Peer Effects, Pupil-Teacher Ratios, and Teacher Incentives: Evidence from a Randomized Evaluation in Kenya

Esther Duflo[1], Pascaline Dupas[2], and Michael Kremer[3, 4]

14 September 2007

This paper reports on a project designed to provide experimental evidence on several questions regarding resource allocation in primary education: the impact of pupil-teacher ratios, tracking, and the institutional environment (teacher contracts and beneficiary control). The project involved 210 primary schools in Western Kenya. We find that in this context, reducing the pupil-teacher ratio (from 80 to 46 on average), in the absence of any other reform, lead to reduced teacher effort, and to small and insignificant increases in test scores. In contrast, combining class size reduction with improved incentives (either by hiring local teachers on short term contract or by increasing parental oversight) leads to significantly larger test score increases. Finally, combining class size reduction with tracking by initial achievement leads to large test score increases, regardless of a child's initial achievement, suggesting that students benefits from homogenous classes. In contrast, we find no evidence that test scores are affected by the average pretest score of their peers.

[1] Abdul Latif Jameel Poverty Action Lab (JPAL), Massachusetts Institute of Technology, Paris School of Economics, NBER, and BREAD (eduflo@mit.edu).
[2] Dartmouth College and JPAL (pascaline.dupas@dartmouth.edu).
[3] Harvard University, Brookings, CGD, JPAL, NBER, and BREAD (mkremer@fas.harvard.edu).
[4] We thank the Kenya Ministry of Education, Science and Technology, International Child Support Africa, and Matthew Jukes for their collaboration. We thank Jessica Morgan, Ian Tomb, Paul Wang, Meghann Curtis, Erin Baggot and especially Willa Friedman for excellent research assistance and Alaka Holla for comments. We are grateful to Grace Makana and her field team for collecting all the data. We thank the World Bank and the Government of the Netherlands for the grant that made this study possible.

1. Introduction

This paper takes advantage of a randomized evaluation of a Kenyan program which provided school committees with funds to hire teachers locally on short term contracts in order to shed light on the impact of peers, pupil-teacher ratios, teacher incentive systems, and parental monitoring of teachers. A key finding is that peer quality and pupil-teacher ratios affect educational outcomes like test scores not only in a direct way but also through their impact on teacher behavior. In the context we examine, tracking students by initial academic achievement leads to large increases in learning for all students, contrary to standard models of peer effects in which students' test scores increase with those of their classmates. Decentralized hiring of extra teachers on a contract basis does not lead to increases in test scores for students randomly assigned to remain with civil service teachers, despite the reduction in pupil-teacher ratios in their classes, but it does lead to large test score increases for students randomly assigned to contract teachers. These effects are intermediated by changes in teacher behavior. Locally-hired contract teachers are much less likely to be absent from the classroom than centrally-hired teachers with civil service protection, but locally hiring contract teachers in a school decreases classroom presence among civil service teachers. Increased parental supervision moderates this effect, however.

The program we evaluate sought to address challenges created by the introduction of free primary education in Kenya and the associated influx of new pupils with varying levels of academic preparation. The program, called "Extra-Teacher Program" or ETP, provided funds to 140 schools randomly selected from a pool of 210 schools to hire an extra teacher for first grade classes. These teachers were hired locally, at perhaps a quarter of the cost of civil service teachers, but had the same academic qualifications. In half of these 140 schools, first grade

students were randomly assigned to either the contract teacher or a civil service teacher (in "non-tracked" ETP schools). In the other half, first grade classes were divided into two classes by initial achievement and then the classes were randomly assigned to either a civil service teacher or a contract teacher (hereafter referred to as "tracked" ETP schools). Finally, half of the 70 non-tracked ETP schools and half of the 70 tracked ETP schools were given funds to empower the local school committee to monitor and train teachers. In the second year of the program, resources were directed at second grade classes, so that the same students were affected over two years.

By comparing all students in the schools that received no treatment (hereafter referred to as "comparison schools") to the students in all non-tracked ETP schools, we can estimate the overall impact of adding a contract teacher to a school on average performance of all the students in the grade in that school. Students in non-tracked ETP schools scored 0.16 standard deviations higher than students in the comparison schools, suggesting that the program as a whole was successful. Students assigned to contract teachers in non-tracked schools were more likely to be present at school and scored 0.18 standard deviations better on achievement tests than their counterparts in the same school randomly assigned to civil service teachers. Contract teachers were more likely to be in class and teaching during random visits than the civil service teachers in either the ETP schools or the comparison schools. Children who were placed with a civil service teacher in non-tracked schools, however, did not score significantly higher than those in the comparison schools. This suggests that the greater average scores in non-tracked ETP schools relative to the comparison schools were primarily due to the effects of contract teachers on students, and not to the smaller pupil-teacher ratio.

Our results also cast doubt on some frequently used models of peer effects in which students with higher-achieving peers learn more. If all children benefit from having peers with high initial achievement, high achieving children would benefit from tracking by initial achievement, while children with low initial achievement would suffer. What we find instead is that children throughout the distribution of initial achievement benefit from tracking by academic initial achievement. After 18 months, point estimates suggest that scores of children in ETP schools in which pupils were tracked by initial achievement were 0.12 standard deviations higher than those of children in the non-tracked ETP schools that assigned children to the two classes randomly, although this difference is not significant. We also find that civil service teachers were more likely to be in class and teaching in schools that were tracked by initial achievement than in the non-tracked schools.

Moreover, a regression discontinuity design approach shows that children who were close to the 50^{th} percentile in their school (and therefore on the margin of being in either the high or low initial achievement groups) did not suffer from being assigned to the lower achievement group. Random variation in the mean or variance of initial test in non-tracked schools also did not influence test scores, perhaps because teachers did not change their pedagogy in response to these more modest changes in class composition.

Finally, giving parents oversight power reduced the negative effect on civil service teachers associated with having an additional contract teacher in the school. Civil-service teachers in ETP schools where the school committee was empowered to monitor teachers were more likely to be in class and teaching during random visits, and their students performed better than those of civil-service teachers in either comparison schools or ETP schools without empowered committees.

Our results contribute to several different debates on education. The results on the impact of tracking students by initial achievement are consistent with earlier studies which suggest that many students are left behind under the standard curriculum and that remedial instruction by locally-hired contract teachers tailored to these students can dramatically improve achievement (Banerjee et al., 2007; Glewwe, Kremer, and Moulin; 2004). One difference is that Banerjee et al. (2007) examined an intervention which combined a curriculum focused on remedial education, the use of para-teachers facing special incentives, and separate instruction for students who had fallen behind the official curriculum, while this paper separates out the effects of peers and of the type of teacher.

These findings are consistent with a "focus" model of peer effects whereby students benefit from having homogenous peers (Hoxby and Weingarth, 2007), and they suggest that teachers can adapt their teaching to the initial academic achievement of students in their classroom. These results also suggest, however, that generalizations based on the popular approach to estimate peer effects based on cohort-to-cohort variation in the composition of a school-grade (see Hoxby, 2000; Lavy and Schlosser, 2007) may yield a misleading picture of the impact of policy-driven larger-scale changes in class composition, which we find generates an endogenous teacher response, since this type of cohort-to-cohort variation generates only limited variation in peer composition

The findings on student performance under contract teachers versus civil service teachers also contribute to the literature on the widespread and growing phenomenon of local hiring of contract teachers in the developing world (see De Laat and Vegas, 2003; and Chaudhury et al., 2005). The previous literature on these programs is mixed. One interpretation is that the impact of such programs may be sensitive to the details of the local circumstances and to program

design (see Olken, 2007; Kremer, De Laat and Vermeersch, 2007; Banerjee and Duflo, 2005; Bjorkman and Svensson, 2007). But our findings are consistent with a recent study of higher education in the United States which suggests that students taught by adjunct professors do as well as students taught by tenure-track professors (Bettinger and Long, 2006).

Our results on the impact of tracking students by initial achievement are consistent with earlier studies which suggest that many students are left behind under the standard curriculum and that remedial instruction by locally-hired contract teachers tailored to these students can dramatically improve achievement (Banerjee et al., 2007; Glewwe, Kremer, and Moulin; 2004). Banerjee et al. (2007), however, examined an intervention which combined a curriculum focused on remedial education, the use of para-teachers facing special incentives, and separate instruction for students who had fallen behind the official curriculum. This paper separates out the effects of peers and teacher type.

The remainder of this paper proceeds as follows: Section 2 provides background and describes the project and data; Section 3 presents results on the intermediate outcomes of teacher and student effort; Section 4 discusses the impact of each treatment arm and of exogenous random variation on peer composition on student achievement; and Section 5 draws conclusions.

2. Background

2.1 Primary Education in Kenya

Kenya and many other developing countries are making rapid progress toward the Millennium Development Goal of universal primary education by 2015, in part due to the elimination of school fees. Between 1999 and 2004 the number of new entrants to primary school in sub-Saharan Africa increased by more than 30 percent (UNESCO, 2007). In Kenya

alone, enrollment in primary schools rose from 5.9 million to 7.6 million between 2002 and 2005, an increase of nearly 30 percent, (UNESCO, 2006).

Yet this progress creates its own new challenges. First, the influx of new students has raised pupil-teacher ratios. Free primary education was put in place in 2003, and by 2005 in the area we examine in Kenya, the average class size in first grade was 83, the median was 74, and 28 percent of first grade classes had more than 100 students. Second, integrating new students, many of whom have had little preparation, can be difficult.

To accommodate the increase in the number of pupils without a proportionate increase in teacher salary budgets, many countries, from Togo to India, have arranged for the local hiring of teachers on short-term contracts (De Laat and Vegas, 2003). Although the hiring of contract teachers is widespread, there is limited evidence to suggest what impact this practice has on student achievement or how contract teachers can best be used to accommodate the influx of relatively unprepared students.

Historically, most teachers in Kenya have been hired centrally through the Ministry of Education Science and Technology (MoEST) Teachers Service Commission (TSC). Teachers hired through the TSC have civil service protection and receive wages and benefits that are considerably above levels that would clear the market, and there is considerable queuing for these positions. For teachers in these positions, promotions, transfers, and disciplinary actions are decided through MoEST, rather than by more local bodies such as the school committees associated with every Kenyan school, which are primarily composed of elected representatives of parents.

Some school committees hire teachers locally using parent contributions to supplement the teachers hired by the Ministry of Education. In the area of study, we observed that these

locally hired teachers, called PTA (Parent-Teacher Association) teachers, received compensation in the range of 2000 Kenyan shillings, or US$ 30 per month. In comparison, the average civil service teacher receives around 7,000 shillings (US$ 120) per month plus generous benefits including housing allowances, provisions for retirement, and medical coverage (Glewwe et al., 2003). New graduates of teacher training colleges often work for several years as PTA teachers and then obtain positions as civil service teachers.

Since the introduction of free primary education, parents can no longer be required to pay fees for their children to attend school, and thus school committees are generally unable to raise the funds necessary to hire PTA teachers. In 2004, 80 percent of primary schools in our Western Kenya sample had no PTA teachers. Free primary education has thus raised the pupil-teacher ratio both by increasing the number of students and by reducing the number of teachers. It also led to a large influx of students who had little or no previous schooling.[5]

2.2 Program Structure

The project we study was modeled on the school committees' experience hiring extra teachers (see Figure 2 for project timeline and implementation). With funding from the World Bank, International Child Support Africa (an NGO that has been working with schools in this region for several years) provided funds to school committees to hire an additional first grade teacher who was local and had the same academic qualifications as a civil service teacher. Schools that participated in this Extra Teachers Program (ETP) received funding to create an additional class in first grade at the start of the second school term in May of 2005 (the school year runs from January to December in Kenya). Most schools had only one first grade class and

[5] Pre-school is not sponsored by the MoEST and therefore not free, since parents have to contribute to the pre-school teacher salary.

used the funds to split it in two. Schools with two or more classes in first grade before the program were also encouraged to create an additional class. In each school, ICS held a meeting with parents and teachers to explain program rules regarding the hiring of an extra teacher. Once a teacher had been hired, ICS disbursed funds to the school committees. School committees then paid extra teachers each month. The monthly allowance totaled 2,500 Kenyan shillings (around US$ 35), putting it at the top of the range of what is typically paid to extra-teachers by school committees in Kenya. When the program continued the following school year, school committees were free to replace or keep the original extra teachers and were encouraged to move the teachers to second grade with the same group of students.[6] The ICS program did not include provisions for giving teachers civil service status, but it did give them experience. By the end of the program 32 percent of the contract teachers hired through the ICS program had been absorbed by the Kenyan Government.

A total of 210 schools were enrolled in the program at the end of the first school term.[7] After compiling a list of all students enrolled in first grade in each of the 210 schools, the schools were randomly assigned to two groups: 140 treatment (or "ETP") schools and 70 comparison schools (see Figure 1).

In 70 of the 140 treatment schools, first grade pupils were randomly assigned to either the contract teacher or a regular, civil service teacher. In each of the other 70 treatment schools, students were tracked according to initial achievement level. ICS field officers collected grade sheets from school exams administered before the allocation of children to classes under the program was announced. The classes were then formed in the office so that the less prepared 50

[6] Students enrolled in Standard Two in 2005 and who repeated Standard Two in 2006 were randomly assigned to either the contract teacher or the civil service teacher in 2006.
[7] The school year in Kenya starts in January and is divided in three terms of three months, with three month-long breaks in April, August, and December.

percent of students were assigned to group A and the more prepared 50 percent were assigned to group B (in schools with only one class initially).[8] After the assignment, transfers from one group to the next were systematically granted for siblings who wanted to stay together. Other transfers were in principle possible upon parental request, but they had to be granted pro-forma approval from the ICS program administrator. In practice, though, there were very few requests of this kind. As a result, in tracked schools there is a sharp discontinuity in the average prior achievement of peers experienced by students at the 50th percentile of the distribution in each school (Figure 3, Panel A), which is not apparent in the untracked schools (Panel B) . The contract teacher was then randomly assigned to one class, and the civil service teacher took the other one.[9] Since the correlation between a pupils' age and her initial rank in the school is 0.18, the tracking by initial attainment generated a gap between the two classes in Tracked Schools, with the students in the higher initial attainment class being older by 6 months on average.

In addition, among the 140 schools sampled to receive funding to locally-hire a contract teacher, 70 schools were randomly selected (after stratification by tracking by initial achievement status) to participate in a School-Based Management (SBM) initiative designed to empower the school committees to monitor teachers' performance. In those schools, school committees held a formal review meeting at the end of the first school year of the program (2005) to assess the contract teacher's performance and decide whether to renew his or her contract or to replace her. To prepare the school committee, ICS provided its members a short, focused training on monitoring the contract teacher's performance. The school committee

[8] If the school initially had two classes in first grade, the students were split into three streams after the introduction of the extra-teacher program, with the less prepared third of the students forming one stream, the more prepared third forming a second stream, and the middle third of the distribution forming the third stream.
[9] 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 comparison school might be endogenous, they are excluded from the analysis.

members were taught techniques for soliciting input from parents and checking teacher attendance. A formal sub-committee of first grade parents was formed to evaluate the contract teacher and deliver a performance report at the end of the first year.

2.3 Data Collection

Sample and Summary Statistics

The sample includes about 21,000 students who were enrolled in first grade at the end of the first school term in one of the 210 primary schools enrolled in the study. Slightly less than half are girls (48.8 percent). On average, students were seven years old at the outset of the program (with a standard deviation of 1.3 years), but ages ranged from 5 to 14.

In addition, we have data on 1,124 teachers who taught lessons in first grade in 2005, in second grade in 2006, or both. The vast majority of these teachers were centrally hired civil service teachers (926 in total). The average age of civil service teachers was 42 (with a standard deviation of 9 years), and 69 percent were female.

In addition, 198 contract teachers were hired through the ETP program by the 140 treatment schools over the course of the program. On average, contract teachers hired with ETP funding were 28 years old (with a standard deviation of 7 years), and 51 percent were female. They all had received a teacher's certificate from a governmental teacher training college, as all civil-service teachers do.

Outcomes

Our key outcome of interest is student achievement, as measured by scores on a standardized math and language test administered in all schools 18 months after the start of the

program. Exams were administered by trained enumerators and graded blindly by data processors. In each school, 60 students were randomly drawn from the initial sample to participate in the tests. Students were asked math and literacy questions ranging from counting and identifying letters to subtracting two-digit numbers and writing words. To limit attrition, enumerators were instructed to go to the homes of students who had dropped out or were absent on the day of the test and bring them to school for the test. This was not always possible, however.

The attrition rate on this endline test was 18 percent on average (see Table 2). It was not significantly different across treatment arms or across deciles of the initial achievement distribution, and it is therefore unlikely to bias our results. Students in ETP schools were as likely to have been transferred to a new school as students in comparison schools (Table 2, column 1).

In addition, data was collected on pupils' attendance and dropout/grade promotion through unannounced school visits made by enumerators on a quarterly basis. Overall, the dropout rate among first grade students in our sample was low (below 0.5 percent).

Baseline data on students' initial achievement was collected from the school records. While internally consistent within schools, this data is not comparable across schools.

2.4 Framework for Thinking about the Program

The following framework may be useful in thinking about the impact of the program. In particular, suppose that educational outcomes for student i, Y_i, are given by:

$$Y_i = f(T, P, X_i, , \underline{X}, , E, W)$$

where T is the number of teachers in the grade, P is the number of pupils, X is the pretest score, bars denote the average in the class, E is an index of teacher effort and behavior, and W represents other factors, possibly stochastic.

Suppose also that teacher effort, E, is given by:

$$E = g\,(T, P, \underline{X}, I, O)$$

where I represents the incentive system in place and O represents other unmodeled factors.

In this framework, to understand the impact of changing peers or of hiring extra teachers on educational outcomes, it is important to look at the total derivative of f with respect to $\underline{X}$, P, and T, not just the partial derivative, to see how the teacher-pupil ratio and average peer quality also affect teacher effort and not just student outcomes directly. Of course, the partial derivative of g with respect to T, P, and $\underline{X}$ can potentially vary with the context and in particular the incentive system. Thus, Section Three examines how teacher (and student) behavior responds to the interventions we study, before Section Four examines the overall impact of these interventions on educational outcomes.

3. Intermediate Outcomes: Teacher and Student Effort

Before looking at student achievement, we investigate the effect of the program on the intermediate outcomes of teacher and student effort to help us interpret the results on achievement. While average class size increased from 79 to 84 in the comparison schools between the beginning of the program and the end of the first school term, it fell from 84 to 46 in ETP schools (see Table 1). This pupil-teacher ratio was maintained the following year in second

grade: in March 2006, the average second grade pupil-teacher ratio was 73 students in comparison schools, and 45 in ETP schools.

Teacher Effort

Table 3 shows regressions estimating the impact of various program variants on teachers' effort, measured by their presence in school on the day of a random visit and whether or not they were found in class teaching when the observers entered the school compound. The omitted category in this regression is civil service teachers in the comparison schools. The first row in the table is thus the difference between civil service teachers in the non-tracked treatment schools that did not receive the SBM intervention and those in the comparison schools. This difference tells us the effect of reducing the pupil-teacher ratio. The second and third row interact the treatment with initial achievement tracking and SBM, respectively, and these coefficients estimate the difference between civil service teachers in between civil service teachers in non-tracked schools and tracked schools, both without SBM (Row 2); and between SBM schools and non-SBM schools, both without tracking (Row 3). Row 4 presents the coefficients of the additional effect of combining tracking and SBM. The next three rows present the coefficients for contract teachers, interacting with each possible treatment combination. The last three rows present the effect of having a rotation scheme among TSC teachers, and its interaction with the ETP program.[10]

[10] There are two different ways in which primary schools can organize learning in lower grades. Some schools assign each teacher to one class exclusively. The teacher is then in charge of the entire curriculum for her class. Some other schools assign a pool of teachers to the three or four lower primary grades, and teachers in the pool rotate from one class to another, teaching only the specific subjects in the curriculum that they have been assigned to teach.

While civil service teachers in ETP schools were as likely to be present in school as teachers in comparison schools, the likelihood that civil service teachers in untracked schools with no SBM were found in class teaching if they were present fell 21 percentage points (column 2) relative to the (civil service) teachers in comparison schools.[11] This effect, significant at the 5 percent confidence level, corresponds to a 30 percent decrease in teacher presence in class compared with civil service teachers in comparison schools. This finding suggests that civil service teachers took advantage of the presence of the extra contract teachers to work less. While this crowding-out effect may be exacerbated by the vulnerable position of the contract teachers, who are presumably easy for civil service teachers to order around, it might have occurred even if the additional hired teachers had civil service protection.

Contract teachers in untracked schools with no SBM were 25 percentage points more likely to be found in class teaching when at school than their civil service counterparts in the same schools (Row 5, Column 2). A natural interpretation of this is that the strong incentives to perform that contract teachers face for contract renewal play a large role.

The SBM initiative reinforced the role of parents (as opposed to only headmasters, who often dominate those committees) in hiring, monitoring, and retaining the contract teachers. Although parents were instructed on how to monitor the contract teacher, the SBM initiative did not have a significant impact on the attendance records or effort of the contract teacher (perhaps because it was already very high) but did improve the effort of civil service teachers. Compared with civil service teachers in untracked non-SBM treatment schools, civil service teachers in untracked SBM schools were 7.8 percentage points more likely to be found in class teaching during random spot checks by the NGO (significant at the 10 percent level; Row 3, Column 3).

[11] When teachers are not in class teaching, they are usually in the teachers' room. When a teacher is absent, it is common for classes to be combined and taught by the teacher who is present.

One potential explanation for the effect of SBM on civil service teachers' effort is that the SBM initiative emphasized the responsibility of the contract teacher with respect to the specific class to which she was assigned and thus made it more difficult for the head teacher or the civil service teachers in those schools to use the extra teacher to relieve themselves of their own duties when they actually did show up to school. The contract teachers in these schools had a greater incentive to please the school committee and less of an incentive to please the other teachers and the headmaster.

Civil service teachers are significantly more likely to be in class teaching in schools tracked by initial achievement than in non-tracked treatment schools: the difference is 8.6 percentage points, significant at the 10% level). One interpretation is that teaching is easier and thus more pleasant when students have less heterogeneous levels of academic preparation. Another is that classes tracked by achievement may be perceived as distinct groups by the school administration and the teacher, and thus are less easy to merge into one class to lessen the teachers' workload.

Student Attendance

The reduction in pupil-teacher ratio had a negative (though not statistically significant) impact on student attendance for students assigned to the civil service teachers in ETP schools (see Table 4, Row 1). However, students of contract teachers were 3 percentage points more likely to be in school than students of civil-service teachers in the same schools. This corresponds to a 22 percent decrease in absenteeism among students of contract teachers, significant at the 1 percent level. Since we found in Table 3 that the ETP program reduced the rate of class presence by civil service teachers, while that of contract teachers was significantly higher than the class presence

of teachers in comparison schools, a plausible interpretation of the effect of the program on student's presence is that students come to school more if their teacher teaches more and students come to school less if their teacher teaches less.

4. Effect on Test Scores

In this section, we will review the overall effect of the program (section 4.1) and the separate effects of pupil-teacher ratio, teacher type, and initial achievement tracking on test scores (see Tables 5 to 7).

Table 5 presents average treatment effects on test scores (panel A) and schooling outcomes (panel B). Test scores were normalized such that the mean and standard deviation of the comparison group are zero and one, respectively. Table 5 presents the average outcomes by treatment group. All standard errors are clustered at the school level.

Tables 6 and 7 present regression estimates with individualized and school-level controls. For each outcome, we present the results of four different regression specifications: the first specification estimates average treatment effects with treatment dummies. The second specification adds a control for the student's initial attainment and interaction terms between initial attainment and treatments. The third specification adds controls for the teachers' characteristics (age and gender). . Lastly, the fourth specification reproduces the second specification, restricting the sample to students initially in the middle of the distribution (5^{th} and 6^{th} deciles on the initial exam).

4.1 Overall Program Impact

Adding contract teachers substantially improved average test scores. Eighteen months into the program, students in all treatment schools had, on average, test scores that were 22 percent of a standard deviation higher than students in the comparison schools (Table 5, column 2). This difference is significant at the 1 percent level.

Simply introducing a new contract teacher and randomly assigning students to either this new teacher or the civil service teacher, without training the school committee and without tracking by initial achievement has a small (13 percent of a standard deviation) but insignificant effect on test scores, despite a reduction in average class size of about 40 students (see Table 5, columns 3 and 4). The effect is larger (19 percent of a standard deviation, significant at the 10 percent level) when the school committees are also trained on how to handle the contract teachers. The effect of the program is much larger and very significant when the class is not divided in two streams randomly but when students are tracked by initial achievement, with or without training of the school committees (between 25 and 31 percent of a standard deviation, significant at the 1 percent level; columns 6 and 7). This suggests that reducing class size alone is not key to improving school quality; the environment in which the size reduction is done appears to be an essential element of the success of such an intervention. We now turn to more detailed evidence to unpack these results.

4. 2 Pupil-Teacher Ratio

In order to isolate the relationship between pupil-teacher ratio reduction and school achievement, we compare the test scores of students assigned to civil service teachers in non-tracked ETP schools with those of pupils in the non-ETP schools. Because students were

randomly assigned to either civil service teachers or contract teachers, this comparison can be interpreted as the causal effect of being assigned to a smaller class with a civil service teacher.

While pupils in non-tracked schools with a reduced pupil-teacher ratio scored somewhat higher on both the math and language tests compared to pupils in the non-ETP schools (9 percent of a standard deviation overall), the null hypothesis of no effect cannot be rejected (see column 8 of Table 5). Likewise, reducing pupil-teacher ratio without tracking did not significantly increase the likelihood that students had reached third grade in 2007, two years after program inception (see column 8 in Panel B). These results are confirmed by the regression results in Tables 6 and 7 (first row).

At the sample mean, these results suggest that, in lower grades, reducing class size from 80 to 40 students without any other change does not lead to a significant increase in test scores. This is a striking result, in particular since the literature in richer countries suggests that there are significant effects of pupil-teacher ratio reduction on test scores in some contexts (see e.g. Krueger and Whitmore, 2002; Angrist and Lavy, 1999; Piketty, 2004). However, this finding does echo the results from India by Banerjee et al. (2007), which showed no impact of the reduction in pupil-teacher ratio achieved through the hiring of a remedial education teacher for students who remained with the regular (civil service) teacher. There are several possible explanations for these results. One is that, at 40, the class size remains too large and too diverse for the teacher to be able to devote any individual attention to each student, so that having 40 rather than 80 students still does not make any difference. Another possible explanation, however, is that in this program, class size reduction was accompanied by a reduction in teacher effort, as discussed in Section 3 where we noted that civil service teachers were less likely to be in class teaching in ETP schools than in non-ETP schools.

We also find some evidence that the SBM initiative was helpful in raising test-scores for the students of civil service teachers just as it was successful in decreasing the classroom absence rates of these teachers. Students with civil service teachers in ETP schools with SBM scored 0.18 to 0.24 standard deviations higher in mathematics than their counterparts in ETP schools without SBM (Table 6, row 3).

It should be noted that these results do not tell us what would have been the impact of achieving the same pupil-teacher ratio reduction by civil service teachers, as opposed to contract teachers. For example, it is possible that if additional civil service teachers had been hired, existing civil service teachers would be more inclined to teach, while with this program they can more easily pass work off to the contract teachers.

4.2 Comparing Contract Teachers and Civil Service Teachers

To measure the impact of being taught by a contract teacher rather than by a civil service teacher, keeping pupil-teacher ratio constant, we compare the achievements of children who have been assigned to the contract teacher hired through the ETP program to that of children who stayed with the civil service teachers in ETP schools. Because children were randomly assigned to either group, the comparison is a direct measure of the impact of being taught by a contract teacher rather than the civil service teacher in a school where there is a contract teacher. It is important to note that because the effort of the civil service teacher was likely affected by the presence of the contract teacher, this comparison cannot be interpreted as the effect of hiring contract teachers rather than civil service teachers to achieve pupil-teacher ratio reductions.

The regression results in Table 6 show that students assigned to a contract teacher in ETP schools score 0.23 standard deviations higher than their schoolmates assigned to civil service

teachers (Row 5) (and 0.23 + 0.07 [Row 1] =0.30 standard deviations higher than students in the non-ETP schools). Together rows 3 and 7 show that the SBM intervention does not affect students assigned to the contract teacher, consistent with the lack of its effect on the probability that contract teachers were found in school teaching. The positive and significant impact of the ETP Teacher does not disappear not reduce when controls for demographics of the teachers are added (column 3), thus suggesting that the contract teacher effect is not a pure age effect.

Students of contract teachers were 5.5 percentage points more likely to have reached third grade in 2007 than students in treatment schools taught by civil service teachers (Table 7, row 5, column 3). This corresponds to an 11 percent increase in grade promotion, significant at the 1 percent level, however we do not focus on this outcome since promotion may reflect teacher and parent preferences.

The superior performance of contract teachers could potentially be due either to better choice of teachers by local school committees or to the stronger incentives faced by contract teachers. It is worth noting that good performance as a contract teacher could also have been perceived as a stepping-stone to a tenured civil service position.

These results thus suggest that teacher incentives do matter in schools, and they are relevant for assessing a policy that is probably more realistic for many developing countries, which is to hire additional teachers first on a temporary contract basis, and then regularize the ones that perform well with a civil service position. However, it cannot necessarily be assumed that the results would remain constant if the status of all teachers were changed to conditional contracts. This is because contract teachers in the ICS program may have worked hard in part because they hoped it would help them obtain civil service teaching positions.

These differences between the contract teachers and the civil service teachers may also result from the different classroom experiences that they faced. Civil service teachers often rotate across classes, so the observed test score advantage that the students of contract teachers display could also be attributed to a potential benefit from remaining with a single teacher for most classes.

4.3 Peer effects: Initial Achievement Tracking

4.3.1 Overall Effect

Changing incentives is not the only way to change working conditions for teachers and students. Another one is to give them different, more focused tasks. Banerjee et al. (2007) conjecture that the large effects of the remedial education programs and computer-assisted programs they examine might be due to the fact that both programs gave students the opportunity to work at their appropriate level.

Indeed, pupil-teacher ratio reduction with tracking by initial achievement led to a large increase in learning, especially in mathematics. Students in schools tracked by initial achievement performed 0.25 (no SBM) and 0.31 (SBM) standard deviations better than students in comparison schools (see Table 5, panel A, columns 6 and 7). And in contrast to the results in non-tracked schools, students in tracked schools benefited from the ETP program even when they were assigned to civil service teachers: the overall test score of students assigned to civil service teachers in schools tracked by initial achievement was 0.18 standard deviations higher than that of students in comparison schools (Table 5, panel A, column 9).

With a specification that allows the effects of tracking to vary with teacher type, the regression results in Table 6 suggest that the estimated overall tracking gains relative to the

comparison schools arise from their positive effect on civil service teachers only. The program effect is 0.25 standard deviations for students assigned to civil service teachers in schools tracked by initial achievement, but the estimated effect for students assigned to contract teachers in tracked schools (computed by adding the coefficient estimates in rows 2 and 16) cannot be distinguished from zero.

The regressions also allow us to see whether the effect of tracking by initial achievement is different for initially strong students (who are paired with other strong students) and initially weak students (who are paired with other weak students). Both lower and higher-achieving students benefited from tracking (in Row 14, the interaction between being in the bottom half and in a tracked school cannot be distinguished from zero in columns 2, 6, and 10). In mathematics, the initially higher achieving students in ETP tracked schools scored 0.25 standard deviations higher than the initially higher achieving students in the comparison schools, and the initially lower achieving pupils in tracked ETP schools display a similar advantage over their counterparts in the comparison schools. Note that the position in the roster of test scores has strong predictive content: children who were in the bottom half of the class score much worse at the endline test than students who were initially in the top half of the class.

The beneficial impact of tracking on students in the bottom half of the initial achievement distribution is confirmed by the grade promotion results (see Table 7): students initially just below the 50^{th} percentile cutoff were 9.4 percentage points (i.e. 19 percent) more likely to have been promoted to third grade by 2007 if they had been tracked by initial achievement, compared with both the initially lower-achieving students in comparison schools and students in treatment schools with untracked class assignment.

One hypothesis consistent with these results is that tracking by initial achievement improves student learning because it allows teachers to focus instruction. Teaching a group of more homogeneous students might allow teachers to adjust the pace of instruction to students' needs. For example, a teacher might instruct at a slower pace, providing more repetition and reinforcement, when students are initially less prepared, but 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. Our results on teacher effort, however, also suggest that tracking increased teacher presence in the classroom for the civil service teachers in the ETP program. Thus, another channel through which tracking can affect student achievement is through its effect on teachers' willingness to show up for class.

We use two additional strategies to further examine the impact of peer tracking on achievement. First, we use a regression discontinuity design to look at large variations in the average initial achievement levels of peers to which students in the middle of the distribution are exposed, depending on the group to which they were assigned. Second, we examine the impact of the random variation in class composition generated by random assignment to classes in the non-tracked school.

4.3.2 Does the quality of peers matter in an initial achievement tracking set up? A regression discontinuity design approach

Figure 4 plots the average achievement of children at the endline test as a function of their position in their class before the program. In schools tracked by initial achievement, there is no discontinuity in the test score at the 50^{th} percentile threshold, suggesting that students near the

median are equally well off being the strongest students in the weaker group or the weakest students in the stronger group.

We examine this result in a regression framework in Table 6 (columns 4, 8, and 12). First, we use the same specification as before but we restrict the sample to children who were initially in the 5th or 6th deciles of the achievement in their grade. Consistent with the graphs, we find that the coefficient of the interaction term (being in the bottom half in a school with tracking by initial achievement) is not significantly lower than zero, which suggests that the tracking by initial achievement did not generate a discontinuity in outcomes between students initially in the 5th and those initially in the 6th deciles. Second, we run a few regression discontinuity design specifications with this sample, following Lee (2001). To start, we regress test scores on a third-order polynomial in initial percentile, separately, for students below the 50th percentile cutoff (typically assigned to the lower initial achievement class) and students above the 50th percentile. The results are plotted in Figure 4 separately for each program variant. Each point is an average of the test scores for each percentile of the initial distribution. The vertical line represents the cutoff line for being assigned to the "lower initial achievement" class. As apparent from the figure, there is no discontinuity around the vertical line in the schools tracked by initial achievement, despite the strong discontinuity in peer attainment observed in Figure 3. The data exhibit a continuous and smooth relationship throughout the distribution in both non-tracked and tracked schools. Figure 5 presents the results with a linear fit, rather than a polynomial. Again, in schools with tracking by initial achievement, we do not see an effect of the group in which the students were placed for students in the middle of the distribution.

Figure 6 presents the smoother average score by initial percentile for all three groups of schools: Non-ETP schools (comparison), Tracked ETP Schools, and Non-Tracked ETP Schools.

The Figure shows that students in the middle of the initial distribution performed better in the Tracked Schools than either the Non-Tracked or Non-ETP schools. These results thus suggest that, at least for the children in the middle of the distribution, what matters more than the average quality of the peers is their homogeneity; children benefit from being in a homogenous group. Note that regardless of the group they are assigned to, the children in the middle of the distribution are likely at the extreme of their class under tracking but they still benefit relative to being in a class where they are exactly at the middle of the distribution. What this may suggest is that it is not only the content of the material taught that matters, but whether most children in the class find it stimulating or comprehensible. What is clear is that a "linear in means" model of peer effects—or even more generally, a model in which all students benefit from the same peers—is clearly inappropriate in this context, since the average quality of peers does not appear to matter, while homogeneity does.

It may also be the case that tracking benefited students in the middle of the distribution because it may be better to have teaching be at a level which is not perfectly suited to one's initial level but is always within some moderate distance of that level. When the class consists of very different initial achievement levels, teachers might have to divide their time between each of these groups leaving large periods of time when teaching is either way too advanced or way too basic for any individual student and therefore causing students to become distracted.

4.4 Peer effects: Random Variation in Peer Composition

The regression discontinuity design approach we have discussed has the advantage of generating large differences in average peer initial achievement, but the drawback is that we can only look at its effect for children in the middle of the distribution. The effect of peer initial achievement could potentially be very different for children with different initial levels of

achievement. Fortunately, the evaluation generated another source of random variation in the average achievement in the peer group: in non-tracked schools, children were randomly assigned to either class and very few re-arrangements between classes took place. As shown in Figure 7, this generates a fair amount of random variation in the composition of the different classes, both in terms of average peer achievement (panel A) and sex ratio (panel B). We can thus implement methods to evaluate the impact of the composition of a class similar to those used by Hoxby (2001) and Lavy and Schlosser (2007), with the difference that, in this case, we use actual random variation in peer group. The results of various specifications are presented in Table 8. In none of the specifications do we find any effect of either the mean or the variance in the quality of the peers, nor of the gender composition of the peers.

5. Conclusion

The results of this impact evaluation shed light on a variety of education policy questions: Can contract teachers help improve learning in areas where primary school enrollment has increased rapidly? How does the tracking of students into classes by initial achievement affect learning? Does empowering parent-run school committees to monitor teachers improve teacher performance?

Contract teachers tend to be present and in class teaching more frequently than tenured civil service teachers in the same type of school. Their students learn more and perform better on cumulative achievement tests; specifically, students of contract teachers performed 0.18 standard deviations higher on achievement exams than their counterparts in classes taught by

civil service teachers. This could potentially be due either to the stronger incentives faced by these teachers or to the fact that they were hired from the local community.

Students in smaller classes in treatment schools taught by civil service teachers did not score significantly higher than students in larger classes taught by civil service teachers in comparison schools, suggesting a limited impact of reductions in pupil-teacher ratios.

Students at both the high and low ends of the initial achievement spectrum benefited from being tracked into classes by initial achievement. After 18 months, point estimates suggest that the average score of a child in a school in which pupils were tracked by initial attainment is 0.12 standard deviations higher than that of a child in an untracked treatment school, although this difference is not significant. Moreover, a regression discontinuity design approach reveals that children who are very close to the 50^{th} percentile in their school, and thus were assigned either to the low initial achievement track or the high initial achievement track, do not suffer from being in a low initial achievement track.

Civil service teachers are more likely to be in class and teaching in schools with tracking by initial achievement classes than in those with random assignment classes. This finding may indicate that teachers are more invested in homogenous classes on which they feel they have a greater impact, and are thus less inclined to be absent. It also suggests that interventions that affect the composition of students in a classroom may contribute to student achievement not just through peer effects and potential changes in pedagogy but also through a change in teachers' incentives to show up to class.

Civil service teachers were more likely to be in class and teaching during random visits in schools where the school committee was empowered to monitor teachers than in treatment schools without the monitoring. Furthermore, we find evidence that suggests the students of civil

service teachers in schools with empowered parent committees performed better on mathematics achievement tests than their counterparts in treatment schools without empowered committees.

In terms of policy implications, this paper suggests that, if appropriately structured, support for school committees to hire contract teachers could be a cost-effective way to respond to the challenges created by free primary education. More generally, the results suggest that the effects of variables such as pupil-teacher ratios and class size may be critically intermediated by the behavioral response of teachers.

References

Angrist, J. and Lavy, V. (1999). "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *Quarterly Journal of Economics* 114, 533-575.

Banerjee, A., Cole, S., Duflo, E. and Linden, L. (2007). "Remedying Education: Evidence from Two Randomized Experiments in India." Forthcoming, *Quarterly Journal of Economics*.

Bettinger, E. and B. Long (2006). "Does Cheaper Mean Better? The Impact of Adjunct Instructors on Student Outcomes." Mimeo, Harvard University.

Devarajan, S., Miller, M.J. and Swanson, E.V. (2002). "Goals for Development: History, Prospect and Costs." *World Bank*, Working Paper # 2819.

De Laat, J., and Vegas, E. (2003). "Do Differences in Teacher Contracts Affect Student Performance? Evidence from Togo." *World Development Report 2004: Making Services Work For Poor People*, Background Paper, 26 June 2003.

Ehrenberg, R., Brewer D., Gamoran A., and Willms J.D. (2001). "Class Size and Student Achievement." *Psychological Science in the Public Interest*. Vol. 2, No. 1.

Finn, J.P. and Achilles, C.M. (1999). "Tennessee's Class Size Study: Findings, Implications and Misconceptions." *Educational Evaluation and Policy Analysis* 21, 97-110.

Fredriksen, B. (2002). "Education for All African Children by 2015: What Will it Take to Keep the Promise?" Paper prepared for the Annual Bank Conference on Development Economics, Oslo, June.

Glewwe, P., Kremer, M., and Naumam. I. (2003). "Teachers Incentives." Mimeo, Harvard University and University of Minnesota.

Hoxby, C. and G. Weingarth (2006). "Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects." Mimeo, Harvard University.

ILO/UNESCO (2007). *Joint ILO/UNESCO Committee of Experts on the Application of the Recommendations concerning Teaching Personnel*. Geneva: ILO Publications.

Krueger, A. and D. Whitmore (2002). "Would Smaller Classes Help Close the Black-White Achievement Gap?" In John E. Chubb and Tom Loveless, eds., *Bridging the Achievement Gap*. Washington: Brookings Institution Press.

Lavy, V. and A. Schlosser (2007). "Mechanisms and Impacts of Gender Peer Effects at School." Mimeo, Hebrew University and Princeton University.

Piketty, T. (2004). « L'Impact de la taille des classes et de la ségrégation sociale sur la réussite scolaire dans les écoles françaises : une estimation à partir du panel primaire 1997. » Mimeo, PSE.

UNESCO (2000). "Assessing Learning Achievement." *Education for All: Status and Trends 2000*. http://unesdoc.unesco.org/images/0011/001198/119823e.pdf

UNESCO (2003). EFA Africa Bulletin Board, July. http://www.dakar.unesco.org/efa/bulletin/2003_07_en.shtml#africa

UNESCO (2006). *Fact Book on Education for All.* UNESCO Nairobi Office.

UNESCO (2007). *Strong Foundations: Early Childhood Care and Education.* Paris: UNESCO Publishing.

Figure 1
Experimental Design: The Extra-Teacher Project

Group	# Schools	Class Size	Peer Grouping	Training on School-Based Management of Teachers (SBM)	Teacher Employer	# Classes
Non-ETP Schools (Comparison)	70	Normal	Unchanged	No	Government	88
Non-Tracked Schools	70	Reduced	Random	No	Government	41
					School Committee	35
				Yes	Government	42
					School Committee	35
Tracked Schools	70	Reduced	Tracking by Initial Achieve-ment	No	Government	41
					School Committee	35
				Yes	Government	41
					School Committee	35

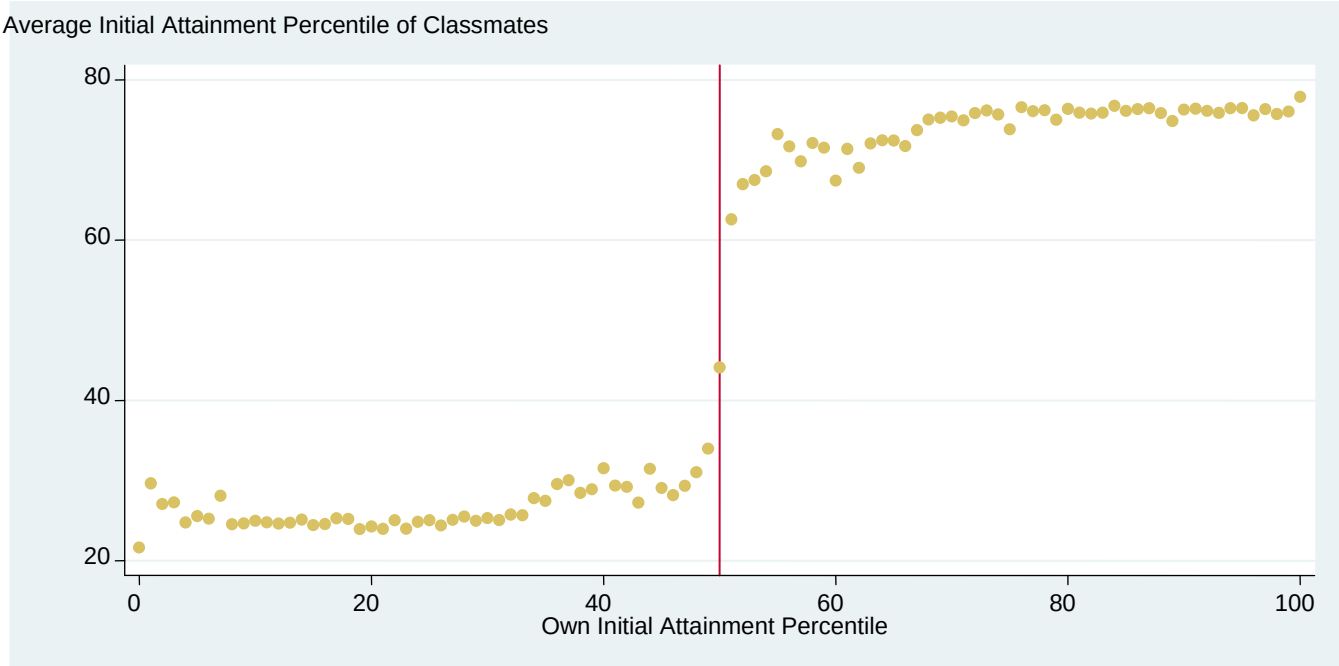
Figure 2
Project Chronology

2005	
February/March	Initial visit to collect baseline enrollment information and Standard 1 mark lists in all 210 program schools
March/April	Parent-teacher meetings in all 210 program schools
	School Committees in all 70 schools sampled for SBM reinforcement receive training on monitoring of contract teacher and contract renewal decisions
May	Streaming done in all 140 schools sampled to receive a contract teacher ("ETP schools")
	Delivery of teachers' salaries for first four months of program to school committees in ETP schools
May/June/July	Pupil and teacher attendance visit in Std 1 in all 210 program schools
	Post-hire surveys administered to the contract teacher, the Head Teacher and a School Committee member in all ETP schools
August	Government of Kenya's Teacher Service Commission hires large number of teachers including 35 working as contract teachers in ETP schools; ETP schools instructed to replace contract teacher
September	Delivery of salaries for second four months of program to school committees in ETP schools
September/October	Pupil and teacher attendance visits in Std 1 in all 210 program schools
	Follow-up surveys administered to contract teacher/Head Teacher/School Committee member/another Std 1 teacher in all 140 ETP schools
November/December	Teacher evaluation/Contract renewal meetings organized in all 70 SBM schools
2006	
January/February	Streaming done in all 140 ETP schools to distribute newcomers (to ensure balanced class-size)
	Delivery of salaries for contract teachers to ETP school committees for first four months of 2006
February/March	Pupil and teacher attendance visit in Standards 1 and 2 in all 210 program schools
	Follow-up survey conducted with contract teacher/Head

	Teacher/School Committee Member in all treatment schools
June	Deposit of contract teacher's salaries for next four months to ETP school committees
June/July/August	Pupil and teacher attendance visits in Standards 1 and 2 in all 210 program schools
	Classroom observations and pupil and teacher questionnaires conducted in all 210 program schools
September	Delivery of salaries for last four months of program to ETP school committees
October/November	Pupil attendance visits in Standards 1 and 2 in all 210 program schools
	Individual Achievement Tests administered in all 210 program schools
2007	
February/March	Pupil and teacher attendance visits in all 210 program schools
	Final follow-up survey conducted in all 140 treatment schools.

Figure 3
Distribution of Average Peer Achievement

Panel A. Tracked ETP Schools



Panel B. Non-Tracked ETP Schools

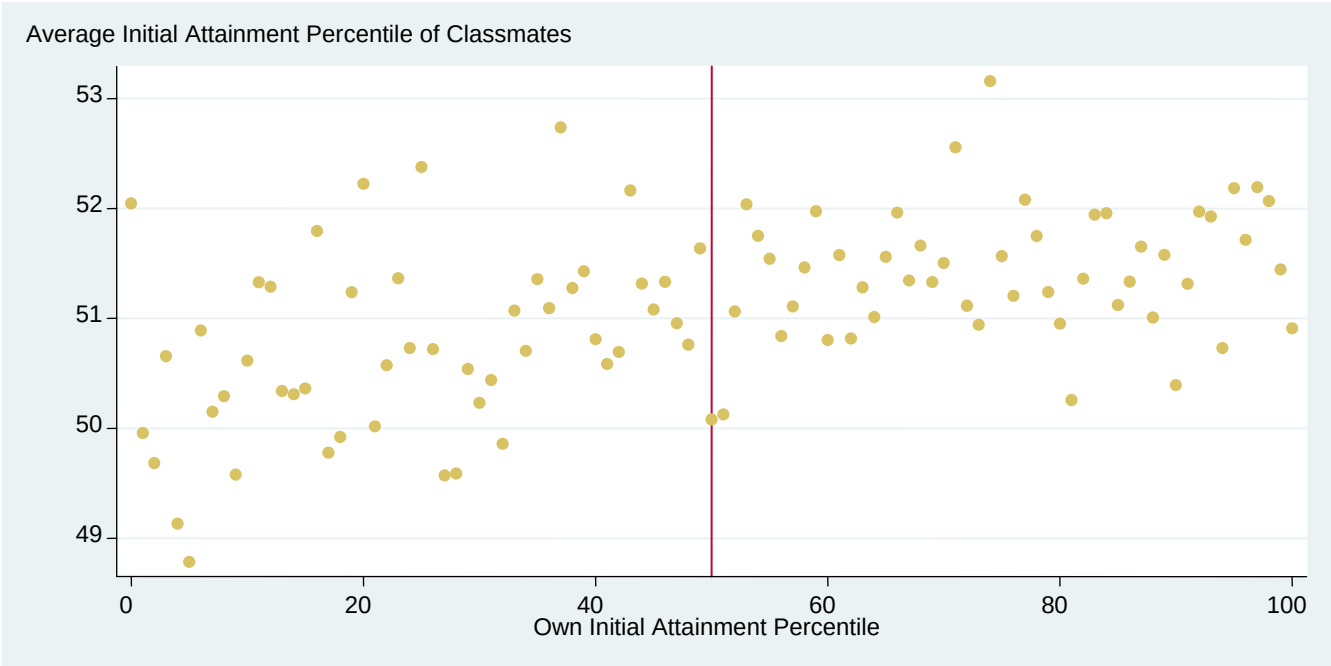
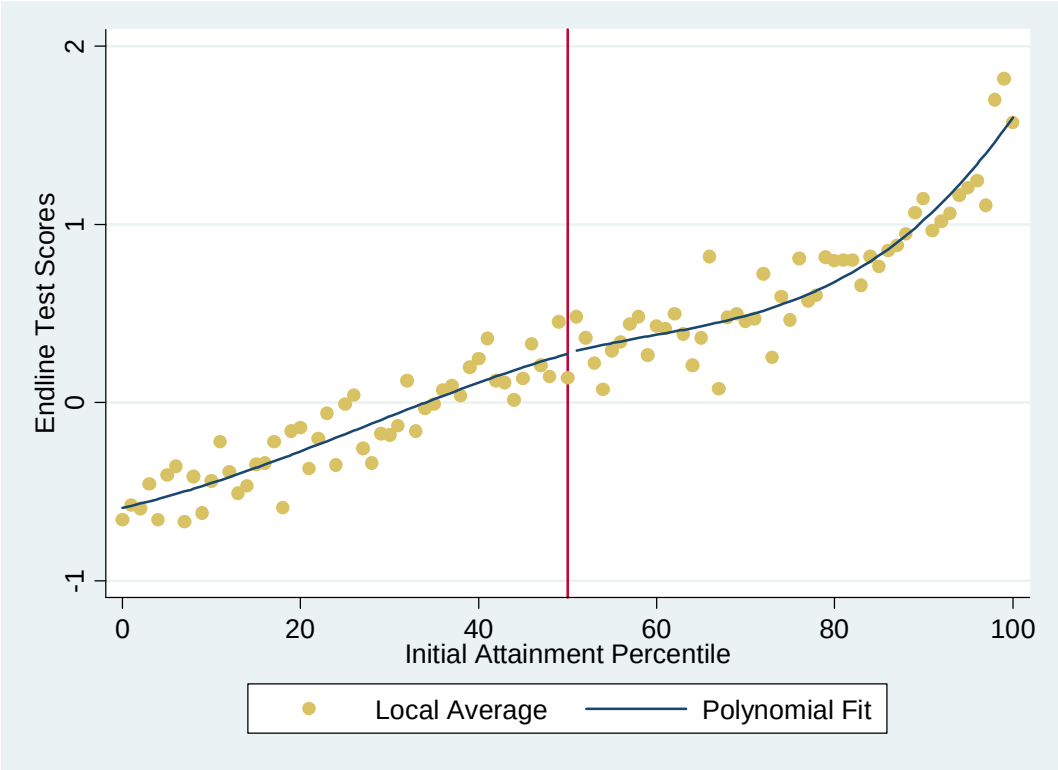


Figure 4
Regression Discontinuity Design

Panel A. Tracked ETP Schools



Panel B. Non-Tracked ETP Schools

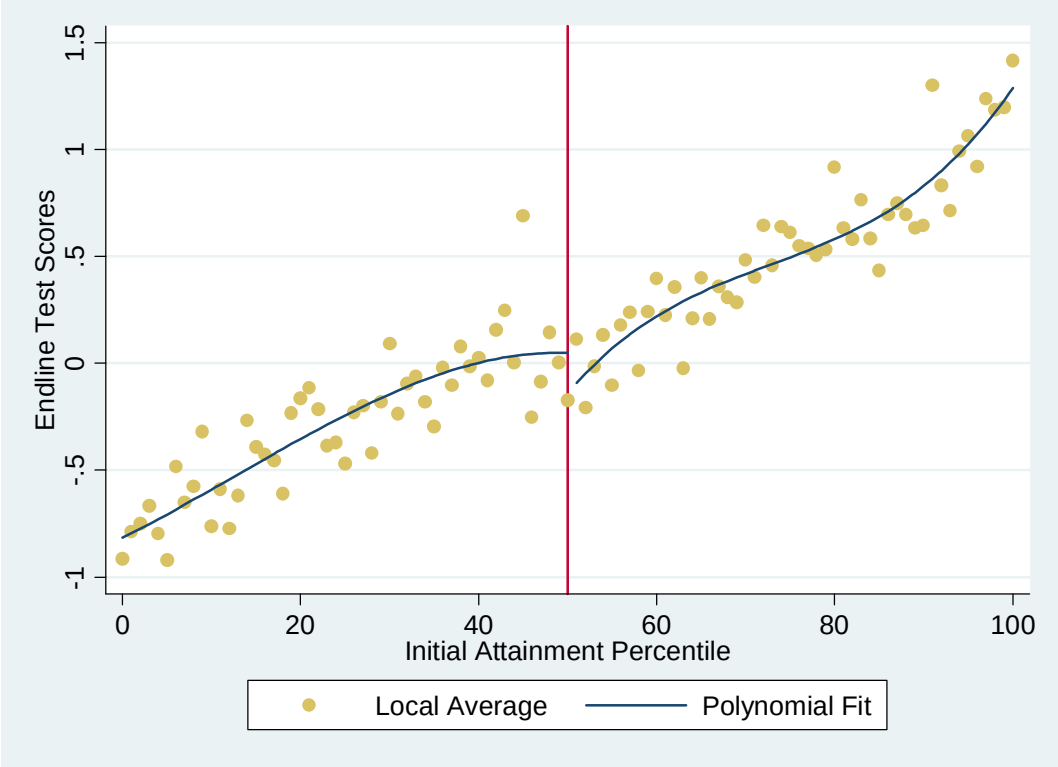
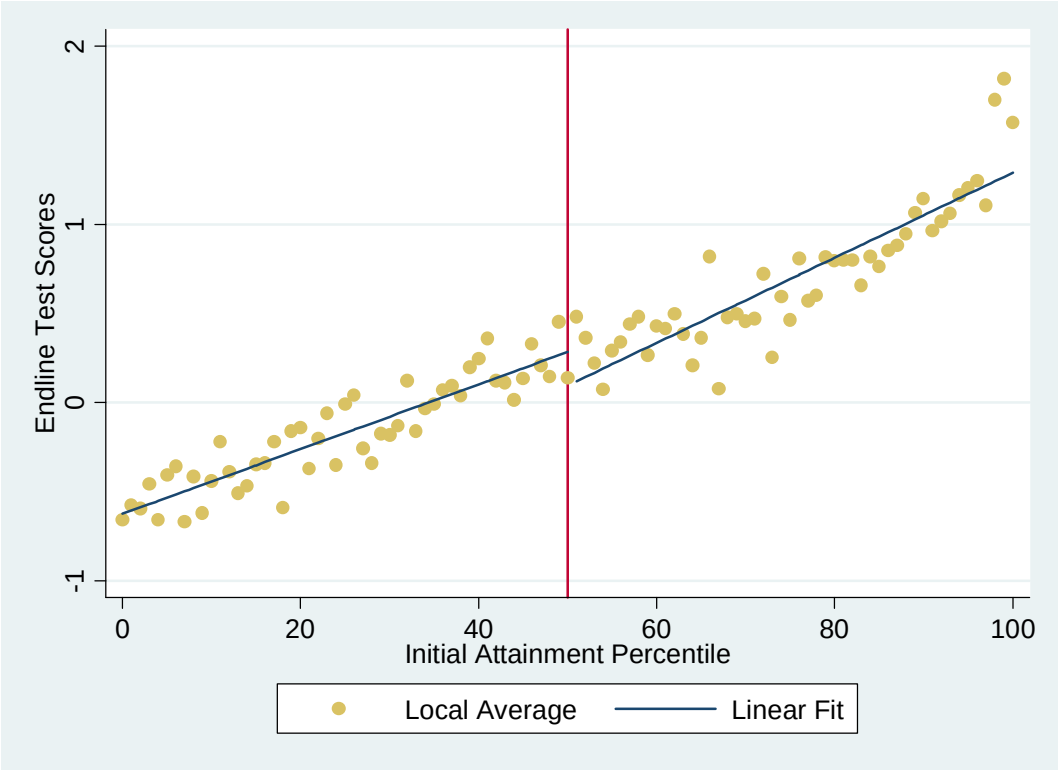


Figure 5
Regression Discontinuity Design

Panel A. Tracked ETP Schools



Panel B. Non-Tracked ETP Schools

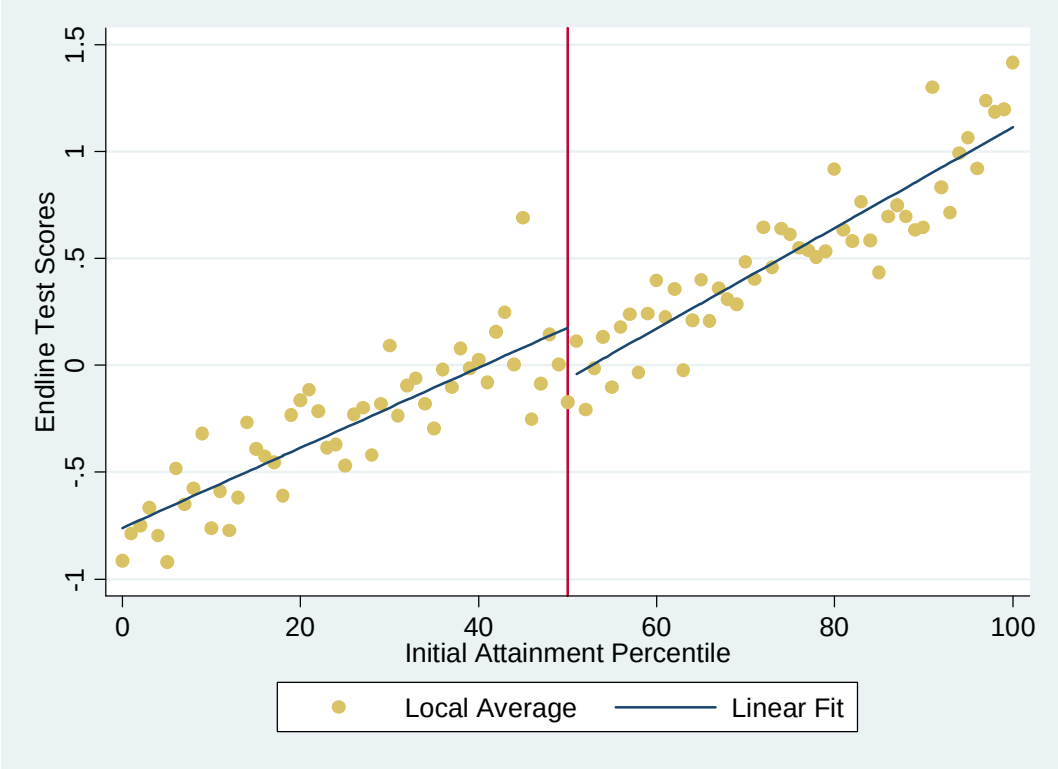


Figure 6
Endline Scores by Initial Attainment

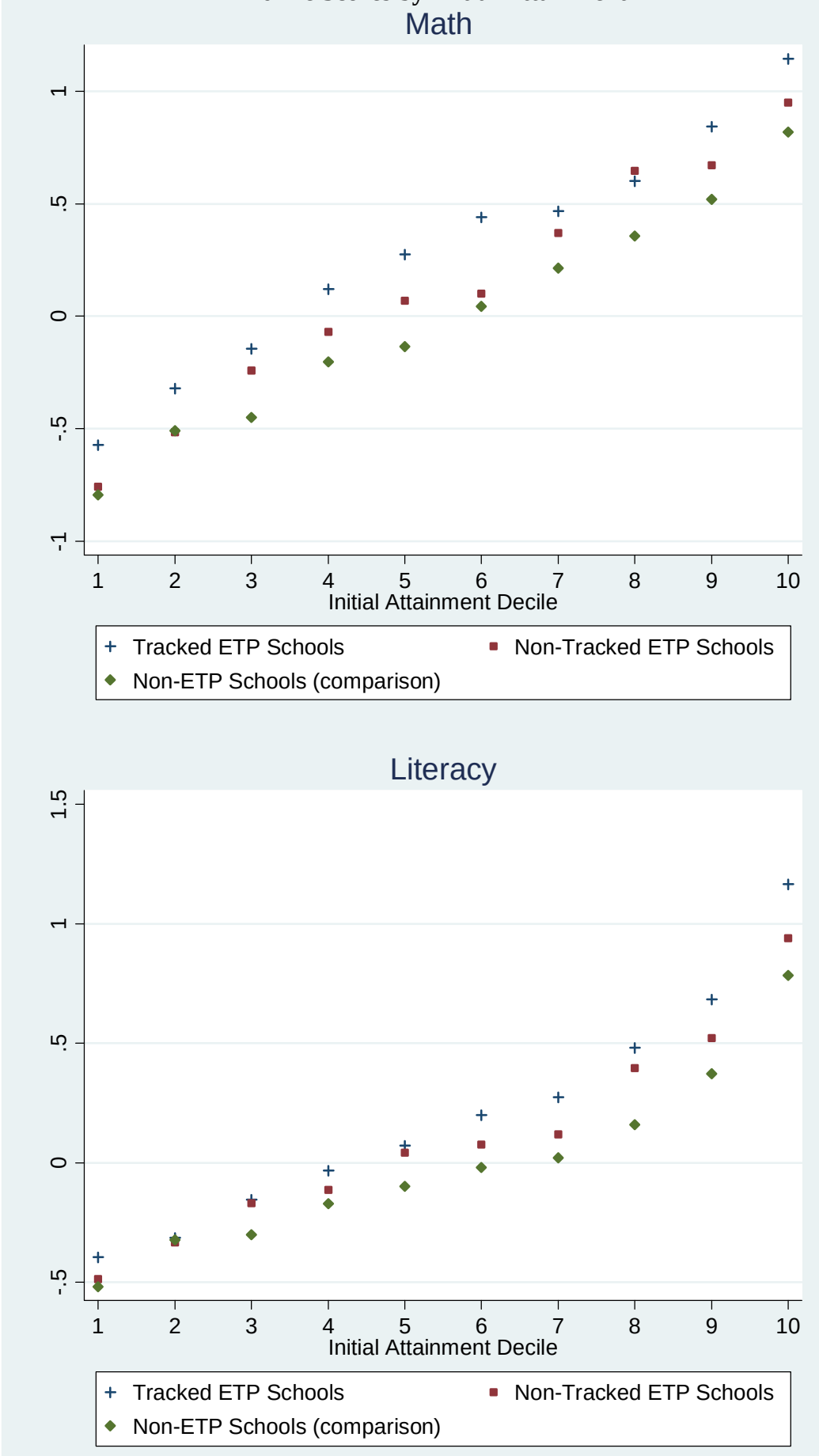
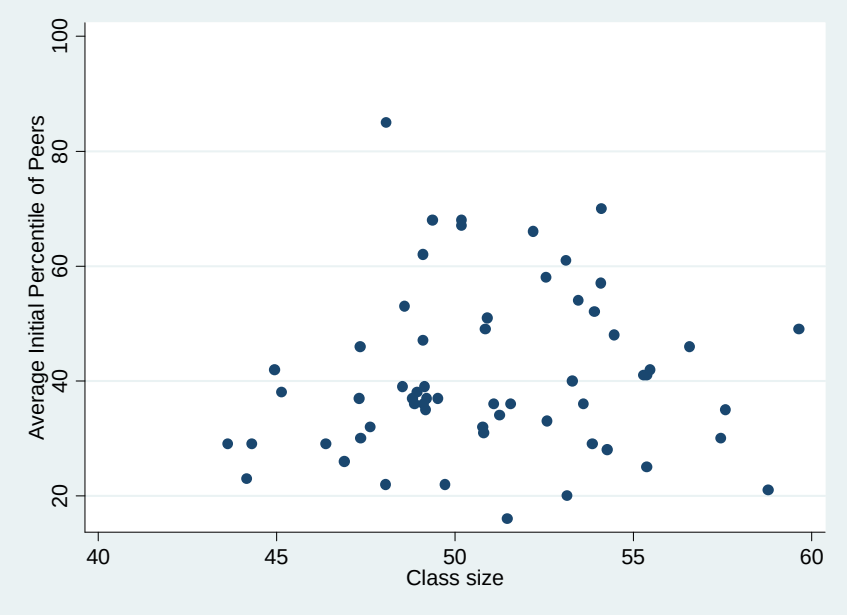


Figure 7
Variation in Peer Composition for Students assigned to Civil-Service Teachers in Schools with Randomized Assignment

Panel A: Variation in Initial Attainment of Peers



Panel B: Variation in Gender Composition of Peers

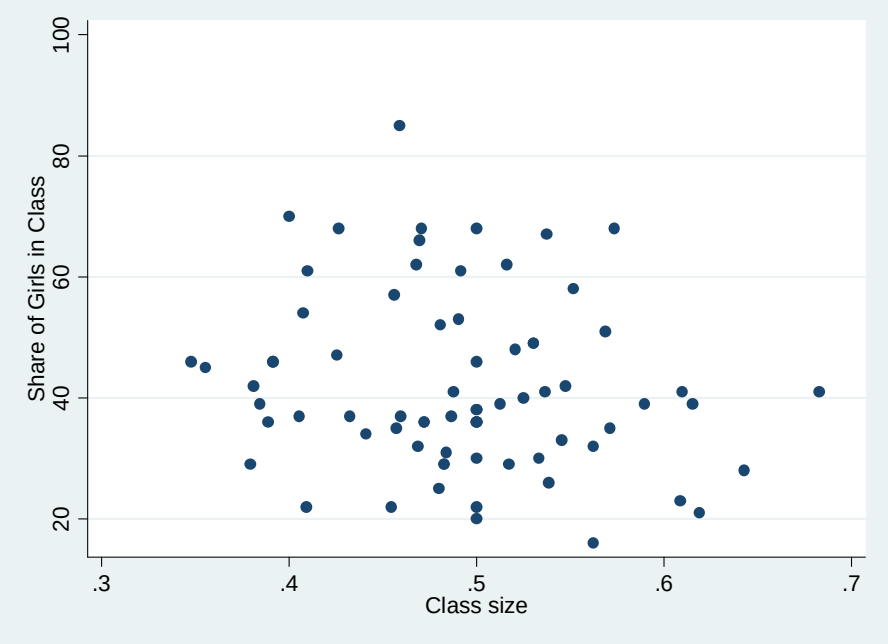


Table 1.
School and Class Characteristics, by Treatment Group, Pre- and Post-Program Inception

	All Schools	Comparison Schools	ETP Schools			Within Tracked Schools	
			All	Non-Tracked	Tracked by Initial Achievement	Lower Achievement Class	Higher Achievement Class
Panel A. Baseline School Characteristics							
Total enrollment in 2004	666 (251)	664 (259)	666 (249)	678 (263)	654 (235)		
Number of government teachers in 2004*	12.4 (3.6)	12.4 (4.1)	11.9 (3.4)	12.0 (3.8)	11.8 (2.9)		
Number of PTA teachers in 2004	1.1 (1.5)	1.2 (1.7)	1.0 (1.4)	1.2 (1.5)	0.8 (1.2)		
School pupil/teacher ratio	52 (24)	49 (15)	53 (28)	54 (36)	52 (14)		
Performance at national exam in 2004 (out of 400)	258 (24)	256 (24)	259 (24)	258 (23)	260 (25)		
Number of classes without a classroom (classes held outside)	0.54 (1.51)	0.54 (1.24)	0.55 (1.63)	0.55 (1.68)	0.55 (1.60)		
Panel B. Class Size Prior to Program Inception (March 2005)							
Average class size in first grade	83 (32)	79 (27)	84 (34)	85 (35)	84 (32)		
Proportion of female first grade students	0.49	0.50	0.49	0.49	0.48		
Proportion of schools with only one class in the first grade	0.82	0.79	0.84	0.84	0.84		
Proportion of schools with a pre-primary (ECD) class	0.93	0.94	0.92	0.89	0.96		
Average class size in second grade	83 (30)	79 (27)	84 (32)	84 (35)	85 (29)		
Panel C. Class Size 6 Months After Program Inception (October 2005)							
Average class size in first grade	59 (27)	84 (27)	46 (16)	47 (17)	45 (15)		
Range of class sizes in sample (first grade)	19-154	43 - 154	19-98	20-97	19-98		
Panel D. Class Size in Year 2 of Program (March 2006)							
Average class size in first grade	77 (29)	78 (26)	77 (30)	78 (32)	75 (28)		
Average class size in second grade	55 (23)	73 (26)	45 (15)	46 (15)	45 (15)		
Range of class sizes in sample (second grade)	18-151	30-151	18-95	18-93	21-95		
Tracking Status							
Proportion Female						0.476	0.494
Average Age						8.76 (1.46)	9.22 (1.44)
Was in preschool in 2004						0.111	0.073
Was in grade 1 in 2004						0.011	0.024
Number of Schools	**210**	**70**	**140**	**70**	**70**	**70**	

**PTA (Parents-Teachers Association) teachers are locally hired by school committees*

Table 2
Does Attrition Vary Across Treatment Arms?

	(1)	(2)	(3)
	Transferred to other school	If not transferred: missed test	Total Attrition
ETP School	-0.01	0.01	0.00
	(0.02)	(0.04)	(0.04)
(decile=2) x ETP School	-0.01	-0.05	-0.07
	(0.03)	(0.05)	(0.05)
(decile=3) x ETP School	0.04	0.02	0.04
	(0.03)	(0.05)	(0.05)
(decile=4) x ETP School	-0.01	-0.03	-0.04
	(0.03)	(0.05)	(0.05)
(decile=5) x ETP School	0.01	0.00	0.00
	(0.03)	(0.05)	(0.05)
(decile=6) x ETP School	0.01	0.03	0.03
	(0.02)	(0.05)	(0.05)
(decile=7) x ETP School	-0.02	-0.01	-0.03
	(0.03)	(0.05)	(0.05)
(decile=8) x ETP School	-0.01	-0.03	-0.03
	(0.03)	(0.05)	(0.05)
(decile=9) x ETP School	0.02	-0.03	-0.02
	(0.03)	(0.05)	(0.05)
(decile=10) x ETP School	0.02	-0.01	0.00
	(0.03)	(0.05)	(0.05)
(decile=1) x ETP School x Initial Achievement Tracking	0.02	0.01	0.03
	(0.02)	(0.03)	(0.04)
(decile=2) x ETP School x Initial Achievement Tracking	-0.02	-0.01	-0.02
	(0.03)	(0.04)	(0.05)
(decile=3) x ETP School x Initial Achievement Tracking	-0.04	-0.04	-0.06
	(0.03)	(0.04)	(0.05)
(decile=4) x ETP School x Initial Achievement Tracking	0.00	-0.01	0.00
	(0.02)	(0.04)	(0.04)
(decile=5) x ETP School x Initial Achievement Tracking	-0.02	-0.01	-0.03
	(0.03)	(0.04)	(0.05)
(decile=6) x ETP School x Initial Achievement Tracking	0.00	-0.03	-0.02
	(0.03)	(0.04)	(0.04)
(decile=7) x ETP School x Initial Achievement Tracking	0.01	-0.03	-0.01
	(0.03)	(0.04)	(0.04)
(decile=8) x ETP School x Initial Achievement Tracking	-0.01	-0.02	-0.03
	(0.03)	(0.04)	(0.04)
(decile=9) x ETP School x Initial Achievement Tracking	0.00	-0.02	-0.02
	(0.03)	(0.04)	(0.05)
(decile=10) x ETP School x Initial Achievement Tracking	-0.02	-0.01	-0.02
	(0.03)	(0.04)	(0.04)
SBM Training	0.01	-0.02	-0.01
	(0.02)	(0.03)	(0.04)
Observations	11332	10535	11332
Mean	0.08	0.13	0.18
Std Dev.	0.27	0.34	0.39
F test	1.77	1.82	1.35
Prob >F	0.01	0.00	0.09

Notes: OLS Regressions; standard errors clustered at school level. Decile dummies included but not shown.

Table 3
Teacher Effort

	(1)	(2)	(3)
	Dep Var: Teacher Found in school on random school day	Dep. Var: If found in school, found in class teaching	Dep. Var: Teacher found in class teaching (unconditional on presence)
ETP School	-0.053	-0.213	-0.233
	(0.088)	(0.106)**	(0.115)**
ETP School x Initial Achievement Tracking	0.035	0.083	0.086
	(0.028)	(0.052)	(0.051)*
ETP School x SBM	0.003	0.097	0.078
	(0.0290)	(0.054)*	(0.046)*
ETP School x Initial Achievement Tracking x SBM	-0.037	-0.029	-0.041
	(0.042)	(0.080)	(0.072)
ETP School x Contract Teacher	0.107	0.253	0.294
	(0.131)	(0.106)**	(0.089)***
ETP School x Contract Teacher x Initial Achievement Tracking	0.029	-0.086	-0.030
	(0.055)	(0.059)	(0.065)
ETP School x Contract Teacher x SBM	0.077	-0.094	-0.011
	(0.050)	(0.066)	(0.065)
ETP School x Contract Teacher x Initial Achievement Tracking x SBM	-0.076	0.009	-0.055
	(0.074)	(0.090)	(0.089)
TSC Teachers Rotation Scheme (*)	0.003	-0.210	-0.179
	(0.062)	(0.063)***	(0.103)*
TSC Teachers Rotation Scheme x ETP School	0.044	0.065	0.102
	(0.088)	(0.106)	(0.116)
TSC Teachers Rotation Scheme x ETP School x ETP Teacher	-0.073	0.077	-0.004
	(0.113)	(0.094)	(0.075)
Individual Controls			
Years of Experience Teaching	0.002	0.000	0.001
	(0.001)**	(0.001)	(0.001)
Female	0.002	0.061	0.049
	(0.0160)	(0.023)***	(0.023)**
Constant	12.460	48.751	41.024
	(7.717)	(14.092)***	(12.611)***
Observations	2951	2502	2951
Mean in non-ETP schools	0.840	0.691	0.582
F test	2.32	14.13	12.75
Prob >F	0.0037	0.0000	0.0000

Notes: Linear probability model regressions. Multiple observations per teacher. Standard errors clustered at school level. Region and date of test dummies were included in all regressions but are not shown.

(*) In schools with a rotation scheme, TSC teachers specialize in teaching a few specific subjects and rotate between classes.

Table 4
Student Attendance

	(1)	(2)	(3)
	Dep Var: Present at school on day of random visit		
ETP School	-0.064	-0.056	-0.06
	(0.042)	(0.046)	(0.046)
ETP School x Initial Achievement Tracking	0.018	0.007	0.008
	(0.016)	(0.016)	(0.016)
ETP School x SBM	0.02	0.03	0.03
	(0.015)	(0.015)**	(0.015)**
ETP School x Initial Achievement Tracking x SBM	-0.035	-0.041	-0.042
	(0.024)	(0.024)*	(0.024)*
ETP School x Contract Teacher	0.031	0.037	0.031
	(0.015)**	(0.014)***	(0.016)*
ETP School x Contract Teacher x Initial Achievement Tracking	0.001	-0.007	0.002
	(0.010)	(0.011)	(0.010)
ETP School x Contract Teacher x SBM	-0.004	-0.004	-0.004
	(0.009)	(0.009)	(0.009)
ETP School x Contract Teacher x Initial Achievement Tracking x SBM	0.014	0.007	0.009
	(0.014)	(0.014)	(0.014)
TSC Teachers Rotation Scheme (*)	-0.008	0.004	0.003
	(0.015)	(0.017)	(0.017)
TSC Teachers Rotation Scheme x ETP School	0.061	0.061	0.061
	(0.042)	(0.046)	(0.046)
TSC Teachers Rotation Scheme x ETP School x ETP Teacher	-0.021	-0.022	-0.022
	(0.013)	(0.015)	(0.013)
Girl	-0.003	-0.004	-0.004
	(0.003)	(0.003)	(0.003)
Decile in initial exam		0.004	0.006
		(0.001)***	(0.001)***
Bottom Half (Initial Attainment)			0.007
			(0.008)
ETP School x Bottom Half			0.008
			(0.010)
ETP School x Tracking with Initial Attainment x Bottom Half			0.000
			(0.008)
Constant	0.904	0.875	0.861
	(0.018)***	(0.020)***	(0.023)***
Observations	77700	68915	68915
Mean in non-ETP Schools	0.861	0.861	0.861
F test	6.085	7.376	6.58
Prob >F	0.000	0.000	0.000

Note: Multiple observations per pupil. Standard errors clustered at school level. Region and date of test dummies were included in all regressions but are not shown. Initial attainment information is missing for 14 schools, hence the decrease in the number of observations when controls for initial attainment are added.

(*) In schools with a rotation scheme, TSC teachers specialize in teaching a few specific subjects and rotate between classes.

Table 5
Average Treatment Effects on Test Scores and Grade Promotion

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
		Students in ETP Schools									
	Students in non-ETP Schools	All	Non-Tracked, All	Non-Tracked, no SBM	Non-Tracked, with SBM	Initial Achievement Tracking, no SBM	Initial Achievement Tracking with SBM	Non-Tracked, Civil-Service Teacher	Initial Achievement Tracking, Civil-Service Teacher	Non-Tracked, Contract Teacher	Initial Achievement Tracking, Contract Teacher
Panel A: Test Scores after 18 months											
Total Score	0.00	0.22	0.16	0.13	0.19	0.25	0.31	0.09	0.18	0.23	0.39
		(0.072)***	(0.081)*	(0.099)	(0.097)*	(0.092)***	(0.105)***	(0.087)	(0.098)*	(0.084)***	(0.102)***
Mathematics	0.00	0.23	0.17	0.13	0.21	0.28	0.31	0.09	0.19	0.24	0.41
		(0.063)***	(0.070)**	(0.09)	(0.078)***	(0.078)***	(0.088)***	(0.07)	(0.085)**	(0.076)***	(0.093)***
Literacy	0.00	0.17	0.12	0.11	0.14	0.16	0.25	0.08	0.13	0.17	0.29
		(0.074)**	(0.09)	(0.11)	(0.10)	(0.095)*	(0.108)**	(0.09)	(0.10)	(0.087)*	(0.098)***
Observations	3308	6678	3222	1656	1566	1643	1813	1623	1788	1599	

Notes: The table presents the mean test scores in the comparison group (Column 1) and various treatment groups (Columns 2 to 10). Scores are normalized such that the mean and standard deviation of the comparison group (non-ETP schools) are zero and one, respectively. Robust standard errors clustered at the school level in parenthesis.

Panel B: Schooling Status after 24 months											
Dropped out	0.003	0.000	0.000	0.000	0.000	-0.001	0.002	0.001	0.002	-0.001	-0.001
	0.06	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.002)	(0.001)	(0.00)
In Standard Three	0.48	0.044	0.028	0.009	0.048	0.053	0.069	-0.003	0.028	0.062	0.098
	0.50	(0.021)**	(0.024)	(0.029)	(0.025)*	(0.028)*	(0.026)**	(0.025)	(0.031)	(0.024)***	(0.032)***
Transferred	0.087	-0.001	0.002	0.009	-0.004	-0.002	-0.009	0.003	-0.010	0.001	-0.001
(grade unknown)	*0.28*	(0.009)	(0.011)	(0.013)	(0.012)	(0.012)	(0.010)	(0.012)	(0.011)	(0.011)	(0.010)
Observations	6731	12881	6576	3453	3123	2924	3381	3480	3341.000	3096.000	

Data collected in March 2007. Sample: Students enrolled in Grade 1 at program inception (March 2005). Program ended in December 2006. Column 1: Standard deviation in italics. Columns 2 to 10: clustered standard errors in parenthesis.

Table 6
Regressions Results: Test Scores

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Dep. Var: Total Score				Dep. Var: Math Score				Dep Var: Literacy Score			
	ALL	ALL	ALL	DECILES 5 & 6	ALL	ALL	ALL	DECILES 5 & 6	ALL	ALL	ALL	DECILES 5 & 6
ETP School	0.084	0.070	0.079	0.015	0.047	0.014	0.018	-0.078	0.101	0.105	0.117	0.095
	(0.094)	(0.108)	(0.101)	(0.113)	(0.075)	(0.080)	(0.079)	(0.106)	(0.106)	(0.128)	(0.119)	(0.124)
ETP School x Initial Achievement Tracking	0.232	0.215	0.209	0.260	0.269	0.233	0.227	0.413	0.152	0.156	0.150	0.070
	(0.116)**	(0.106)**	(0.101)**	(0.121)**	(0.099)***	(0.085)***	(0.083)***	(0.124)***	(0.123)	(0.127)	(0.121)	(0.130)
ETP School x SBM	0.125	0.160	0.163	0.140	0.179	0.245	0.245	0.236	0.052	0.052	0.056	0.026
	(0.120)	(0.128)	(0.120)	(0.155)	(0.092)*	(0.099)**	(0.096)**	(0.138)*	(0.133)	(0.143)	(0.133)	(0.165)
ETP School x Initial Achievement Tracking x SBM	-0.283	-0.158	-0.174	-0.133	-0.343	-0.259	-0.263	-0.244	-0.173	-0.035	-0.060	-0.008
	(0.183)	(0.169)	(0.164)	(0.187)	(0.149)**	(0.137)*	(0.136)*	(0.176)	(0.195)	(0.186)	(0.178)	(0.199)
ETP School x Assigned to Contract Teacher	0.229	0.226	0.204	0.186	0.260	0.235	0.195	0.179	0.155	0.172	0.172	0.155
	(0.067)***	(0.067)***	(0.101)**	(0.130)	(0.062)***	(0.059)***	(0.089)**	(0.133)	(0.070)**	(0.076)**	(0.109)	(0.137)
ETP School x Initial Achievement Tracking x Assigned to Contract Teacher	-0.239	-0.083	-0.060	0.135	-0.252	-0.084	-0.073	0.080	-0.179	-0.065	-0.038	0.157
	(0.163)	(0.095)	(0.099)	(0.149)	(0.158)	(0.091)	(0.092)	(0.164)	(0.150)	(0.107)	(0.112)	(0.152)
ETP School x SBM x Assigned to Contract Teacher	-0.168	-0.178	-0.112	0.089	-0.215	-0.245	-0.208	-0.113	-0.093	-0.084	-0.005	0.252
	(0.101)*	(0.107)*	(0.109)	(0.161)	(0.102)**	(0.101)**	(0.100)**	(0.156)	(0.100)	(0.108)	(0.114)	(0.174)
ETP School x Initial Achievement Tracking x SBM x Assigned to Contract Teacher	0.581	0.284	0.239	0.017	0.590	0.332	0.307	0.090	0.457	0.185	0.131	-0.051
	(0.238)**	(0.155)*	(0.158)	(0.226)	(0.227)***	(0.147)**	(0.148)**	(0.226)	(0.221)**	(0.162)	(0.167)	(0.240)
TSC Teachers Rotation Scheme (*)	-0.266	-0.288	-0.292	-0.387	-0.173	-0.193	-0.192	-0.314	-0.295	-0.316	-0.323	-0.374
	(0.197)	(0.194)	(0.184)	(0.229)*	(0.172)	(0.170)	(0.165)	(0.199)	(0.193)	(0.192)	(0.178)*	(0.236)
Teacher's Age			-0.003	0.003			-0.004	-0.002			-0.002	0.007
			(0.005)	(0.006)			(0.005)	(0.006)			(0.005)	(0.006)
Ratio of Female Among Teachers (**)			0.286	0.308			0.168	0.163			0.334	0.375
			(0.071)***	(0.096)***			(0.064)***	(0.090)*			(0.074)***	(0.095)***
Bottom Half (Initial Attainment)		-0.757	-0.756	-0.141		-0.797	-0.797	-0.170		-0.569	-0.568	-0.086
		(0.040)***	(0.040)***	(0.056)**		(0.037)***	(0.036)***	(0.056)***		(0.044)***	(0.044)***	(0.060)
ETP School x Bottom Half		-0.068	-0.070	0.096		-0.038	-0.040	0.123		-0.080	-0.082	0.052
		(0.059)	(0.058)	(0.088)		(0.052)	(0.052)	(0.093)		(0.068)	(0.067)	(0.094)
ETP School x Initial Achievement Tracking x Bottom Half		-0.027	-0.051	-0.132		0.037	0.024	-0.139		-0.079	-0.108	-0.099
		(0.066)	(0.066)	(0.097)		(0.061)	(0.061)	(0.106)		(0.075)	(0.075)	(0.107)
School Size	-0.070	-0.083	-0.114	-0.032	-0.177	-0.190	-0.211	-0.145	0.039	0.028	-0.005	0.075
	(0.085)	(0.089)	(0.089)	(0.105)	(0.073)**	(0.076)**	(0.076)***	(0.100)	(0.091)	(0.096)	(0.095)	(0.107)
Girl	0.116	0.081	0.081	0.007	0.060	0.023	0.023	-0.062	0.143	0.116	0.116	0.068
	(0.023)***	(0.022)***	(0.022)***	(0.041)	(0.021)***	(0.020)	(0.020)	(0.039)	(0.024)***	(0.024)***	(0.024)***	(0.046)
Age	0.023	-0.029	-0.030	-0.079	0.061	0.009	0.009	-0.045	-0.015	-0.057	-0.058	-0.094
	(0.010)**	(0.010)***	(0.010)***	(0.014)***	(0.009)***	(0.009)	(0.009)	(0.015)***	(0.010)	(0.011)***	(0.011)***	(0.015)***
Observations	9931	9225	9200	1868	9931	9225	9200	1868	9934	9228	9203	1868
F test	9.20	84.01	76.51	7.33	14.25	114.87	102.61	7.15	7.73	39.05	35.70	5.85
Prob >F	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

*Notes: OLS regressions. Standard errors clusters at school level (number of clusters: 208). *** indicates significance at 1%. Region and date of test dummies were included in all regressions but are not shown. Initial attainment information is missing for 14 schools, hence the decrease in the number of observations when controls for initial attainment are added.*

Scores are normalized such that the mean and standard deviation of the comparison group are zero and one, respectively.

(*) In schools with a rotation scheme, TSC teachers specialize in teaching a few specific subjects and rotate between classes.

(**) In schools where the ETP teacher did not change, the sex ratio is either 0 or 1. In schools with a rotation scheme, the average ratio of female among rotating teachers is 0.67.

Table 7
Regressions Results: Grade Promotion

	(1)	(2)	(3)	(4)
	Dep Var: In Standard Three in 2007			
	ALL	ALL	ALL	DECILES 5 & 6
ETP School	-0.028	-0.006	-0.010	-0.013
	(0.026)	(0.030)	(0.029)	(0.050)
ETP School x Initial Achievement Tracking	0.070	0.004	0.002	-0.003
	(0.040)*	(0.033)	(0.033)	(0.059)
ETP School x SBM	0.052	0.028	0.032	-0.007
	(0.030)*	(0.038)	(0.036)	(0.064)
ETP School x Initial Achievement Tracking x SBM	-0.076	0.008	0.005	0.046
	(0.057)	(0.050)	(0.048)	(0.085)
ETP School x Assigned to Contract Teacher	0.069	0.054	0.002	0.041
	(0.019)***	(0.023)**	(0.035)	(0.061)
ETP School x Initial Achievement Tracking x Assigned to Contract Teacher	-0.045	0.019	0.018	-0.058
	(0.060)	(0.034)	(0.035)	(0.055)
ETP School x SBM x Assigned to Contract Teacher	-0.013	0.006	0.002	-0.022
	(0.030)	(0.033)	(0.033)	(0.061)
ETP School x Initial Achievement Tracking x SBM x Assigned to Contract Teacher	0.100	-0.045	-0.034	-0.022
	(0.088)	(0.049)	(0.048)	(0.084)
TSC Teachers Rotation Scheme (*)	-0.022	-0.010	-0.006	-0.076
	(0.043)	(0.042)	(0.041)	(0.050)
Teacher's Age			-0.004	-0.005
			(0.002)**	(0.003)*
Ratio of Female Among Teachers (**)			0.032	0.027
			(0.020)	(0.032)
Bottom Half (Initial Attainment)		-0.378	-0.378	-0.072
		(0.016)***	(0.016)***	(0.027)***
ETP School x Bottom Half		0.008	0.008	-0.043
		(0.027)	(0.027)	(0.039)
ETP School x Initial Achievement Tracking x Bottom Half		0.030	0.024	0.088
		(0.027)	(0.027)	(0.041)**
School Size		-0.054	-0.063	-0.067
		(0.033)	(0.031)**	(0.052)
Girl	-0.048	-0.054	-0.063	-0.067
	(0.033)	(0.033)	(0.031)**	(0.052)
Constant	0.554	0.740	0.896	0.912
	(0.055)***	(0.052)***	(0.083)***	(0.140)***
Observations	19358	16919	16874	3412
F test	5.819	99.771	5.815	4.643
Prob >F	0.000	0.000	0.001	0.000
Mean in non-ETP schools	0.483	0.483	0.483	0.483

*Notes: OLS regressions. Standard errors clusters at school level. *** indicates significance at 99%. Region dummies were included in all regressions but are not shown.*

(*) In schools with a rotation scheme, TSC teachers specialize in teaching a few specific subjects and rotate between classes.

(**) In schools where the ETP teacher did not change, the sex ratio is either 0 or 1. In schools with a rotation scheme, the average ratio of female among rotating teachers is 0.67.

Table 8
Estimating Peer Effects Using Variation in Peer Composition
(Sample: Students assigned to Civil-Service Teachers in Non-Tracked ETP schools)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Dep. Var: Total Score				Dep. Var: Math Score				Dep. Var: Literacy Score			
Average Initial Percentile of Classmates	-0.010	-0.010		-0.010	0.000	0.000		0.010	-0.010	-0.010		-0.020
	(0.010)	(0.010)		(0.020)	(0.010)	(0.010)		(0.020)	(0.020)	(0.020)		(0.020)
Std Dev. of Initial Percentile of Classmates			0.000	0.010			-0.010	-0.020			0.010	0.030
			(0.020)	(0.030)			(0.020)	(0.030)			(0.010)	(0.030)
% of Girls among Classmates	-3.830	-3.880	-3.590	-3.930	-3.430	-3.470	-3.600	-3.280	-3.420	-3.470	-2.850	-3.720
	(2.810)	(2.790)	(2.570)	(2.780)	(2.960)	(2.950)	(2.670)	(2.940)	(2.820)	(2.820)	(2.620)	(2.750)
Girl x % of Girls among Classmates		-0.820	-0.800	-0.820		-0.560	-0.570	-0.560		-0.880	-0.840	-0.880
		(0.820)	(0.830)	(0.820)		(0.840)	(0.840)	(0.840)		(0.890)	(0.900)	(0.890)
Girl	-0.060	0.330	0.330	0.330	-0.150	0.110	0.110	0.120	0.040	0.460	0.460	0.450
	(0.080)	(0.420)	(0.420)	(0.420)	(0.08)*	(0.410)	(0.410)	(0.410)	(0.080)	(0.460)	(0.460)	(0.460)
Own Initial Decile	0.160	0.160	0.160	0.160	0.170	0.170	0.170	0.170	0.120	0.120	0.130	0.120
	(0.01)***	(0.01)***	(0.01)***	(0.01)***	(0.01)***	(0.01)***	(0.01)***	(0.01)***	(0.01)***	(0.01)***	(0.01)***	(0.01)***
Age	-0.060	-0.060	-0.060	-0.060	-0.030	-0.020	-0.020	-0.020	-0.080	-0.080	-0.080	-0.080
	(0.02)***	(0.02)***	(0.02)***	(0.02)***	(0.020)	(0.020)	(0.020)	(0.020)	(0.02)***	(0.02)***	(0.02)***	(0.02)***
Class Size	0.010	0.010	0.020	0.000	0.060	0.060	0.070	0.090	-0.040	-0.040	-0.020	-0.070
	(0.040)	(0.040)	(0.040)	(0.070)	(0.050)	(0.050)	(0.050)	(0.060)	(0.050)	(0.050)	(0.040)	(0.080)
Constant	1.370	1.490	0.440	1.810	-1.700	-1.620	-1.460	-2.720	3.830	3.970	2.050	5.480
	(2.950)	(2.900)	(1.640)	(3.700)	(3.190)	(3.200)	(1.610)	(3.600)	(3.570)	(3.490)	(2.270)	(4.330)
Observations	1322	1322	1322	1322	1322	1322	1322	1322	1322	1322	1322	1322
F test	47.01	41.55	40.54	39.45	97.05	82.43	60.03	66.81	15.11	16.02	18.30	17.47
Prob >F	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Notes: The table reports results of linear regressions with school fixed effects. Robust standard errors clustered at the school level are reported in parenthesis. The estimations are based on the 57 non-tracked ETP schools for which data on initial achievement is available. Region and date of test dummies were included in all regressions but are not shown.

The average initial percentile of classmates ranges from 41.6 to 63.4, with a mean of 51.0 and a median of 50.7. The standard deviation of the initial percentile of classmates ranges from 24 to 32, with a mean of 29 and a median of 29. The share of girls among classmates ranges from 0.33 to 0.70, with a mean of 0.49 and a median of 0.49.

Scores are normalized such that the mean and standard deviation are zero and one, respectively.