Peer Effects and Returns to Scale in the Education Production Function

Marc Luppino*

July 23rd, 2010

---

* UCLA Department of Economics. mluppino@ucla.edu. This paper has benefited from many helpful insights from my dissertation committee: Leah Platt Boustan, and Moshe Buchinsky (chair). All remaining errors are mine.

1) Introduction

There has been a proliferation of papers investigating peer effects in educational settings. While much attention has been paid to the question of whether peer effects exist, recent studies have pushed to examine the underlying mechanisms that determine why and how students are affected by their peers. The importance of peer effects is a central point of contention in the debate surrounding a number of education policy areas, ranging from school choice and ability tracking to "mainstreaming" of special education students and racial and economic desegregation. Greater evidence on peer effect mechanisms is needed to inform decision makers, so that they are better able to formulate effective education policies. This paper is one of the first to investigate the potential peer effects induced by grade retention in elementary school. It makes a number of contributions to the literature by exploring the possibility of heterogeneous and non-linear peer effects. Specifically, I examine whether the sign and magnitude of grade repeater induced peer effects vary by gender, socio-economic background, and school location. I also test whether these peer effects vary non-linearly as one increases the number of grade repeaters assigned to a student's classroom.

The identification of peer effects is associated with a number of difficult econometric challenges. First, it is necessary to address the obvious selection problem stemming from non-random assignment of students of varying ability into schools and classrooms. In particular, one would expect student ability to also be strongly correlated with school and teacher quality, which are also likely strong determinants of student outcomes in addition to peer quality. In order to overcome this selection problem, I take advantage of data generated from Project STAR, which randomly allocated grade-repeaters across classrooms as part of its experimental protocol, to identify how sharing a classroom with grade-repeaters affects the achievement of non-repeaters.

Second, researchers need to choose an appropriate measure of peer ability in order to estimate peer effects. One common strategy is to define peer ability as the mean of contemporaneous or lagged peer test scores. Using contemporaneous scores generates estimates that are likely to suffer from the reflection problem (Manski, 1993). Lagged peer test scores are also unlikely to be exogenous to own current achievement if peer groups are largely consistent over time. An alternative strategy is to use peer socio-economic background characteristics (such as gender or race) as proxies for ability. In this paper, I use the proportion of grade repeaters assigned to the classroom which I argue is a more appropriate measure of peers' ability. Specifically, grade retention is highly correlated with student academic performance. In addition, grade retention decisions made at the end of a given grade are unlikely to be based on characteristics of the incoming cohort of students.

In an effort to explore non-linear peer effects, most researchers have focused on trying to more fully characterize the distribution of student ability within the classroom. This has translated into researchers expanding on the basic linear-in-mean empirical specification by including both the mean and standard deviation of peer ability or the share of students in distinct ability tiers. In contrast, I consider non-linear peer effects as they relate to returns to scale in the education production function.[1] Specifically because I use the number of grade repeaters assigned to the classroom as a measure of peer ability, I am able to investigate how the marginal effect of adding an additional low ability students to the classroom changes as the number of grade repeaters assigned to the classroom increases.

My findings suggest that grade repeater induced peer effects are both heterogeneous and non-linear. In the majority of specifications that I estimate, I reject that such peer effects exhibit constant or decreasing returns to scale. Instead, I find that clustering of grade repeaters into one classroom generates large negative effects on higher performing students (e.g. female and non-poor students). This effect is particular pronounced in Project STAR schools located outside of center cities. Specifications that assume linear peer effects fail to pick up such effects. In addition, these results provide further evidence that the nature of educational production varies greatly by school location. Specifically, I find that having a higher proportion of grade repeaters as classmates has a sizable positive effect on weaker students (e.g. boys and poor students) in center city schools. Conversely, I find that grade repeaters generate negative peer effects on weaker students in Project STAR schools located outside of center cities. The strength of the ability peer effect generated by grade repeaters and student sensitivity to peer continuity are

1 In equating the terms "linear effect" and "constant returns to scale," I am implicitly assuming that the education production function has additively separably inputs (which is the standard assumption in the empirical literature). In a production function where inputs enter additively and linearly, a doubling of inputs implies a doubling of output. In other words, production exhibits constant returns to scale.

likely to vary by school location. Such differences potentially explain this contrast in findings across distinct school locations.[2]

Papers that try to capture heterogeneous and/or non-linear peer effects often produce results that are not clear cut and thus have limited use in informing policy debate. By investigating peer effects induced by grade repeaters and by focusing on how such effects are potentially mediated by scale, I am able to produce findings that have more straightforward policy implications. In particular, my results indicate that in most instances school administrators should refrain from clustering or tracking weaker students (particularly early grade repeaters and students with similar academic and behavioral issues). At the same time, the findings of this paper reinforce the notion that it is important to take into account school contexts when formulating optimal school policies. In center city schools, where behavioral problems are typically more abundant, my own prior research suggests that stronger friendship bonds (formed by sharing successive classes) actually hinder learning. Administrator in center city schools might aim to reduce disciplinary problems by taking measures, such as grade retention, to promote peer turnover.

The remainder of the paper is organized as follows. Section 2 contains background information on grade retention and the peer effects literature. Section 3 presents the data. In Section 4, I discuss my empirical strategy for identifying how the proportion of grade repeaters in the classroom affects student achievement. Section 5 presents my main findings. Finally, in Section 6 I conclude.

2) Background

There has been a proliferation of recent studies that have attempted to use natural or quasi-experimental strategies to identify peer effects in the classroom. Given this ample body of work, I review the studies most relevant to this paper. Since I examine peer effects in kindergarten, I first focus on studies that examine peer effects in primary school. Using data from the Early Childhood Longitudinal Survey (ECLS-K), Aizer (2008) finds that children with undiagnosed ADD generate negative externalities in the classroom. Using the same data, Neidell

2 Luppino (2010) presents a conceptual framework to explain heterogeneous peer effects by school location.

and Waldfogel (2008) find that students who attended preschool have positive effects on their peers in terms of learning and behavioral outcomes. Carrell and Hoekstra (2009) use administrative data from Florida to analyze the spillovers caused by children who are exposed to domestic violence. They find that children from troubled families significantly decrease their peers' reading and math test scores and increase misbehavior in the classroom. Finally, Cooley (2010) uses the introduction of student accountability policies in North Carolina, which led low achieving students to increase their study effort, to identify peer effects in third through fifth grade classrooms. Most of these studies find moderately sized peer effects, with a one standard deviation increase in peer quality typically improving outcomes by less than 10 percent of a standard deviation.

This paper is one of the first to investigate the potential peer effects induced by grade retention in elementary school. Not surprisingly, Cooley et al. (2010) find that primary school students that are retained have lower general and cognitive abilities than those who are promoted to the next grade. In addition, these authors find that kindergarten retention lowers the reading scores of retained students between 24 and 28 percent. The general consensus from the peer effects literature suggests that these lower ability students are likely to generate negative externalities for their school peers.[3] Lavy et al. (2009) find that the proportion of repeaters in a school has a negative effect on the performance of regular middle and high school students in Israel. Such peer effects are likely to be particularly important at earlier ages when environmental factors are so vital to development.

Non-random assignment of students into schools and classrooms makes identifying peer effects a difficult challenge for researchers. In this paper, I take advantage of data generated from Project STAR, which random allocated grade-repeaters across classrooms as part of its experimental protocol, to identify how sharing a classroom with grade-repeaters affects the achievement of non-repeaters. While their have been a number of prior studies that have used the STAR data to examine peer effects, I deviate from this previous research in a number of

3 Alternatively, grade repeaters might have a negative effect on students that normally progress through school if relative age matters for student success. As older grade repeaters enter their new academic cohort the rest of the students will become relatively younger. However, the findings of Cascio and Whitmore-Schazenbach (2008) suggest that such relative age effects are not important and that is a student's absolute age that matters for academic achievement.

important ways. In particular, Sojourner (2008) studies peer effects on student achievement in first grade using lagged outcome measures as the key peer characteristic. Because early age test scores are particularly noisy measures of ability, classical measurement error will downward bias peer effect estimates based on lagged test scores. I avoid this problem by using the proportion of grade repeaters found in the classroom as a measure of peers' ability. Graham (2008) and Boozer and Cacciola (2001) use variation in peer quality across small and regular size classrooms to identify peer effects in kindergarten and the later grades. However as Lazear (2001) notes, peer effects are likely to be mediated by classroom size. Therefore, I investigate kindergarten peer effects only using students randomly assigned to regular size classrooms. In addition since Graham bases his identification strategy on conditional variance restrictions, he is only able to identify the square of the social multiplier and so is unable to determine the sign of peer effects. Finally, Graham's approach only provides consistent estimates of peer effects that take the "linear-in-means" form (Manski, 1993). In my analysis, I explore the possibility of non-linear peer effects.[4]

Most studies that have attempted to evaluate non-linear peer effects have tended to focus on the specific effects of peer ability dispersion. The insights gained from these efforts have been mixed. Vigdor and Nechyba (2004) find that classroom ability dispersion has a positive effect on student achievement, while Burke and Sass (2008) and Duflo et al. (2008) find negative effects. Alternatively, I consider non-linear peer effects as they relate to returns to scale in the education production function. Specifically since I use the number of grade repeaters assigned to the classroom as a measure of peer ability, I am able to investigate how the marginal effect of adding an additional low ability student to the classroom changes as the number of grade repeaters assigned to the classroom increases. This type of analysis allows one to examine whether critical thresholds exist over which small changes in peer composition lead to particularly large or distinct changes in student outcomes.

4 The current study is also similar to the work of Cascio and Whitmore-Schazenbach (2008), who investigate relative age effects. Using the STAR data, Cascio and Whitmore-Schazenbach estimate the effect of absolute age and relative age at the start of kindergarten on kindergarten and eighth grade reading and math achievement. They find no evidence that relative age matters for test scores or for the likelihood of taking a college-entrance exam. As these authors note, their estimation strategy identifies the average effect of relative age for children who comply with school-entry regulations. Conversely, this paper utilizes a different source of variation in the age composition of classrooms.

3) Data

In order to assess the possible peer effects resulting from kindergarten retention, 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. 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.

I restrict my sample to kindergärtners without missing data on test scores and personal characteristics. In addition, I remove from my sample students that were randomly assigned to small classrooms. As Lazear (2001) suggests, peer effects are likely to be mediated by class size. Limiting my analysis to the subsample of students assigned to regular size classrooms should make the corresponding findings more generalizable since they should not be driven by interactive effects with class size treatment. My results are qualitatively similar if I analyze the full sample of Project STAR kindergärtners, instead of the regular student subsample.

Summary statistics for all the variables used in the subsequent analysis are reported in **Table I**. Consistent with random assignment of students and teachers to classrooms, student and teacher characteristics appear to be balanced across small and regular size classrooms. In addition, randomized assignment means that the subsample of students assigned to regular size classrooms is representative of all students in Project STAR schools. As Whitmore-Schanzenbach (2007) documents, Project STAR schools have a higher proportion of poor and minority students relative to average elementary schools both statewide and nationally. These schools also receive fewer resources in terms of per pupil expenditures compared to the national average. It is important to bear in mind these features of the Project STAR sample when drawing inferences from this data.

Grade repeaters comprise around 4 percent of all Project STAR kindergärtners. Fifteen schools did not have any repeaters in their kindergarten cohorts. Within the remaining schools, 32 percent of regular size classrooms were assigned no repeaters. Thirty-seven percent of regular size classrooms were assigned one repeater, 19 percent were assigned two repeaters, and the remaining regular size classrooms were assigned between 3 and 5 repeaters. Grade repeaters represent a smaller share of all kindergärtners in center city schools relative to other schools (3 percent compared to 5 percent). As a result, center city schools have a higher portion of classrooms with no repeaters (41 percent compared to 26 percent).

The summary statistics also demonstrate how students in center city schools are very different from those in the other participating schools. Students in these schools are disproportionately African-American (64 percent) and poor (58 percent received free lunch) relative to students from schools outside of center cities. Teachers in center city schools also have lower credentials in terms of experience, education, and whether they are on a career ladder compared to their counterparts in the other Project STAR schools. In addition to these highlighted differences, the findings of Luppino (2010) suggest that center and non-center city schools differ in their schooling inputs and educational practices (or technology). As such, it is likely that the peer effects generated by kindergarten retention will differ by school location.

4) Empirical Strategy

My identification strategy relies on the fact that all students participating in Project STAR (including grade repeaters) were randomly assigned to their kindergarten classroom. This induced random variation within schools in the proportion of repeaters assigned to a particular classroom. I use this variation to identify how sharing a classroom with these repeaters affects student performance. Specifically, I estimate the peer effects generated by grade repeaters using the basic specification:

$$T_{ics} = \beta_0 rep_{ics} + \beta_1 \overline{rep_{(-i)cs}} + \beta_2 rank_c(a_i) + \beta_3 x_{ics} + \beta_4 \overline{x_{(-i)cs}} + \beta_5 z_{cs} + \alpha_s + \delta_{cs} + \varepsilon_{ics}$$

where the outcome of interest, T_{ics} is the reading or math Stanford Achievement Test (SAT) score of student *i*. I standardize test scores within sample to facilitate the comparison of my results to those of other related studies.[5] *rep* indicates whether the child is repeating kindergarten and $\overline{rep}_{(-i)cs}$ represents the class mean of *rep* (e.g. it reflects the proportion of repeaters in classroom *c* not including the index child). $rank_c(a_i)$ is student *i*'s rank in his or her kindergarten classroom age distribution, normalized to lie between zero (for the youngest child) and one (for the oldest child).[6] I include this particular control variable to ensure that β_1 more cleanly captures ability peer effects instead of effects driven by shifts in relative age. X_{ics}, is a vector of individual characteristics including measures of gender, race, family income status (as measured by whether or not the student is a free lunch recipient), and age and $\overline{X}_{(-1)cs}$ represents the class mean of *X*. Z_{cs} is a vector of classroom and teacher specific characteristics. It includes measures of class size, whether the classroom has a teacher's aide, the teacher's years of experience, whether the teacher is new, whether the teacher has more than a bachelor's degree, and whether the teacher is on a step of the career ladder (rather than being an apprentice, on probation, or not-on-ladder). α_s represents a school fixed effect. The error term consists of a group specific component (δ_{cs}) and an individual, idiosyncratic component (ε_{ics}). I cluster all standard errors at the classroom level.

5 Specifically, I take the mean and standard deviation of each test score for the entire sample of children randomly assigned to regular-size classrooms and calculated z-scores for each student.

6 I calculate a child's age rank ($rank_{ic}$) in his or her classroom by ordering children from youngest ($rank_{ic}$=1) to oldest ($rank_{ic}=n_c$). The normalized rank measure is then $rank_c(a_i)=(rank_{ic}-1)/(n_c-1)$ where n_c is the number of students in classroom *c* with non-missing age variables.

Due to the random assignment of students and teachers to classrooms, $\overline{rep_{(-i)cs}}$ should be exogenous. As a check to confirm this random assignment, I perform a number of statistical tests. Specifically following Clotfelter, Ladd, and Vigdor (2006), I conduct a series of chi-square tests to compare the balance of observable characteristics of students across kindergarten classrooms within each school. These tests examine whether students' classroom assignments are statistically independent of a set of four student characteristics: gender, race, participation in the Federal subsidized school lunch program, and whether the student is a grade repeater. According to these authors, the overall distribution of p-values for each set of tests should be close to uniform under random assignment.[7] 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 distribution skewed towards low p-values. Informally the distribution of p-values for each student characteristic appears to be roughly uniform. 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: 0.29 for race, 0.33 for gender, 0.78 for income status, and 0.13 for whether a student repeated kindergarten. This evidence is largely consistent with random assignment to classroom.

Alternatively, I run a regression of whether a student is a grade repeater on classroom and teacher characteristics. If the assignment of grade repeaters to classrooms in Project STAR schools is truly random, repeater status should not be strongly correlated with the characteristics of the classroom and teacher. The results from this regression are reported in **Table II**. Being a grade repeater is positively correlated with being a boy and with being poorer, which is consistent with what we would expect given the profile of grade retention outlined in the literature. Conversely, I find no statistically significant association between grade retention and race. Most importantly, I also find no statistically significant correlation between being a grade repeater and observable classroom and teacher characteristics. Additionally, I find that classroom fixed effects have no additional predictive power of repeater status once I control for

7 The p-value is the probability of rejecting the null hypothesis when it is true (i.e. of committing a Type I error). The null hypothesis in this case is that schools randomly assigned students across classrooms. Even if randomization occurred we should expect that these tests would produce p-values less that 0.05 roughly five percent of the time and p-values less than 0.10 roughly ten percent of the time. This implies that if the null is true, the distribution of p-values resulting from distinct trials should be roughly uniform.

school fixed effects. Specifically, I conduct a likelihood ratio test where I compare the specification with class fixed effects (i.e. the unrestricted model) with the specification with only school fixed effects (i.e. the restricted model). The difference in the predictive power on the unrestricted and restricted models is not statistically significant. Again this evidence is consistent with random assignment of students and teachers to classrooms and suggests that the proportion of grade repeaters found in a classroom is exogenous. Therefore, OLS estimation should produce an unbiased estimate of β_1, the peer effect of kindergarten grade retention on kindergarten test scores.

5) Results

I begin by examining the peer effects of grade repeaters on the academic performance of non-repeaters using the full sample of Project STAR schools. **Table III** presents results for models in which both reading and math SAT score are the dependent variable. If peer effects play an important role in learning, then we would expect students whom share a classroom with a higher number of repeaters to perform worse academically. Analyzing the full sample of schools, I find that the proportion of a repeaters in a student's kindergarten classroom has only a slight negative effect on his or her achievement. This effect is also not statistically significant.

It is important to note that since center city schools are overrepresented in the STAR sample, it is difficult to interpret estimates of grade repeater peer effects based on the full, unweighted sample. Table III also presents results separately for schools in center city school districts and schools outside of these districts. Kindergärtners in center city schools appear to benefit from having grade repeaters in their classrooms. Specifically, I find that increasing the proportion of grade repeaters in the classroom by 4.5 percent (e.g. approximately one child) increases both reading and math test scores by approximately 0.064 of a standard deviation. However, this effect is not statistically significant at conventional levels.

In schools outside of center city districts, I find that students are negatively affected by sharing classrooms with grade repeaters. In these schools, I find that increasing the number of grade repeaters in the classroom by approximately one child lowers math scores by 0.084 of a standard deviation. This effect is highly statistically significant. Repeater driven peer effects on reading scores are slightly smaller and not statistically significant.

These results are further evidence that the nature of educational production varies greatly by school location. Analyzing the entire sample of Project STAR schools, Whitmore (2005) finds that kindergärtners assigned to a predominately female class score higher on the SAT than those students assigned to classes with a higher fraction of boys. Here we see that her results are largely driven by the STAR students that attended center city schools. In these schools classroom gender composition has a large and statistically significant effect on student performance, while in schools outside of center city I find effects that are considerably smaller and marginally statistically significant effect. Since teachers consistently report that girls are better behaved than boys, these findings suggest that behavioral disruption is potentially of greater concern in center city schools relative to other Project STAR schools.

The distinct findings by school location reported in this paper are consistent with those reported in Luppino (2010). In this earlier work I found that first graders in center city Project STAR schools perform worse academically when they are assigned to classrooms with a higher proportion of familiar peers. Conversely, I find that peer group consistency has a positive effect on student performance in the other Project STAR schools. Given this tendency, we might expect grade retention in center city schools to have less of a negative shock on students than in other school contexts since directly affected students are exposed to an entirely new cohort of peers upon repeating kindergarten. Therefore, we would expect any ability peer effects arising from grade repeaters to be weaker in center city schools compared to other schools.

In addition, if children of the same age gain familiarity with each other before starting kindergarten (possibly by attending preschool or day care) then the proportion of grade repeaters assigned to a students kindergarten class may also affect the degree of peer continuity that they experience in school. It follows that differences in the strength of ability peer effects generated by grade repeaters and differences in sensitivity to peer continuity might explain why I find that grade repeater peer effects are positive in center city schools and negative in the other Project STAR schools.

The magnitude of these estimated effects is consistent with those found in the related literature. Again, I find that increasing the proportion of grade repeaters in the classroom by 4.5 percent (e.g. approximately one child) increases test scores in center city schools and lowers scores in other schools by 0.064 and 0.084 of a standard deviation, respectively. Lavy et al. find that a similar increase in the proportion of grade repeaters leads to a decrease in middle and high school test scores between 0.021 and 0.051 of a standard deviation. We can also compare the findings on the present study to those of Neidell and Waldfogel (2008) who conduct a similar analysis that looks at the spillover effects from preschool attendance. They find that adding one of these better prepared students to a kindergarten classrooms increases the math scores of their classmates by 0.006 of a standard deviation. This effect is an order of magnitude smaller than the effects I find from assigning an additional grade repeater to the classroom.

5.1) Non-linear Peer Effects

Most related research assume that peer effects are linear (e.g. adding an above average student should have the same affect on peer learning as removing a below average student, assuming the above and below average students are relatively equidistant from the average students in terms of ability). Specifying a linear peer effect is equivalent to assuming that peer inputs exhibit constant returns to scale within the education production function. However, I am unable to test this type of linearity assumption since the STAR experiment did not collect any pre-kindergarten ability measures. While I am unable to test this assumption over the entire distribution of peer ability, I am able to test whether the number of grade repeaters generates peer effects with constant returns to scale. To test this assumption, I consider a more flexible specification that includes a set of dummy variables indicating the number of grade repeaters found in a student's kindergarten classroom. This number ranges from zero to five, where I designate zero as the omitted category. Given this specification, I am able to conduct a Wald test to examine whether the peer effect is linear in the number of grade repeaters assigned to the classroom.

It is not clear a priori what type of returns to scale we should expect for this type of peer effect. According to Lavy et al., teachers tend to shift a good portion of their attention to grade repeaters (at the expense of the other students) once they are introduced into the classroom. However since instruction is at least partially a public good, we might expect this negative peer effect to exhibit diminishing returns as the number of grade repeaters in the classroom increases.

Conversely, one could also envision a scenario in which reaching a critical mass of grade repeaters leads to a substantial increase in the level of classroom disruption. In addition to testing the constant returns to scale assumption, I also test for a specific form of diminishing returns where the null hypothesis assumes that all the grade repeater dummy variables have the same effect. In other words, I test the assumption that additional grade repeaters do not matter once a single repeater has been assigned to the classroom.

Tables IV and **V** presents results for models that allow for non-linear effects for suburban and rural schools and for center city schools, respectively. In schools outside of center city districts, I again find that students are negatively affected by sharing classrooms with grade repeaters. However, now I find that these effects are highly statistically significant for both reading and math. For the both reading and math specifications, I reject that grade repeater induced peer effects exhibit constant returns to scale and I also reject that the grade repeater dummy variables have the same effect. The coefficients for the majority of classroom grade repeater dummies are highly statistically significant and most of them are negative.[8] Increasing the number of grade repeaters in the classroom from one to four leads to a generally increasing, though diminishing negative effect on student achievement. However, I find that adding a fifth grade repeaters to the classroom produces an extremely large jump in this negative effect on student achievement. Specifically, increasing the number of grade repeaters assigned to the classroom from four to five lowers math scores by 0.75 of a standard deviation. The results for the specifications with reading as the dependent variable are qualitatively similar.

Focusing on the sample of center city school children, I now find that the number of grade repeaters in the classroom has a statistically significant effect on both reading and math scores. Moreover, I reject that these effects exhibit constant returns to scale and that different numbers of grade repeaters have the same effect. However, examining individual coefficients it would appear that differing the number of grade repeaters in the classroom generally leads to a very similarly positive effect. The lone exception is the case where three grade repeaters are

8 For both reading and math specifications, the effect of sharing a classroom with two grade repeaters is smaller and not statistically significant compared to the effect of sharing a classroom with one grade repeater. Additionally for the reading specification, the effect of sharing a classroom with four grade repeaters is smaller and not statistically significant compared to the effect of sharing a classroom with three grade repeaters. This pattern possibly suggests that the ability to pair-up students of like ability (which is easier with even numbers) for group learning activities (which are more prevalent during reading instruction than math instruction) might at least partially mitigate negative peer effects generated from weaker students.

assigned to the class, in which case the estimated effect is negative and highly statistically significant. Only one classroom in all of the center city schools was randomly assigned three grade repeaters. It could be the case that unique characteristics of this classroom (that are not captured by the set of control variables) are driving this result.

In order to further explore the influence of outliers on the results of my non-linear analysis, I drop from my sample schools that assign a large number of grade repeaters to a particular classroom. For the center city sample, I drop schools that assign three or more grade repeaters to a particular classrooms. In schools outside of center cities, there are relatively more grade repeaters and classrooms that are assigned three grade repeaters are sufficiently more prevalent. Therefore for the non-center city sample, I drop schools that assign four or more grade repeaters to a particular classrooms. Based on these restricted samples I find that the findings of non-linear effects are not robust for center city schools. These results are presented in columns (3) and (4) of Table V. Conversely, the results in columns (3) and (4) of Table IV suggest that the estimated non-linear peer effects are extremely robust for the other Project STAR schools.

5.2) Heterogeneous Effects by Student Socio-economic Status

The findings of Lavy et al. suggest that grade repeaters have a negative impact mainly on students from a low socio-economic background. In particular, one might expect poorer students with fewer family resources to be especially sensitive to school environmental factors. To investigate this possibility, I estimate models separately for children that receive a subsidized school lunch and for children that do not. When I restrict my sample to poorer children, I find that grade repeaters induce large positive peer effects on children in center city schools and large negative peer effects on children in the other Project STAR schools. These effects are all highly statistically significant in both specifications that assume linear peer effects and in those that allow for non-linear effects. The corresponding results for center city and other Project STAR schools are reported in **Tables VI** and **VII**, respectively. According to the specifications that assume a linear effect, adding a grade repeater to the classroom increases the math scores of the poorer children in center city schools by 0.181 of a standard deviation. Conversely, I find that assigning an additional grade repeater to the classroom lowers the math scores of poorer children in the other Project STAR schools by 0.144 of a standard deviation. The magnitude of these peer effects on reading scores is similar. These findings suggest that is the most vulnerable students who are particularly affected by grade repeating peers.

The results for the poorer student subsample also suggest that peer effects induced by grade repeaters are non-linear. For poor students in both center city and non-center city schools, I reject that grade repeater induced peer effects are linear in specifications with both reading and math score as the outcome. I also generally reject that different numbers of grade repeaters have the same effect. When I restrict these subsamples to limit the influence of outliers, I find that the findings of non-linear effects are not as robust for center city schools and highly robust for the other Project STAR schools.

Based on estimates from specifications assuming both linear and non-linear peer effects, it appears that grade repeaters have a negative effect on non-poor students in both the center city and other Project STAR schools. However, these effects are generally not statistically significant. Turning to specifications that allow for non-linear peer effects, I find a statistically significant effect of the number of grade repeaters assigned to the classroom on the academic performance of non-poor students in center city and non-center city schools. I also both strongly reject that this effect exhibits constant returns to scale and that different numbers of grade repeaters have the same effect. Examining individual coefficients, I find that only the effect of being in a classroom with four or five grade repeater is statistically significant. When I restrict these subsamples to limit the influence of outliers, these results appear to be more robust for specifications with reading score as the outcome and less so for specifications with math score as the dependent variable.

5.3) Heterogeneous Effects by Gender

Because grade retention disproportionately affects boys and peer groups are largely gender specific at this age, one might expect the impact to be greater among other boys in the class. In related work, Aizer (2009) finds that the peer effects generated from undiagnosed ADD are largely experienced by boys. ADD is a disorder that also disproportionately directly affects boys. I find that grade repeaters induce large positive peer effects on boys in center city schools and large negative peer effects on boys in the other Project STAR schools. These effects are all highly statistically significant in both specifications that assume linear peer effects and in those that allow for non-linear effects. The corresponding results for center city and other Project STAR schools are reported in **Tables VIII** and **IX**, respectively. According to the specifications that assume a linear effect, adding a grade repeater to the classroom increases the reading scores of boys in center city schools by 0.131 of a standard deviation. Conversely, I find that assigning an additional grade repeater to the classroom lowers the reading scores of boys in the other Project STAR schools by 0.093 of a standard deviation. The magnitude of these peer effects on math scores is similar. These estimates are somewhat larger than the peer effects estimated by Aizer.

The results for boys also suggest that peer effects induced by grade repeaters are non-linear. For male students in center city schools, I reject that grade repeater induced peer effects are linear in specifications with both reading and math score as the outcome. I also reject that different numbers of grade repeaters have the same effect. For boys in the other Project STAR schools, I find stronger evidence for non-linear peer effects in specifications with reading scores as the outcome. Reading lessons at this age are typically more group oriented compared to math lessons, which might partially explain this discrepancy in findings. When I restrict these subsamples to limit the influence of outliers, I find that the findings of non-linear effects for boys are generally robust for both center city and non-center city schools.

Based on estimates from specifications assuming linear peer effects, I find no statistically significant effect of the number of grade repeaters assigned to the classroom on the academic performance of girls. However, I do find a statistically significant peer effect for girls in the specifications that allow for non-linear peer effects, particularly outside of center cities. The magnitude of these effects for girls is similar to those found for non-poor students. Over most specifications, I strongly reject that grade repeater induced peer effects exhibit constant returns to scale for girls. Also, I generally strongly reject that different numbers of grade repeaters have the same effect. Results from the restricted samples suggest that these results are not highly sensitive to outliers.

6) Conclusion

In this paper, I investigate the peer effects experienced by kindergärtners as a results of grade retention policies. Taking advantage of data generated from Project STAR, I find consistent evidence that such policies generate both heterogeneous and non-linear peer effects. In the majority of specifications that I estimate, I reject that grade repeater induced peer effects exhibit constant or decreasing returns to scale. Instead, I find that clustering of grade repeaters into one classroom generates large negative effects on higher performing students (e.g. female and non-poor students). This effect is particular pronounced in Project STAR schools located outside of center cities. Specifications that assume linear peer effects fail to identify these large effects. In addition, my results provide further evidence that the nature of educational production varies greatly by school location. Specifically, I find that having a higher proportion of grade repeaters as classmates has a sizable positive effect on weaker students (e.g. boys and poor students) in center city schools. Conversely, I find that grade repeaters generate negative peer effects on weaker students in Project STAR schools located outside of center cities.

Papers that try to capture heterogeneous and/or non-linear peer effects often produce results that are not clear cut and thus have limited use in informing policy debate. By investigating peer effects induced by grade repeaters and by focusing on how such effects are potentially mediated by scale, I am able to produce findings that have more straightforward policy implications. In particular, my results indicate that in most instances school administrators should refrain from clustering or tracking weaker students (particularly early grade repeaters and students with similar academic and behavioral issues) in early grades. At the same time, the findings of this paper reinforce the notion that it is important to take into account school contexts when formulating optimal school policies.

It is important to note the potential limitations of this study. Since especially young children are likely to be particularly sensitive to environmental factors, the findings of this study may not generalize to children in later grades. In addition while this paper uses grade retention as a means of identifying peer effects, its results are not sufficient to fully evaluate the externalities caused by such policies. Specifically, I am unable to evaluate how retention of some children in kindergarten potentially affected those students who were not retained and progressed regularly to first grade. Further research that fully accounts for compositional changes and resulting peer effects caused by grade retention policies would be of great value in assessing the costs and benefits of such policies.

Finally, in this paper I evaluate the returns to scale of adding weaker students to the classroom. It is highly likely that such results would not hold when analyzing peer effects generated from students at the other end of the ability distribution. Future research is needed to better characterize the returns to scale of peer ability over the full range of the ability distribution.

7) References

Aizer, A. (2009). “Peer Effects and Human Capital Accumulation: the Externalities of ADD.” NBER Working Paper No. 145354.

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.

Carrell, S. E. and M. L. Hoekstra (2008). “Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids.” NBER Working Paper No. 1 4246.

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.

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).

Deming, D. and S. Dynarksi (2008). “The Lengthening of Childhood.” Journal of Economic

Perspectives, 22(3): 71-92.

Duflo, E., P. Dupas, and M. Kremer (2008). “Peer Effects and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya.” Unpublished manuscript.

Figlio, D. N. (2007). “Boys Named Sue: Disruptive Children and Their Peers.” Education Finance and Policy: 376-394.

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).

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

Lavy, V., M. D. Paserman, and A. Schlosser (2009). “Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom.” NBER Working Paper No. 14415.

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

Luppino, M. (2010). “Classmate Turnover and Student Achievement: New Findings from Project STAR.” Job Market Paper.

Manski, C. F. (1993), “Identification of Endogenous Social Effects: The Reflection Problem.” Review of Economic Studies, 60: 531-542.

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

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

Vigdor, J. and T. Nechyba (2004). “Peer Effects in Elementary School: Learning from “Apparent” Random Assignment.” Unpublished manuscript.

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.

8) Appendix

Table I: Summary Statistics – Mean (Standard Deviation)

Variables	All Schools Classes	Center City Schools	Other Schools
Student Characteristics:			
Repeater	0.04	0.03	0.05
Female	0.49	0.49	0.48
Black	0.33	0.65	0.07
Free Lunch Recipient	0.49	0.59	0.40
Age in Years	5.51	5.48	5.53
	(0.35)	(0.33)	(0.36)
Classroom Characteristics:			
Regular Class w/ Aide Treatment	0.50	0.52	0.49
Class Size	22	23	22
	(2)	(2)	(2)
Teacher Characteristics:			
New	0.06	0.08	0.05
Experience in Years	9	9	9
	(6)	(7)	(5)
Missing Experience Info	0.01	0.00	0.01
Has at least a Master's Degree	0.36	0.28	0.43
On Career Ladder	0.78	0.68	0.86
Missing Career Ladder Info	0.09	0.14	0.05
SAT Scores:			
Reading	435	433	437
	(31)	(31)	(31)
Math	483	480	486
	(47)	(48)	(46)

Table II: Regression of Whether a Student is a Grade Repeater on Classroom and Teacher Characteristics

Variables	All STAR Schools School Fixed Effects (FE) [1]	[2]	Classroom FE [3]	Rural Schools School Fixed Effects (FE) [1]	[2]	Classroom FE [3]	Urban Schools School Fixed Effects (FE) [1]	[2]	Classroom FE [3]
Female	-0.0267***	-0.0266***	-0.0274***	-0.0308***	-0.0306***	-0.0321***	-0.0218***	-0.0218***	-0.0217***
	[0.00475]	[0.00477]	[0.00501]	[0.00729]	[0.00731]	[0.00774]	[0.00565]	[0.00568]	[0.00587]
Black	0.0052	0.0050	0.0099	-0.0033	-0.0033	0.0026	0.0163	0.0164	0.0193
	[0.0123]	[0.0124]	[0.0125]	[0.0184]	[0.0184]	[0.0183]	[0.0151]	[0.0152]	[0.0158]
Free Lunch Recipient	0.0548***	0.0549***	0.0543***	0.0565***	0.0565***	0.0561***	0.0513***	0.0514***	0.0506***
	[0.00709]	[0.00711]	[0.00729]	[0.00877]	[0.00877]	[0.00903]	[0.0121]	[0.0122]	[0.0124]
Classroom Treatment:									
Small		-0.0090			-0.0078			-0.0103	
		[0.00621]			[0.00939]			[0.00790]	
Regular w/ Aide		-0.0060			-0.0035			-0.0101	
		[0.00603]			[0.00922]			[0.00758]	
Joint Test P-value		0.34			0.71			0.33	
Teacher Characteristics:									
New		-0.0090			-0.0196			-0.0028	
		[0.0139]			[0.0384]			[0.0142]	
Experience in Years		-0.0003			-0.0006			-0.0003	
		[0.000483]			[0.000798]			[0.000610]	
Has at least a Master's Degree		-0.0026			-0.0058			0.0045	
		[0.00692]			[0.00987]			[0.00985]	
Education Missing		-0.0405			-0.0556				
		[0.0331]			[0.0528]				
On Career Ladder		0.0047			-0.0037			0.0107	
		[0.00937]			[0.0257]			[0.00811]	
Missing Career Ladder Info		-0.0040			0.0075			-0.0083	
		[0.0109]			[0.0289]			[0.0111]	
Joint Test P-value		0.55			0.49			0.29	
Likelihood Ratio Test P-value*			0.77			0.79			0.49
Observations	6,299	6,299	6,299	3,472	3,472	3,472	2,827	2,827	2,827
R-squared	0.06	0.06	0.09	0.05	0.05	0.09	0.06	0.06	0.09

* Unrestricted model = [3], Restricted model = [1].
** Robust standard errors in brackets. "***", "**", and "*" signify statistical significance at the 1%, 5%, and 10% significance levels, respectively.

Table III: Estimates of Linear Peer Effects

Variables	Reading All	Reading Rural	Reading Urban	Math All	Math Rural	Math Urban
Repeated Kindergarten	-0.288***	-0.293***	-0.251**	-0.352***	-0.404***	-0.271**
	[0.0767]	[0.105]	[0.102]	[0.0808]	[0.104]	[0.123]
Proportion of Repeaters in Classroom	**-0.209**	**-0.913**	**1.418**	**-0.592**	**-1.867*** **	**1.501**
	[0.583]	[0.592]	[1.133]	[0.587]	[0.592]	[1.170]
Normalized Rank in Classroom Age Distribution	0.527***	0.582***	0.451**	0.727***	0.889***	0.566***
	[0.122]	[0.172]	[0.176]	[0.130]	[0.188]	[0.176]
Female	0.182***	0.208***	0.159***	0.180***	0.193***	0.176***
	[0.0305]	[0.0400]	[0.0471]	[0.0285]	[0.0373]	[0.0452]
Black	-0.269***	-0.265***	-0.214**	-0.330***	-0.341***	-0.238**
	[0.0661]	[0.0896]	[0.0974]	[0.0685]	[0.0889]	[0.105]
Free Lunch Recipient	-0.462***	-0.509***	-0.357***	-0.397***	-0.432***	-0.322***
	[0.0337]	[0.0400]	[0.0601]	[0.0346]	[0.0409]	[0.0644]
Age	-0.159	-0.211	-0.099	-0.162	-0.231	-0.126
	[0.123]	[0.178]	[0.161]	[0.123]	[0.179]	[0.162]
Proportion of Females in Classroom	0.351	0.006	1.095***	0.688***	0.496*	1.178***
	[0.238]	[0.271]	[0.369]	[0.238]	[0.289]	[0.378]
Proportion of Blacks in Classroom	-0.740	-1.429**	1.572**	0.019	-0.430	2.255**
	[0.647]	[0.670]	[0.753]	[0.635]	[0.652]	[0.866]
Proportion of Free Lunch Recipients in Classroom	-0.031	-0.320	0.535	-0.127	-0.518	0.676
	[0.219]	[0.227]	[0.417]	[0.264]	[0.315]	[0.414]
Average Age of Classmates	0.413*	0.443	0.140	0.736***	1.094**	0.168
	[0.223]	[0.402]	[0.401]	[0.259]	[0.420]	[0.374]
Regular Size Classroom w/ Aide Treatment	0.038	-0.010	0.100	0.012	0.024	0.009
	[0.0395]	[0.0421]	[0.0725]	[0.0415]	[0.0466]	[0.0697]
Class Size	-0.020	-0.041	0.011	0.003	-0.0467*	0.0530*
	[0.0247]	[0.0277]	[0.0317]	[0.0208]	[0.0241]	[0.0280]
New Teacher	-0.087	-0.134	0.017	-0.483***	-0.377*	-0.473**
	[0.161]	[0.189]	[0.226]	[0.151]	[0.220]	[0.196]
Teacher Experience in Years	0.0129**	0.002	0.0167**	0.008	0.002	0.006
	[0.00616]	[0.00706]	[0.00823]	[0.00558]	[0.00670]	[0.00719]
Teacher Experience Missing	-0.087	-0.285		-0.268	-0.794***	
	[0.224]	[0.235]		[0.223]	[0.269]	
Teacher has at least a Master's Degree	0.049	-0.004	0.131*	-0.028	-0.111	0.105
	[0.0631]	[0.0819]	[0.0750]	[0.0580]	[0.0680]	[0.0876]
Teacher on Career Ladder	-0.099	-0.019	-0.159	-0.232*	-0.150	-0.227
	[0.164]	[0.153]	[0.216]	[0.123]	[0.151]	[0.158]
Teacher Missing Career Ladder Info	-0.058	0.088	-0.030	-0.071	0.192	-0.103
	[0.140]	[0.173]	[0.175]	[0.147]	[0.177]	[0.173]
Observations	4,035	2,268	1,767	4,094	2,285	1,809
R-squared	0.31	0.30	0.34	0.33	0.29	0.40

Robust standard errors, clusted by classroom in brackets. Model includes school fixed effects.
*** p<0.01, ** p<0.05, * p<0.1

Table IV: Estimates of Non-Linear Peer Effects (Center City Schools)

Variables	All Reading (1)	All Math (2)	Restricted Sample Reading (3)	Restricted Sample Math (4)
Linear Specification:				
Proportion of Repeaters in Class	1.418	1.501	1.187	1.877
	[1.133]	[1.170]	[1.757]	[1.770]
Non-linear Specification:				
1 Repeater in Classroom	0.160*	0.176*	0.157*	0.137
	[0.0884]	[0.0986]	[0.0924]	[0.0987]
2 Repeaters in Classroom	0.141	0.164	-0.019	0.119
	[0.170]	[0.163]	[0.209]	[0.193]
3 Repeaters in Classroom	-0.380	-0.529**		
	[0.265]	[0.249]		
4 Repeaters in Classroom	0.169	0.144		
	[0.185]	[0.186]		
Joint Significance Test P-values:				
H0: All coefficients equal zero	0.002	0.000	0.226	0.380
H0: Effect of # Repeaters Linear	0.005	0.000	0.195	0.489
H0: Effect of 1 = … = Effect of n	0.012	0.001	0.413	0.919
Observations	1,767	1,809	1,553	1,592

Notes: Each set of estimates is from a separate regression. Robust standard errors in brackets, clustered by classroom. All models include school fixed effects, as well as controls for individual characteristics, peer characteristics, and teacher characteristics. The corresponding coefficient estimates for these controls are omitted from the table.
*** p<0.01, ** p<0.05, * p<0.1

Table V: Estimates of Non-Linear Peer Effects (Other Project STAR Schools)

Variables	All		Restricted Sample	
	Reading	Math	Reading	Math
	(1)	(2)	(3)	(4)
Linear Specification:				
Proportion of Repeaters in Class	-0.913	-1.867***	-0.719	-2.037***
	[0.592]	[0.592]	[0.732]	[0.728]
Non-linear Specification:				
1 Repeater in Classroom	-0.152**	-0.234***	-0.145**	-0.212***
	[0.0680]	[0.0679]	[0.0704]	[0.0698]
2 Repeaters in Classroom	0.036	-0.074	0.048	-0.065
	[0.0947]	[0.0952]	[0.0960]	[0.0957]
3 Repeaters in Classroom	-0.326***	-0.425***	-0.304**	-0.523***
	[0.0965]	[0.132]	[0.117]	[0.141]
4 Repeaters in Classroom	-0.122	-0.429***		
	[0.0995]	[0.145]		
5 Repeaters in Classroom	-1.209***	-1.182***		
	[0.237]	[0.312]		
Joint Significance Test P-values:				
H0: All coefficients equal zero	0.000	0.000	0.007	0.001
H0: Effect of # Repeaters Linear	0.000	0.036	0.011	0.025
H0: Effect of 1 = … = Effect of n	0.000	0.007	0.013	0.014
Observations	2,268	2,285	2,037	2,053

Notes: Each set of estimates is from a separate regression. Robust standard errors in brackets, clustered by classroom. All models include school fixed effects, as well as controls for individual characteristics, peer characteristics, and teacher characteristics. The corresponding coefficient estimates for these controls are omitted from the table.
*** p<0.01, ** p<0.05, * p<0.1

Table VI: Estimates of Non-Linear Peer Effects by Socio-economic Status (Center City Schools)

Variables	All Reading (1)	All Math (2)	Restricted Sample Reading (3)	Restricted Sample Math (4)
Poor – Linear Specification:				
Proportion of Repeaters in Class	3.946***	4.032***	4.486***	4.348**
	[1.475]	[1.351]	[1.689]	[1.683]
Poor – Non-linear Specification:				
1 Repeater in Classroom	0.228**	0.210**	0.203**	0.192*
	[0.0941]	[0.0988]	[0.0923]	[0.0994]
2 Repeaters in Classroom	0.437**	0.447***	0.453**	0.434**
	[0.187]	[0.159]	[0.196]	[0.169]
3 Repeaters in Classroom	-0.536	-0.507		
	[0.370]	[0.327]		
4 Repeaters in Classroom	0.369	0.406*		
	[0.318]	[0.220]		
Joint Significance Test P-values:				
H0: All coefficients equal zero	0.000	0.000	0.027	0.028
H0: Effect of # Repeaters Linear	0.001	0.002	0.829	0.804
H0: Effect of 1 = … = Effect of n	0.000	0.000	0.182	0.125
Observations	1,055	1,074	1,024	1,043
Non-poor – Linear Specification:				
Proportion of Repeaters in Class	-0.130	-1.296	-2.971*	-1.786
	[1.118]	[1.106]	[1.607]	[1.893]
Non-poor – Non-linear Specification:				
1 Repeater in Classroom	0.027	0.032	0.090	-0.100
	[0.103]	[0.129]	[0.145]	[0.163]
2 Repeaters in Classroom	-0.225	-0.180	-0.464**	-0.151
	[0.149]	[0.190]	[0.182]	[0.237]
3 Repeaters in Classroom	-1.051***	-1.030***		
	[0.302]	[0.330]		
4 Repeaters in Classroom	0.051	-0.227		
	[0.186]	[0.202]		
Joint Significance Test P-values:				
H0: All coefficients equal zero	0.000	0.002	0.043	0.654
H0: Effect of # Repeaters Linear	0.000	0.001	0.078	0.910
H0: Effect of 1 = … = Effect of n	0.000	0.003	0.027	0.863
Observations	712	735	529	549

Notes: Each set of estimates is from a separate regression. Robust standard errors in brackets, clustered by classroom. All models include school fixed effects, as well as controls for individual, peer, and teacher characteristics. The corresponding coefficient estimates for these controls are omitted from the table. *** p<0.01, ** p<0.05, * p<0.1

Table VII: Estimates of Non-Linear Peer Effects by Socio-economic Status (Other Project STAR Schools)

Variables	All Reading (1)	All Math (2)	Restricted Sample Reading (3)	Restricted Sample Math (4)
Poor – Linear Specification:				
Proportion of Repeaters in Class	-2.213***	-3.199***	-3.441***	-5.163***
	[0.635]	[0.826]	[0.871]	[0.953]
Poor – Non-linear Specification:				
1 Repeater in Classroom	-0.376***	-0.450***	-0.395***	-0.430***
	[0.0806]	[0.0739]	[0.0820]	[0.0744]
2 Repeaters in Classroom	-0.211**	-0.367***	-0.211**	-0.354***
	[0.0916]	[0.101]	[0.0940]	[0.103]
3 Repeaters in Classroom	-0.595***	-0.760***	-0.605***	-0.897***
	[0.127]	[0.167]	[0.143]	[0.171]
4 Repeaters in Classroom	-0.378**	-0.482***		
	[0.168]	[0.149]		
5 Repeaters in Classroom	-0.925***	-1.443***		
	[0.248]	[0.363]		
Joint Significance Test P-values:				
H0: All coefficients equal zero	0.000	0.000	0.000	0.000
H0: Effect of # Repeaters Linear	0.000	0.000	0.000	0.001
H0: Effect of 1 = ... = Effect of n	0.019	0.064	0.007	0.006
Observations	913	919	793	798
Non-poor – Linear Specification:				
Proportion of Repeaters in Class	-0.167	-0.932	0.794	-0.237
	[0.935]	[0.752]	[1.027]	[0.909]
Non-poor – Non-linear Specification:				
1 Repeater in Classroom	-0.019	-0.090	-0.006	-0.066
	[0.0846]	[0.0873]	[0.0854]	[0.0901]
2 Repeaters in Classroom	0.198	0.108	0.209	0.111
	[0.129]	[0.122]	[0.129]	[0.122]
3 Repeaters in Classroom	-0.178	-0.210	-0.150	-0.301*
	[0.132]	[0.147]	[0.148]	[0.159]
4 Repeaters in Classroom	-0.098	-0.509***		
	[0.118]	[0.153]		
5 Repeaters in Classroom	-1.524***	-1.130***		
	[0.344]	[0.360]		
Joint Significance Test P-values:				
H0: All coefficients equal zero	0.000	0.001	0.058	0.123
H0: Effect of # Repeaters Linear	0.000	0.017	0.025	0.082
H0: Effect of 1 = ... = Effect of n	0.000	0.001	0.026	0.062
Observations	1,355	1,366	1,244	1,255

Notes: Each set of estimates is from a separate regression. Robust standard errors in brackets, clustered by classroom. All models include school fixed effects, as well as controls for individual, peer, and teacher characteristics. The corresponding coefficient estimates for these controls are omitted from the table. *** p<0.01, ** p<0.05, * p<0.1

Table VIII: Estimates of Non-Linear Peer Effects by Gender (Center City Schools)

Variables	All		Restricted Sample	
	Reading	Math	Reading	Math
	(1)	(2)	(3)	(4)
Males – Linear Specification:				
Proportion of Repeaters in Class	2.905**	3.328***	4.315**	5.159***
	[1.215]	[1.026]	[1.683]	[1.528]
Males – Non-linear Specification:				
1 Repeater in Classroom	0.400***	0.372***	0.393***	0.335***
	[0.0886]	[0.0876]	[0.0902]	[0.0879]
2 Repeaters in Classroom	0.130	0.235*	0.138	0.368***
	[0.148]	[0.140]	[0.169]	[0.132]
3 Repeaters in Classroom	-0.210	-0.151		
	[0.281]	[0.218]		
4 Repeaters in Classroom	0.439**	0.500***		
	[0.204]	[0.145]		
Joint Significance Test P-values:				
H0: All coefficients equal zero	0.000	0.000	0.000	0.001
H0: Effect of # Repeaters Linear	0.000	0.000	0.007	0.064
H0: Effect of 1 = … = Effect of n	0.000	0.000	0.165	0.785
Observations	896	919	778	798
Females – Linear Specification:				
Proportion of Repeaters in Class	0.243	-0.404	-1.476	-1.253
	[1.497]	[1.559]	[1.951]	[1.962]
Females – Non-linear Specification:				
1 Repeater in Classroom	-0.075	-0.026	-0.067	-0.061
	[0.104]	[0.118]	[0.106]	[0.116]
2 Repeaters in Classroom	0.210	0.115	-0.132	-0.117
	[0.253]	[0.231]	[0.256]	[0.249]
3 Repeaters in Classroom	-0.479	-0.920**		
	[0.371]	[0.366]		
4 Repeaters in Classroom	-0.048	-0.244		
	[0.252]	[0.276]		
Joint Significance Test P-values:				
H0: All coefficients equal zero	0.346	0.027	0.754	0.805
H0: Effect of # Repeaters Linear	0.234	0.013	0.998	0.987
H0: Effect of 1 = … = Effect of n	0.221	0.016	0.801	0.829
Observations	871	890	775	794

Notes: Each set of estimates is from a separate regression. Robust standard errors in brackets, clustered by classroom. All models include school fixed effects, as well as controls for individual, peer, and teacher characteristics. The corresponding coefficient estimates for these controls are omitted from the table. *** p<0.01, ** p<0.05, * p<0.1

Table IX: Estimates of Non-Linear Peer Effects by Gender (Other Project STAR Schools)

Variables	All Reading (1)	All Math (2)	Restricted Sample Reading (3)	Restricted Sample Math (4)
Males – Linear Specification:				
Proportion of Repeaters in Class	-2.066***	-2.322***	-1.752**	-2.857***
	[0.769]	[0.747]	[0.870]	[0.906]
Males – Non-linear Specification:				
1 Repeater in Classroom	-0.239***	-0.244***	-0.226***	-0.225***
	[0.0742]	[0.0777]	[0.0773]	[0.0793]
2 Repeaters in Classroom	-0.079	-0.203*	-0.061	-0.192
	[0.110]	[0.118]	[0.110]	[0.118]
3 Repeaters in Classroom	-0.518***	-0.501***	-0.414***	-0.566***
	[0.129]	[0.148]	[0.148]	[0.178]
4 Repeaters in Classroom	-0.291**	-0.373**		
	[0.123]	[0.180]		
5 Repeaters in Classroom	-1.619***	-1.137***		
	[0.330]	[0.305]		
Joint Significance Test P-values:				
H0: All coefficients equal zero	0.000	0.001	0.004	0.002
H0: Effect of # Repeaters Linear	0.000	0.158	0.032	0.257
H0: Effect of 1 = ... = Effect of n	0.000	0.045	0.056	0.131
Observations	1,173	1,183	1,054	1,064
Females – Linear Specification:				
Proportion of Repeaters in Class	0.514	-1.020	0.593	-0.737
	[0.729]	[0.722]	[0.982]	[0.904]
Females – Non-linear Specification:				
1 Repeater in Classroom	-0.041	-0.173**	-0.039	-0.152*
	[0.0831]	[0.0812]	[0.0853]	[0.0831]
2 Repeaters in Classroom	0.174	0.114	0.184	0.117
	[0.131]	[0.115]	[0.133]	[0.116]
3 Repeaters in Classroom	-0.092	-0.285*	-0.160	-0.440***
	[0.126]	[0.157]	[0.141]	[0.153]
4 Repeaters in Classroom	0.104	-0.380**		
	[0.149]	[0.163]		
5 Repeaters in Classroom	-0.626**	-1.328***		
	[0.258]	[0.340]		
Joint Significance Test P-values:				
H0: All coefficients equal zero	0.042	0.001	0.132	0.007
H0: Effect of # Repeaters Linear	0.022	0.002	0.069	0.007
H0: Effect of 1 = ... = Effect of n	0.022	0.001	0.063	0.004
Observations	1,095	1,102	983	989

Notes: Each set of estimates is from a separate regression. Robust standard errors in brackets, clustered by classroom. All models include school fixed effects, as well as controls for individual, peer, and teacher characteristics. The corresponding coefficient estimates for these controls are omitted from the table. *** p<0.01, ** p<0.05, * p<0.1