Classmate Turnover and Student Achievement: New Findings from Project STAR

Marc Luppino*

[Job Market Paper]

November 2nd, 2009

Abstract

New classroom assignment at the onset of each grade means that a student's peer group regularly changes over an academic career. This is one of the first papers to examine how peer turnover impacts student outcomes. I use the random variation in classroom composition induced by Tennessee's Student Teacher Achievement Ratio (STAR) experiment to identify the causal effect of classmate turnover on academic achievement. In schools outside of center cities, I find that first graders who experience greater classmate turnover perform worse in reading and math. Conversely, peer turnover is found to be beneficial for young students in center city schools. These results are consistent with a model of classroom learning in which familiar classmates can either be a resource or a distraction. They suggest that a richer understanding of peer continuity effects is essential for designing optimal classroom assignment policies.

---

* UCLA Department of Economics. mluppino@ucla.edu. This paper has benefited from many helpful insights from my dissertation committee: Sandra Black, Leah Platt Boustan, and Moshe Buchinsky (chair). Thanks also to Pascaline Dupas and the participants in UCLA's applied micro seminar for thoughtful comments and suggestions. All remaining errors are mine.

1) Introduction

There is a large literature on the role of peers on student performance. While much of this has examined the role of student composition by race, gender, and ability, there has been surprisingly little about the role of consistency of peers. New classroom assignment at the onset of each grade means that a student's peer group regularly changes over an academic career. This is one of the first papers that attempts to estimate the effect of classmate turnover for student achievement.

Most research on turnover effects has arisen because of concerns over the consequences of student mobility. A 1994 study by the United States General Accounting Office found that slightly over 40 percent of children changed schools at least once between first and third grade. Student mobility, in turn, leads to high rates of classroom turnover. Alexander et al. (1996) report that classroom turnover of 50 percent or more during the school year is not unusual in many inner-city schools. High pupil turnover creates classroom management challenges for teachers. In particular, there is a tendency for teachers to respond to high student mobility rates by reducing their instructional pace to accommodate the variation and uncertainty in students' prior learning (Kerbow, 1996). Student mobility and turnover also leads to the break up of peer groups which may increase students' adjustment problems (Temple & Reynolds, 1999).

Administrators and policymakers have limited scope to influence inward and outward school mobility, since this process is largely driven by parental choices. However, school administrators do have control over student mobility *within* schools. Classroom assignment decisions also generate peer turnover as students progress from grade to grade. Better understanding classmate turnover effects is essential for designing optimal classroom assignment policies, especially for young children.

In the United States, much of the research and policy debate on optimal classroom assignment centers around the merits of ability tracking (see Betts and Shkolnik, 2000). Conversely in Israel, education policy stresses peer continuity and the promotion of social capital

(see Katriel and Nesher, 1986). Israeli first graders are arbitrarily assigned to classes. These initial class groupings are kept intact as students advance through all grades of elementary school. It is an open question as to how, if at all, these differing classroom assignment policies affect child development.

Because of the endogenous nature of school choice and school turnover rates, it is difficult to evaluate the consequences of peer turnover. There is a persistent concern that estimates of turnover effects, or externalities, are driven by omitted variable bias (e.g. that unobserved factors lead to both higher student turnover and worse academic performance and, therefore, generate spurious correlation between these two variables). Such concerns are still relevant when we shift focus to turnover induced by classroom assignment. To overcome these issues, this paper takes a unique approach. Specifically, I use random variation in classroom composition generated by Tennessee's Student Teacher Achievement Ratio (STAR) experiment to identify the causal effect of classmate turnover on student achievement.

Importantly for the purpose of the present study, all students participating in the STAR experiment who were initially assigned to regular size classes in kindergarten were again randomly assigned to regular size classrooms in first grade. This induced random variation within first grade classes in the percentage of children who previously shared the same class in kindergarten. Students assigned to a first grade classroom with a larger share of prior classmates experienced relatively less classmate turnover. In my analysis, I compare the performance of students who experience different rates of classmate turnover. To account for other potentially endogenous sources of variation in class composition, such as sample attrition or deviations from random assignment, I also control for kindergarten and first grade class fixed effects. Using the re-randomization and classroom fixed effects, I am able to identify the impact of classmate turnover on reading and math achievement.

I find that classmate turnover has contrasting effects on student performance depending on school location. Outside of center city school districts, first graders perform better when they experience less classmate turnover. Peer turnover has an especially large negative effect on above-average students in schools outside of center cities. In these schools, assigning a teacher's aide to the class partially mitigates the negative impact of classmate turnover on reading and

math achievement. Conversely, I find that turnover has a positive effect on students in center city school districts (whom are predominantly African-American and poor). Above-average students in center city schools benefit the most from peer turnover. My findings are robust to a number of alternative specifications and sample restrictions.

These results are consistent with a model of classroom learning in which familiar classmates can either be a resource or a distraction. In particular, they suggest that peer turnover possibly benefits disadvantaged children in center city schools by reducing their susceptibility to distraction or to peer pressure to misbehave. In other schools where the incidence of behavioral problems is not as high, classmate turnover has a negative effect on student performance. This is likely the case if sharing a class with previous classmates aids students in influencing the difficulty level and pace of classroom discussion and instruction. Future research needs to further investigate the mechanisms by which peer turnover affects student achievement. The findings of this study suggest that a richer understanding of peer continuity effects is essential for designing optimal classroom assignment policies.

The remainder of the paper is organized as follows. In Section 2, I review the related literature on student mobility and turnover externalities. Section 3 presents the data. In Section 4, I discuss my empirical strategy for identifying the effect of classmate turnover on student achievement. Section 5 presents my main findings. In Section 6, I test the robustness of these results. Finally in Section 7 I conclude.

2) Related Literature

The large literature on student mobility highlights the potential importance of classmate turnover. While most studies focus on how mobility affects movers, authors such as Alexander et al. (1996) and Kerbow (1996) raise the possibility that the turnover induced by student mobility may also affect non-movers. In particular, Hanushek et al. (2004) suggest that high student turnover can disrupt orderly teaching and curriculum development, imposing serious externalities from mobility.

While the present study is the first to examine the potential effects of student mobility

within a school, researchers have investigated the effects of *between* school mobility on student outcomes. Hanushek et al. (2004) find that about a third of all students in Texas move at least once in elementary and middle school and that these moves adversely affect the academic performance of both movers and students in the receiving schools. Imberman, Kugler, and Sacerdote (2009) similarly find that the large influx of new students in Texas and Louisiana resulting from the Hurricane Katrina evacuation reduced the test scores of native students in those states. Angrist and Lang (2004) study the Metropolitan Council for Educational Opportunity (Metco), which introduced urban minorities into affluent suburban schools in the Boston area, and conversely find that this desegregation program had modest and short lived negative effects on suburban students.

One limitation of previous studies is that they do not disentangle the pure effect of turnover from any effects resulting from changes in the composition of peer quality. For example, the Hurricane Katrina evacuation introduced a large influx of new students into schools in Texas and Louisiana, potentially producing a turnover effect. In addition, the evacuee children came from some of the worst-performing schools in the country. As a result, we might expect them to have a negative peer effect on native students in the schools that received evacuee children. In the end, these types of "natural experiments" do not allow us to separately investigate these distinct effects. A key contribution of this paper is that I ensure identification of a pure turnover effect by controlling for classroom fixed-effects, which directly account for changes in class peer composition.

Additionally because of the endogenous nature of student mobility, it is very difficult to identify how student turnover or integration at the school level affect learning. Potential endogeneity is also an issue when trying to identify the effect of peer turnover at the classroom level. Specifically, children with more family resources may systematically sort into smaller schools with fewer classrooms. In such schools, there would be less scope for classmate turnover. Unobservable parental inputs that drive this type of sorting are also likely to directly influence student outcomes and would otherwise bias estimates of the effect of classmate continuity. In order to avoid this problem, I take advantage of the randomized assignment of students to first grade classrooms conducted as part of the STAR experiment.

This is the first study to investigate the effect of classmate turnover on achievement using

the Project STAR data. However, it joins a number of studies that use this experiment's randomization to investigate other aspects of educational production not directly related to class size (see, for example, Dee, 2004; Nye et al., 2005; Whitmore, 2005; Graham, 2008; and Cascio and Whitmore-Schazenbach, 2008). While most of these studies exploit the initial random assignment to class type, the present study is one of the first to take advantage of the re-randomization before first grade. The only other paper to do so, Sojourner (2008), uses the variation resulting from the re-randomization to examine the problem of missing data when estimating peer effects.

Of the existing research based on this data, the current study is most similar to the work of Boozer and Cacciola (2001), who examine peer effects. Specifically, these authors attempt to identify peer effects by instrumenting the contemporaneous average of classmate test scores with the fraction of the class previously assigned to a small class. As part of the experimental protocol, students initially assigned to small classes were to stay in such classes throughout the course of the STAR study. Therefore, this instrument measures the degree of classmate turnover experienced by students in small classes. Variation in the fraction of classmates assigned to small classes largely results from the introduction of new students into these classes and the potentially endogenous attrition of students from these classes. Conversely, in this study I rely on exogenous variation in classroom composition induced by the randomized assignment of students to first grade regular sized classrooms. In addition by using kindergarten and first grade class fixed-effects, it is less likely that my results are driven by endogenous variation resulting from attrition or deviations from random assignment.

3) Data

In order to assess the effect of classmate turnover on student achievement, I use data from the Project STAR experiment. Project STAR was a large-scale randomized experiment commissioned by the state of Tennessee to exam the effect of classroom size on student achievement. The experiment followed a cohort of students from kindergarten through third grade at seventy-nine public schools. Those starting kindergarten in 1985 were randomly assigned to a small-size class (with a target of thirteen to seventeen students), a regular-size class

(with a target of twenty-two to twenty-five students), or a regular-size class with a full-time teacher's aide. Teachers were also randomly assigned to one of the three class types. All randomization was conducted within schools.

Because of requirements imposed by the Tennessee legislature for geographic diversity, schools in cities are overrepresented in the STAR data. Thirty-one of the seventy-nine original participating schools were drawn from the central school districts of Chattanooga, Knoxville, Memphis, and Nashville. Of these schools, fifteen were located in inner city Memphis. Students in the corresponding center city school districts accounted for approximately 29 percent of students in Tennessee public schools. Conversely, children from center cities make up roughly 45 percent of STAR students. As a result, students participating in the experiment were more economically disadvantaged and more likely to be African-American than those in the state overall.

The sampling of schools is especially important if classmate turnover has a differential effect on particular types of students. Specifically, Hanushek et al. (2004) suggest that low income and minority students (who are more heavily concentrated in large urban districts) are more sensitive to classmate turnover. One would need appropriate sampling weights to estimate the average effect of classmate turnover using the STAR data. As Hanushek (1999) notes such weights are not available. Therefore, one must be careful interpreting results based on the full sample of schools. In an attempt to mitigate this issue, I conduct separate analyses for center city and non-center city schools.

My sample consists of first graders who began the STAR program in kindergarten and who remained in the same school and in a regular size classroom in both kindergarten and first grade. While there were 76 schools who participated in Project STAR during the first grade phase of the project, one school is dropped from the sample because it contained only a single regular size class. The sample is also limited to students without missing data on test scores and personal characteristics. With these restrictions, the initial sample covers students from 211 classes in 75 schools.

Summary statistics for all the variables used in the subsequent analysis are reported in **Table I**. Within the full sample, the average student experienced a classmate turnover rate of 77 percent between kindergarten and first grade. This means that in a classroom of 23 students,

returning first graders knew on average five of their current classmates from their kindergarten class. Classmate turnover was particularly high since kindergarten was not mandatory in Tennessee at the time that Project STAR was implemented. As a result, many new students entered school in first grade. The large inflow of new students is not particularly problematic for my analysis for two reasons. First, I do not directly consider new students in my analysis since I limit my sample to those students who were in Project STAR schools in both kindergarten and first grade. Second, I account for any indirect effects of new students by controlling for first grade class fixed effects. All my results on the impact of classmate turnover are based on within-first grade classroom variation.

The summary statistics demonstrate how students in center city schools are very different from those in the other participating schools. Center city first graders experience higher rates of classmate turnover. This is partly the case because center city schools experienced higher levels of attrition after kindergarten relative to other schools that participated in Project STAR (44 percent relative to 24 percent in schools outside of center cities). Students in these schools are disproportionately African-American (65 percent) and poor (62 percent received free lunch in either kindergarten or first grade) relative to students from schools outside of center cities.

While relatively non-existent at the end of kindergarten, a sizable test score gap between students in center city and non-center city schools emerges by the end of first grade. On average, returning first graders in center city schools score around half a standard deviation lower on the Stanford Achievement Test (SAT) reading and math examinations than students in schools outside of the center cities. In addition to the highlighted differences in student demographics, this finding suggests that center and non-center city schools potentially differ in their schooling inputs and educational practices (or technology). As such, it is likely that the effect of classmate turnover on student achievement will differ by school location.

4) Empirical Strategy

My identification strategy relies on the fact that all students initially assigned to regular size classes (either with or without a teacher's aide) were again randomly assigned to the two regular size class types before the start of first grade (see Whitmore Schanzenbach, 2007). This

induced random variation within 1st grade classes in the percentage of children who previously shared the same class in kindergarten. I use this variation to identify how being in classes with varying proportions of prior classmates affects student performance.

The re-assignment process generated random variation in the peer composition of first grade classrooms. For illustration, **Table II** presents the classroom transition matrix for regular size classroom students in one of the Project STAR schools (School ID #168214). Students in both kindergarten classrooms (K-1 and K-2) were randomly assigned to one of the two first grade classrooms in this school. In addition, new students were also randomly assigned to their first grade class. On average, each first grade classroom will contain an equal number of students from the two kindergarten classes. However, there is a very high probability that the random assignment process will generate within-classroom variation in the percentage of children whom previously shared the same class in kindergarten. In the analysis that follows, I compare the performance of students with a larger share of prior classmates (e.g. K-2 students in class F-1) relative to students in the same classroom with a smaller share of prior classmates (e.g. K-1 students in class F-1).

There has been some contention in the literature regarding the exact nature of the randomized assignment conducted as part of Project STAR. While it is clear that the experiment was designed to randomly assign students to class type, the project documentation does not specify whether students were randomly assigned to classrooms within class type. However according to accounts by the original Project STAR investigators, the project team took each school's list of teachers and new students, randomly assigned them to specific classes within type, provided the principals with these class assignments, and monitored their compliance through frequent site visits during the school year (see Sojourner, 2008).[1]

1 Krueger (1999) highlights, students may not have maintained their random class assignments. Approximately 10 percent of students appear to have switched between small and regular classes between grades, with reports listing behavioral problems or parental complaints as the primary reasons for these deviations from the experimental protocol. The Project STAR data does not allow us to determine how many students potentially switched classrooms within their assigned class types. If departures from random assignment were limited and idiosyncratic, then they should not bias the main findings of this paper. However if the execution of Project STAR left any room for systematic sorting across classrooms, then my findings would be put into question.

Such concerns are unlikely to affect my results for two reasons. First, it is important to note that the relatively small size of Project STAR schools left less scope for within school non-random sorting of students. Specifically, Zabel (2008) argues that school-grades with four or fewer classes have quasi-random classroom assignment since this leaves less scope for sorting than in school-grades with more classes. Of the 75 schools in the full sample, only one (center city) school has more than 4 regular size 1st grade classes (of either type).

As noted previously, I restrict my sample to returning students. This is done because the randomized assignment to first grade classrooms does not affect the degree of classmate turnover experienced by new students. New students did not share a prior class with any returning students and we cannot observe whether they shared a prior class with any of each other. Since new students were randomly assigned to first grade classrooms, their distribution should not affect the analysis of returning students. In addition, even if new students deviated from random assignment, classroom fixed effects should account for their allocation across classrooms within the same school.

Following Clotfelter, Ladd, and Vigdor (2006), I conduct a series of chi-square tests to compare the balance of observable characteristics of returning students across first grade classrooms within each school. These tests examine whether students' first grade classroom assignments are statistically independent of a set of six student characteristics: gender, race, participation in the Federal subsidized school lunch program, whether the student's kindergarten reading SAT score is observed, whether the student's kindergarten reading score is above or below average (relative to the full STAR sample), and whether the student had a teacher's aide in kindergarten. For each test the critical value is based on a significance level of 10 percent.

As shown in **Table III**, I fail to reject the null hypothesis of random assignment for all six of these tests in 41 of the 75 schools. If we assumed that the six chi-squared tests are independent, we would expect about 33 percent of schools to fail at least one test at the 10 percent significance level under the null hypothesis of random assignment. This is very close to the actual proportion of schools that fail at least one test.

In addition, I examine the overall distribution of p-values for each set of tests. Under random assignment, this distribution should be close to uniform. **Figure 1** presents histograms of the schools' p-values discretized into 10 bins. Informally they appear roughly uniform. If administrators intentionally created demographic balance across classrooms, we would expect to observe a distribution skewed towards high p-values. Conversely, if administrators deliberately assigned different kinds of students to different classrooms, we would expect to observe a

Second, the statistical evidence suggests that any deviations from random assignment did not change the balance of observable covariates across classes within the same school. Sojourner conducts a variety of statistical tests using the full sample of first grade students (both returning and new students) and finds strong, though not absolute, evidence of random assignment of regular size classroom kindergärtners to their first grade classes.

distribution skewed towards low p-values. The distribution for the chi-square test for gender looks the most skewed towards too much balance. This is confirmed by computing an overall chi-square test statistic for each variable assuming independence across schools. The p-values of these overall test statistics are: 0.40 for race, 0.01 for gender, 0.82 for ever low income, 0.32 for below average reading score, 0.60 for missing reading score, and 0.29 for having a teacher aid in kindergarten. This evidence is largely consistent with random assignment to classroom.

Finally, Sojourner also provides evidence that teachers were randomly assigned to first grade classrooms in Project STAR schools. However, there is more limited information in the STAR data on teacher quality relative to student (or peer) quality. As a result, it is particularly difficult to detect any deviations of the random assignment of teacher's to classrooms. This is not a large concern for this study since classroom fixed effects should account for any non-random assignment of teachers within schools.

4.1) Conceptual Framework

The following framework may be useful in thinking about how peer turnover affects student performance. Suppose that the academic achievement of student i in classroom c, y_{ic}, is given by:

$$y_{ic} = f(x_{ic}, k_c, m_{ic}, e_{ic})$$

where x_{ic} is the student's preexisting human capital stock, k_c is an index of classroom inputs, m_{ic} is the match quality between student i and classroom c, and e_{ic} is the student's effort. All of these factors are assumed to have a positive impact on the student's academic performance.

Peer turnover, p_{ic}, is likely to affect achievement through two distinct channels.[2] First, one would expect turnover to have a negative effect on student-classroom match quality. Classroom instruction can be thought of as a public good where students try to influence the level of difficulty and scope of course material by asking questions in class. Class time dedicated toward answering a question from a particular student is likely to be more informative for students with similar classroom histories (e.g. similar stocks of acquired knowledge).

2 The discussion that follows builds on ideas put forth by Lazear (2001).

Therefore, students who experience greater classmate continuity should exhibit greater match quality (and, therefore, greater learning) relative to the students in the same class who experience a higher degree of peer turnover.

Second, young children are also more likely to be distracted by children with whom they have a prior history. The influence of peers grows as students spend more time together. It follows that students will be less resistant to peer pressure to misbehave or to not pay attention in class when they are surrounded by a greater number of familiar classmates. Holding all else equal, we might expect peer turnover to have a positive effect on a student's effort or attention to classroom instruction.

Therefore, we can decompose the impact of peer turnover on achievement into two distinct and countervailing effects:

$$\frac{\partial y_{ic}}{\partial p_{ic}} = \underbrace{\frac{\partial f}{\partial m_{ic}}\frac{\partial m}{\partial p_{ic}}}_{\textit{Match Quality Effect}} + \underbrace{\frac{\partial f}{\partial e_{ic}}\frac{\partial e}{\partial p_{ic}}}_{\textit{Effort Effect}}$$

In particular, an increase in peer turnover leads to a negative match quality effect and a positive effort effect. Since it is not clear a priori which effect should dominate, peer turnover may have either a positive or negative effect on student achievement.

However, this framework does suggest in which way we might expect peer turnover to have differential effects for students in center city schools relative to other students. The incidence of disruption and behavioral problems is likely to be higher in schools that serve center city students. Therefore, we might expect the effect of peer turnover on effort to be more pronounced in schools in center cities compared to other schools. This suggests that peer turnover with either have a less negative or more positive effect on the achievement of center city students relative to non-center city students.

4.2) Empirical Specification

I estimate the effect of classmate turnover on achievement using the basic specification:

$$T_{ikf1} = \alpha T_{ikf0} + \beta X_{kf1} + Z'_{ikf0}\Gamma + C_{f1} + C_{k1} + \varepsilon_{ikf1}$$

where the outcome of interest, T_{ikf1} is either the reading or math SAT score of student i in 1st grade.[3] I standardize these measures within sample to facilitate the comparison of my results to those of other related studies. The main independent variable of interest, X_{kf1}, is the classmate turnover rate experienced by students in first grade class f who were also students in kindergarten class k. The classmate turnover rate is calculated as the proportion of first grade classmates who shared kindergarten class k with student i. Since X is the same for all students with the same sequence of kindergarten and first grade classes, I cluster my standard errors by kindergarten-first grade classroom pairings (k, f).

Within this specification, I control both directly and indirectly for the preexisting human capital of each student. To account for innate ability and for knowledge acquired in kindergarten, I include a lagged test score, T_{ikf0}, as a regressor. I also consider a number of personal characteristic control variables, Z'_{ikf0}. These include measures of gender, race, family income status (as measured by whether or not the student is a free lunch recipient), age, and the number of days the student was absent in kindergarten.[4] In order to flexibly account for any age effects I include age and age-squared in my specifications. Finally, I include a kindergarten class fixed effect, C_{k1}, to account for differences in acquired human capital due to differences in kindergarten class quality.

Finally, I indirectly control for classroom inputs by including a first grade classroom fixed effect, C_{f1}. This fixed effect accounts for teacher quality and other classroom specific resources, as well as for traditional peer effect inputs such as the mean and variance of class

3 This specification is referred to as a value-added model in the education literature (see Todd and Wolpin, 2003 and 2007) and typically relates an achievement outcome to contemporaneous (e.g. school and family) inputs and to a lagged (baseline) achievement measure.

4 As Sojourner (2008) notes, this measure is likely to reflect aspects of student health, family stability, and/or educational orientation that are likely to affect academic achievement.

ability. Conditional on lagged test score, other control variables, and classroom fixed effects, X_{kf1} is exogenous due to the re-randomization across classrooms and my sample inclusion criteria discussed previously. As a result, OLS estimation should produce an unbiased estimate of the effect of classmate turnover on first grade test scores.

One might be concerned that differences in kindergarten class quality drives both sample attrition and first grade academic achievement. In particular, we could imagine that the students who were randomly assigned to the worst kindergarten teachers would be more likely to exit the study (since if parents take elementary teacher quality as a signal of school quality, they would be more likely to transfer their child to another school given a negative initial signal). Those students who were originally assigned to poor teachers and remained in the sample would have a lower probability of sharing a first grade classroom with their kindergarten classmates (as a result of attrition). These students would also be more likely to perform poorly in first grade as a result of the quality of their kindergarten education (which in this example is negatively correlated with the degree of classmate turnover experienced by these students).

I test the sensitivity of my findings to this form of selective attrition by comparing estimates from specifications that take varying measures to account for differences in kindergarten class quality. First, I estimate the model without any controls relating to kindergarten class. Second, I consider a set of control variables including kindergarten class size, the proportion of the kindergarten class that exited the study, and whether or not the kindergarten class had an aid. Finally, since it is possible that these control variables are poor measures of kindergarten class quality, I also estimate specifications with kindergarten class fixed effects. The lagged class fixed effect should account for all observed and unobservable characteristics of the prior class that might otherwise potentially bias my results.

If the classmate turnover experienced by first graders is truly exogenous, it should not be strongly correlated with observable student characteristics. To test this, I run a regression of classmate turnover on the set of control variables previously highlighted. These results are reported in **Table IV**. I find extremely small and statistically insignificant associations between the majority of other control variables and classmate turnover. A key exception is race, where I find a highly statistically significant association with classmate turnover. However the magnitude of this relationship is relatively small. Controlling for kindergarten and first grade

class fixed effects, the degree of classmate turnover experienced by black students is 1.3 percentage points higher than non-black students. In a classroom of 23 students, this equates to black students on average knowing 0.3 fewer students in their first grade classroom than non-black students in the same class. This result is also not a large concern since my main results hold when I look within race groups.

Similarly, I find a statistically significant association between number of days absent and classmate turnover. Again, the magnitude of this relationship is extremely small. A standard deviation increase in the number of days absent in kindergarten (approximately 10 days) is associated with an increase in classmate turnover of 0.3 percent. Conversely, the standard deviation of classmate turnover in the full STAR sample is 10.9 percent. Of course, we cannot test whether or not classmate turnover is correlated with any unobservable determinants of student achievement. However, these findings lend credence to the assumption that the variation that I use to identify classmate turnover effects is exogenous.

5) Results

I begin by examining the effect of classmate turnover on the academic performance of the full sample of returning first graders. **Table V** presents results where reading SAT score is the dependent variable. These results suggest that classmate turnover has a negative and marginally statistically significant impact on first grade reading scores. They are robust to a number of alternative specifications. Controlling for both personal characteristics and lagged kindergarten inputs only leads to a marginal decline in the absolute magnitude of the estimated turnover effect. Based on the specification with the fullest set of control variables, I find that a reduction in classmate turnover by 10 percent leads to an increase in reading achievement by 0.03 of a standard deviation. Analysis based on the full sample provides weaker evidence for an effect of classmate turnover on math performance. These estimated effects, presented in **Table VI**, are more sensitive to different specifications. They are also smaller than those for reading and are not statistically significant.

Since center city schools are overrepresented in the STAR sample, it is difficult to interpret estimates of classmate turnover effects based on the full, unweighted sample. This is

particularly problematic since schools in Tennessee's center cities have a higher proportion of African-American and poor students and the literature suggests that turnover effects differ substantially by income and race. **Tables VII** and **VIII** present results broken down by income-level, race, and school location for reading and math scores, respectively. Classmate turnover appears to have a large and positive effect on the poor, African-Americans, and students who live in center cities. Conversely, I find a large and negative effect for better-off students, non-minorities, and those living outside of center cities. According to these estimates, a 10 percent increase in classmate turnover improves math scores of center-city students by 0.12 of a standard deviation, while it reduces scores of non-center city students by 0.07 of a standard deviation. The results for reading performance are qualitatively very similar.

To put the magnitude of these estimated effects in perspective, consider the impact of small class sizes on the performance of Project STAR students. Using the same data, Rouse (2000) found that African-American students assigned to small classes improved their achievement test scores by 0.22 of a standard deviation. In addition, she finds small class have a less pronounced effect on the performance of non-minority students. While the effect of classmate turnover is not as great, my findings do suggest that changing the classroom assignment mechanism may be a significantly more cost effective means of improving learning relative to hiring more teachers and building more classrooms.

These results are further evidence that the nature of educational production varies greatly by school location. Within classrooms in center city schools, first graders that are more familiar with their classmates do worse. In these schools, it appears that the positive effect of peer turnover on student attention or effort dominates any negative effects on student-classroom match quality. Conversely, students in schools outside of center cities perform better in reading and math when classroom peer groups are kept more consistent from grade to grade. For non-center city school children the match quality between student and classroom appears to be more important relative to concerns over student attentiveness or effort.

My findings stand in sharp contrast to those reported in Hanushek et al. (2004), who find strong negative effects of student turnover for minorities and poor students and negligible effects for whites and the non-poor. Specifically, they find that raising a school's turnover rate by 10 percentage points reduces the achievement gains of low income and African-American students

by 0.014 and 0.018 of a standard deviation, respectively. There are a number of ways to interpret this discrepancy. Since these authors use non-random variation in student turnover, it could be that their study suffers from omitted variable bias.

Alternatively, these seemly contradictory results may stem from the differences in age of the students examined by each study. In particular, we would expect first graders to be less able to focus on school tasks than older students (such as the 4th through 7th graders examined by Hanushek et al.). It follows that any positive effects of classmate turnover on student attention or effort would be more pronounced in the earliest grades. Conversely, the negative effect of classmate turnover on student-classroom match quality should be relatively more important in later grades. If age mediates how peer turnover affects students, the findings of this study are potentially consistent with those of other studies in the literature.

5.1) Heterogeneous Effects by Student Ability

To further exam the effects of classmate turnover within different types of school, I split my sample into above-average and below-average students based on kindergarten test scores. In schools outside of center cities, above-average students are the most negatively affected by classmate turnover. The corresponding results are presented in **Table IX**. A 10 percent increase in classmate turnover reduces math scores of above-average students by 0.11 of a standard deviation. This result is highly statistically significant. Conversely, a similar change only reduces the scores of below-average students by 0.03 of a standard deviation. The effect on below-average students is also not statistically significant.

While above-average students are the most susceptible to turnover in both types of school, I find that better students in center city schools benefit the most from classmate turnover. These results are presented in **Table X**. For schools in center cities, a 10 percent increase in classmate turnover raises math scores of above-average students by 0.16 of a standard deviation. This result is highly statistically significant. Conversely, a similar change only increases the scores of below-average students by 0.045 of a standard deviation, which is not a statistically significant effect. This pattern is not as pronounced for reading achievement.

These later findings are consistent with the idea that peers within economically disadvantaged areas impose especially high costs on students that strive to put forth effort in their studies or 'act white' (Fryer and Torelli, 2005). Fryer (2006) presents a model of social interaction in which children face a trade off between investing in their own general human capital or in group-specific capital. Since peer turnover destroys group-specific capital, it causes the opportunity cost of investing in general human capital to decrease in center city schools. Since the return to investing in their own human capital is higher for above-average students, it follows that better students in center cities schools would gain the most from peer turnover.

5.2) The Role of Teachers' Aides

One way in which schools can try to deal with issues of ability mis-match and behavioral problems in classrooms is by assigning a teacher's aide. One potential role of a teacher's aide is to provide assistance to those students that would otherwise not be able to keep up with the pace and difficulty of classroom material. Teacher's aides might also serve as class disciplinarians, allowing the teacher to focus on instruction instead of regulating student behavior. Aides are most likely to focus their attention on addressing the most pressing needs of their class. My findings suggest that student-classroom match quality is a relatively larger issue in schools outside of center cities. Therefore in these schools, we would expect the assignment of teachers' aides to mitigate the negative effect of classmate turnover. Conversely, I find that student effort or attentiveness is a relatively larger concern in center city classrooms. It follows that we would expect teachers' aides in these schools to lessen the negative effect of peer continuity. In other words, assignment of a teacher's aide should reduce the positive effect of peer turnover in center city schools.

To test these predictions, I further split my center city and non-center city school samples into students who were randomly assigned a teacher's aide in first grade and those who were not. **Table XI** presents these results for schools outside of center cities. In these schools, teachers' aides appear to be effective in lessening the negative effect of classmate turnover on reading achievement, but not for math. Moving from a class without an aid to a class with an aid reduces the estimated effect of classmate turnover on reading achievement by nearly half.

Conversely, teachers' aides in center city schools severely mitigate the negative effect of peer continuity on effort. In center city classrooms with a teacher's aide, students' reactions to classmate turnover are very similar to what we see for students outside of center cities. As **Table XII** highlights, a 10 percent reduction in turnover in classes with aides raises the test scores of center city students in reading and math by 0.148 and 0.085 of a standard deviation respectively. When these students are left without a teacher's aide to regulate classroom behavior, center city students appear to have a harder time avoiding distraction instigated by familiar classmates. Under such circumstances, they benefit from greater classmate turnover.

6) Robustness Checks

6.1) Random Assignment to Classroom

Statistical tests have limited power to detect non-random patterns in class assignment given that Project STAR schools have relatively few students. Therefore, as a robustness check I consider subsamples of cases where students had less scope to change classrooms within their randomly assigned first grade class type.

As Graham (2008) first noted, thirty-three of the Project STAR schools only offered two regular-size classrooms in 1st grade (one with an aide and one without). For these small schools, assignment to class type directly equated to assignment to classroom. In addition, there were 14 schools that only offered one regular size class with a teacher's aide (while these schools had multiple regular size classes without teachers' aides). In these schools, students assigned to a class with an aid had no option to switch to another classroom of the same type. Similarly, nine other schools offered only one regular size class without a teacher's aide (and multiple regular size classes with aides). For these 23 schools, random assignment to class type in first grade is most likely to have ensured random assignment to classroom for only a subset of returning first graders. Given the classroom composition of STAR schools, the re-randomization to class type in 1st grade generated eighty-nine "random" classes (from fifty-five schools) for which we can be the most confident that students were randomly assigned to their particular classroom.

Tables XIII and **XIV** report results based on restricted samples for schools outside and

within center cities, respectively. Following Graham (2008) and Sojourner (2008), I further restrict my sample to only students from regular classes in small schools (e.g. schools that had only one regular size class with a teacher's aide and one without). For non-center city schools, the estimated effect of classmate turnover on reading achievement is very consistent comparing the full sample to these restricted samples. For math scores, the estimated effect of classmate turnover is less pronounced and not statistically significant for the restricted samples of non-center city school children. Conversely, the positive effect of classmate turnover found for students in center city schools becomes even larger using the restricted samples.

However, it is important to note that these restricted samples have proportionally more small schools relative to the full sample. As such, differences in the magnitudes of the estimated effects between the full and restricted samples could suggest that school size mediates the effect of classmate turnover. Even if this is the case, it does not appear that deviations from random assignment to classroom are driving my main results.

6.2) Attrition

One might be concerned that attrition would introduce non-random variation in the peer composition of first grade classrooms. In particular, Hanushek (1999) voices a number of concerns about how attrition might potentially bias the main classroom size findings generated by Project STAR. Approximately a third of the students originally assigned to regular size classrooms in kindergarten did not transition into first grade in Project STAR schools. For my analysis, students that exited the sample include those held back, transfers to new schools, and those that non-randomly switched to small classes.

To further examine the nature of attrition before the start of first grade, I estimate a linear probability model (LPM) using a host of kindergarten observables as predictor variables for attrition from the first grade regular classroom sample. These results are reported in **Table XV**. Based on F-tests, it does not appear that kindergarten SAT test information (either scores or whether the student took the test) predict sample attrition. The strongest predictor of attrition is whether the number of days a student was absent in kindergarten is missing. A student with this information missing was approximately 11 percent less likely to exit the study after kindergarten.

This possibly suggests that the STAR researchers were less vigilant or had more difficulty obtaining or updating their data for students who left the study schools relative to those who did not. The corresponding coefficient becomes marginally statistically significant once kindergarten class fixed effects are accounted for. These results are qualitatively very similar if I separately consider center city and non-center city students.

As I previously noted, we might fear that attrition is driven by differences in kindergarten quality. The evidence reported in Table XV suggests that within-school differences in kindergarten quality is a statistically significant predictor of attrition. Specifically, I reject the null hypothesis that attrition is driven by differences in school quality alone by conducting a likelihood ratio test where I compare the LPM of attrition with kindergarten class fixed effects (i.e. the unrestricted model) with the LPM with only school fixed effects (i.e. the restricted model).[5] This difference is statistically significant at the 10 percent level for the full sample and at the 1 percent level for the non-center city school sample. However while attrition may be driven by differences in kindergarten class quality, this does not appear to have a strong bias on my results. Particularly, my estimates of the classmate turnover effect do not change very much when I directly control for measures of kindergarten quality or indirectly control for them by including a kindergarten class fixed effect in my specifications. Both of these results suggest that attrition does not bias my results through this particular channel.

6.3) Specification

A potential issue with the value-added model is that OLS estimates of the return to prior learning are likely to be biased as a result of measurement error. Specifically, consider the case where the lagged test score is a noisy measure of true kindergarten learning. If the first grade test score is a function of true kindergarten learning and we instead use the noisy measure, then the error term in the main estimating equation will include both the measurement error of the first grade test score, as well as the measurement error of the kindergarten test score. Since the kindergarten test score is correlated with its own measurement error, OLS will lead to a downward biased estimate of the return to prior learning (α). Therefore, as a robustness check I instrument for the lagged test score with lagged test scores in other subjects. This

5 The p-value for this test is 0.07.

procedure, outlined by Cunha and Heckman (2008), should correct for any measurement error bias. The results generated by this procedure, reported in **Table XVI**, are very similar to those produced by simple OLS. Specifically, I still find that classmate turnover has a large positive effect on student performance in center city schools and a large negative impact on students from other schools.

7) Conclusion

The re-randomization after kindergarten conducted during the STAR experiment provides a useful context for investigating the effects of classmate turnover on student achievement. Using this experimental variation, I find that classmate turnover has an negative effect on the achievement of students in schools outside of center cities (who are predominantly non-poor and non-minorities). Conversely, turnover has the opposite effect on students in center city schools. It does not appear that potential deviations from random assignment nor attrition drive my main results. These results are robust to a number of alternative specifications and sample restrictions.

My findings are consistent with a model of classroom learning in which familiar classmates can either improve student-classroom match quality or distract student attention away from class instruction. In schools where the incidence of behavioral problems is relatively low, young students appear to benefit from stability in their class grouping as they advance from grade to grade. Such stability in student grouping might allow teachers to better coordinate their curriculum across grades.

In center city schools, where behavioral problems are typically more abundant, it appears that stronger friendship bonds (formed by sharing successive classes) actually hinder learning. Assignment of a teacher's aide to classes within these schools seems to mitigate this negative effect. However, schools in center cities have particularly limited means to hire more teachers' aides. As a more cost-effective alternative, administrator in center city schools might aim to reduce disciplinary problems by taking measures to promote peer turnover. These findings also suggest that programs such as the Metropolitan Council for Educational Opportunity (Metco) desegregation program and the Chicago Public Schools (CPS) lotteries would be even more

effective if they placed greater emphasis on providing school mobility opportunities to the youngest of inner city children (see Angrist and Lang, 2004; and, Cullen, Jacob, and Levitt, 2006).

As previously noted, the results of this paper stand in contrast to the findings of other related studies that examine older children. These differences in findings potentially highlight some of the limitations of this study. Particular, the cost of exploiting the STAR re-randomization after kindergarten is that I am only able to identify classmate turnover effects for first graders. As Cunha and Heckman (2008) note, educational production over the life cycle is likely to be marked with distinct critical periods. The findings of this study may not generalize to children in later grades.

Heretofore, most of the discussion on classroom assignment policies has focused on traditional peer effects. Moving forward, researchers need to further evaluate potential classmate turnover effects at the different stages of the education life cycle. Finally, future work should aim to weigh the relative importance of classroom composition and classmate turnover when considering optimal tracking and/or mixing policies.

8) References

Alexander, K.L., D.R. Entwisle, and S.L. Dauber (1996). "Children in Motion: School Transfers and Elementary School Performance." Journal of Educational Research, 90: 3-12.

Angrist, J. and K. Lang (2004). "Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program." The American Economic Review, 94:1613-34.

Betts, J. R. and J. L. Shkolnik (1999) "Key Difficulties in Identifying the Effects of Ability Grouping on Student Achievement." Economics of Education Review, 19(1): 21-26.

Boozer, M. and S. C. Cacciola (2001). “Inside the 'Black Box' of Project STAR: Estimation of Peer Effects Using Experimental Data,” Yale Economic Growth Center Discussion Paper No. 832.

Cascio, E. and D. Whitmore-Schazenbach (2008). “Age and the Education Production Function.” NBER Working Paper 13663.

Clotfelter, C.T., H.F. Ladd and J.L. Vigdor (2006). "Teacher-Student Matching and the Assessment of Teacher Effectiveness." Journal of Human Resources, 41(4): 778-820.

Cullen, J. B., B. A. Jacob, and S. Levitt (2006). “The Effect of School Choice on Participants: Evidence from Randomized Lotteries.” Econometrica, 74(5): 1191–1230.

Cunha, F. and J. J. Heckman (2008). "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation," Journal of Human Resources, 43(4).

Dee, T. S. (2004). “Teachers, Race, and Student Achievement in a Randomized Experiment.” The Review of Economics and Statistics, 86(1).

Fryer, R. G. (2006). “A Model of Social Interactions and Endogenous Poverty Traps.” NBER Working paper No. 12363.

Fryer, R. G. and P. Torelli (2005). “An Empirical Analysis of 'Acting White'.” NBER Working paper No. 11334.

Graham, B. S. (2008). “Identifying Social Interactions Through Conditional Variance Restrictions.” Econometrica, 76 (3): 643-660.

Hanushek, E. A. (1999). “Some Findings from an Independent Investigation of the Tennessee STAR Experiment and from Other Investigations of Class Size Effects,” Educational Evaluation and Policy Analysis, 21(2).

Hanushek, E. A., J. F. Kain, and S. G. Rivkin (2004). “Disruption versus Tiebout improvement: the costs and benefits of switching schools.” Journal of Public Economics, 88: 1721-1746.

Imberman, S., A. Kugler, and B. Sacerdote (2009). “Katrina's Children: A Natural Experiment in Peer Effects from Hurricane Evacuees.” Working Paper.

Katriel, T. and P. Nesher (1986). “Gibush: The Rhetoric of Cohesion in Israeli School Culture.” Comparative Education Review, 30(2): 216-231.

Kerbow, D. (1996). “Patterns of Urban Student Mobility and Local School Reform.” Journal of Education for Students Placed at Risk. 1(2): 147-169.

Krueger, A. B. (1999). “Experimental Estimates of Education Production Functions.” Quarterly Journal of Economics, 114(2).

Lazear, E. P. (2001). “Educational Production.” Quarterly Journal of Economics, 116(3): 777-803.

Nye, B. L., S. Konstantopoulos, and L. V. Hedges (2005). “How Large Are Teacher Effects?” Educational Evaluation and Policy Analysis, 26(3).

Rouse, C. E. (2000). “School Reform in the 21st Century: The Effect of Class Size and School Vouchers on African-American and Hispanic Students.” Princeton Univesity Industrial Relations Section Working Paper No. 440.

Sojourner, A. (2008). “Inference on Peer Effects with Missing Peer Data: Evidence from Project STAR.” Job Market Paper.

Temple, J. A. and A. J. Reynolds (1999). “School Mobility and Achievement: Longitudinal Findings From an Urban Cohort.” Journal of School Psychology, 37(4): 355-377.

Todd, P. E. and K. I. Wolpin (2003). “On the Specification and Estimation of the Production Function for Cognitive Achievement.” The Economic Journal, 113: F3-F33.

Todd, P. E. and K. I. Wolpin (2007). “The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps.” Journal of Human Capital, 1(1): 91-136.

United States General Accounting Office (1994). “Elementary School Children: Many Change Schools Frequently, Harming Their Education.” Report no. GAO/HEHS-94-95.

Whitmore, D. (2005). “Resource and Peer Impacts on Girls' Academic Achievement: Evidence from a Randomized Experiment.” AEA Papers and Proceedings, 95(2): 199-203.

Whitmore-Schanzenbach, D. (2007). “What Have Researchers Learned from Project STAR?” Brookings Papers on Education Policy: 2006/2007.

Zabel, J. E. (2008). “The Impact of Peer Effects on Student Outcomes in New York Public Schools.” Education Finance and Policy, 3(2): 197-249.

9) Appendix

Table I: Summary Statistics for Center City and Non-center City STAR Schools

	Total	Center City (1)	Non-city (2)	Diff. (1) – (2)	P-value
Classmate Turnover Rate	0.77	0.81	0.75	0.06	0.000
Proportion of Kindergarten Classmates that Attrited	0.31	0.44	0.24	0.19	0.000
Female	0.50	0.49	0.50	-0.01	0.605
Black	0.28	0.65	0.08	0.57	0.000
Ever Low Income	0.51	0.62	0.44	0.18	0.000
Age	6.44	6.42	6.46	-0.04	0.005
Days Absent:					
Kindergarten	9.89	8.22	10.81	-2.60	0.000
Reading SAT Score:					
Kindergarten	439.27	437.50	440.22	-2.72	0.035
1st Grade	524.87	506.68	535.22	-28.54	0.000
Math SAT Score:					
Kindergarten	489.91	487.30	491.35	-4.06	0.027
1st Grade	532.22	520.07	538.89	-18.82	0.000
Aide in Class:					
Kindergarten	0.52	0.55	0.50	0.06	0.006
1st Grade	0.48	0.46	0.50	-0.04	0.053
Class Size:					
Kindergarten	22.45	23.05	22.12	0.93	0.000
1st Grade	23.04	22.83	23.15	-0.32	0.001
Sample Size:					
Reading	2,465	887	1,578	-	-
Math	2,571	912	1,659	-	-

* Ever low income is a dummy variable coded as one if a student was a participant in the school free lunch program in either kindergarten or first grade. Center city refers to schools located within the center city school districts of Chattanooga, Knoxville, Memphis, and Nashville. Non-center city refers to schools outside of these school districts.

Table II: Transition Matrix for Students in School ID# 168214

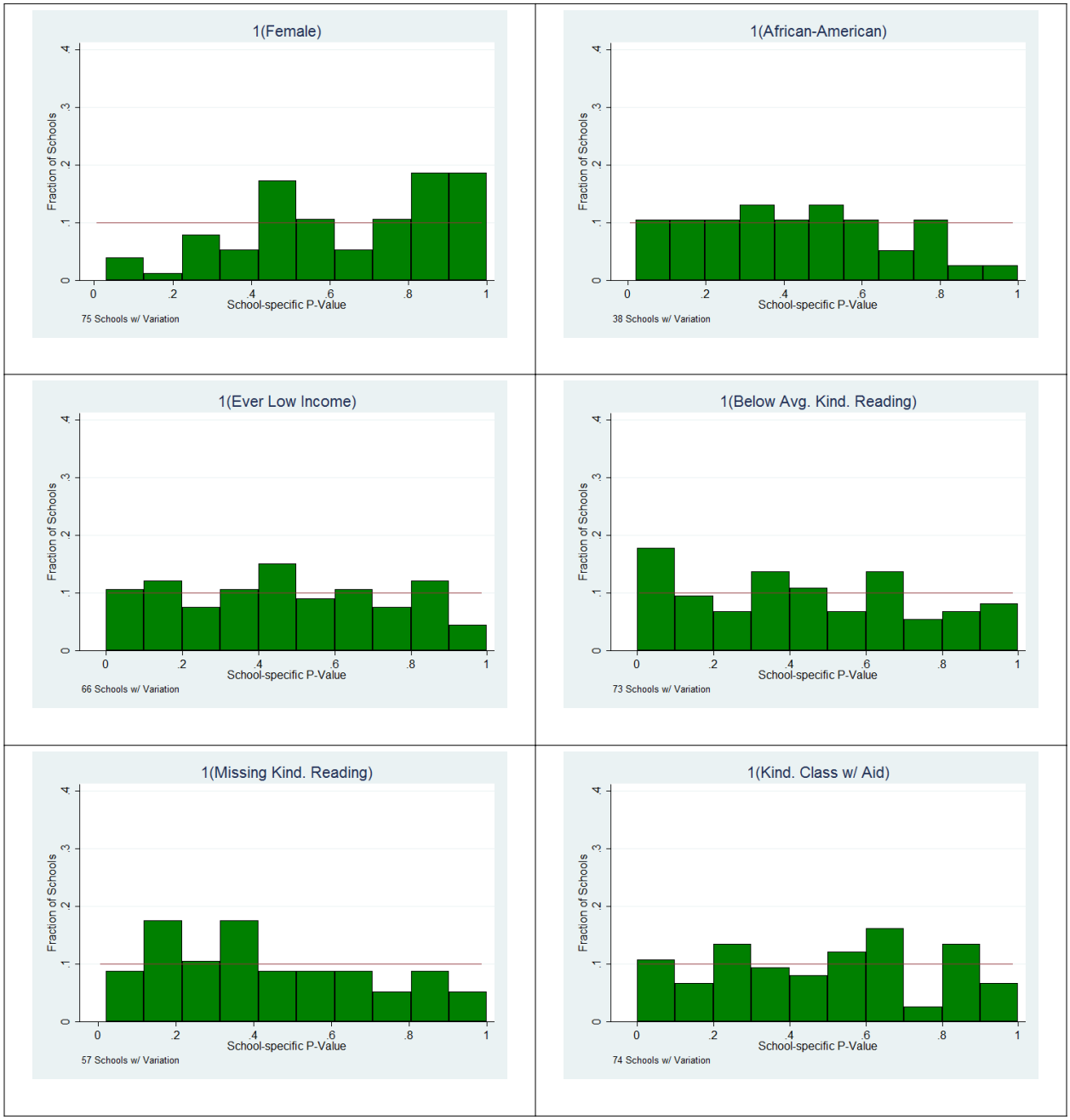
Kindergarten Class	1st Grade Class F-1	F-2	Not in 1st Grade	Total # Students
K-1	8	11	4	23
K-2	11	9	2	22
Not in Kindergarten	4	5	-	9
Total # Students	23	25	6	-

Table III: Summary of Chi-Squared Tests of Random Assignment of Students Across 1st Grade Classrooms within Project STAR Schools

	Observed		Expected Under Random Assignment*	
# of Tests Failed	# of Schools	% of Schools	# of Schools	% of Schools
0 of 6	41	54.7	44.0	58.7
1 of 6	27	36.0	24.5	32.7
2 of 6	7	9.3	5.7	7.6
3 of 6	0	0.0	0.7	1.0
4 of 6	0	0.0	0.1	0.1
5 of 6	0	0.0	0.0	0.0
6 of 6	0	0.0	0.0	0.0
Total	75	100	75	100

Note: This table reports the results of Chi-squared tests of the null hypothesis that students are randomly distributed across classrooms within schools along six observable student characteristics: race, gender, subsidized lunch receipt, previous year reading test score, missingness of prior reading score, and whether or not their kindergarten class had an aid. The tests are based on data on returning students. The corresponding significance level is 10 percent.
* Expected values are computed assuming that the six chi-squared tests are independent. They also take into account that not all schools have variation to perform all of the chi-squared tests. See text for further details.

Figure I: Histograms of Chi-Squared Tests of Random Assignment of Students Across 1st Grade Classrooms within Project STAR Schools



* In each school, the null hypothesis of random assignment predicts that the class distribution of each variable match the school distribution with finite-sample variance. A chi-square statistic measures the degree of match. Under the null, the distribution of the p-values of these test statistics across school should be uniform. Testing whether these p-values are uniformly-distributed using auxiliary chi-square tests returns overall p-values of: 0.40 for race, 0.01 for gender, 0.82 for ever low income, 0.32 for below average reading score, 0.60 for missing reading score, and 0.29 for having a teacher aid in kindergarten. This evidence is largely consistent with random assignment to classroom.

Table IV: Partial Regression Estimates for Classmate Turnover on Other Control Variables

	OLS	School FE	Class FE		
			(1)	(2)	(3)
Normalized Kindergarten	0.0071 **	0.0002	0.0000	0.0009	0.0013
Reading SAT Score	(0.0036)	(0.0019)	(0.0016)	(0.0016)	(0.0013)
Female	-0.0052	-0.0023	-0.0017	-0.0013	-0.0012
	(0.0037)	(0.0024)	(0.0022)	(0.0021)	(0.0018)
Black	0.0286 ***	0.0053	0.0081	0.0106 **	0.0132 ***
	(0.0107)	(0.0055)	(0.0051)	(0.0045)	(0.0041)
Ever Low Income	-0.0087	-0.0013	-0.0022	0.0014	0.0006
	(0.0064)	(0.0037)	(0.0035)	(0.0027)	(0.0025)
Age	-0.1875	-0.0282	-0.0696	0.0368	0.0185
	(0.1363)	(0.1147)	(0.0998)	(0.0699)	(0.0584)
Age^2	0.0150	0.0022	0.0055	-0.0026	-0.0011
	(0.0103)	(0.0087)	(0.0076)	(0.0053)	(0.0044)
Number of Days Absent	-0.0001	0.0001	0.0001	0.0002 **	0.0003 ***
in Kindergarten	(0.0003)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Class w/ Aid:					
in 1st Grade	-0.0089	-0.0069	-0.0041		
	(0.0101)	(0.0072)	(0.0068)		
in Kindergarten	-0.0159	-0.0059		-0.0052	
	(0.0100)	(0.0069)		(0.0058)	
Class Size:					
in 1st Grade	0.0049 **	-0.0023	-0.0035		
	(0.0025)	(0.0038)	(0.0035)		
in Kindergarten	0.0021	-0.0089 ***		-0.0083 ***	
	(0.0021)	(0.0032)		(0.0027)	
Fraction of Kindergarten	0.2290 ***	0.3566 ***		0.3438 ***	
Classmates that	(0.0281)	(0.0386)		(0.0345)	
Exited the Sample					
Class Fixed Effect:					
1st Grade	No	No	No	Yes	Yes
Kindergarten	No	No	Yes	No	Yes
Observations	2,465	2,465	2,465	2,465	2,465
R-squared	0.17	0.59	0.62	0.72	0.74

~ Robust standard errors , clustered by kindergarten-first grade class pairings, are reported in parentheses. Statistical significance is denoted by "***" when significant at the 1% level, "**" at the 5% level, and "*" at the 10% level. "School FE" refers to school level fixed effects.

Table V: Effect of Classmate Turnover on Normalized First Grade SAT Reading Score (Full Sample)

	(1)		(2)		(3)		(4)	
Normalized Kindergarten	0.608	***	0.567	***	0.568	***	0.599	***
Reading SAT Score	(0.023)		(0.022)		(0.022)		(0.026)	
Classmate	**-0.375**	**	**-0.318**	*	**-0.301**		**-0.279**	*
Turnover Rate	**(0.175)**		**(0.165)**		**(0.185)**		**(0.148)**	
Female			0.133	***	0.133	***	0.140	***
			(0.026)		(0.026)		(0.027)	
Black			-0.078		-0.079		-0.057	
			(0.053)		(0.054)		(0.054)	
Ever Low Income			-0.217	***	-0.217	***	-0.212	***
			(0.035)		(0.035)		(0.036)	
Age			4.195	***	4.188	***	3.892	***
			(0.805)		(0.807)		(0.846)	
Age^2			-0.332	***	-0.331	***	-0.307	***
			(0.061)		(0.061)		(0.064)	
Number of Days Absent			-0.005	**	-0.004	**	-0.004	**
in Kindergarten			(0.002)		(0.002)		(0.002)	
Class w/ Aid					-0.014			
in Kindergarten					(0.023)			
Class Size in Kindergarten					0.013			
					(0.012)			
Proportion of Kindergarten					-0.033			
Classmates that Exited the Sample					(0.156)			
Class Fixed Effect:								
1st Grade	Yes		Yes		Yes		Yes	
Kindergarten	No		No		No		Yes	
Observations	2,465		2,465		2,465		2,465	
R-squared	0.61		0.63		0.63		0.65	

* Robust standard errors, clustered by kindergarten-first grade class pairing in parentheses. Statistical significance is denoted by "***" when significant at the 1% level, "**" at the 5% level, and "*" at the 10% level.

Table VI: Effect of Classmate Turnover on Normalized First Grade SAT Math Score (Full Sample)

	(1)		(2)		(3)		(4)	
Normalized Kindergarten	0.639	***	0.589	***	0.589	***	0.638	***
Math SAT Score	(0.018)		(0.018)		(0.018)		(0.020)	
Classmate	**-0.238**		**-0.137**		**-0.201**		**-0.057**	
Turnover Rate	**(0.189)**		**(0.181)**		**(0.202)**		**(0.163)**	
Female			-0.010		-0.010		-0.016	
			(0.024)		(0.024)		(0.025)	
Black			-0.315	***	-0.314	***	-0.295	***
			(0.064)		(0.065)		(0.067)	
Ever Low Income			-0.235	***	-0.235	***	-0.226	***
			(0.032)		(0.032)		(0.032)	
Age			5.892	***	5.886	***	5.629	***
			(0.869)		(0.871)		(0.915)	
Age^2			-0.453	***	-0.453	***	-0.433	***
			(0.066)		(0.066)		(0.070)	
Number of Days Absent			-0.003		-0.003		-0.002	
in Kindergarten			(0.002)		(0.002)		(0.002)	
Class w/ Aid					0.005			
in Kindergarten					(0.022)			
Class Size in Kindergarten					0.001			
					(0.012)			
Proportion of Kindergarten					0.147			
Classmates that Exited the Sample					(0.170)			
Class Fixed Effect:								
1st Grade	Yes		Yes		Yes		Yes	
Kindergarten	No		No		No		Yes	
Observations	2,571		2,571		2,571		2,571	
R-squared	0.63		0.65		0.65		0.69	

* Robust standard errors, clustered by kindergarten-first grade class pairing in parentheses. Statistical significance is denoted by "***" when significant at the 1% level, "**" at the 5% level, and "*" at the 10% level.

Table VII: Effect of Classmate Turnover on Normalized 1st Grade SAT Reading Score By Income, Race, and School Location

	Ever low income				Race				Center City			
	Yes		No		Black		Non-black		Yes		No	
Normalized Kindergarten	0.615	***	0.573	***	0.412	***	0.634	***	0.497	***	0.642	
Reading SAT Score	(0.045)		(0.030)		(0.049)		(0.031)		(0.037)		(0.033)	
Classmate	**0.081**		**-0.628**	**	**0.454**	*	**-0.451**	**	**0.295**		**-0.565**	
Turnover Rate	**(0.219)**		**(0.256)**		**(0.264)**		**(0.176)**		**(0.261)**		**(0.186)**	
Female	0.149	***	0.150	***	0.131	***	0.139	***	0.131	***	0.143	
	(0.039)		(0.045)		(0.049)		(0.033)		(0.041)		(0.035)	
Black	-0.098		0.054						-0.071		-0.052	
	(0.070)		(0.115)						(0.112)		(0.063)	
Ever Low Income					-0.300	***	-0.192	***	-0.220	***	-0.196	
					(0.102)		(0.040)		(0.079)		(0.039)	
Age	3.980	***	4.920	***	3.654	***	4.329	***	2.743	**	4.766	
	(1.153)		(1.764)		(1.315)		(1.081)		(1.100)		(1.141)	
Age^2	-0.312	***	-0.388	***	-0.273	***	-0.344	***	-0.216	**	-0.375	
	(0.087)		(0.135)		(0.100)		(0.082)		(0.084)		(0.086)	
Number of Days Absent	-0.006	**	-0.001		-0.004	**	-0.004	**	-0.006	***	-0.003	
in Kindergarten	(0.003)		(0.003)		(0.002)		(0.002)		(0.002)		(0.002)	
Observations	1,243		1,222		701		1,764		887		1,578	
R-squared	0.67		0.63		0.64		0.62		0.66		0.62	

~ Robust standard errors , clustered by kindergarten-first grade class pairings, are reported in parentheses. Statistical significance is denoted by "***" when significant at the 1% level, "**" at the 5% level, and "*" at the 10% level. All models include first grade and kindergarten class fixed effects.

Table VIII: Effect of Classmate Turnover on Normalized 1st Grade SAT Math Score By Income, Race, and School Location

	Ever low income				Race				Center City			
	Yes		No		Black		Non-black		Yes		No	
Normalized Kindergarten	0.609	***	0.648	***	0.482	***	0.679	***	0.575	***	0.667	***
Math SAT Score	(0.033)		(0.029)		(0.046)		(0.021)		(0.038)		(0.023)	
Classmate	**0.306**		**-0.427**	*	**1.158**	***	**-0.444**	**	**1.212**	***	**-0.700**	***
Turnover Rate	**(0.231)**		**(0.244)**		**(0.346)**		**(0.186)**		**(0.281)**		**(0.180)**	
Female	0.015		-0.014		0.077		-0.048	*	0.030		-0.048	
	(0.038)		(0.038)		(0.051)		(0.028)		(0.043)		(0.030)	
Black	-0.259	***	-0.256						-0.271	**	-0.298	***
	(0.079)		(0.177)						(0.125)		(0.080)	
Ever Low Income					-0.302	***	-0.194	***	-0.223	***	-0.215	***
					(0.072)		(0.035)		(0.062)		(0.037)	
Age	6.681	***	4.564	***	7.869	***	4.519	***	6.013	***	5.280	***
	(1.110)		(1.589)		(1.526)		(1.021)		(1.740)		(1.059)	
Age^2	-0.507	***	-0.357	***	-0.580	***	-0.355	***	-0.453	***	-0.411	***
	(0.083)		(0.121)		(0.116)		(0.078)		(0.132)		(0.080)	
Number of Days Absent	-0.005	**	0.001		0.000		-0.003		0.000		-0.004	**
in Kindergarten	(0.002)		(0.003)		(0.003)		(0.002)		(0.002)		(0.002)	
Observations	1,300		1,271		725		1,846		912		1,659	
R-squared	0.70		0.66		0.69		0.67		0.70		0.66	

~ Robust standard errors , clustered by kindergarten-first grade class pairings, are reported in parentheses. Statistical significance is denoted by "***" when significant at the 1% level, "**" at the 5% level, and "*" at the 10% level. All models include first grade and kindergarten class fixed effects.

Table IX: Effect of Classmate Turnover on Normalized 1st Grade SAT Scores By Ability Groupings Based on Kindergarten Test Scores (Schools Outside of Center Cities)

	Reading				Math			
	Below Avg.		Above Avg.		Below Avg.		Above Avg.	
Normalized Kindergarten SAT Score	0.756	***	0.527	***	0.646	***	0.558	***
	(0.080)		(0.044)		(0.041)		(0.034)	
Classmate Turnover Rate	**-0.008**		**-0.973**	***	**-0.330**		**-1.145**	***
	(0.220)		**(0.285)**		**(0.245)**		**(0.287)**	
Female	0.109	**	0.076		0.003		-0.110	**
	(0.054)		(0.054)		(0.048)		(0.043)	
Black	0.129		-0.298	***	-0.179	**	-0.419	**
	(0.082)		(0.105)		(0.086)		(0.164)	
Ever Free Lunch Recipient	-0.214	***	-0.113	**	-0.219	***	-0.189	***
	(0.061)		(0.057)		(0.054)		(0.058)	
Age	5.280	***	5.158	**	4.566	***	5.258	**
	(1.732)		(2.344)		(1.233)		(2.191)	
Age^2	-0.411	***	-0.411	**	-0.356	***	-0.409	**
	(0.130)		(0.180)		(0.092)		(0.169)	
Number of Days Absent in Kindergarten	0.001		-0.005		-0.005	*	-0.003	
	(0.003)		(0.004)		(0.003)		(0.004)	
Observations	767		811		807		852	
R-squared	0.52		0.57		0.64		0.62	

~ Robust standard errors , clustered by kindergarten-first grade class pairings, are reported in parentheses. Statistical significance is denoted by "***" when significant at the 1% level, "**" at the 5% level, and "*" at the 10% level. All models include first grade and kindergarten class fixed effects.

Table X: Effect of Classmate Turnover on Normalized 1st Grade SAT Scores By Presence of Aide in 1st Grade Classroom (Schools Outside of Center Cities)

	Reading				Math			
	No Aid		Aid		No Aid		Aid	
Normalized Kindergarten SAT Score	0.644	***	0.657	***	0.654	***	0.688	***
	(0.055)		(0.047)		(0.035)		(0.036)	
Classmate Turnover Rate	**-0.917**	**	**-0.522**		**-0.862**	***	**-0.864**	*
	(0.366)		**(0.382)**		**(0.329)**		**(0.481)**	
Female	0.077		0.207	***	-0.095	**	-0.009	
	(0.047)		(0.060)		(0.045)		(0.049)	
Black	0.005		-0.084		-0.409	***	-0.174	
	(0.094)		(0.105)		(0.112)		(0.143)	
Ever Free Lunch Recipient	-0.207	***	-0.209	***	-0.206	***	-0.198	***
	(0.059)		(0.065)		(0.056)		(0.060)	
Age	4.884	***	4.701	**	4.855	***	4.934	***
	(1.662)		(2.171)		(1.611)		(1.765)	
Age^2	-0.389	***	-0.365	**	-0.385	***	-0.379	***
	(0.126)		(0.164)		(0.121)		(0.135)	
Number of Days Absent in Kindergarten	-0.006		-0.001		-0.005	*	-0.002	
	(0.004)		(0.003)		(0.003)		(0.003)	
Observations	801		777		836		823	
R-squared	0.65		0.62		0.67		0.68	

~ Robust standard errors , clustered by kindergarten-first grade class pairings, are reported in parentheses. Statistical significance is denoted by "***" when significant at the 1% level, "**" at the 5% level, and "*" at the 10% level. All models include first grade and kindergarten class fixed effects.

Table XI: Effect of Classmate Turnover on Normalized 1st Grade SAT Scores By Ability Groupings Based on Kindergarten Test Scores (Center City Schools)

	Reading				Math			
	Below Avg.		Above Avg.		Below Avg.		Above Avg.	
Normalized Kindergarten SAT Score	0.674	***	0.343	***	0.577	***	0.508	***
	(0.086)		(0.044)		(0.056)		(0.053)	
Classmate Turnover Rate	**0.131**		**0.268**		**0.458**		**1.606**	***
	(0.435)		**(0.413)**		**(0.358)**		**(0.605)**	
Female	0.045		0.203	***	0.046		-0.003	
	(0.050)		(0.074)		(0.054)		(0.074)	
Black	0.152		0.071		0.207		-0.259	
	(0.170)		(0.236)		(0.164)		(0.231)	
Ever Free Lunch Recipient	-0.301	***	0.007		-0.297	***	-0.114	
	(0.112)		(0.136)		(0.083)		(0.091)	
Age	1.004		2.456		3.384		4.190	
	(1.519)		(2.639)		(2.419)		(3.480)	
Age^2	-0.077		-0.198		-0.247		-0.308	
	(0.114)		(0.206)		(0.187)		(0.267)	
Number of Days Absent in Kindergarten	-0.006	**	-0.003		-0.002		0.005	
	(0.003)		(0.006)		(0.003)		(0.004)	
Observations	457		430		459		453	
R-squared	0.66		0.70		0.73		0.72	

~ Robust standard errors , clustered by kindergarten-first grade class pairings, are reported in parentheses. Statistical significance is denoted by "***" when significant at the 1% level, "**" at the 5% level, and "*" at the 10% level. All models include first grade and kindergarten class fixed effects.

Table XII: Effect of Classmate Turnover on Normalized 1st Grade SAT Scores By Presence of Aide in 1st Grade Classroom (Center City Schools)

	Reading				Math			
	No Aid		Aid		No Aid		Aid	
Normalized Kindergarten SAT Score	0.503	***	0.458	***	0.654	***	0.487	***
	(0.056)		(0.062)		(0.056)		(0.061)	
Classmate Turnover Rate	**0.493**		**-1.481**	**	**2.008**	***	**-0.852**	
	(0.474)		**(0.715)**		**(0.493)**		**(0.616)**	
Female	0.202	***	0.081		0.016		0.073	
	(0.060)		(0.068)		(0.061)		(0.076)	
Black	-0.028		-0.106		-0.374	**	-0.057	
	(0.208)		(0.153)		(0.168)		(0.223)	
Ever Free Lunch Recipient	-0.113		-0.491	***	-0.217	***	-0.265	**
	(0.115)		(0.135)		(0.077)		(0.123)	
Age	2.342		1.496		6.641	***	2.761	
	(1.746)		(2.097)		(2.382)		(3.839)	
Age^2	-0.192		-0.111		-0.503	***	-0.198	
	(0.132)		(0.161)		(0.181)		(0.295)	
Number of Days Absent in Kindergarten	-0.006		-0.007	*	-0.002		0.002	
	(0.004)		(0.004)		(0.003)		(0.004)	
Observations	476		411		496		416	
R-squared	0.68		0.70		0.75		0.70	

~ Robust standard errors , clustered by kindergarten-first grade class pairings, are reported in parentheses. Statistical significance is denoted by "***" when significant at the 1% level, "**" at the 5% level, and "*" at the 10% level. All models include first grade and kindergarten class fixed effects.

Table XIII: Effect of Classmate Turnover on Normalized 1st Grade SAT Scores with Restricted Samples (Schools Outside of Center Cities)

	Reading			Math		
	Full Sample	"Random" Classes	Small Schools	Full Sample	"Random" Classes	Small Schools
Normalized Kindergarten	0.642 ***	0.640 ***	0.656 ***	0.667 ***	0.715 ***	0.716 ***
SAT Score	(0.033)	(0.055)	(0.061)	(0.023)	(0.029)	(0.033)
Classmate	**-0.565 *** **	**-0.519 ** **	**-0.468 * **	**-0.700 *** **	**-0.146 **	**-0.160 **
Turnover Rate	**(0.186)**	**(0.240)**	**(0.240)**	**(0.180)**	**(0.179)**	**(0.189)**
Female	0.143 ***	0.093 *	0.117 **	-0.048	-0.039	-0.031
	(0.035)	(0.050)	(0.056)	(0.030)	(0.041)	(0.046)
Black	-0.052	-0.180 *	-0.178 *	-0.298 ***	-0.361 ***	-0.287 ***
	(0.063)	(0.095)	(0.102)	(0.080)	(0.105)	(0.104)
Ever Low Income	-0.196 ***	-0.223 ***	-0.169 ***	-0.215 ***	-0.242 ***	-0.199 ***
	(0.039)	(0.059)	(0.061)	(0.037)	(0.054)	(0.057)
Age	4.766 ***	3.155 *	1.401	5.280 ***	4.281 ***	4.430 ***
	(1.141)	(1.822)	(1.998)	(1.059)	(1.527)	(1.650)
Age^2	-0.375 ***	-0.260 *	-0.133	-0.411 ***	-0.347 ***	-0.356 ***
	(0.086)	(0.137)	(0.151)	(0.080)	(0.116)	(0.126)
Number of Days Absent	-0.003	-0.007 *	-0.007 *	-0.004 **	-0.003	-0.002
in Kindergarten	(0.002)	(0.004)	(0.004)	(0.002)	(0.004)	(0.004)
Observations	1,578	750	603	1,659	793	631
R-squared	0.62	0.61	0.61	0.66	0.67	0.67

~ Robust standard errors , clustered by kindergarten-first grade class pairings, are reported in parentheses. Statistical significance is denoted by "***" when significant at the 1% level, "**" at the 5% level, and "*" at the 10% level. All models include first grade and kindergarten class fixed effects.

Table XIV: Effect of Classmate Turnover on Normalized 1st Grade SAT Scores with Restricted Samples (Center City Schools)

	Reading			Math		
	Full Sample	"Random" Classes	Small Schools	Full Sample	"Random" Classes	Small Schools
Normalized Kindergarten SAT Score	0.497 ***	0.470 ***	0.471 ***	0.575 ***	0.628 ***	0.617 ***
	(0.037)	(0.072)	(0.088)	(0.038)	(0.079)	(0.104)
Classmate Turnover Rate	**0.295**	**0.931 ** **	**0.995 ** **	**1.212 *** **	**2.529 *** **	**2.590 *** **
	(0.261)	**(0.453)**	**(0.490)**	**(0.281)**	**(0.829)**	**(0.872)**
Female	0.131 ***	0.129 *	0.197 **	0.030	0.093	0.140
	(0.041)	(0.077)	(0.093)	(0.043)	(0.076)	(0.102)
Black	-0.071	0.027	-0.021	-0.271 **	-0.105	-0.008
	(0.112)	(0.183)	(0.238)	(0.125)	(0.243)	(0.296)
Ever Low Income	-0.220 ***	-0.284 *	-0.331	-0.223 ***	-0.181	-0.298 *
	(0.079)	(0.147)	(0.202)	(0.062)	(0.116)	(0.158)
Age	2.743 **	-0.265	1.204	6.013 ***	2.549	9.037 **
	(1.100)	(1.964)	(3.153)	(1.740)	(3.716)	(3.598)
Age^2	-0.216 **	0.017	-0.094	-0.453 ***	-0.187	-0.686 **
	(0.084)	(0.154)	(0.242)	(0.132)	(0.288)	(0.275)
Number of Days Absent in Kindergarten	-0.006 ***	-0.007	-0.005	0.000	-0.007 *	-0.004
	(0.002)	(0.005)	(0.005)	(0.002)	(0.004)	(0.005)
Observations	887	322	218	912	326	222
R-squared	0.66	0.65	0.66	0.70	0.7	0.66

~ Robust standard errors , clustered by kindergarten-first grade class pairings, are reported in parentheses. Statistical significance is denoted by "***" when significant at the 1% level, "**" at the 5% level, and "*" at the 10% level. All models include first grade and kindergarten class fixed effects.

Table XV: Linear Probability Model of Attrition from STAR Schools At the End of Kindergarten

	OLS		School FE		Class FE	
Kindergarten Reading	-0.002		0.009		0.005	
Standardized Score	(0.008)		(0.008)		(0.008)	
Missing Kindergarten Reading Score	-0.021		-0.007		-0.005	
	(0.045)		(0.045)		(0.046)	
Kindergarten Math Standardized Score	-0.003		-0.014	*	-0.013	
	(0.008)		(0.008)		(0.008)	
Missing Kindergarten Math Score	0.052		0.041		0.042	
	(0.051)		(0.050)		(0.051)	
Female	0.005		0.001		0.004	
	(0.011)		(0.010)		(0.011)	
Black	0.038	**	-0.005		-0.011	
	(0.015)		(0.020)		(0.020)	
Free Lunch Recipient in Kindergarten	-0.002		0.007		0.005	
	(0.013)		(0.013)		(0.013)	
Missing Kindergarten Free Lunch Status	0.241	**	0.214	**	0.163	
	(0.117)		(0.098)		(0.111)	
Age	-0.456		-0.388		-0.352	
	(0.346)		(0.322)		(0.314)	
Age^2	0.039		0.033		0.030	
	(0.031)		(0.029)		(0.028)	
Missing Age	-0.432		-0.259		-0.112	
	(0.965)		(0.899)		(0.873)	
Number of Days Absent in Kindergarten	0.002	**	0.001		0.001	
	(0.001)		(0.001)		(0.001)	
Missing Number of Days	-0.158	***	-0.148	***	-0.117	*
Absent in Kindergarten	(0.047)		(0.051)		(0.063)	
Class w/ Aid in Kindergarten	-0.017		-0.016		-	
	(0.011)		(0.011)			
Class Size in Kindergarten	0.001		0.006		-	
	(0.002)		(0.005)			
Observations	3,087		3,087		3,087	
R-squared	0.01		0.11		0.15	

~ Robust standard errors , clustered by kindergarten-first grade class pairings, are reported in parentheses. Statistical significance is denoted by "***" when significant at the 1% level, "**" at the 5% level, and "*" at the 10% level. "School FE" refers to school level fixed effects and "Class FE" refers to kindergarten class fixed effects.

Table XVI: Effect of Classmate Turnover on Normalized 1st Grade SAT Scores by School Location (Instrumenting Lagged Test Score w/ Other Lagged SAT Subject Tests)

	Reading				Math			
	Center City		Non-center City		Center City		Non-center City	
Normalized Kindergarten	0.603	***	0.698	***	0.907	***	0.927	***
SAT Score	(0.046)		(0.035)		(0.074)		(0.044)	
Classmate	**0.245**		**-0.572**	***	**1.110**	***	**-0.728**	***
Turnover Rate	**(0.273)**		**(0.194)**		**(0.312)**		**(0.204)**	
Female	0.121	***	0.131	***	-0.003		-0.070	**
	(0.045)		(0.036)		(0.052)		(0.033)	
Black	-0.035		-0.045		-0.230		-0.209	
	(0.114)		(0.067)		(0.123)		(0.090)	
Ever Low Income	-0.180		-0.162		-0.132	*	-0.119	***
	(0.082)		(0.041)		(0.076)		(0.040)	
Age	2.227	**	4.516	***	3.365	**	3.241	***
	(1.135)		(1.195)		(1.684)		(1.165)	
Age^2	-0.178	**	-0.357	***	-0.255	**	-0.261	***
	(0.086)		(0.090)		(0.128)		(0.088)	
Number of Days Absent	-0.005	**	-0.003		0.003		-0.002	
in Kindergarten	(0.002)		(0.002)		(0.003)		(0.002)	
Observations	881		1,570		887		1,640	
R-squared	0.66		0.62		0.64		0.63	

~ Robust standard errors , clustered by kindergarten-first grade class pairings, are reported in parentheses. Statistical significance is denoted by "***" when significant at the 1% level, "**" at the 5% level, and "*" at the 10% level. All models include first grade and kindergarten class fixed effects.