Mbiti, Karthik Muralidharan, Mauricio Romero, Youdi Schipper, Constantine Manda, Rakesh Rajani. 2019. Inputs, Incentives, and Complementarities in Education: Experimental Evidence from Tanzania*. *The Quarterly Journal of Economics* **134**:3, 1627-1673. [Crossref]

18. Prashant Loyalka, Sean Sylvia, Chengfang Liu, James Chu, Yaojiang Shi. 2019. Pay by Design: Teacher Performance Pay Design and the Distribution of Student Achievement. *Journal of Labor Economics* **37**:3, 621-662. [Crossref]
19. Ying Shi. 2019. Who benefits from selective education? Evidence from elite boarding school admissions. *Economics of Education Review* 101907. [Crossref]
20. Tanika Chakraborty, Simone Schüller, Klaus F. Zimmermann. 2019. Beyond the average: Ethnic capital heterogeneity and intergenerational transmission of education. *Journal of Economic Behavior & Organization* **163**, 551-569. [Crossref]
21. Luc Behaghel, Karen Macours, Julie Subervie. 2019. How can randomised controlled trials help improve the design of the common agricultural policy?. *European Review of Agricultural Economics* **46**:3, 473-493. [Crossref]
22. Leonardo Bursztyn, Georgy Egorov, Robert Jensen. 2019. Cool to be Smart or Smart to be Cool? Understanding Peer Pressure in Education. *The Review of Economic Studies* **86**:4, 1487-1526. [Crossref]
23. Moussa Pouguinimpo Blimpo, David K. Evans, Muthoni Ngatia. Developing Universal Foundational Skills in Sub-Saharan Africa 129-173. [Crossref]
24. Omar Arias, David K. Evans, Indhira Santos. Executive Summary 1-16. [Crossref]
25. Omar Arias, David K. Evans, Indhira Santos. Overview 17-70. [Crossref]
26. Amy Damon, Paul Glewwe, Suzanne Wisniewski, Bixuan Sun. 2019. What education policies and programmes affect learning and time in school in developing countries? A review of evaluations from 1990 to 2014. *Review of Education* **7**:2, 295-387. [Crossref]
27. Michael Kaganovich, Xuejuan Su. 2019. College curriculum, diverging selectivity, and enrollment expansion. *Economic Theory* **67**:4, 1019-1050. [Crossref]
28. Jia Wu, Xiangdong Wei, Hongliang Zhang, Xiang Zhou. 2019. Elite schools, magnet classes, and academic performances: Regression-discontinuity evidence from China. *China Economic Review* **55**, 143-167. [Crossref]
29. Marcel Fafchamps. Engines of Growth and Africa's Economic Performance Revisited 77-92. [Crossref]
30. Manisha Shah, Bryce Steinberg. 2019. The Right to Education Act: Trends in Enrollment, Test Scores, and School Quality. *AEA Papers and Proceedings* **109**, 232-238. [Abstract] [View PDF article] [PDF with links]
31. Paulina Aguinaga, Alessandra Cassar, Jennifer Graham, Lauren Skora, Bruce Wydick. 2019. Raising achievement among microentrepreneurs: An experimental test of goals, incentives, and support groups in Medellin, Colombia. *Journal of Economic Behavior & Organization* **161**, 79-97. [Crossref]
32. Ariel BenYishay, A Mushfiq Mobarak. 2019. Social Learning and Incentives for Experimentation and Communication. *The Review of Economic Studies* **86**:3, 976-1009. [Crossref]
33. Karthik Muralidharan, Abhijeet Singh, Alejandro J. Ganimian. 2019. Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India. *American Economic Review* **109**:4, 1426-1460. [Abstract] [View PDF article] [PDF with links]
34. Shi Min, Zhouhang Yuan, Xiaobing Wang, Lingling Hou. 2019. Do peer effects influence the academic performance of rural students at private migrant schools in China?. *China Economic Review* **54**, 418-433. [Crossref]
35. Tommaso Frattini, Elena Meschi. 2019. The effect of immigrant peers in vocational schools. *European Economic Review* **113**, 1-22. [Crossref]
36. Joseph Deutsch, Audrey Dumas, Jacques Silber. 2019. School externalities and scholastic performance: an efficiency analysis. *International Journal of Manpower* **40**:1, 102-119. [Crossref]

37. Youjin Hahn, Liang Choon Wang. 2019. The Effectiveness of Single-Sex Schools through Out-of-School Activities: Evidence from South Korea. *Oxford Bulletin of Economics and Statistics* **81**:2, 369-393. [Crossref]
38. S. Joel Warrican, Melissa L. Alleyne, Patriann Smith, Jehanzeb Cheema, James R. King. 2019. Peer Effects in the Individual and Group Literacy Achievement of High-School Students in a Bi-dialectal Context. *Reading Psychology* **40**:2, 117-148. [Crossref]
39. Facundo Albornoz, María Victoria Anauati, Melina Furman, Mariana Luzuriaga, María Eugenia Podestá, Inés Taylor. 2019. Training to Teach Science: Experimental Evidence from Argentina. *The World Bank Economic Review* **49**. . [Crossref]
40. Vincenzo Andrietti, Xuejuan Su. 2019. Education curriculum and student achievement: theory and evidence. *Education Economics* **27**:1, 4-19. [Crossref]
41. Johannes Hoogeveen, Marcello Matranga, Mariacristina Rossi. Student Learning and Teacher Competence 63-86. [Crossref]
42. Manjistha Banerji, Mansi Nanda. Till What Age Is "Age" Relevant? Examining the Effect of Age on Early Learning 35-62. [Crossref]
43. Yan Lau. 2019. The dragon cohort of Hong Kong: traditional beliefs, demographics, and education. *Journal of Population Economics* **32**:1, 219-246. [Crossref]
44. Asad Islam. 2019. Parent–teacher meetings and student outcomes: Evidence from a developing country. *European Economic Review* **111**, 273-304. [Crossref]
45. Gabrielle Wills, Heleen Hofmeyr. 2019. Academic resilience in challenging contexts: Evidence from township and rural primary schools in South Africa. *International Journal of Educational Research* **98**, 192-205. [Crossref]
46. Hans Rawhouser, Michael Cummings, Scott L. Newbert. 2019. Social Impact Measurement: Current Approaches and Future Directions for Social Entrepreneurship Research. *Entrepreneurship Theory and Practice* **43**:1, 82-115. [Crossref]
47. Laura Panza. 2019. The Impact of Ethnic Segregation on Schooling Outcomes in Mandate Palestine. *SSRN Electronic Journal* . [Crossref]
48. Tianheng Wang. 2019. Classroom Composition and Student Academic Achievement: The Impact of Peers' Parental Education and Peers' Gender. *SSRN Electronic Journal* . [Crossref]
49. Baeksan Yu, Sean Kelly. 2019. The non-cognitive returns to vocational school tracking: South Korean evidence. *International Journal of Educational Research* **98**, 379-394. [Crossref]
50. Guy Grossman, Melina R. Platas, Jonathan Rodden. 2018. Crowdsourcing accountability: ICT for service delivery. *World Development* **112**, 74-87. [Crossref]
51. Harry Bett, Lazarus Makewa. 2018. Can Facebook groups enhance continuing professional development of teachers? Lessons from Kenya. *Asia-Pacific Journal of Teacher Education* **3**, 1-15. [Crossref]
52. Sangyoon Park. 2018. Coeducation, academic performance, and subject choice: evidence from quasi-random classroom assignments. *Education Economics* **26**:6, 574-592. [Crossref]
53. Ciro Avitabile, Rafael de Hoyos. 2018. The heterogeneous effect of information on student performance: Evidence from a randomized control trial in Mexico. *Journal of Development Economics* **135**, 318-348. [Crossref]
54. Fan Li, Prashant Loyalka, Hongmei Yi, Yaojiang Shi, Natalie Johnson, Scott Rozelle. 2018. Ability tracking and social trust in China's rural secondary school system. *School Effectiveness and School Improvement* **29**:4, 545-572. [Crossref]
55. Stephen S Lim, Rachel L Updike, Alexander S Kaldjian, Ryan M Barber, Krycia Cowling, Hunter York, Joseph Friedman, R Xu, Joanna L Whisnant, Heather J Taylor, Andrew T Leever, Yesenia

Roman, Miranda F Bryant, Joseph Dieleman, Emmanuela Gakidou, Christopher J L Murray. 2018. Measuring human capital: a systematic analysis of 195 countries and territories, 1990–2016. *The Lancet* **392**:10154, 1217-1234. [Crossref]

56. Duncan McVicar, Julie Moschion, Chris Ryan. 2018. Achievement effects from new peers: Who matters to whom?. *Economics of Education Review* **66**, 154-166. [Crossref]
57. Christian Dustmann, Hyejin Ku, Do Won Kwak. 2018. Why Are Single-Sex Schools Successful?. *Labour Economics* **54**, 79-99. [Crossref]
58. Pieter Gautier, Paul Muller, Bas van der Klaauw, Michael Rosholm, Michael Svarer. 2018. Estimating Equilibrium Effects of Job Search Assistance. *Journal of Labor Economics* **36**:4, 1073-1125. [Crossref]
59. Chao Fu, Nirav Mehta. 2018. Ability Tracking, School and Parental Effort, and Student Achievement: A Structural Model and Estimation. *Journal of Labor Economics* **36**:4, 923-979. [Crossref]
60. Sajitha Bashir, Marlaine Lockheed, Elizabeth Ninan, Jee-Peng Tan. Deploying the Budget to Improve Quality 341-404. [Crossref]
61. Steve Cicala, Roland G. Fryer, Jörg L. Spenkuch. 2018. Self-Selection and Comparative Advantage in Social Interactions. *Journal of the European Economic Association* **16**:4, 983-1020. [Crossref]
62. Arlen Guarín, Carlos Medina, Christian Posso. 2018. Calidad, cobertura y costos ocultos de la educación secundaria pública y privada en Colombia. *Revista Desarrollo y Sociedad* :81, 61-114. [Crossref]
63. Donald R. Baum, Husein Abdul-Hamid, Hugo T. Wesley. 2018. Inequality of educational opportunity: the relationship between access, affordability, and quality of private schools in Lagos, Nigeria. *Oxford Review of Education* **44**:4, 459-475. [Crossref]
64. Robert Garlick. 2018. Academic Peer Effects with Different Group Assignment Policies: Residential Tracking versus Random Assignment. *American Economic Journal: Applied Economics* **10**:3, 345-369. [Abstract] [View PDF article] [PDF with links]
65. Silvia Mendolia, Alfredo R Paloyo, Ian Walker. 2018. Heterogeneous effects of high school peers on educational outcomes. *Oxford Economic Papers* **70**:3, 613-634. [Crossref]
66. Marcel Fafchamps, Di Mo. 2018. Peer effects in computer assisted learning: evidence from a randomized experiment. *Experimental Economics* **21**:2, 355-382. [Crossref]
67. Jamie Emerson, Brian Hill. 2018. Peer effects in marathon racing: The role of pace setters. *Labour Economics* **52**, 74-82. [Crossref]
68. Benjamin Piper, Stephanie Simmons Zuilkowski, Margaret Dubeck, Evelyn Jepkemei, Simon J. King. 2018. Identifying the essential ingredients to literacy and numeracy improvement: Teacher professional development and coaching, student textbooks, and structured teachers' guides. *World Development* **106**, 324-336. [Crossref]
69. Timothy Powell-Jackson, Calum Davey, Edoardo Masset, Shari Krishnaratne, Richard Hayes, Kara Hanson, James R Hargreaves. 2018. Trials and tribulations: cross-learning from the practices of epidemiologists and economists in the evaluation of public health interventions. *Health Policy and Planning* **33**:5, 702-706. [Crossref]
70. Francesca Borgonovi, Maciej Jakubowski, Artur Pokropek. 2018. Will the last be first and the first last? The role of classroom registers in cognitive skill acquisition. *PLOS ONE* **13**:5, e0197746. [Crossref]
71. David Kiss. 2018. How do ability peer effects operate? Evidence on one transmission channel. *Education Economics* **26**:3, 253-265. [Crossref]
72. Youjin Hahn, Liang Choon Wang, Hee-Seung Yang. 2018. Does greater school autonomy make a difference? Evidence from a randomized natural experiment in South Korea. *Journal of Public Economics* **161**, 15-30. [Crossref]

73. Yulia Kuzmina, Alina Ivanova. 2018. The effects of academic class composition on academic progress in elementary school for students with different levels of initial academic abilities. *Learning and Individual Differences* **64**, 43-53. [Crossref]

74. Todd Pugatch, Elizabeth Schroeder. 2018. Teacher pay and student performance: evidence from the Gambian hardship allowance. *Journal of Development Effectiveness* **10**:2, 249-276. [Crossref]

75. Haining Wang, Zhiming Cheng, Russell Smyth. 2018. Do migrant students affect local students' academic achievements in urban China?. *Economics of Education Review* **63**, 64-77. [Crossref]

76. Facundo Albornoz, Samuel Berlinski, Antonio Cabrales. 2018. Motivation, resources, and the organization of the school system. *Journal of the European Economic Association* **16**:1, 199-231. [Crossref]

77. Jörg Peters, Jörg Langbein, Gareth Roberts. 2018. Generalization in the Tropics – Development Policy, Randomized Controlled Trials, and External Validity. *The World Bank Research Observer* **33**:1, 34-64. [Crossref]

78. Segismundo S. Izquierdo, Luis R. Izquierdo, Dunia López-Pintado. 2018. Mixing and diffusion in a two-type population. *Royal Society Open Science* **5**:2, 172102. [Crossref]

79. Jiaying Gu, Shu Shen. 2018. Oracle and adaptive false discovery rate controlling methods for one-sided testing: theory and application in treatment effect evaluation. *The Econometrics Journal* **21**:1, 11-35. [Crossref]

80. Claudio Allende, Rocío Díaz, Juan Pablo Valenzuela. School Segregation in Chile 5544-5554. [Crossref]

81. Claudio Allende, Rocío Díaz, Juan Pablo Valenzuela. School Segregation in Chile 1-11. [Crossref]

82. Martin Foureaux Koppensteiner. 2018. Relative Age, Class Assignment, and Academic Performance: Evidence from Brazilian Primary Schools. *The Scandinavian Journal of Economics* **120**:1, 296-325. [Crossref]

83. Federico Gonzalez. 2018. Impactos De Corto Plazo Del Programa Extracurricular De Refuerzo Escolar 'Con Las Manos' En Un Colegio De Bogott (The Short-Term Effects of the Extracurricular Program Con Las Manos in a School in Bogott, Colombia). *SSRN Electronic Journal* . [Crossref]

84. Rebecca DizonnRoss. 2018. Parents' Beliefs About Their Children's Academic Abilities: Implications for Educational Investments. *SSRN Electronic Journal* . [Crossref]

85. Brendan Houng, Chris Ryan. 2018. Achievement Gains from Attendance at Selective High Schools. *SSRN Electronic Journal* . [Crossref]

86. Youjin Hahn, Asadul Islam, Eleonora Patacchini, Yves Zenou. 2018. Friendship and Female Education: Evidence From a Field Experiment in Bangladeshi Primary Schools. *SSRN Electronic Journal* . [Crossref]

87. Michela Carlana, Eliana La Ferrara, Paolo Pinotti. 2018. Goals and Gaps: Educational Careers of Immigrant Children. *SSRN Electronic Journal* . [Crossref]

88. Eugen Dimant. 2018. Contagion of Pro- and Anti-Social Behavior Among Peers and the Role of Social Proximity. *SSRN Electronic Journal* . [Crossref]

89. Sönke Hendrik Matthewes. 2018. Better Together? Heterogeneous Effects of Tracking on Student Achievement. *SSRN Electronic Journal* . [Crossref]

90. Rosangela Bando, Francisco Gallego, Paul Gertler, Dario Romero Fonseca. 2017. Books or laptops? The effect of shifting from printed to digital delivery of educational content on learning. *Economics of Education Review* **61**, 162-173. [Crossref]

91. Simon Lange, Marten von Werder. 2017. Tracking and the intergenerational transmission of education: Evidence from a natural experiment. *Economics of Education Review* **61**, 59-78. [Crossref]

92. ###. 2017. Peer effects on academic achievement in high school Society#Culture course: focusing on the difference between strong pupils and weak ones. *SECONDARY EDUCATION RESEARCH* **65**:4, 933-956. [Crossref]
93. Abhijit Banerjee, Rukmini Banerji, James Berry, Esther Duflo, Harini Kannan, Shobhini Mukerji, Marc Shotland, Michael Walton. 2017. From Proof of Concept to Scalable Policies: Challenges and Solutions, with an Application. *Journal of Economic Perspectives* **31**:4, 73-102. [Abstract] [View PDF article] [PDF with links]
94. Lee Crawfurd. 2017. School Management and Public–Private Partnerships in Uganda. *Journal of African Economies* **26**:5, 539-560. [Crossref]
95. . Teacher skills and motivation both matter (though many education systems act like they don't) 131-144. [Crossref]
96. . Overview: Learning to realize education's promise 1-35. [Crossref]
97. Timothy C. Salmon, Danila Serra. 2017. Corruption, social judgment and culture: An experiment. *Journal of Economic Behavior & Organization* **142**, 64-78. [Crossref]
98. Katharine M. Conn. 2017. Identifying Effective Education Interventions in Sub-Saharan Africa: A Meta-Analysis of Impact Evaluations. *Review of Educational Research* **87**:5, 863-898. [Crossref]
99. Taeko Okitsu, D. Brent Edwards. 2017. Policy promise and the reality of community involvement in school-based management in Zambia: Can the rural poor hold schools and teachers to account?. *International Journal of Educational Development* **56**, 28-41. [Crossref]
100. Christian Dustmann, Patrick A. Puhani, Uta Schönberg. 2017. The Long-Term Effects of Early Track Choice. *The Economic Journal* **127**:603, 1348-1380. [Crossref]
101. Thurston Domina, Andrew Penner, Emily Penner. 2017. Categorical Inequality: Schools As Sorting Machines. *Annual Review of Sociology* **43**:1, 311-330. [Crossref]
102. David Kiss. 2017. A Model about the Impact of Ability Grouping on Student Achievement. *The B.E. Journal of Economic Analysis & Policy* **17**:3. . [Crossref]
103. Julian R. Betts, Youjin Hahn, Andrew C. Zau. 2017. Can testing improve student learning? An evaluation of the mathematics diagnostic testing project. *Journal of Urban Economics* **100**, 54-64. [Crossref]
104. Dennis Epple, Richard E. Romano, Miguel Urquiola. 2017. School Vouchers: A Survey of the Economics Literature. *Journal of Economic Literature* **55**:2, 441-492. [Abstract] [View PDF article] [PDF with links]
105. Nienke Ruijs. 2017. The impact of special needs students on classmate performance. *Economics of Education Review* **58**, 15-31. [Crossref]
106. Patricia Navarro-Palau. 2017. Effects of differentiated school vouchers: Evidence from a policy change and date of birth cutoffs. *Economics of Education Review* **58**, 86-107. [Crossref]
107. Ricardo Estrada, Jérémie Gignoux. 2017. Benefits to elite schools and the expected returns to education: Evidence from Mexico City. *European Economic Review* **95**, 168-194. [Crossref]
108. Bénédicte de la Brière, Deon Filmer, Dena Ringold, Dominic Rohner, Karelle Samuda, Anastasiya Denisova. Key Investments to Build the Foundations of Human Capital 143-181. [Crossref]
109. Jan Feld, Ulf Zölitz. 2017. Understanding Peer Effects: On the Nature, Estimation, and Channels of Peer Effects. *Journal of Labor Economics* **35**:2, 387-428. [Crossref]
110. Tommaso Agasisti, Patrizia Falzetti. 2017. Between-classes sorting within schools and test scores: an empirical analysis of Italian junior secondary schools. *International Review of Economics* **64**:1, 1-45. [Crossref]

111. Robert Akerlof. 2017. Value Formation: The Role of Esteem. *Games and Economic Behavior* **102**, 1-19. [Crossref]

112. Chris Ryan. 2017. Measurement of Peer Effects. *Australian Economic Review* **50**:1, 121-129. [Crossref]

113. Kalinca Léia Becker. 2017. O efeito da interação social entre os jovens nas decisões de consumo de álcool, cigarros e outras drogas ilícitas. *Estudos Econômicos (São Paulo)* **47**:1, 65-92. [Crossref]

114. Sok Chul Hong, Jungmin Lee. 2017. Who is sitting next to you? Peer effects inside the classroom. *Quantitative Economics* **8**:1, 239-275. [Crossref]

115. Joseph R. Cummins. 2017. Heterogeneous treatment effects in the low track: Revisiting the Kenyan primary school experiment. *Economics of Education Review* **56**, 40-51. [Crossref]

116. Olga Yakusheva. 2017. Health Spillovers among Hospital Patients: Evidence from Roommate Assignments. *American Journal of Health Economics* **3**:1, 76-107. [Crossref]

117. Luc Behaghel, Clément de Chaisemartin, Marc Gurgand. 2017. Ready for Boarding? The Effects of a Boarding School for Disadvantaged Students. *American Economic Journal: Applied Economics* **9**:1, 140-164. [Abstract] [View PDF article] [PDF with links]

118. K. Muralidharan. Field Experiments in Education in Developing Countries 323-385. [Crossref]

119. Pelin Akyol, Kala Krishna. 2017. Preferences, selection, and value added: A structural approach. *European Economic Review* **91**, 89-117. [Crossref]

120. Luana Marotta. 2017. Peer effects in early schooling: Evidence from Brazilian primary schools. *International Journal of Educational Research* **82**, 110-123. [Crossref]

121. Bastian Ravesteijn, Hans van Kippersluis, Pekka Martikainen, Hannu Vessari. 2017. The Impact of Later Tracking on Mortality by Parental Income in Finland. *SSRN Electronic Journal* . [Crossref]

122. Erik O. Kimbrough, Hitoshi Shigeoka. 2017. How Do Peers Impact Learning? An Experimental Investigation of Peer-to-Peer Teaching and Ability Tracking. *SSRN Electronic Journal* . [Crossref]

123. Marcus Roller. 2017. The Distributional Effects of Early School Stratification - Non-Parametric Evidence from Germany. *SSRN Electronic Journal* . [Crossref]

124. Kathrin Thiemann, Niklas Wallmeier. 2017. An Experiment on Peer Effects Under Different Relative Performance Feedback and Grouping Procedures. *SSRN Electronic Journal* . [Crossref]

125. Paul T. von Hippel, Joseph Workman, Douglas B. Downey. 2017. Are Schools (Still) a Great Equalizer? Replicating a Summer Learning Study Using Better Test Scores and a New Cohort of Children. *SSRN Electronic Journal* . [Crossref]

126. Andrew McEachin, Thurston Domina, Andrew M. Penner. 2017. Understanding the Effects of Middle School Algebra: A Regression Discontinuity Approach. *SSRN Electronic Journal* . [Crossref]

127. Mauricio Romero, Justin Sandefur, Wayne Aaron Sandholtz. 2017. Can Outsourcing Improve Liberia's Schools? Preliminary Results from Year One of a Three-Year Randomized Evaluation of Partnership Schools for Liberia. *SSRN Electronic Journal* . [Crossref]

128. Tomasz Obloj, Cedric Gutierrez, Frank Douglas. 2017. Better to Have Led and Lost than Never to Have Led at All? Competitive Dethronement, the Endowment Effect, and Risk Taking. *SSRN Electronic Journal* . [Crossref]

129. Tommaso Frattini, Elena Meschi. 2017. The Effect of Immigrant Peers in Vocational Schools. *SSRN Electronic Journal* . [Crossref]

130. Han Feng, Jiayao Li. 2016. Head teachers, peer effects, and student achievement. *China Economic Review* **41**, 268-283. [Crossref]

131. Edwin Leuven, Erik Plug, Marte Rønning. 2016. Education and cancer risk. *Labour Economics* **43**, 106-121. [Crossref]

132. Ben Kelcey, Zuchao Shen, Jessaca Spybrook. 2016. Intraclass Correlation Coefficients for Designing Cluster-Randomized Trials in Sub-Saharan Africa Education. *Evaluation Review* **40**:6, 500-525. [Crossref]
133. Marcel Fafchamps, Simon Quinn. 2016. Networks and Manufacturing Firms in Africa: Results from a Randomized Field Experiment. *The World Bank Economic Review* **3**, lhw057. [Crossref]
134. . Reductions in Inequality: A Policy Perspective 129-170. [Crossref]
135. David Card, Laura Giuliano. 2016. Can Tracking Raise the Test Scores of High-Ability Minority Students?. *American Economic Review* **106**:10, 2783-2816. [Abstract] [View PDF article] [PDF with links]
136. Anders Böhlmark, Helena Holmlund, Mikael Lindahl. 2016. Parental choice, neighbourhood segregation or cream skimming? An analysis of school segregation after a generalized choice reform. *Journal of Population Economics* **29**:4, 1155-1190. [Crossref]
137. Hongliang Zhang. 2016. The role of testing noise in the estimation of achievement-based peer effects. *Economics of Education Review* **54**, 113-123. [Crossref]
138. Hongliang Zhang. 2016. Identification of treatment effects under imperfect matching with an application to Chinese elite schools. *Journal of Public Economics* **142**, 56-82. [Crossref]
139. Ernesto Treviño, Juan Pablo Valenzuela, Cristóbal Villalobos. 2016. Within-school segregation in the Chilean school system: What factors explain it? How efficient is this practice for fostering student achievement and equity?. *Learning and Individual Differences* **51**, 367-375. [Crossref]
140. Shu Shen, Xiaohan Zhang. 2016. Distributional Tests for Regression Discontinuity: Theory and Empirical Examples. *Review of Economics and Statistics* **98**:4, 685-700. [Crossref]
141. Adam S. Booij, Edwin Leuven, Hessel Oosterbeek. 2016. Ability Peer Effects in University: Evidence from a Randomized Experiment. *The Review of Economic Studies* rdw045. [Crossref]
142. Kathryn Anderson, Xue Gong, Kai Hong, Xi Zhang. 2016. Do selective high schools improve student achievement? Effects of exam schools in China. *China Economic Review* **40**, 121-134. [Crossref]
143. Alejandro J. Ganimian, Richard J. Murnane. 2016. Improving Education in Developing Countries. *Review of Educational Research* **86**:3, 719-755. [Crossref]
144. Isaac M. Mbiti. 2016. The Need for Accountability in Education in Developing Countries. *Journal of Economic Perspectives* **30**:3, 109-132. [Abstract] [View PDF article] [PDF with links]
145. M. Caridad Araujo, Pedro Carneiro, Yyannú Cruz-Aguayo, Norbert Schady. 2016. Teacher Quality and Learning Outcomes in Kindergarten. *The Quarterly Journal of Economics* **131**:3, 1415-1453. [Crossref]
146. Lars J. Kirkeboen, Edwin Leuven, Magne Mogstad. 2016. Field of Study, Earnings, and Self-Selection. *The Quarterly Journal of Economics* **131**:3, 1057-1111. [Crossref]
147. David K. Evans, Anna Popova. 2016. What Really Works to Improve Learning in Developing Countries? An Analysis of Divergent Findings in Systematic Reviews. *The World Bank Research Observer* **31**:2, 242-270. [Crossref]
148. Stephen Gibbons, Shqiponja Telhaj. 2016. Peer Effects: Evidence from Secondary School Transition in England. *Oxford Bulletin of Economics and Statistics* **78**:4, 548-575. [Crossref]
149. Shawanda Stockfelt. 2016. Economic, social and embodied cultural capitals as shapers and predictors of boys' educational aspirations. *The Journal of Educational Research* **109**:4, 351-359. [Crossref]
150. James Cowan, Dan Goldhaber. 2016. National Board Certification and Teacher Effectiveness: Evidence From Washington State. *Journal of Research on Educational Effectiveness* **9**:3, 233-258. [Crossref]
151. Elizabeth U. Cascio, Diane Whitmore Schanzenbach. 2016. First in the Class? Age and the Education Production Function. *Education Finance and Policy* **11**:3, 225-250. [Crossref]

152. Surendrakumar Bagde, Dennis Epple, Lowell Taylor. 2016. Does Affirmative Action Work? Caste, Gender, College Quality, and Academic Success in India. *American Economic Review* **106**:6, 1495-1521. [Abstract] [View PDF article] [PDF with links]
153. Edwin Leuven, Marte Rønning. 2016. Classroom Grade Composition and Pupil Achievement. *The Economic Journal* **126**:593, 1164-1192. [Crossref]
154. Takako Nomi, Stephen W. Raudenbush. 2016. Making a Success of "Algebra for All". *Educational Evaluation and Policy Analysis* **38**:2, 431-451. [Crossref]
155. Alexia Delfino, Luigi Marengo, Matteo Ploner. 2016. I did it your way. An experimental investigation of peer effects in investment choices. *Journal of Economic Psychology* **54**, 113-123. [Crossref]
156. Erica Field, Seema Jayachandran, Rohini Pande, Natalia Rigol. 2016. Friendship at Work: Can Peer Effects Catalyze Female Entrepreneurship?. *American Economic Journal: Economic Policy* **8**:2, 125-153. [Abstract] [View PDF article] [PDF with links]
157. Sam Jones. 2016. How does classroom composition affect learning outcomes in Ugandan primary schools?. *International Journal of Educational Development* **48**, 66-78. [Crossref]
158. Jun Xiao. 2016. Asymmetric all-pay contests with heterogeneous prizes. *Journal of Economic Theory* **163**, 178-221. [Crossref]
159. Michèle Belot, Jonathan James. 2016. Partner selection into policy relevant field experiments. *Journal of Economic Behavior & Organization* **123**, 31-56. [Crossref]
160. Adrien Bouguen. 2016. Adjusting content to individual student needs: Further evidence from an in-service teacher training program. *Economics of Education Review* **50**, 90-112. [Crossref]
161. William C. Horrace, Xiaodong Liu, Eleonora Patacchini. 2016. Endogenous network production functions with selectivity. *Journal of Econometrics* **190**:2, 222-232. [Crossref]
162. Ryan Yeung, Phuong Nguyen-Hoang. 2016. Endogenous peer effects: Fact or fiction?. *The Journal of Educational Research* **109**:1, 37-49. [Crossref]
163. Damon Clark, Emilia Del Bono. 2016. The Long-Run Effects of Attending an Elite School: Evidence from the United Kingdom. *American Economic Journal: Applied Economics* **8**:1, 150-176. [Abstract] [View PDF article] [PDF with links]
164. M. Urquiola. Competition Among Schools 209-237. [Crossref]
165. P. Glewwe, K. Muralidharan. Improving Education Outcomes in Developing Countries 653-743. [Crossref]
166. Behrang Kamali-Shahdadi. 2016. Sorting and Peer Effects. *SSRN Electronic Journal* . [Crossref]
167. Renata Rabovic, Pavel Cizek. 2016. Estimation of Spatial Sample Selection Models: A Partial Maximum Likelihood Approach. *SSRN Electronic Journal* . [Crossref]
168. Duncan McVicar, Julie Moschion, Chris Ryan. 2016. Achievement Effects from New Peers: Who Matters to Whom?. *SSRN Electronic Journal* . [Crossref]
169. Vincenzo Andrietti, Xuejuan Su. 2016. Education Curriculum and Student Achievement: Theory and Evidence. *SSRN Electronic Journal* . [Crossref]
170. Prashant Kumar Loyalka, Sean Sylvia, Chengfang Liu, James Chu, Yaojiang Shi. 2016. Pay by Design: Teacher Performance Pay Design and the Distribution of Student Achievement. *SSRN Electronic Journal* . [Crossref]
171. Robert Garlick. 2016. Academic Peer Effects with Different Group Assignment Policies: Residential Tracking versus Random Assignment. *SSRN Electronic Journal* . [Crossref]
172. Abhijit V. Banerjee, Rukmini Banerji, James Berry, Harini Kannan, Shobhini Mukerji, Michael Walton. 2016. Mainstreaming an Effective Intervention: Evidence from Randomized Evaluations of 'Teaching at the Right Level' in India. *SSRN Electronic Journal* . [Crossref]

173. Abhijit V. Banerjee, Rukmini Banerji, James Berry, Harini Kannan, Shobhini Mukerji, Michael Walton. 2016. From Proof of Concept to Scalable Policies: Challenges and Solutions, with an Application. *SSRN Electronic Journal* . [Crossref]
174. Simon Lange, Marten von Werder. 2016. Tracking and the Intergenerational Transmission of Education: Evidence from a Natural Experiment. *SSRN Electronic Journal* . [Crossref]
175. Facundo Albornoz, Samuel Berlinski, Antonio Cabrales. 2016. Motivation, Resources and the Organization of the School System. *SSRN Electronic Journal* . [Crossref]
176. Miguel Urquiola. 2015. Progress and challenges in achieving an evidence-based education policy in Latin America and the Caribbean. *Latin American Economic Review* **24**:1. . [Crossref]
177. Arna Vardardottir. 2015. The impact of classroom peers in a streaming system. *Economics of Education Review* **49**, 110-128. [Crossref]
178. Feng Hu. 2015. Do girl peers improve your academic performance?. *Economics Letters* **137**, 54-58. [Crossref]
179. Albert Park, Xinzheng Shi, Chang-tai Hsieh, Xuehui An. 2015. Magnet high schools and academic performance in China: A regression discontinuity design. *Journal of Comparative Economics* **43**:4, 825-843. [Crossref]
180. James Fenske. 2015. African polygamy: Past and present. *Journal of Development Economics* **117**, 58-73. [Crossref]
181. Seth Richards-Shubik. 2015. Peer effects in sexual initiation: Separating demand and supply mechanisms. *Quantitative Economics* **6**:3, 663-702. [Crossref]
182. Klára Gurzó, Dániel Horn. 2015. A korai iskolai szelekció hosszú távú hatása. *Közgazdasági Szemle* **62**:10, 1070-1096. [Crossref]
183. Lant Pritchett, Martina Viarengo. 2015. Does public sector control reduce variance in school quality?. *Education Economics* **23**:5, 557-576. [Crossref]
184. Patrick J. McEwan. 2015. Improving Learning in Primary Schools of Developing Countries. *Review of Educational Research* **85**:3, 353-394. [Crossref]
185. Thomas Dee, Xiaohuan Lan. 2015. The achievement and course-taking effects of magnet schools: Regression-discontinuity evidence from urban China. *Economics of Education Review* **47**, 128-142. [Crossref]
186. Leonardo Bursztyn, Robert Jensen. 2015. How Does Peer Pressure Affect Educational Investments?. *The Quarterly Journal of Economics* **130**:3, 1329-1367. [Crossref]
187. Rachel Heath, A. Mushfiq Mobarak. 2015. Manufacturing growth and the lives of Bangladeshi women. *Journal of Development Economics* **115**, 1-15. [Crossref]
188. Yann Girard, Florian Hett, Daniel Schunk. 2015. How individual characteristics shape the structure of social networks. *Journal of Economic Behavior & Organization* **115**, 197-216. [Crossref]
189. Takako Nomi. 2015. "Double-dose" English as a strategy for improving adolescent literacy: Total effect and mediated effect through classroom peer ability change. *Social Science Research* **52**, 716-739. [Crossref]
190. Andrew M. Penner, Thurston Domina, Emily K. Penner, AnneMarie Conley. 2015. Curricular policy as a collective effects problem: A distributional approach. *Social Science Research* **52**, 627-641. [Crossref]
191. Grant Miller, A. Mushfiq Mobarak. 2015. Learning About New Technologies Through Social Networks: Experimental Evidence on Nontraditional Stoves in Bangladesh. *Marketing Science* **34**:4, 480-499. [Crossref]

192. Maciej Jakubowski. 2015. Latent variables and propensity score matching: a simulation study with application to data from the Programme for International Student Assessment in Poland. *Empirical Economics* **48**:3, 1287-1325. [Crossref]

193. Fabian T. Pfeffer. 2015. Equality and quality in education. A comparative study of 19 countries. *Social Science Research* **51**, 350-368. [Crossref]

194. Jon Marius Vaag Iversen, Hans Bonesrønning. 2015. Conditional gender peer effects?. *Journal of Behavioral and Experimental Economics* **55**, 19-28. [Crossref]

195. Esther Duflo, Pascaline Dupas, Michael Kremer. 2015. School governance, teacher incentives, and pupil–teacher ratios: Experimental evidence from Kenyan primary schools. *Journal of Public Economics* **123**, 92-110. [Crossref]

196. Tarun Jain, Mudit Kapoor. 2015. The Impact of Study Groups and Roommates on Academic Performance. *Review of Economics and Statistics* **97**:1, 44-54. [Crossref]

197. Amrei M. Lahno, Marta Serra-Garcia. 2015. Peer effects in risk taking: Envy or conformity?. *Journal of Risk and Uncertainty* **50**:1, 73-95. [Crossref]

198. Liang Choon Wang. 2015. All work and no play? The effects of ability sorting on students' non-school inputs, time use, and grade anxiety. *Economics of Education Review* **44**, 29-41. [Crossref]

199. Kangoh Lee. 2015. Higher education expansion, tracking, and student effort. *Journal of Economics* **114**:1, 1-22. [Crossref]

200. Lant Pritchett, Amanda Beatty. 2015. Slow down, you're going too fast: Matching curricula to student skill levels. *International Journal of Educational Development* **40**, 276-288. [Crossref]

201. Qian Tang, Andrew B. Whinston. Improving Internet Security through Mandatory Information Disclosure 4813-4823. [Crossref]

202. Jun Xiao. 2015. Asymmetric All-Pay Contests with Heterogeneous Prizes. *SSRN Electronic Journal* . [Crossref]

203. Jun Xiao. 2015. Ability Grouping in All-Pay Contests. *SSRN Electronic Journal* . [Crossref]

204. Youjin Hahn, Liang Choon Wang. 2015. The Mediating Effects of Students' Out-of-School Activities on Academic Achievement in Single-Sex and Private High Schools: Evidence from Korea. *SSRN Electronic Journal* . [Crossref]

205. Maresa Sprietsma, Lisa Pfeil. 2015. Peer Effects in Language Training for Migrants. *SSRN Electronic Journal* . [Crossref]

206. Kathryn H. Anderson, Xue Gong, Kai Hong, Xi Zhang. 2015. Do Selective High Schools Improve Student Achievement? Effects of Exam Schools in China. *SSRN Electronic Journal* . [Crossref]

207. Antonia Grohmann, Sahra Sakha. 2015. The Effect of Peer Observation on Consumption Choices: Experimental Evidence. *SSRN Electronic Journal* . [Crossref]

208. Markus Baldauf, Joshua Mollner. 2015. Peloton Peers: The Effect of Targets on Performance. *SSRN Electronic Journal* . [Crossref]

209. Marinho Bertanha. 2015. Regression Discontinuity Design with Many Thresholds. *SSRN Electronic Journal* . [Crossref]

210. Tisorn Songsermsawas. 2015. Factors Affecting Farm Productivity in Rural India: Social Networks and Market Access. *SSRN Electronic Journal* . [Crossref]

211. Andrew J. Hill. 2014. The costs of failure: Negative externalities in high school course repetition. *Economics of Education Review* **43**, 91-105. [Crossref]

212. Ivy A. Kodzi, Moses Oketch, Moses W. Ngware, Maurice Mutisya, Evangeline N. Nderu. 2014. Social relations as predictors of achievement in math in Kenyan primary schools. *International Journal of Educational Development* **39**, 275-282. [Crossref]

213. Moussa P. Blimpo. 2014. Team Incentives for Education in Developing Countries: A Randomized Field Experiment in Benin. *American Economic Journal: Applied Economics* **6**:4, 90-109. [Abstract] [View PDF article] [PDF with links]
214. Marc Piopiunik. 2014. The effects of early tracking on student performance: Evidence from a school reform in Bavaria. *Economics of Education Review* **42**, 12-33. [Crossref]
215. Joshua D. Angrist. 2014. The perils of peer effects. *Labour Economics* **30**, 98-108. [Crossref]
216. Eric Taylor. 2014. Spending more of the school day in math class: Evidence from a regression discontinuity in middle school. *Journal of Public Economics* **117**, 162-181. [Crossref]
217. Sa A. Bui, Steven G. Craig, Scott A. Imberman. 2014. Is Gifted Education a Bright Idea? Assessing the Impact of Gifted and Talented Programs on Students. *American Economic Journal: Economic Policy* **6**:3, 30-62. [Abstract] [View PDF article] [PDF with links]
218. Todd Pugatch, Elizabeth Schroeder. 2014. Incentives for teacher relocation: Evidence from the Gambian hardship allowance. *Economics of Education Review* **41**, 120-136. [Crossref]
219. Adrienne M. Lucas, Isaac M. Mbiti. 2014. Effects of School Quality on Student Achievement: Discontinuity Evidence from Kenya. *American Economic Journal: Applied Economics* **6**:3, 234-263. [Abstract] [View PDF article] [PDF with links]
220. Will Dobbie, Roland G. Fryer, Jr.. 2014. The Impact of Attending a School with High-Achieving Peers: Evidence from the New York City Exam Schools. *American Economic Journal: Applied Economics* **6**:3, 58-75. [Abstract] [View PDF article] [PDF with links]
221. Fangwen Lu. 2014. Testing peer effects among college students: evidence from an unusual admission policy change in China. *Asia Pacific Education Review* **15**:2, 257-270. [Crossref]
222. Kalena E. Cortes, Joshua S. Goodman. 2014. Ability-Tracking, Instructional Time, and Better Pedagogy: The Effect of Double-Dose Algebra on Student Achievement. *American Economic Review* **104**:5, 400-405. [Abstract] [View PDF article] [PDF with links]
223. Abigail Barr, Truman Packard, Danila Serra. 2014. Participatory accountability and collective action: Experimental evidence from Albania. *European Economic Review* **68**, 250-269. [Crossref]
224. Alejandro Drexler, Greg Fischer, Antoinette Schoar. 2014. Keeping It Simple: Financial Literacy and Rules of Thumb. *American Economic Journal: Applied Economics* **6**:2, 1-31. [Abstract] [View PDF article] [PDF with links]
225. Mingming Ma, Xinzheng Shi. 2014. Magnet Classes and Educational Performance: Evidence from China. *Economic Development and Cultural Change* **62**:3, 537-566. [Crossref]
226. Nicole Schneeweis, Martina Zweimüller. 2014. Early Tracking and the Misfortune of Being Young. *The Scandinavian Journal of Economics* **116**:2, 394-428. [Crossref]
227. Tao Li, Li Han, Linxiu Zhang, Scott Rozelle. 2014. Encouraging classroom peer interactions: Evidence from Chinese migrant schools. *Journal of Public Economics* **111**, 29-45. [Crossref]
228. Guillaume Roels, Xuanming Su. 2014. Optimal Design of Social Comparison Effects: Setting Reference Groups and Reference Points. *Management Science* **60**:3, 606-627. [Crossref]
229. Christian Helmers, Manasa Patnam. 2014. Does the rotten child spoil his companion? Spatial peer effects among children in rural India. *Quantitative Economics* **5**:1, 67-121. [Crossref]
230. Julide Yildirim, Servet Ozdemir, Ferudun Sezgin. 2014. A Qualitative Evaluation of a Conditional Cash Transfer Program in Turkey: The Beneficiaries' and Key Informants' Perspectives. *Journal of Social Service Research* **40**:1, 62-79. [Crossref]
231. F. Avvisati, M. Gurgand, N. Guyon, E. Maurin. 2014. Getting Parents Involved: A Field Experiment in Deprived Schools. *The Review of Economic Studies* **81**:1, 57-83. [Crossref]

232. Liang Choon Wang. 2014. All Work and No Play? The Effects of Ability Sorting on Students' Non-School Inputs, Time Use, and Grade Anxiety. *SSRN Electronic Journal* . [Crossref]
233. Tim Salmon, Danila Serra. 2014. Does Social Judgment Diminish Rule Breaking?. *SSRN Electronic Journal* . [Crossref]
234. Kirill Borusyak. 2014. Incidence Rather than Magnitude: A Lower Bound on the Share of Responders. *SSRN Electronic Journal* . [Crossref]
235. Ju Hyun Kim. 2014. Identifying the Distribution of Treatment Effects under Support Restrictions. *SSRN Electronic Journal* . [Crossref]
236. Tianshu Sun, Siva Viswanathan, Elena Zheleva. 2014. Creating Social Contagion Through Firm Mediated Message Design: Evidence from a Randomized Field Experiment. *SSRN Electronic Journal* . [Crossref]
237. Jun Xiao. 2014. Ability Grouping in All-Pay Contests. *SSRN Electronic Journal* . [Crossref]
238. David Kiss. 2013. The impact of peer achievement and peer heterogeneity on own achievement growth: Evidence from school transitions. *Economics of Education Review* **37**, 58-65. [Crossref]
239. Sharique Hasan, Surendrakumar Bagde. 2013. The Mechanics of Social Capital and Academic Performance in an Indian College. *American Sociological Review* **78**:6, 1009-1032. [Crossref]
240. Arna Vardardottir. 2013. Peer effects and academic achievement: a regression discontinuity approach. *Economics of Education Review* **36**, 108-121. [Crossref]
241. T. Bold, M. S. Kimenyi, J. Sandefur. 2013. Public and Private Provision of Education in Kenya. *Journal of African Economies* **22**:suppl 2, ii39-ii56. [Crossref]
242. Demetra Kalogrides, Susanna Loeb. 2013. Different Teachers, Different Peers. *Educational Researcher* **42**:6, 304-316. [Crossref]
243. Sari Pekkala Kerr, Tuomas Pekkarinen, Roope Uusitalo. 2013. School Tracking and Development of Cognitive Skills. *Journal of Labor Economics* **31**:3, 577-602. [Crossref]
244. Cristian Pop-Eleches,, Miguel Urquiola. 2013. Going to a Better School: Effects and Behavioral Responses. *American Economic Review* **103**:4, 1289-1324. [Abstract] [View PDF article] [PDF with links]
245. Aaron Sojourner. 2013. Identification of Peer Effects with Missing Peer Data: Evidence from Project STAR. *The Economic Journal* **123**:569, 574-605. [Crossref]
246. Diether W. Beuermann, Emma Naslund-Hadley, Inder J. Ruprah, Jennelle Thompson. 2013. The Pedagogy of Science and Environment: Experimental Evidence from Peru. *Journal of Development Studies* **49**:5, 719-736. [Crossref]
247. M. Kremer, C. Brannen, R. Glennerster. 2013. The Challenge of Education and Learning in the Developing World. *Science* **340**:6130, 297-300. [Crossref]
248. Jishnu Das,, Stefan Dercon,, James Habyarimana,, Pramila Krishnan,, Karthik Muralidharan,, Venkatesh Sundararaman. 2013. School Inputs, Household Substitution, and Test Scores. *American Economic Journal: Applied Economics* **5**:2, 29-57. [Abstract] [View PDF article] [PDF with links]
249. Norbert Bach, Marcel Battré, Joachim Prinz. Mündliche Gruppenprüfungen als Instrument der Bewertung universitärer Leistungen 87-109. [Crossref]
250. Giuseppe Bertola, Paolo Sestito. 2013. A Comparative Perspective on Italy's Human Capital Accumulation. *SSRN Electronic Journal* . [Crossref]
251. Kalena E. Cortes, Joshua Goodman, Takako Nomi. 2013. Intensive Math Instruction and Educational Attainment: Long-Run Impacts of Double-Dose Algebra. *SSRN Electronic Journal* . [Crossref]
252. Fernando Martel García. 2013. When and Why is Attrition a Problem in Randomized Controlled Experiments and How to Diagnose it. *SSRN Electronic Journal* . [Crossref]

253. Duncan McVicar, Julie Moschion, Chris Ryan. 2013. Right Peer, Right Now? Endogenous Peer Effects and Achievement in Victorian Primary Schools. *SSRN Electronic Journal* . [Crossref]

254. Fernando Martel García. 2013. Definition and Diagnosis of Problematic Attrition in Randomized Controlled Experiments. *SSRN Electronic Journal* . [Crossref]

255. Katja Kaufmann, Matthias Messner, Alex Solis. 2013. Returns to Elite Higher Education in the Marriage Market: Evidence from Chile. *SSRN Electronic Journal* . [Crossref]

256. Youjin Hahn, Liang Choon Wang, Hee-Seung Yang. 2013. Do Greater School Autonomy and Accountability Make a Difference? Evidence from the Random Assignment of Students into Private and Public High Schools in Seoul. *SSRN Electronic Journal* . [Crossref]

257. Jun Xiao. 2013. Ability Grouping in All-Pay Contests. *SSRN Electronic Journal* . [Crossref]

258. Luke C. Miller, Joel Mittleman. 2012. High Schools That Work and college preparedness: Measuring the model's impact on mathematics and science pipeline progression. *Economics of Education Review* **31**:6, 1116-1135. [Crossref]

259. Ritwik Banerjee, Elizabeth M. King, Peter F. Orazem, Elizabeth M. Paterno. 2012. Student and teacher attendance: The role of shared goods in reducing absenteeism. *Economics of Education Review* **31**:5, 563-574. [Crossref]

260. Scott E. Carrell, Mark Hoekstra. 2012. Family Business or Social Problem? The Cost of Unreported Domestic Violence. *Journal of Policy Analysis and Management* **31**:4, 861-875. [Crossref]

261. Scott A. Imberman,, Adriana D. Kugler,, Bruce I. Sacerdote. 2012. Katrina's Children: Evidence on the Structure of Peer Effects from Hurricane Evacuees. *American Economic Review* **102**:5, 2048-2082. [Abstract] [View PDF article] [PDF with links]

262. Katherine Grace Carman, Lei Zhang. 2012. Classroom peer effects and academic achievement: Evidence from a Chinese middle school. *China Economic Review* **23**:2, 223-237. [Crossref]

263. Michele Battisti, Jane Friesen, Ross Hickey. 2012. How Student Disability Classifications and Learning Outcomes Respond to Special Education Funding Rules: Evidence from British Columbia. *Canadian Public Policy* **38**:2, 147-166. [Crossref]

264. Victor Lavy, Olmo Silva, Felix Weinhardt. 2012. The Good, the Bad, and the Average: Evidence on Ability Peer Effects in Schools. *Journal of Labor Economics* **30**:2, 367-414. [Crossref]

265. Rachel Glennerster. 2012. The Power of Evidence: Improving the Effectiveness of Government by Investing in More Rigorous Evaluation. *National Institute Economic Review* **219**:1, R4-R14. [Crossref]

266. Timothy Powell-Jackson, Kara Hanson, Christopher J. M. Whitty, Evelyn K. Ansah. 2012. Who Benefits from Removing User Fees for Health Care? Evidence from a Randomised Experiment in Ghana. *SSRN Electronic Journal* . [Crossref]

267. Esther Duflo, Pascaline Dupas, Michael Kremer. 2012. School Governance, Teacher Incentives, and Pupil-Teacher Ratios: Experimental Evidence from Kenyan Primary Schools. *SSRN Electronic Journal* . [Crossref]

268. Román Andrés Zárate. 2012. Peer Effects, Cooperation and Competition in Human Capital Formation. *SSRN Electronic Journal* . [Crossref]

269. Tarun Jain, Mudit Kapoor. 2012. The Impact of Formal and Informal Peers on Academic Performance. *SSRN Electronic Journal* . [Crossref]

270. Lant Pritchett, Amanda Beatty. 2012. The Negative Consequences of Overambitious Curricula in Developing Countries. *SSRN Electronic Journal* . [Crossref]

271. Steve Cicala, Roland G. Fryer, Jörg L. Spenkuch. 2011. A Roy Model of Social Interactions. *SSRN Electronic Journal* . [Crossref]