Making the Grade:
The Impacts of Classroom Disruption and Class Size on Academic Achievement

Kalena E. Cortes*
Texas A&M University
kcortes@bushschool.tamu.edu

Wael S. Moussa
Syracuse University
wsmoussa@maxwell.syr.edu

Jeffrey M. Weinstein
Syracuse University
jmwein04@maxwell.syr.edu

This version: January 20, 2012

Abstract

This paper examines empirical implications of Lazear's (2001) educational production model. Using exogenous variation on course scheduling for 9th graders in Chicago Public Schools and policy-driven variation in class size, we study heterogeneity in the impact of class size on student achievement. Our research design allows us to isolate overall class size effects from an underlying mechanism by which class size affects student performance, classroom disruption. Our empirical results are consistent with the Lazear framework. For black students, we estimate that an additional non-disruptive student in attendance increases the probability of passing regular Algebra I by 2.98 percentage points, while an additional student enrolled in the class decreases this probability by 0.96 percentage points. The estimated effects are larger in magnitude for regular English I. For regular Algebra I, we also observe that an additional non-disruptive student in attendance increases own mathematics test score by 0.0249 student-level standard deviations, while an additional enrolled student decreases this score by 0.0431 student-level standard deviations. Again in line with the Lazear model, the estimated impacts on course passing are larger in magnitude for remedial classes.

JEL Classifications: H75, I21, J13
Keywords: Educational Production, Classroom Composition, Class Size, Student Achievement

---

* We are indebted to Sue Sporte, Associate Director for Evaluation and Data Resources, the Consortium on Chicago School Research (CCSR), University of Chicago, and Peter Godard, Director of Data Management in the Office of Research, Evaluation and Accountability, Chicago Public Schools, for making the data available for this project. We are thankful to the Maxwell School of Citizenship and Public Affairs at Syracuse University for providing an internal grant in support of this project. Institutional support from the Bush School of Government and Public Service at Texas A&M University and the Center of Policy Research at the Maxwell School of Citizenship and Public Affairs is also gratefully acknowledged. Research results, conclusions, and all errors are naturally our own.

I. Introduction

Lazear (2001) posits an elegant theoretical model of class size, in which students enrolled in smaller classes learn more because these students experience fewer student disruptions during their class instruction. The Lazear framework hypothesizes that the mechanism behind the effect of class size on achievement is classroom disruptions, whereby adding more students to a classroom increases the number of disruptions and consequently decreases the amount of time during which learning can take place because a teacher's time is spent dealing with the students who are causing the disruptions. In other words, classroom education inherently has properties of a *public good*, in that if one student disrupts his or her class, the learning of all other students within the class is also harmed. Recent evidence suggests that there is considerable variation across students in the propensity to disrupt class and that this propensity is correlated with measurable student-level characteristics, such as socioeconomic status (Segal 2008).

Our paper examines empirical implications of the Lazear educational production model. Using exogenous variation on course scheduling for 9th graders in Chicago Public Schools (CPS) and policy-driven variation in class size, we study heterogeneity in the impact of class size on student achievement in mathematics and reading. Our identification strategy allows us to isolate overall class size effects from an underlying mechanism by which class size affects student performance, classroom composition. Our measure of classroom composition is constructed as the average number of non-disruptive students in attendance on a given school day; we assume that the probability of classroom disruption is, on average, higher for students with the largest propensity to drop out of school, as measured by whether the student drops out of high school in any grade.

As one of the largest urban public school districts in the United States, currently serving over 400,000 students, CPS provides a unique opportunity for analyzing heterogeneity in class size effects for a population of predominantly racial/ethnic minority students that are largely from lower-income families. Approximately 85 percent of CPS students receive federal lunch subsidies, and the racial/ethnic makeup of the student body is approximately 55 percent black; 35 percent Hispanic; and the remaining 10 percent white, American Indian, or Asian.

Our data are taken from CPS administrative student transcript files, which include the ordering of classes over the day, student absences, course titles, grades, scores from standardized tests in mathematics and reading, and demographic characteristics for the universe of CPS high school students from the 1993-94 to the 2005-06 school years. To study the effects of classroom composition and class size on academic achievement, we use an instrument variables approach, exploiting exogenous variation in the period of the day a course is offered and class size caps adopted by the CPS Board of Education. Our analysis focuses on students' course passing and test scores in four 9th grade course subjects: regular Algebra I, remedial Algebra I, regular English I, and remedial English I.

The Lazear theoretical model of education production suggests that classroom composition and class size are interdependent, implying that both variables should be included as explanatory variables when estimating an educational production function. That is, the behavioral composition of a classroom is influenced by the size of the class, and the number of students ultimately assigned to a classroom is endogenously determined by the composition of the potential pool of students. Our two-stage least squares (2SLS) results are in line with the Lazear framework, as we find that classroom composition and class size are both important determinants of student achievement. Specifically, our 2SLS estimates indicate that an additional non-disruptive student in attendance on a given school day increases the probability that a black student passes regular Algebra I by 2.98 percentage points, relative to the mean passing rate of 72.2 percent, while an additional student enrolled in the class decreases this probability by 0.96 percentage points (the latter effect is statistically insignificant). For black students enrolled in regular Algebra I, we also observe that an additional non-disruptive student in attendance increases own mathematics test score by 0.0249 student-level standard deviations, while an additional enrolled student decreases this test score by 0.0431 student-level standard deviations.

The aforementioned effect size of 0.0431 student-level standard deviations for a one-student reduction in class size is similar in magnitude to the effect sizes reported in Finn and Achilles (1990) for 1st grade minority students that participated in Tennessee's Project STAR class-size randomized experiment. For the Stanford Achievement Test (SAT) and the Basic Skills First (BSF) test in mathematics, the effect sizes estimated by Finn and Achilles (1990) are 0.31 and 0.16 student-level standard deviations, respectively, for an approximately eight-student reduction in class size (Table 6, page 567). These translate into effect sizes of 0.0388 and 0.0200 student-level standard deviations, respectively, for a one-student reduction in class size. The effect size based on the SAT is almost identical to our estimated effect size, while the effect size based on the BSF test is about half the size of our estimated effect size.

The estimated impact of classroom composition on course passing for black students enrolled in regular English I is larger than the corresponding effect for regular Algebra I. We find that an additional non-disruptive student in attendance increases the probability that a black student passes regular English I by 4.77 percentage points, relative to the mean passing rate of 77.3 percent, while an additional enrolled student decreases this probability by 2.77 percentage points.

The Lazear model also suggests that the impacts of classroom composition and class size on student outcomes should be larger for students enrolled in remedial versus regular courses because the baseline academic performance of students in remedial classes is lower, on average, than that of students in regular classes. Consistent with this, we find larger effects of classroom composition and class size on Algebra I and English I course passing for black students enrolled in remedial versus regular classes. For remedial Algebra I, our 2SLS results show that an additional non-disruptive student in attendance increases the probability that a black student passes this course by 4.79 percentage points, relative to the mean passing rate of 68.4 percent, while an additional enrolled student decreases this probability by 6.33 percentage points. The estimated impacts on course passing are larger for remedial English I; we observe that an additional non-disruptive student in attendance increases the probability that a black student passes this course by 6.60 percentage points, relative to the mean passing rate of 73.5 percent, while an additional enrolled student decreases this probability by 6.08 percentage points.

For Hispanic students, we find that an additional non-disruptive student in attendance leads to a 5.37 percentage point increase in the probability of passing remedial English I, relative to the mean passing rate of 78.5 percent, while an additional enrolled student leads to a 5.16 percentage point decrease in this probability.

The remainder of this paper is organized as follows. Section II overviews the empirical literature on class size effects and mechanisms. Section III presents our empirical specification, discussing the strategy we use to identify the effects of classroom composition and class size on student achievement. Section IV describes our data, providing descriptive statistics and empirical support for our instrumental variables estimation strategy. Section V reports our empirical findings and compares our estimated class size effects to those in related studies. Section VI concludes.

II. Overview of Empirical Literature on Class Size Effects and Mechanisms

Identifying the causal impact of class size on student attainment is difficult in observational studies due to nonrandom sorting of students across schools and classrooms by students, parents, teachers, and administrators, as well as heterogeneity in financial and educational resources. As a result, studies that estimate class size effects have generally used experimental or quasi-experimental research designs. For example, many papers have used data from Tennessee's Project STAR class-size randomized experiment to examine the effect of smaller class sizes on student achievement, whereby students and teachers in participating elementary schools were randomly assigned to one of three class types: small (13-17 student) classes, regular (22-25 student) classes, and regular classes with a teacher aide. Finn and Achilles (1990), Word et al. (1990), Krueger (1999), Nye, Hedges, and Konstantopoulos (1999, 2000), Finn, Gerber, Achilles, and Boyd-Zaharias (2001), Krueger and Whitmore (2001), and McKee, Rivkin, and Sims (2010) find statistically significant effects of attending a smaller class on student achievement and educational attainment.

Other work has used quasi-experimental research, isolating plausibly exogenous variation in class sizes in earlier grades (elementary and/or middle) from non-linear relationships between enrollment and class sizes (class-size rules) in a regression discontinuity design framework and/or idiosyncratic population compositions due to random variation in the timing of births. Such studies have been conducted using data from Israel (Angrist and Lavy (1999)), Connecticut (Hoxby (2000)), Texas ((Rivkin, Hanushek, and Kain (2005)), and California (Babcock and Betts (2009) and Jepsen and Rivkin (2009)). Other than Hoxby (2000), this research finds statistically significant impacts of smaller class sizes on student outcomes primarily for elementary school students.[1] While our main focus is on the effect of classroom composition on student achievement, to control for the overall effect of class size on outcomes, we use a similar identification strategy as in, for example, Angrist and Lavy (1999), exploiting caps in class sizes for high school students in CPS.[2]

[1] Many of the studies above have also examined whether the impact of attending a larger class is heterogeneous across student demographics. They generally find evidence of heterogeneity, with larger class size effects for black and lower-income students. Using quantile regression analysis, other research has looked at whether there is heterogeneity across the distribution of prior student achievement. Three non-experimental studies, Eide and Showalter (1998) with class size data from the United States, and Levin (2001) and Ma and Koenker (2006) with data from The Netherlands, find little to no heterogeneity across the prior achievement distribution in the benefits of attending a smaller class. However, Konstantopoulos (2008), McKee, Rivkin, and Sims (2010), and Ding and Lehrer (2011), which use data from Tennessee's Project STAR class-size randomized experiment, do find that smaller class sizes yield larger benefits to students with higher past achievement.

[2] This is also similar to the instrumental variables approach taken by Angrist and Lang (2004) in their study of the effects of the Metco desegregation program in Boston Public Schools on students in a suburban public school district (Brookline Public Schools) that received Metco students.

A recent study, McKee, Rivkin, and Sims (2010), extends Lazear's (2001) theoretical framework with the goal of empirically investigating heterogeneity in class size effects by income and prior achievement. As with Lazear (2001), McKee, Rivkin, and Sims (2010) assume that the amount of time available for teaching depends on the level of classroom disruption, implying that class size effects are largest in classrooms with students that have higher propensities to disrupt. This would lead to larger benefits of reduced class size in poorer schools if the likelihood of disruption were larger at the lower end of the income distribution. The amount of time available for learning also depends on the quality of learning, which is itself a function of baseline academic achievement and class size. The authors then discuss how smaller class sizes may be more or less beneficial to higher-achieving versus lower-achieving students, concluding that the magnitude of class size effects across the achievement distribution is ambiguous. Using data from Tennessee's Project STAR class-size randomized experiment, the authors empirically test the predictions of their model, finding that greater benefits from reduced class sizes accrue to students with higher baseline achievement, as well as to students in lower-income schools.

Babcock and Betts (2009) also examine mechanisms, investigating whether class size effects for elementary school students in San Diego vary depending on two separately identified student classifications: baseline student effort, as measured by teacher assessments of students' conduct in the classroom, and baseline achievement, as calculated by letter grades in academic subjects. Exogenous variation in class size follows from a state policy that legislatively lowered class sizes in kindergarten through third grade only, allowing the authors to study the impact of class size on test scores using the transition from third to fourth grade. The results indicate that class size effects are larger for students with lower baseline effort, consistent with an implication of the Lazear model that students with more behavioral problems benefit more from smaller classes, while there is no evidence of heterogeneity across high- and low-achieving students.

III. Empirical Strategy

We model the effects of classroom composition and class size on course passing and test scores in four 9th grade CPS course subjects: regular Algebra I, remedial Algebra I, regular English I, and remedial English I. We define a classroom to be composed of two types of students, disruptive and non-disruptive, where a student is defined as disruptive if he or she ever left high school due to one of the following reasons reported in CPS administrative records: legally committed to a correctional facility, lost (truant officer cannot locate), and uniform discipline code violation (infringement of the CPS code of conduct). Consider the following linear specification:

(1) $Y_{ict} = \beta_0 + \beta_1 Composition_{ct} + \beta_2 Size_{ct} + \delta' X_{it} + \lambda Fall_{ct} + \eta_{jkt} + \varepsilon_{ict}$,

where Y_{ict} denotes one of two outcome variables for student *i* enrolled in class *c* in academic year *t*: the receipt of a grade of D or better in a particular course or the test score in a subject-relevant (mathematics or reading) standardized exam.[3] The main explanatory variable of interest is $Composition_{ct}$, which denotes the average number of non-disruptive students in attendance on a given school day in the classroom that student *i* is enrolled in:

(2) $Composition_{ct} = \frac{1}{90}\sum_{i \in ct}(90 - Absences_{ict})(1 - Disruptive_i)$,

where $Disruptive_i$ is an indicator for whether student *i* is disruptive, and $Absences_{ict}$ is the number of days student *i* was absent in a particular semester in year *t* for class *c*. With 90 school days in each semester, $(90 - Absences_{ict})$ is the number of days student *i* was in attendance in a particular semester in year *t* for class *c*. This variable is then summed over every non-disruptive student in year *t* for class *c*. The sum is divided by 90 to obtain the daily average number of non-disruptive students in attendance in class *c* in time *t*, $Composition_{ct}$.

[3] An alternative specification adds to equation (1) an interaction term between classroom composition and class size. Interpreting the coefficient on such an interaction term is difficult because classroom composition and class size are codetermined, as discussed earlier. Hence, we do not include this interaction term in our estimating equations.

$Size_{ct}$ is the roster class size for a given classroom, which is simply the number of students enrolled in that class. X_{it} is a vector of observable student-specific characteristics, which includes the subject-specific 8th grade Iowa Test of Basic Skills (ITBS) test score measured in student-level standard deviations, demographic variables, and neighborhood variables measured at the census block level. Also included is an indicator variable for whether the course is taken in the fall semester, $Fall_{ct}$. η_{jkt} represents our fixed effects, where subscript *j* is for a high school or a high school teacher, and *k* is for a middle school. Four different sets of fixed effects are used to capture time-variant unobserved high school (school attended in 9th grade), middle school (school attended in 8th grade), and/or high school teacher (9th grade teacher) quality. Specifically, we include, in separate specifications, high school-by-year fixed effects, middle school-by-high school-by-year fixed effects, teacher-by-year fixed effects, or middle school-by-teacher-by-year fixed effects. Finally, ε_{ict} represents the idiosyncratic error term.

Ordinary Least Squares (OLS) estimation of equation (1) leads to biased estimates of the effects of classroom composition and class size on student achievement because our main explanatory variable of interest, $Composition_{ct}$, is a function of the number of absences, which are not randomly assigned across students. For example, number of absences is negatively correlated with prior student achievement when students with lower past achievement have higher probabilities of being absent on a given school day.[4] This implies that the coefficient on classroom composition from an OLS regression has a downward bias. Additionally, roster class size, $Size_{ct}$, is endogenous due to, for instance, school administrators' grouping students with lower past achievement into

[4] Our data lend support to this hypothesis. For students enrolled in regular Algebra I, the raw correlation between 8th grade ITBS (baseline) mathematics test score and number of 9th grade absences (averaged across all periods) is -0.1857. For students enrolled in regular English I, the raw correlation between 8th grade ITBS reading test score and number of 9th grade absences is -0.2087. The corresponding raw correlations for students enrolled in remedial Algebra I and remedial English I are -0.1127 and -0.1687, respectively.

smaller classes. This causes the coefficient on class size from an OLS regression to have an upward bias.

To eliminate these biases, and hence to estimate the causal effects of classroom composition and class size on student outcomes, we use an instrumental variables regression framework, where $Composition_{ct}$ and $Size_{ct}$ are endogenous variables. One set of excluded instrumental variables is $Period_{ct}$, a vector of indicator variables for the period of the school day in which a class is scheduled; first period is the omitted category, and second through seventh periods are indicator variables. This identification strategy is similar to the one used by Cortes, Bricker, and Rohlfs (forthcoming), in which they exploit variation in the ordering of classes over the day to measure how the returns to classroom learning vary by course subject and how attendance in one class spills over into learning in other subjects. As we show later, student absences in a particular class vary depending on the period of the day in which the class is offered, implying that $Period_{ct}$ is correlated with the $Composition_{ct}$ variable. In addition, we would expect that the different periods of the day affect course passing and test scores only through classroom composition.

Furthermore, the CPS Board of Education caps class sizes at a maximum of 28 students for regular classes and 25 students for remedial classes.[5] These policies determine the minimum number of classrooms into which students are assigned; for instance, students in schools with regular Algebra I enrollment of one to 28 students would be placed into one or more classrooms, students in schools with regular Algebra I enrollment of 29 to 56 students would be split into two or more classrooms, and so on. To create an instrument for roster class size, we employ the regression discontinuity design approach taken by, for example, Angrist and Lavy (1999). Specifically, we use observed course enrollment to predict class size according to the CPS policy cutoffs of 28 students for regular classes and 25 students for remedial classes:

[5] Chicago Public Schools Policy Manual Board Report 95-0814-PO1 (August 14, 1995).

(3) $\widehat{Size}_{ct} = E_{jt} / \left[int\left(\frac{E_{jt}-1}{Cutoff}\right) + 1 \right]$,

where E_{jt} is the total number of students enrolled in a given course subject at a given high school in a given semester in a given school year, $Cutoff$ is either 28 or 25, and $int(x)$ gives the largest integer less than or equal to *x* (i.e., the floor function).[6] In this case, equation (1) is now the second-stage equation, and the two first-stage equations are expressed as follows:

(4) $Composition_{ct} = \alpha_0 + \alpha_1 \widehat{Size}_{ct} + \sum_{p=2}^{7} \alpha_p Period_{ct} + \theta' X_{it} + \pi Fall_{ct} + \eta_{jkt} + \upsilon_{ict}$

(5) $Size_{ct} = \alpha_0 + \tilde{\alpha}_1 \widehat{Size}_{ct} + \sum_{p=2}^{7} \tilde{\alpha}_p Period_{ct} + \tilde{\theta}' X_{it} + \tilde{\pi} Fall_{ct} + \eta_{jkt} + \xi_{ict}$,

where the endogenous variables, $Composition_{ct}$ and $Size_{ct}$, are functions of the two excluded sets of instruments, as well as the control variables and fixed effects that appear in equation (1).

IV. Data and Sample Characteristics

A. Data Source

The data for this study come from CPS administrative student records. Our data cover the universe of 9th grade students in CPS from the 1993-94 to the 2005-06 school years. We link each student's record to his or her individual transcript file. The transcript data include course title and number, period of the day, absences by class period, and unique teacher identifiers for each class taken by students. The CPS data also include multiple standardized test scores, a detailed set of descriptive variables about each student, and 1990 and 2000 neighborhood characteristics for the Census block in which each student resides.

The standardized tests that were administered, and the grades of the students who took them, vary from year to year in our sample. Consequently, the samples for the test score regressions are smaller than the samples for the course passing regressions. For the majority of students, 8th

[6] For example, if 51 students are enrolled in remedial Algebra I at a given high school in a given semester in a given school year, then predicted class size equals 17 students because three teachers are required to respect the class size cap of 25 students for remedial classes.

grade mathematics and reading scores are available from the Iowa Test of Basic Skills (ITBS) for each year of our data. The 9th grade test score data for mathematics and reading/English are taken from the TAP (Test of Achievement and Proficiency) or the EXPLORE test. To compare observations from different years in our sample, each test score is converted into a z-score, whereby each student's raw test score is standardized using the mean and standard deviation across all students in CPS that took the relevant examination in a given year.

B. Descriptive Statistics

The summary statistics of the analytic samples are represented in Table 1.[7] The student-level outcome variables of interest are an indicator for passing regular or remedial Algebra I, an indicator for passing regular or remedial English I, the score on the standardized mathematics exam, and the score on the standardized reading exam. In accordance with CPS policy, we defined a student as having passed a course if she received a grade of D or better in that course.[8] Students enrolled in remedial classes have lower course passing rates in both Algebra I and English I, as compared to students in regular classes. As panel A of Table 1 shows, the passing rates of students enrolled in regular Algebra I and English I are 74 and 79 percent, respectively, while the passing rates of students enrolled in remedial Algebra I and English I are 70 and 76 percent, respectively.

The lower academic performance of students in remedial classes can also be observed in both their mathematics and reading test scores. For example, the average TAP or EXPLORE mathematics test score for students enrolled in remedial Algebra I is -0.430 student-level standard deviations, and the average reading test score for students taking remedial English I is -0.322 student-level standard deviations. These averages are much lower than the corresponding averages

[7] Our sample consists of all first-time 9th grade students in CPS from the 1993-94 to the 2005-06 school years. For consistency, if a student that repeated 9th grade one or more times, we only use her first instance of 9th grade in our data.
[8] Chicago Public Schools Policy Manual Board Report 04-0128-PO1 (January 28, 2004).

for students enrolled in regular classes; the average TAP or EXPLORE mathematics test score for students enrolled in regular Algebra I is 0.069 student-level standard deviations, and the average reading test score for students taking regular English I is 0.085 student-level standard deviations.

This difference in academic performance is also seen in the students' baseline performance, as measured by 8th grade ITBS mathematics and reading test scores, which is shown in the last two rows of panel C. The average baseline mathematics test score for students taking remedial Algebra I is -0.543 student-level standard deviations, and the average baseline reading score for students enrolled in remedial English I is -0.340 student-level standard deviations. On the other hand, the average baseline mathematics test score for students taking regular Algebra I is 0.078 student-level standard deviations, and the average baseline reading test score for students enrolled in regular English I is 0.106 student-level standard deviations.

While student achievement is much lower for students enrolled in remedial classes, as compared to students in regular classes, it is important to note that student-specific characteristics do not differ in a systematic manner across observables for the different course subjects. The mean age of 9th graders in all course subjects is approximately 14.3 years, and classes are comprised of 50 percent male. The racial composition is stable across course subjects. Black students account for between 52 and 57 percent of the students enrolled in any given class; Hispanic students account for between 35 and 37 percent; and white, American Indian, and Asian students together account for the remaining eight to 12 percent. Approximately 85 percent of all 9th grade students receive free or reduced lunch, and their proportion across course subjects is fairly stable, ranging between 83 and 88 percent. We find that the proportion of students in special education is higher in remedial classes relative to regular classes (21 percent compared to 14 percent). Lastly, the neighborhood characteristics of a student's residence (as shown in panel D) are similar for both students enrolled in regular classes and for students enrolled in remedial classes.

C. Instrumental Variables

To measure the causal effects of classroom composition and class size on student academic performance, we now make the case for the two proposed sets of excluded instrumental variables: period of the day and predicted class size.

1. Period of the Day

For the first set of instrumental variables, period of the day, we assert that students enrolled in Algebra I and English I in a particular period are otherwise similar to students who take those classes at another time during the day. That is, after students select the courses that they will take in a semester, the ordering of classes over the day is a computerized and essentially random process that is determined based on scheduling constraints. A testable implication of this assumption is that student and classroom characteristics should be similar between classes that meet in first period and those that meet later in the day. Tables 2A and 2B present strong evidence of this premise, lending credibility to the use of differences in course scheduling in CPS as an exogenous source of variation in classroom composition to identify the effect of classroom composition on student achievement.

Table 2A shows, separately by period of the day, the fraction of courses offered in each subject. Though we are only interested in Algebra I and English I courses for this study, it is still instructive to look at all course subjects to validate period of the day as a viable instrument. Table 2A is calculated from unweighted student-level data, and the fractions in each column sum to one, excluding the rows in italics. The numbers in italics denote subgroups of their respective course subjects: regular English I, remedial English I, regular Algebra I, and remedial Algebra I. As Table 2A shows, the breakdown of classes by subject is generally stable over the course of the day, implying that schools do not appear to systematically schedule academic subjects in certain periods, such as those with low absence rates (we return to this point in the next paragraph). Yet even stronger evidence of the validity of using period of the day as an instrument is provided by regressing an indicator for the period of the day that the student took (regular or remedial) Algebra I or English I on all control variables used in our outcome regressions, as well as middle school-by-teacher-by-year fixed effects. Table 2B reports coefficients from such linear probability models for the sample of students enrolled in regular Algebra I; each column is for a different period of the day.[9] Almost all coefficients are statistically insignificant, and the coefficients that are significant show no apparent pattern across model specifications.

We next examine the raw, reduced-form (i.e., first-stage) effects of having Algebra I and English I in first period on the average number of non-disruptive students in attendance in each period. Panels A and B of Figure 1 show the relationship between period of the day and one of the key endogenous variables of our model, classroom composition. Panel A shows the average number of non-disruptive students in attendance for regular Algebra I and English I as functions of the period of the day during which these classes meet. The corresponding graphs for remedial classes are presented in panel B. In both panels, the solid lines show the means for the full (pooled black, Hispanic, white, Asian, and American Indian) sample, the dashed-dotted lines show the means for black students, and the dashed lines show the means for Hispanic students.

As both panels show, the average numbers of non-disruptive students in attendance for Algebra I and English I are at their lowest levels in first period and then gradually rise over the remainder of the day, though not monotonically. This pattern is seen for Hispanic students, for whom the average number of non-disruptive students in attendance varies across periods from 17.6 to 19.6 for regular Algebra I and from 17.7 to 19.0 for regular English I. The pattern is also detectable among blacks and the full sample of students, for whom the average numbers of non-disruptive students in regular Algebra I vary across periods of the day from 15.8 to 16.9 for black students and from 16.7 to 18.4 for the full sample. The corresponding ranges for regular English I

[9] See Appendix Tables A1-A3 for the regression results for the other samples: Appendix Table A1 for regular English I, Appendix Table A2 for remedial Algebra I, and Appendix Table A3 for remedial English I.

are 15.1 to 17.1 for blacks and 16.2 to 18.1 for the full sample. The same pattern is also seen for the remedial Algebra I and remedial English I.

2. Predicted Class Size

For our second instrumental variable, predicted class size, we exploit CPS enrollment policies, which cap class sizes at 28 and 25 students for regular and remedial classes, respectively. Figure 2 graphs the actual (observed) and predicted average class sizes on the vertical axes and course enrollment counts by school and semester on the horizontal axes. Actual average class size is calculated as the mean class size for each enrollment count in the data, and predicted average class size is calculated using the formula given in equation (3). As expected, predicted class size has a piecewise linear relationship with enrollment size, where predicted class size experiences a sharp drop when enrollment is a multiple of 28 for regular Algebra I (or regular English I) and 25 for remedial Algebra I (or remedial English I).

As shown in the line graphs in both panels of Figure 2, for low levels of enrollment actual class size is predicted well using enrollment size. However, we do find certain cases where actual class size exceeds predicted class size. For instance, an enrollment size of 29 students for a regular course should, in theory, predict two classrooms with a mean class size of 14.5 students; however, we sometimes observe one classroom with 29 students. The reason for this is that some schools, in special circumstances, may be granted an exemption from the CPS class size policy rules.[10] Clearly, predicted class size is not the only factor influencing actual class sizes. That said, the graphs in Figure 2 show that only in a few cases do class sizes exceed what the policy rule predicts. Other factors that could determine class sizes are the numbers of available classrooms and teachers, which are accounted for when using high school-by-year and teacher-by-year fixed effects in the regression

[10] Chicago Public Schools Policy Manual Board Report 95-0814-PO1 (August 14, 1995).

specifications. In other words, if a school has more available classrooms and teachers than needed, then students can be divided into smaller class sizes than predicted by the class size rules.

V. Empirical Results and Discussion

In this section we present our empirical analysis of Lazear's (2001) model, which suggests that classroom composition and class size are codependent, implying that both variables should appear as explanatory variables in our estimating equations. Tables 3 and 4 report the OLS and the two-stage least squares (2SLS) regression results for equation (1), which gives the effects of classroom composition and class size on the outcome variables of interest, course passing and test scores. These tables only report the estimated coefficients on classroom composition and class size, as well as the F-statistics for the test of the predictive power of the excluded instruments in both first-stage regressions.

Table 3 presents the results for students in regular Algebra I classes; the dependent variable in panel A is an indicator for whether a student received a passing grade in her Algebra I course, and the dependent variable in panel B is her z-score on the standardized mathematics examination. Table 4 reports the findings for students in regular English I classes; the dependent variable in panel A is an indicator for whether a student received a passing grade in her English I class, and the dependent variable in panel B is her z-score on the standardized reading examination. Within each panel of each table, the topmost set of results is for the black student sample, the next set is for the Hispanic student sample, and the last set of results is for the full (pooled black, Hispanic, white, Asian, and American Indian) sample of students. Because over 50 percent of our sample consists of black students, we pay particular attention to the black student sample when reporting our main results.

Each column within a panel reports the results for a different regression specification. All specifications include student and neighborhood characteristics as baseline controls (listed in the footnotes to Tables 3 and 4, as well as summarized in Table 1); however, different sets of fixed effects are included in columns (1)-(4). Specifically, the column layouts are as follows: column (1) includes high school-by-year fixed effects, column (2) replaces the previous set of fixed effects with middle school-by-high school-by-year fixed effects, column (3) includes only teacher-by-year fixed effects, and column (4) replaces the previous set of fixed effects with middle school-by-teacher-by-year fixed effects. Since teachers have discretion in determining course grades, for the course passing results we focus on the empirical specification in column (4) because it contains teacher fixed effects and controls for unobserved characteristics of the student's middle school, hence making the specification in column (4) more conservative than that in column (3). For the test score results we focus on the empirical specification in column (2) because it contains high school fixed effects and controls for unobserved characteristics of the student's middle school, therefore making the specification in column (2) more conservative than that in column (1).

A. Algebra I Course Passing and Mathematics Test Score Results for Regular Classes

Turning first to the OLS results in panel A of Table 3, we observe in all four columns a positive association between classroom composition and the probability of passing Algebra I regardless of student race or ethnicity. For black students, we find that an additional non-disruptive student in attendance on a given school day is associated with an increase in the probability of passing Algebra I of less than one half of one percentage point (0.21 to 0.36 percentage points). The OLS estimated coefficients for class size are also as expected; we find a negative association between class size and passing Algebra I. For black students, we find that an additional student enrolled in a class is associated with a decrease in the probability of passing Algebra I of less than one percentage point (0.28 to 0.74 percentage points). All OLS results are stable across all model specifications.

Classroom composition and class size results for Hispanic students are larger in magnitude than the estimates for black students.

As discussed in Section III, the endogeneity of classroom composition and class size implies that OLS estimation of equation (1) leads to biased estimates of the effects of classroom composition and class size on student achievement. As a result, we now turn to our 2SLS estimation results. For each regression, we report the F-statistics for the test of the predictive power of the excluded instruments, period of the day and predicted class size, for both first stages. For almost all of the black student sample and full sample specifications, we observe high first-stage F-statistics for the test of the excluded instruments for both the classroom composition and class size first stages. Due to the smaller samples sizes, the F-statistics for the Hispanic student sample are either marginally statistically significant or statistically insignificant.[11]

In general, after instrumenting with period of the day and predicted class size, all of the 2SLS coefficients are larger in magnitude for classroom composition and for class size, as compared to the corresponding OLS coefficients. Many coefficients are statistically significant, showing that classroom composition and class size are both important determinants of student academic outcomes, which is consistent with the Lazear framework.

Recall that the OLS coefficients for classroom composition are expected to have a downward bias, while the OLS coefficients for class size are expected to have an upward bias. For the black student sample, we see that an additional non-disruptive student in attendance increases the probability of passing Algebra I by 2.98 to 4.07 percentage points. We also find that an additional enrolled student decreases the probability of passing Algebra I for black students by 0.96 to 3.18 percentage points. As stated earlier, our preferred model specification is column (4), which

[11] Appendix Table B1 reports the coefficients on the excluded instruments for four first-stage regressions. Each set of coefficients corresponds to a full sample regression in column (4) of Table 3 or Table 4. The first set is for panel A of Table 3, the second set is for panel B of Table 3, the third set is for panel A of Table 4, and the fourth set is for panel B of Table 4.

controls for student and neighborhood characteristics and middle school-by-teacher-by-year fixed effects. Relative to the mean passing rate of 72.2 percent for black students, the estimated coefficient of 0.0298 for classroom composition translates into an increase of 4.13 percent (0.0298/0.722=0.0413) in the probability of passing Algebra I. Two-stage least squares results for the Hispanic student sample are generally in line with findings for the black student sample, though the estimates are in many cases statistically insignificant because of the smaller sample sizes.

As shown in panel B of Table 3, we also observe in all four OLS specifications a positive association between classroom composition and the student scores on the mathematics examination for all races or ethnicities. For black students, we find that an additional non-disruptive student in attendance is associated with an increase of 0.0033 to 0.0075 student-level standard deviations in the student's mathematics score, which is less than one percentile point. The OLS estimated coefficients for class size are also as expected; we find a negative association between class size and own mathematics test score. For black students, we find that an additional student enrolled in a class is associated with a decrease of 0.0034 to 0.0068 student-level standard deviations in the student's mathematics score, which is again less than one percentile point. All OLS results are stable across all model specifications. OLS classroom composition and class size results for Hispanic students generally have the same magnitudes as the black student findings.

Moving to the 2SLS estimation results, for the black student sample and the full sample we observe high first-stage F-statistics for the test of the excluded instruments for first-stage regressions for classroom composition but not for class size; the F-statistics for class size indicate weak relationships between the excluded instruments and class size. As with the Algebra I course passing regressions, the F-statistics for the Hispanic student sample are statistically insignificant (except for one F-statistic in column (2)) due to the smaller samples sizes.

After instrumenting with period of the day and predicted class size, the 2SLS coefficients for the black student sample are larger in magnitude for classroom composition and for class size, as compared to the corresponding OLS coefficients and are almost always statistically significant, but the 2SLS coefficients for the Hispanic student sample and the full sample are statistically insignificant. For black students, we observe that an additional non-disruptive student in attendance increases the student's mathematics test score by 0.0166 to 0.0289 student-level standard deviations, approximately one half to one percentile point. We also find that an additional enrolled student decreases the student's mathematics test score by 0.0322 to 0.0431 student-level standard deviations, approximately one percentile point. As stated earlier, our preferred model specification is column (2), which controls for student and neighborhood characteristics and middle school-by-high school-by-year fixed effects; we observe that this model specification gives the largest class size effect estimate (0.0431 student-level standard deviations) and the second-largest estimated effect of classroom composition (0.0249 student-level standard deviations).

B. English I Course Passing and Reading Test Score Results for Regular Classes

We focus on the 2SLS results in this subsection because the OLS estimates for the regular English I sample are approximately the same magnitude as the OLS estimates presented in the previous subsection for the regular Algebra I sample. The course passing results are reported in panel A of Table 4. We find that all of the 2SLS coefficients are larger in magnitude for classroom composition and for class size, as compared to the corresponding OLS coefficients, and are statistically significant (with the exception of two cases for the Hispanic student sample), again consistent with the Lazear model. As shown for the black students, we see that an additional non-disruptive student in attendance increases the probability of passing English I by 4.77 to 7.08 percentage points. We also find that an additional enrolled student decreases the probability of passing English I for black students by 2.77 to 6.68 percentage points. For our preferred model specification (column (4)), relative to the mean passing rate of 77.3 percent, the estimated coefficient of 0.0477 for classroom composition translates into an increase of 6.17 percent (0.0477/0.773=0.0617) in the probability of passing English I. As for the 2SLS class size results for black students, we observe in column (4) a decrease in the probability of passing English I by 2.77 percentage points for each additional student enrolled in the class. Relative to the mean English I passing rate, this estimated coefficient translates into a decrease in the probability of passing English I by 3.58 percent (0.0277/0.773=0.0358).

Comparing the 2SLS estimates for black students in column (4) from panel A of Table 4 to the corresponding estimates in Table 3, we find that the effects of classroom composition and class size on the probability of passing English I are larger in magnitude than the impacts on the probability of passing Algebra I. For Hispanic students, the opposite is true; however, the effects are imprecisely estimated in both subjects, so it is difficult to draw comparisons for this sample.

The 2SLS results for reading examination scores are shown in panel B of Table 4. Although our reported F-statistics are high in most cases, virtually all of the estimated coefficients on classroom composition and class size are statistically insignificant, and some have unexpected signs.

C. Course Passing and Test Score Results for Remedial Classes

Another empirical implication of the Lazear theoretical framework is that the effects of classroom composition and class size on student achievement should be larger for students enrolled in remedial classes because students in remedial classes have, on average, lower baseline academic performance than students in regular classes. Table 5 presents the results for students in remedial courses. The layout of this table is somewhat different than that of Tables 3 and 4 because we only present the results for the two most conservative sets of fixed effects. That is, column (1) includes middle school-by-high school-by-year fixed effects (the fixed effects used in column (2) of Tables 3 and 4), and column (2) includes middle school-by-teacher-by-year fixed effects (the fixed effects used in column (4) of Tables 3 and 4). Moreover, we only report the results for the black and Hispanic student samples, since over 92 percent of the students enrolled in remedial courses are black or Hispanic (as shown in panel C of Table 1). For the reasons discussed earlier, our preferred specification is column (2) for the course passing analysis and column (1) for the test score analysis.

We again concentrate on the 2SLS estimates in this subsection because the OLS results for the remedial sample are about the same magnitude as the OLS results analyzed earlier. We begin with the 2SLS course passing results for Algebra I (panel A). For black students, the 2SLS coefficients are larger in magnitude for classroom composition and for class size, as compared to the corresponding OLS coefficients, and are statistically significant. For the specification in column (2), we observe that an additional non-disruptive student in attendance increases the probability of passing Algebra I by 4.79 percentage points, and an additional enrolled student decreases this probability by 6.33 percentage points. Relative to the mean passing rate of 68.4 percent, the coefficient on classroom composition translates into an increase of 7.00 percent (0.0479/0.684=0.0700) in the probability of passing remedial Algebra I. This increased probability is approximately 70 percent larger than the estimated increase in the probability of passing regular Algebra I (column (4) in panel A of Table 3).

Relative to the mean passing rate, the coefficient on class size translates into a decrease of 9.25 percent (0.0633/0.684=0.0925) in the probability of passing remedial Algebra I. This decreased probability is much larger than the statistically insignificant estimated decrease in the probability of passing regular Algebra I (column (4) in panel A of Table 3). The larger estimated effects for remedial Algebra I are consistent with the Lazear model.

For Hispanic students, all 2SLS results are statistically insignificant. As for the 2SLS mathematics test score results (panel B), the estimated coefficients on classroom composition and class size are statistically insignificant for both the black and Hispanic student samples.

The 2SLS course passing results for English I are shown in panel C. For both black and Hispanic students, the 2SLS coefficients are larger in magnitude for classroom composition and for class size, as compared to the corresponding OLS coefficients, and are statistically significant. For black students, again focusing on the specification in column (2), we find that an additional non-disruptive student in attendance increases the probability of passing English I by 6.60 percentage points, and an additional enrolled student decreases this probability by 6.08 percentage points. Relative to the mean passing rate of 73.5 percent, the coefficient on classroom composition translates into an increase of 8.98 percent (0.0660/0.735=0.0898) in the probability of passing remedial English I. This increased probability is approximately 45 percent larger than the estimated increase in the probability of passing regular English I (column (4) in panel A of Table 4).

Relative to the mean passing rate, the coefficient on class size translates into a decrease of 8.27 percent (0.0608/0.735=0.0827) in the probability of passing remedial English I. This decreased probability is over twice as large as the estimated decrease in the probability of passing regular English I (column (4) in panel A of Table 4). The larger estimated effects for remedial English I are again in line with the Lazear theoretical framework.

For Hispanic students, classroom composition and class size effects are similar to those for black students. The estimated coefficients in column (2) in panel C of Table 5 are again much larger in magnitude than the corresponding coefficients in column (4) in panel A of Table 4. As for the 2SLS reading test score results (panel D), the estimated coefficients on classroom composition and class size are statistically insignificant for both the black and Hispanic student samples.

D. Comparisons of Class Size Effects

It is instructive to compare our estimated effects of class size on student test scores with those from Tennessee's Project STAR class-size randomized experiment, as reported in Finn and

Achilles (1990). We focus on effect sizes for black students in regular Algebra I, as the 2SLS coefficients on class size in the other test score regressions are statistically insignificant. For black students in regular Algebra I, the estimated 2SLS coefficient on class size for the mathematics test score regressions is -0.0431 (column (2) in panel B of Table 3), which implies an effect size of 0.0431 student-level standard deviations for a one-student reduction in class size.

Finn and Achilles report effect sizes for 1st grade students, disaggregated by examination subject. Because the reported effect sizes are based on an approximately eight-student reduction in class size (moving from a regular class or a regular class with an aide to a small class), we divide these effect sizes by eight before comparing them with our effect sizes. Finn and Achilles present results for two mathematics tests, the Stanford Achievement Test (SAT) and the Basic Skills First (BSF) test, and find effect sizes of 0.27 and 0.13 student-level standard deviations, respectively, for an approximately eight-student reduction in class size (Table 5, page 566). Dividing these effect sizes by eight gives 0.0338 and 0.0163 student-level standard deviations for an approximately one-student reduction in class size.

Finn and Achilles also report effect sizes by minority status (Table 6, page 567), finding larger effect sizes for minority students than for white students. For the SAT and BSF mathematics tests, the effect sizes for minority students are 0.31 and 0.16 student-level standard deviations, respectively, for an eight-student reduction in class size, translating into 0.0388 and 0.0200 student-level standard deviations for a one-student reduction in class size. The effect size for minority students based on the SAT is almost identical in magnitude to our estimated effect size for black students, while the effect size based on the BSF test is about half the size of our estimated effect size, though it is important to keep in mind that the grade level of the students in our study differs from the grade level of the students in Finn and Achilles. For the SAT and BSF mathematics tests, the effect sizes for white students are 0.22 and 0.09 student-level standard deviations, respectively, for an eight-student reduction in class size, translating into 0.0275 and 0.0113 student-level standard deviations for a one-student reduction in class size.

Because Angrist and Lavy (1999) use an identification strategy similar to ours to identify the effect of class size on academic achievement for Israeli students, we also compare our estimated effects for mathematics test scores with those reported in this paper. Angrist and Lavy report that a one-student reduction in class size leads to 0.230 point increase in mathematics test scores for 5th graders (Table IV, column (8), page 554); estimates for 3rd and 4th graders are much smaller and statistically insignificant. Dividing this estimated coefficient by the *class-level* standard deviation of 9.6 points on the mathematics test for 5th graders, as Angrist and Lavy's analysis for 5th grade students uses class-level data, gives an effect size of 0.0240 class-level standard deviations. The corresponding student-level effect size is approximately 0.0149 student-level standard deviations.[12] While our effect size is approximately three times as large as that of Angrist and Lavy, it is again important to keep in mind that the grade level of the students in our study differs from the grade levels of the students in Angrist and Lavy. In addition, approximately 14 percent of Angrist and Lavy's analytical sample consists of students from disadvantaged backgrounds (Table I, page 539), whereas approximately 85 percent of our analytic sample consists of students who receive federal lunch subsidies (Table 1). Thus, we believe that our estimated class size effects are more comparable to the larger estimates for minority students that are presented in Finn and Achilles (1990).

[12] The student-level effect size was obtained as follows. For 5th grade reading scores, Angrist and Lavy report an effect size of approximately 0.29 class-level standard deviations for an eight-student reduction in class size (page 567). The authors estimate that this class-level effect size corresponds to an approximately 0.18 student-level standard deviation effect size for an eight-student reduction in class size. By multiplying the 0.0240 class-level effect size for mathematics by 0.18/0.29, we can roughly convert it into a student-level standard deviation effect size. Doing this multiplication gives us an effect size of approximately 0.0149 student-level standard deviations.

VI. Conclusion

Using administrative student transcript files from Chicago Public Schools (CPS), we analyze empirical implications of the Lazear educational production model. The Lazear framework suggests that classroom composition and class size are codetermined, indicating that both variables should appear as explanatory variables in the estimation of an educational production function. In other words, the size of a particular class is affected by the characteristics of the potential pool of students enrolled in the appropriate course, and the composition of a classroom is influenced by the size of the class. To that end, we exploit exogenous variation on course scheduling and policy-driven variation in class size in CPS to study heterogeneity in the effect of class size on student achievement in mathematics and reading. Most importantly, our research design allows us to isolate not only overall class size effects but also an underlying mechanism by which class size affects student performance, classroom composition.

In line with the theoretical predictions of the Lazear model, we find that an additional non-disruptive student in attendance increases the probability that a black student passes regular Algebra I by 2.98 percentage points, while an additional student enrolled in the class decreases this probability by 0.96 percentage points (the latter effect is statistically insignificant). The estimated impacts are larger in magnitude for black students enrolled in regular English I. For black students enrolled in regular Algebra I, we also observe that an additional non-disruptive student in attendance increases own mathematics test score by 0.0249 student-level standard deviations, while an additional enrolled student decreases this score by 0.0431 student-level standard deviations. Also consistent with the Lazear framework, the estimated impacts on course passing are larger in magnitude for black students enrolled in remedial classes.

References

Angrist, Joshua D., and Kevin Lang. (2004). Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program. *American Economic Review, 94*(5), 1613-1634.

Angrist, Joshua D., and Victor Lavy. (1999). Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement. *Quarterly Journal of Economics, 114*(2), 533-575.

Babcock, Philip, and Julian R. Betts. (2009). Reduced-Class Distinctions: Effort, Ability, and the Education Production Function. *Journal of Urban Economics, 65*(3), 314-322.

Chicago Public Schools. (1995, August 14). *Class Size* (Chicago Public Schools Policy Manual, Section: 301.2, Board Report: 95-0814-PO1). Retrieved from http://policy.cps.k12.il.us/documents/301.2.pdf.

Chicago Public Schools. (2004, January 28). *High School Promotion Policy* (Chicago Public Schools Policy Manual, Section: 605.1, Board Report: 04-0128-PO1). Retrieved from http://policy.cps.k12.il.us/documents/605.1.pdf.

Cortes, Kalena E., Jesse Bricker, and Christopher Rohlfs. (forthcoming). The Role of Specific Subjects in Education Production Functions: Evidence from Morning Classes in Chicago Public High Schools. *The B.E. Journal of Economic Analysis & Policy (Contributions).*

Ding, Weili, and Steven F. Lehrer. (2011). Experimental Estimates of the Impacts of Class Size on Test Scores: Robustness and Heterogeneity. *Canadian Labour Market and Skills Researcher Network Working Paper No. 77*.

Eide, Eric, and Mark H. Showalter. (1998). The Effect of School Quality on Student Performance: A Quantile Regression Approach. *Economics Letters, 58*(3), 345-350.

Finn, Jeremy D., and Charles M. Achilles. (1990). Answers and Questions About Class Size: A Statewide Experiment. *American Educational Research Journal, 27*(3), 557-577.

Finn, Jeremy D., Susan B. Gerber, Charles M. Achilles, and Jayne Boyd-Zaharias. (2001). The Enduring Effects of Small Classes. *Teachers College Record, 103*(2), 145-183.

Hoxby, Caroline M. (2000). The Effects of Class Size on Student Achievement: New Evidence from Population Variation. *Quarterly Journal of Economics, 115*(4), 1239-1285.

Jepsen, Christopher, and Steven Rivkin. (2009). Class Size Reduction and Student Achievement: The Potential Tradeoff between Teacher Quality and Class Size. *Journal of Human Resources, 44*(1), 223-250.

Konstantopoulos, Spyros. (2008). Do Small Classes Reduce the Achievement Gap between Low and High Achievers? Evidence from Project STAR. *Elementary School Journal, 108*(4), 275-291.

Krueger, Alan B. (1999). Experimental Estimates of Education Production Functions. *Quarterly Journal of Economics, 114*(2), 497-532.

Krueger, Alan B., and Diane M. Whitmore. (2001). The Effect of Attending a Small Class in the Early Grades on College-Test Taking and Middle School Test Results: Evidence from Project STAR. *Economic Journal, 111*(468), 1-28.

Lazear, Edward P. (2001). Educational Production. *Quarterly Journal of Economics, 116*(3), 777-803.

Levin, Jesse. (2001). For Whom the Reductions Count: A Quantile Regression Analysis of Class Size and Peer Effects on Scholastic Achievement. *Empirical Economics, 26*(1), 221-246.

Ma, Lingjie, and Roger Koenker. (2006). Quantile Regression Methods for Recursive Structural Equation Models. *Journal of Econometrics, 134*(2), 471-506.

McKee, Graham J., Steven G. Rivkin, and Katharine R. E. Sims. (2010). Disruption, Achievement and the Heterogeneous Benefits of Smaller Classes. *NBER Working Paper No. 15812*.

Nye, Barbara, Larry V. Hedges, and Spyros Konstantopoulos. (1999). The Long-Term Effects of Small Classes: A Five-Year Follow-Up of the Tennessee Class Size Experiment. *Educational Evaluation and Policy Analysis, 21*(2), 127-142.

Nye, Barbara, Larry V. Hedges, and Spyros Konstantopoulos. (2000). The Effects of Small Classes on Academic Achievement: The Results of the Tennessee Class Size Experiment. *American Educational Research Journal, 37*(1), 123-151.

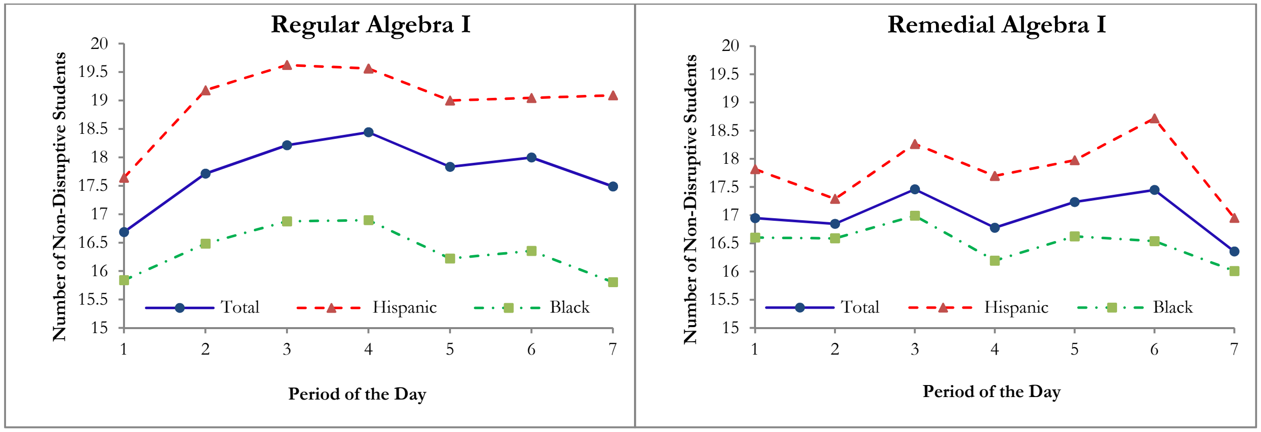
Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. (2005). Teachers, Schools, and Academic Achievement. *Econometrica, 73*(2), 417-458.

Segal, Carmit. (2008). Classroom Behavior. *Journal of Human Resources, 43*(4), 783-814.

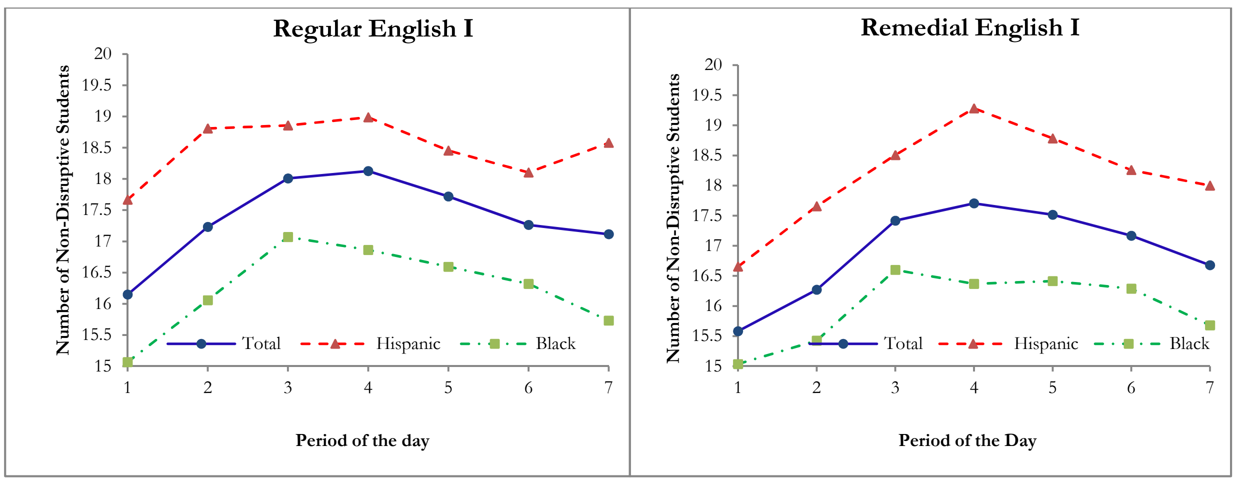
Word, Elizabeth, John Johnston, Helen Pate Bain, B. DeWayne Fulton, Jayne Boyd-Zaharias, Charles M. Achilles, Martha Nannette Lintz, John Folger, and Carolyn Breda. (1990). *The State of Tennessee's Student/Teacher Achievement Ratio (STAR) Project: Final Summary Report 1985-1990*. Nashville, TN: Tennessee State Department of Education. Retrieved from http://www.heros-inc.org/summary.pdf.

Figure 1: Average Number of Non-Disruptive Students in Attendance by Algebra and English Period

Panel A: Regular Algebra I and Remedial Algebra I



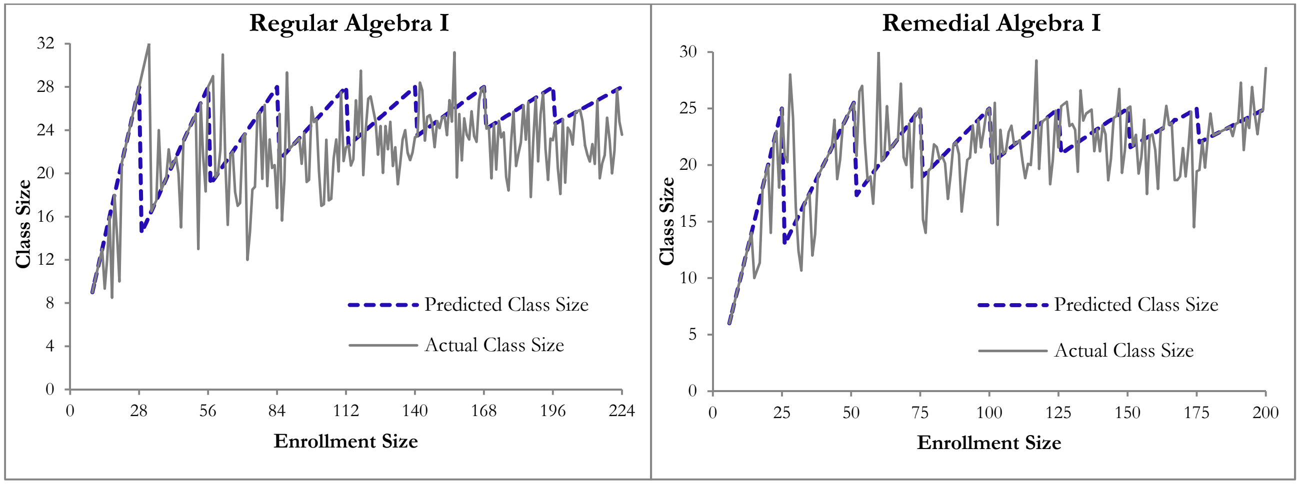
Panel B: Regular English I and Remedial English I



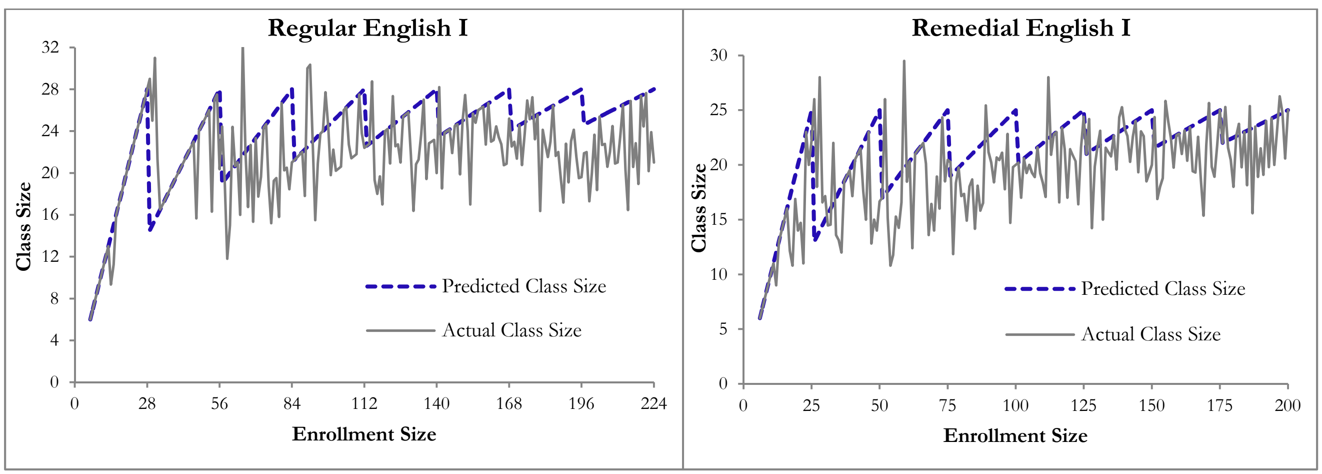
Notes: Figures represent the average number of non-disruptive students in attendance in a given classroom during a given period of the day for regular Algebra I, regular English I, and their respective remedial courses.

Figure 2: Actual Average Class Size and Predicted Average Class Size using Chicago Public Schools Policy for Class Size Cutoffs, by Enrollment Size

Panel A: Regular Algebra I and Remedial Algebra I



Panel B: Regular English I and Remedial English I



Notes: Enrollment sizes were calculated as the average number of students enrolled in a specific 9th grade course subject (regular Algebra I, regular English I, and their respective remedial courses) by school and semester. For the purposes of these graphs, enrollment sizes have been capped at 224 students for regular classes and 200 students for remedial classes.

Table 1: Summary Statistics for 9th Grade Chicago Public School Students by Course Subject

	Regular Algebra I		Remedial Algebra I		Regular English I		Remedial English I	
Panel A: Outcome Variables								
Course pass rate	0.743	(0.437)	0.697	(0.460)	0.786	(0.410)	0.759	(0.427)
Math test score (z-score)	0.069	(0.919)	-0.430	(0.791)	0.056	(0.950)	-0.270	(0.887)
Reading test score (z-score)	0.023	(0.911)	-0.387	(0.718)	0.085	(0.975)	-0.322	(0.761)
Panel B: Classroom Characteristics								
Classroom composition	18.02	(6.26)	17.44	(6.19)	17.89	(6.36)	17.39	(6.37)
Class size	25.23	(5.54)	24.56	(5.36)	25.02	(5.36)	23.76	(5.78)
Panel C: Student Characteristics								
Age	14.21	(0.514)	14.32	(0.569)	14.20	(0.538)	14.32	(0.548)
Male	0.492	(0.500)	0.504	(0.500)	0.486	(0.500)	0.511	(0.500)
White	0.084	(0.277)	0.062	(0.241)	0.091	(0.288)	0.059	(0.236)
Black	0.519	(0.500)	0.569	(0.495)	0.532	(0.499)	0.554	(0.497)
Hispanic	0.370	(0.483)	0.356	(0.479)	0.350	(0.477)	0.370	(0.483)
Asian	0.026	(0.160)	0.013	(0.113)	0.025	(0.155)	0.016	(0.127)
American Indian	0.001	(0.038)	0.001	(0.031)	0.002	(0.040)	0.001	(0.033)
Receives free or reduced lunch	0.855	(0.352)	0.858	(0.349)	0.828	(0.377)	0.877	(0.329)
Classified as disruptive	0.156	(0.363)	0.128	(0.334)	0.176	(0.381)	0.127	(0.333)
Enrolled in bilingual education	0.460	(0.604)	0.422	(0.607)	0.429	(0.590)	0.423	(0.578)
Lives with biological parent	0.822	(0.382)	0.828	(0.377)	0.844	(0.363)	0.800	(0.400)
Special education student	0.141	(0.348)	0.199	(0.400)	0.136	(0.343)	0.212	(0.408)
8th grade ITBS math test score (z-score)	0.078	(0.883)	-0.543	(0.638)	0.058	(0.923)	-0.302	(0.819)
8th grade ITBS reading test score (z-score)	0.031	(0.879)	-0.396	(0.764)	0.106	(0.918)	-0.340	(0.789)
Panel D: Neighborhood (Census-Block-Level) Characteristics								
Median family income	32,732	(15838)	32,700	(14600)	31,398	(15108)	34,276	(16238)
Percent school age (5-18)	0.237	(0.074)	0.241	(0.071)	0.235	(0.074)	0.243	(0.072)
Percent Hispanic	0.280	(0.326)	0.274	(0.329)	0.263	(0.317)	0.289	(0.336)
Percent black	0.471	(0.450)	0.510	(0.449)	0.479	(0.451)	0.502	(0.448)
Mean education	11.80	(1.246)	11.72	(1.272)	11.91	(1.139)	11.61	(1.369)
Percent in poverty	0.246	(0.194)	0.250	(0.183)	0.246	(0.197)	0.255	(0.187)
Observations (student-by-year)	252,771		128,424		216,715		276,200	

Source: Chicago Public Schools High School Transcript Data, 1993-94 through 2005-06

Notes: Regular Algebra I (or regular English I) represents the sample of students enrolled in only one regular Algebra I (or regular English I) course per semester. Remedial Algebra I (or remedial English I) represents the sample of students enrolled in at least one remedial Algebra I (or remedial English I) course per semester; some students take regular Algebra I (or regular English I) in addition to the remedial class. Numbers in parentheses indicate standard deviations.

Table 2A: Distribution of 9th Grade Course Offerings by Period of the Day

	Period of the day is …						
Course Subject	1st	2nd	3rd	4th	5th	6th	7th
English	0.29	0.30	0.28	0.28	0.28	0.28	0.28
Regular English I	***0.20***	***0.20***	***0.18***	***0.18***	***0.18***	***0.18***	***0.18***
Remedial English I	***0.09***	***0.10***	***0.10***	***0.10***	***0.10***	***0.10***	***0.10***
Mathematics	0.21	0.22	0.21	0.21	0.22	0.22	0.21
Regular Algebra I	***0.13***	***0.13***	***0.13***	***0.13***	***0.13***	***0.14***	***0.13***
Remedial Algebra I	***0.05***	***0.05***	***0.05***	***0.05***	***0.06***	***0.05***	***0.05***
Social Studies	0.15	0.16	0.14	0.15	0.15	0.15	0.17
Science	0.11	0.10	0.10	0.12	0.11	0.11	0.11
Foreign Language	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Shop	0.00	0.01	0.00	0.00	0.01	0.00	0.00
Business	0.05	0.06	0.06	0.07	0.06	0.07	0.07
Vocational	0.01	0.01	0.01	0.01	0.01	0.01	0.02
Art, Music, and Physical Education	0.14	0.12	0.16	0.12	0.14	0.12	0.12
Other	0.01	0.01	0.02	0.01	0.01	0.01	0.01

Source: Chicago Public Schools High School Transcript Data, 1993-94 through 2005-06

Notes: All columns sum to one excluding the rows in italics. Numbers in italics denote subgroups of their respective course subjects.

Table 2B: Determinants of Period of the Day
Instrument Validity Regressions for Period of the Day Instruments (Regular Algebra I Sample)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Dependent Variable: Period of the day is …						
	1st	2nd	3rd	4th	5th	6th	7th
Age	-0.0018	0.0025	0.0018	-0.0021	-0.0006	-0.0003	0.0013
	(0.0013)	(0.0016)	(0.0016)	(0.0014)	(0.0013)	(0.0014)	(0.0016)
Male	0.0011	0.0028	-0.0006	0.0015	-0.0012	-0.0014	-0.0021
	(0.0017)	(0.0020)	(0.0017)	(0.0019)	(0.0016)	(0.0017)	(0.0019)
White	0.0024	-0.0004	0.0044	-0.0009	-0.0046	0.0012	-0.0026
	(0.0027)	(0.0034)	(0.0032)	(0.0033)	(0.0035)	(0.0033)	(0.0032)
Black	0.0023	-0.0042	0.0009	-0.0070*	-0.0018	0.0015	0.0058
	(0.0032)	(0.0041)	(0.0035)	(0.0039)	(0.0034)	(0.0036)	(0.0040)
American Indian	-0.0044	0.0257	-0.0086	0.0049	-0.0063	0.0154	-0.0356**
	(0.0142)	(0.0206)	(0.0170)	(0.0180)	(0.0189)	(0.0182)	(0.0179)
Asian	-0.0001	0.0078	-0.0085	-0.0020	-0.0029	0.0053	-0.0026
	(0.0046)	(0.0066)	(0.0054)	(0.0050)	(0.0061)	(0.0067)	(0.0073)
Free or reduced lunch	0.0060***	-0.0032	0.0010	0.0013	-0.0012	-0.0024	-0.0006
	(0.0020)	(0.0024)	(0.0022)	(0.0021)	(0.0020)	(0.0023)	(0.0020)
Bilingual education	-0.0009	0.0010	-0.0021	-0.0022	0.0010	0.0015	0.0041*
	(0.0020)	(0.0024)	(0.0020)	(0.0026)	(0.0022)	(0.0021)	(0.0023)
Lives with biological parent	0.0031*	-0.0009	-0.0010	0.0017	0.0006	-0.0002	-0.0023
	(0.0017)	(0.0022)	(0.0021)	(0.0020)	(0.0018)	(0.0019)	(0.0021)
Special education	0.0033	-0.0051	0.0134	-0.0133	-0.0032	0.0204*	-0.0070
	(0.0087)	(0.0110)	(0.0115)	(0.0104)	(0.0097)	(0.0116)	(0.0114)
Special education x year trend	-0.0008	0.0010	-0.0017	0.0022*	0.0012	-0.0020	-0.0001
	(0.0011)	(0.0013)	(0.0014)	(0.0013)	(0.0013)	(0.0015)	(0.0014)
8th grade ITBS math z-score	-0.0027	-0.0015	0.0025	-0.0017	0.0032*	0.0023	-0.0037*
	(0.0018)	(0.0022)	(0.0022)	(0.0021)	(0.0019)	(0.0020)	(0.0021)
8th grade ITBS reading z-score	-0.0022	0.0004	-0.0003	-0.0008	0.0005	-0.0006	0.0024*
	(0.0015)	(0.0016)	(0.0015)	(0.0014)	(0.0014)	(0.0014)	(0.0013)
Classified as disruptive	-0.0013	-0.0013	-0.0015	-0.0037*	0.0057***	-0.0025	0.0022
	(0.0019)	(0.0024)	(0.0023)	(0.0021)	(0.0020)	(0.0021)	(0.0021)
Neighborhood median family income	-0.0000	0.0000	0.0000**	-0.0000	-0.0000	-0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Neighborhood percent school age (5-18)	-0.0092	-0.0116	-0.0179	0.0207	-0.0155	0.0295**	-0.0060
	(0.0133)	(0.0175)	(0.0156)	(0.0164)	(0.0140)	(0.0148)	(0.0158)
Neighborhood percent Hispanic	-0.0020	0.0063	0.0154**	0.0012	0.0095	-0.0020	-0.0140*
	(0.0058)	(0.0074)	(0.0075)	(0.0068)	(0.0063)	(0.0066)	(0.0074)
Neighborhood percent black	-0.0054	0.0062	0.0106*	0.0083	0.0057	-0.0035	-0.0053
	(0.0050)	(0.0065)	(0.0059)	(0.0054)	(0.0054)	(0.0053)	(0.0054)
Neighborhood mean education	0.0013	0.0002	-0.0001	0.0004	0.0014	0.0001	-0.0026**
	(0.0011)	(0.0015)	(0.0013)	(0.0012)	(0.0012)	(0.0012)	(0.0013)
Neighborhood percent in poverty	-0.0023	0.0023	0.0102	-0.0094	-0.0008	-0.0050	0.0042
	(0.0070)	(0.0082)	(0.0085)	(0.0078)	(0.0070)	(0.0069)	(0.0080)
Observations (student-by-year)	269,937	269,937	269,937	269,937	269,937	269,937	269,937

Notes: Each regression specification corresponds to the OLS specification in column (4) of Table 3, which includes middle school-by-teacher-by-year fixed effects. Also included is a quadratic function in course subject enrollment by school and semester, as well as an indicator for the fall semester. Numbers in parentheses represent standard errors clustered at the high school-by-year level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table 3: OLS and 2SLS Results for Classroom Composition and Class Size on Course Passing and Math Test Scores for Regular Algebra I Sample

		(1)		(2)		(3)		(4)	
		Panel A: Course Passing in Regular Algebra I							
	Sample Mean	**OLS**	**2SLS**	**OLS**	**2SLS**	**OLS**	**2SLS**	**OLS**	**2SLS**
Black Student Sample									
Classroom Composition		0.0029***	0.0377***	0.0036***	0.0350***	0.0021***	0.0407***	0.0026***	0.0298***
	0.722	(0.0007)	(0.0087)	(0.0008)	(0.0076)	(0.0008)	(0.0100)	(0.0008)	(0.0099)
Class Size		-0.0074***	-0.0236*	-0.0047***	-0.0163	-0.0066***	-0.0318**	-0.0028***	-0.0096
		(0.0007)	(0.0126)	(0.0007)	(0.0122)	(0.0007)	(0.0138)	(0.0008)	(0.0129)
First-stage F-Statistic (Composition)		3.52***		5.35***		3.66***		4.40***	
First-stage F-Statistic (Size)		1.87*		1.76*		1.89*		2.03**	
Observations (student-by-year)		130,931		130,931		130,931		130,931	
Hispanic Student Sample									
Classroom Composition		0.0063***	0.0487*	0.0065***	0.0397*	0.0057***	0.0476	0.0057***	0.0456
	0.759	(0.0010)	(0.0249)	(0.0010)	(0.0224)	(0.0010)	(0.0304)	(0.0009)	(0.0417)
Class Size		-0.0098***	-0.0116	-0.0072***	-0.0301	-0.0093***	-0.0122	-0.0062***	-0.0325
		(0.0009)	(0.0234)	(0.0010)	(0.0298)	(0.0010)	(0.0244)	(0.0010)	(0.0491)
First-stage F-Statistic (Composition)		1.22		2.01		1.04		1.57	
First-stage F-Statistic (Size)		1.77*		1.90*		1.96*		2.00*	
Observations (student-by-year)		92,915		92,915		92,915		92,915	
Full Student Sample									
Classroom Composition		0.0044***	0.0411***	0.0049***	0.0294***	0.0038***	0.0453***	0.0038***	0.0261**
	0.743	(0.0006)	(0.0093)	(0.0006)	(0.0087)	(0.0006)	(0.0106)	(0.0006)	(0.0117)
Class Size		-0.0084***	-0.0106	-0.0057***	0.0035	-0.0078***	-0.0157	-0.0041***	0.0079
		(0.0006)	(0.0127)	(0.0006)	(0.0148)	(0.0006)	(0.0136)	(0.0006)	(0.0169)
First-stage F-Statistic (Composition)		3.72***		6.23***		3.64***		4.94***	
First-stage F-Statistic (Size)		2.23**		2.24**		2.42**		2.56**	
Observations (student-by-year)		252,771		252,771		252,771		252,771	
Controls:									
Student and Neighborhood Characteristics		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
High School-by-Year F.E.		Yes	Yes	-	-	-	-	-	-
Middle School-by-High School-by-Year F.E.		-	-	Yes	Yes	-	-	-	-
Teacher-by-Year F.E.		-	-	-	-	Yes	Yes	-	-
Middle School-by-Teacher-by-Year F.E.		-	-	-	-	-	-	Yes	Yes

Table 3: (Continued)

	(1)		(2)		(3)		(4)	
	Panel B: Math Test Scores							
	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS
Black Student Sample								
Classroom Composition	0.0075***	0.0166*	0.0072***	0.0249**	0.0052***	0.0228*	0.0033***	0.0289**
	(0.0011)	(0.0093)	(0.0011)	(0.0118)	(0.0010)	(0.0118)	(0.0009)	(0.0127)
Class Size	-0.0045***	-0.0356**	-0.0068***	-0.0431**	-0.0034***	-0.0394**	-0.0037***	-0.0322
	(0.0010)	(0.0178)	(0.0010)	(0.0204)	(0.0010)	(0.0191)	(0.0010)	(0.0218)
First-stage F-Statistic (Composition)	4.71***		6.35***		4.62***		5.08***	
First-stage F-Statistic (Size)	1.26		1.64		1.35		1.40	
Observations (student-by-year)	90,654		90,654		90,654		90,654	
Hispanic Student Sample								
Classroom Composition	0.0103***	0.0178	0.0050***	0.0029	0.0088***	0.0195	0.0028**	0.0315
	(0.0013)	(0.0200)	(0.0015)	(0.0279)	(0.0013)	(0.0240)	(0.0013)	(0.0430)
Class Size	-0.0066***	-0.0215	-0.0041***	-0.0103	-0.0062***	-0.0106	-0.0015	-0.0541
	(0.0013)	(0.0271)	(0.0015)	(0.0449)	(0.0013)	(0.0287)	(0.0015)	(0.0634)
First-stage F-Statistic (Composition)	1.54		2.33**		1.19		1.62	
First-stage F-Statistic (Size)	1.24		1.37		1.36		1.13	
Observations (student-by-year)	67,576		67,576		67,576		67,576	
Full Student Sample								
Classroom Composition	0.0091***	0.0142	0.0070***	0.0118	0.0071***	0.0133	0.0035***	0.0033
	(0.0009)	(0.0091)	(0.0009)	(0.0125)	(0.0008)	(0.0108)	(0.0007)	(0.0149)
Class Size	-0.0058***	-0.0252	-0.0062***	-0.0105	-0.0049***	-0.0143	-0.0036***	0.0097
	(0.0008)	(0.0195)	(0.0009)	(0.0243)	(0.0008)	(0.0195)	(0.0008)	(0.0291)
First-stage F-Statistic (Composition)	4.65***		7.34***		4.11***		5.88***	
First-stage F-Statistic (Size)	1.23		1.88*		1.37		1.90*	
Observations (student-by-year)	179,734		179,734		179,734		179,734	
Controls:								
Student and Neighborhood Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
High School-by-Year F.E.	Yes	Yes	-	-	-	-	-	-
Middle School-by-High School-by-Year F.E.	-	-	Yes	Yes	-	-	-	-
Teacher-by-Year F.E.	-	-	-	-	Yes	Yes	-	-
Middle School-by-Teacher-by-Year F.E.	-	-	-	-	-	-	Yes	Yes

Notes: The samples include students enrolled in only one regular Algebra I course per semester. Student controls include age, male, race/ethnicity, free or reduced lunch, classified as disruptive, bilingual education, guardian status, special education, special education x year trend, 8th grade ITBS score, and enrolled in 8th+ period. Neighborhood characteristics include median family income, percent school age (5-18), percent Hispanic, percent black, mean education, and percent in poverty. Also included is a quadratic function in course subject enrollment by school and semester, as well as an indicator for the fall semester. Numbers in parentheses represent standard errors clustered at the high school-by-year level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. First-stage F-statistics represent the F-statistics on the excluded instruments for both first-stages (Composition and Size) using standard errors clustered at the high school-by-year level.

Table 4: OLS and 2SLS Results for Classroom Composition and Class Size on Course Passing and Reading Test Scores for Regular English I Sample

		(1)		(2)		(3)		(4)	
		Panel A: Course Passing in Regular English I							
	Sample Mean	**OLS**	**2SLS**	**OLS**	**2SLS**	**OLS**	**2SLS**	**OLS**	**2SLS**
Black Student Sample									
Classroom Composition	0.773	0.0046***	0.0708***	0.0052***	0.0547***	0.0039***	0.0680***	0.0037***	0.0477***
		(0.0009)	(0.0156)	(0.0008)	(0.0116)	(0.0009)	(0.0165)	(0.0009)	(0.0162)
Class Size		-0.0063***	-0.0668***	-0.0050***	-0.0344***	-0.0060***	-0.0566***	-0.0034***	-0.0277**
		(0.0008)	(0.0193)	(0.0007)	(0.0110)	(0.0008)	(0.0195)	(0.0009)	(0.0128)
First-stage F-Statistic (Composition)		4.10***		4.08***		4.33***		2.82***	
First-stage F-Statistic (Size)		2.47**		3.17***		2.27**		2.16**	
Observations (student-by-year)		114,848		114,848		114,848		114,848	
Hispanic Student Sample									
Classroom Composition	0.791	0.0076***	0.0763**	0.0071***	0.0550***	0.0075***	0.0587**	0.0054***	0.0281
		(0.0011)	(0.0304)	(0.0011)	(0.0187)	(0.0011)	(0.0259)	(0.0012)	(0.0177)
Class Size		-0.0109***	-0.0723*	-0.0082***	-0.0358*	-0.0106***	-0.0560**	-0.0066***	-0.0161
		(0.0010)	(0.0394)	(0.0011)	(0.0187)	(0.0010)	(0.0279)	(0.0012)	(0.0147)
First-stage F-Statistic (Composition)		1.64		2.11**		1.29		1.10	
First-stage F-Statistic (Size)		0.81		1.66		1.18		1.62	
Observations (student-by-year)		75,450		75,450		75,450		75,450	
Full Student Sample									
Classroom Composition	0.786	0.0051***	0.0725***	0.0056***	0.0563***	0.0049***	0.0692***	0.0042***	0.0439***
		(0.0007)	(0.0142)	(0.0007)	(0.0102)	(0.0007)	(0.0146)	(0.0007)	(0.0122)
Class Size		-0.0075***	-0.0634***	-0.0059***	-0.0356***	-0.0075***	-0.0552***	-0.0047***	-0.0244**
		(0.0006)	(0.0178)	(0.0006)	(0.0112)	(0.0006)	(0.0165)	(0.0007)	(0.0113)
First-stage F-Statistic (Composition)		5.33***		5.68***		5.01***		3.87***	
First-stage F-Statistic (Size)		2.96***		3.71***		3.02***		3.06***	
Observations (student-by-year)		216,715		216,715		216,715		216,715	
Controls:									
Student and Neighborhood Characteristics		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
High School-by-Year F.E.		Yes	Yes	-	-	-	-	-	-
Middle School-by-High School-by-Year F.E.		-	-	Yes	Yes	-	-	-	-
Teacher-by-Year F.E.		-	-	-	-	Yes	Yes	-	-
Middle School-by-Teacher-by-Year F.E.		-	-	-	-	-	-	Yes	Yes

Table 4: (Continued)

	(1)		(2)		(3)		(4)	
	Panel B: Reading Test Scores							
	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS
Black Student Sample								
Classroom Composition	0.0133***	-0.0189	0.0101***	0.0037	0.0086***	-0.0073	0.0050***	0.0020
	(0.0020)	(0.0219)	(0.0019)	(0.0162)	(0.0016)	(0.0212)	(0.0014)	(0.0177)
Class Size	-0.0052***	0.0587*	-0.0056***	0.0304	-0.0016	0.0578	-0.0031*	0.0315
	(0.0017)	(0.0319)	(0.0018)	(0.0220)	(0.0015)	(0.0358)	(0.0016)	(0.0307)
First-stage F-Statistic (Composition)	4.61***		3.89***		4.48***		2.60**	
First-stage F-Statistic (Size)	2.01*		2.38**		1.69		1.54	
Observations (student-by-year)	83,895		83,895		83,895		83,895	
Hispanic Student Sample								
Classroom Composition	0.0149***	0.0096	0.0124***	0.0051	0.0136***	0.0004	0.0087***	0.0285
	(0.0020)	(0.0378)	(0.0024)	(0.0252)	(0.0023)	(0.0311)	(0.0026)	(0.0245)
Class Size	-0.0095***	0.0296	-0.0110***	0.0341	-0.0077***	0.0215	-0.0065**	-0.0032
	(0.0022)	(0.0498)	(0.0023)	(0.0391)	(0.0025)	(0.0374)	(0.0026)	(0.0296)
First-stage F-Statistic (Composition)	2.60**		2.13**		2.40**		1.27	
First-stage F-Statistic (Size)	1.60		1.27		1.98*		1.57	
Observations (student-by-year)	56,942		56,942		56,942		56,942	
Full Student Sample								
Classroom Composition	0.0221***	-0.0233	0.0113***	-0.0079	0.0109***	-0.0168	0.0072***	0.0038
	(0.0023)	(0.0220)	(0.0016)	(0.0152)	(0.0014)	(0.0201)	(0.0014)	(0.0165)
Class Size	-0.0084***	0.0722**	-0.0084***	0.0369*	-0.0050***	0.0635**	-0.0053***	0.0342
	(0.0020)	(0.0305)	(0.0015)	(0.0206)	(0.0014)	(0.0277)	(0.0014)	(0.0237)
First-stage F-Statistic (Composition)	6.05***		5.16***		5.58***		3.29***	
First-stage F-Statistic (Size)	3.13***		3.11***		3.14***		2.86***	
Observations (student-by-year)	160,837		160,837		160,837		160,837	
Controls:								
Student and Neighborhood Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
High School-by-Year F.E.	Yes	Yes	-	-	-	-	-	-
Middle School-by-High School-by-Year F.E.	-	-	Yes	Yes	-	-	-	-
Teacher-by-Year F.E.	-	-	-	-	Yes	Yes	-	-
Middle School-by-Teacher-by-Year F.E.	-	-	-	-	-	-	Yes	Yes

Notes: The samples include students enrolled in only one regular English I course per semester. Student controls include age, male, race/ethnicity, free or reduced lunch, classified as disruptive, bilingual education, guardian status, special education, special education x year trend, 8th grade ITBS score, and enrolled in 8th+ period. Neighborhood characteristics include median family income, percent school age (5-18), percent Hispanic, percent black, mean education, and percent in poverty. Also included is a quadratic function in course subject enrollment by school and semester, as well as an indicator for the fall semester. Numbers in parentheses represent standard errors clustered at the high school-by-year level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. First-stage F-statistics represent the F-statistics on the excluded instruments for both first-stages (Composition and Size) using standard errors clustered at the high school-by-year level.

Table 5: OLS and 2SLS Results for Classroom Composition and Class Size on Course Passing and Test Scores for Students in Remedial Classes

		(1)		(2)		(1)		(2)	
		Panel A: Course Passing in Remedial Algebra I				Panel B: Math Test Scores			
	Sample Mean	OLS	2SLS	OLS	2SLS	OLS	2SLS	OLS	2SLS
Black Student Sample									
Classroom Composition		0.0052***	0.0618***	0.0035***	0.0479***	0.0026***	0.0128	0.0008	0.0093
	0.684	(0.0012)	(0.0147)	(0.0010)	(0.0153)	(0.0010)	(0.0197)	(0.0009)	(0.0161)
Class Size		-0.0063***	-0.0535***	-0.0045***	-0.0633***	-0.0018	0.0130	-0.0005	0.0094
		(0.0011)	(0.0155)	(0.0011)	(0.0172)	(0.0013)	(0.0163)	(0.0013)	(0.0169)
First-stage F-Statistic (Composition)		1.55		1.60		1.38		1.55	
First-stage F-Statistic (Size)		1.89*		1.32		1.70		1.34	
Observations (student-by-year)		72,747		72,747		61,835		61,835	
Hispanic Student Sample									
Classroom Composition		0.0063***	0.0152	0.0048***	0.0032	0.0027	-0.0011	0.0015	-0.0246
	0.711	(0.0016)	(0.0168)	(0.0016)	(0.0185)	(0.0019)	(0.0284)	(0.0019)	(0.0292)
Class Size		-0.0062***	0.0036	-0.0047***	0.0104	-0.0031*	0.0096	-0.0007	0.0320
		(0.0013)	(0.0173)	(0.0013)	(0.0204)	(0.0017)	(0.0309)	(0.0018)	(0.0359)
First-stage F-Statistic (Composition)		3.05***		2.90***		2.51**		2.28**	
First-stage F-Statistic (Size)		1.86*		1.75*		1.54		1.33	
Observations (student-by-year)		45,533		45,533		38,721		38,721	

		Panel C: Course Passing in Remedial English I				Panel D: Reading Test Scores			
Black Student Sample									
Classroom Composition		0.0057***	0.0694***	0.0039***	0.0660***	0.0020***	-0.0085	0.0009	-0.0019
	0.735	(0.0010)	(0.0119)	(0.0010)	(0.0160)	(0.0008)	(0.0176)	(0.0006)	(0.0247)
Class Size		-0.0057***	-0.0601***	-0.0048***	-0.0608***	-0.0010	-0.0013	-0.0003	0.0055
		(0.0009)	(0.0090)	(0.0009)	(0.0093)	(0.0009)	(0.0100)	(0.0008)	(0.0099)
First-stage F-Statistic (Composition)		1.83*		1.25		1.32		0.93	
First-stage F-Statistic (Size)		2.99***		3.34***		2.21**		2.31**	
Observations (student-by-year)		152,836		152,836		108,161		108,161	
Hispanic Student Sample									
Classroom Composition		0.0070***	0.0434***	0.0064***	0.0537***	0.0046***	-0.0149	0.0029*	-0.0072
	0.785	(0.0015)	(0.0135)	(0.0016)	(0.0198)	(0.0018)	(0.0229)	(0.0016)	(0.0239)
Class Size		-0.0079***	-0.0406***	-0.0069***	-0.0516**	-0.0037**	-0.0019	-0.0020	0.0099
		(0.0014)	(0.0140)	(0.0014)	(0.0211)	(0.0017)	(0.0251)	(0.0015)	(0.0278)
First-stage F-Statistic (Composition)		3.09***		2.66**		2.75***		2.36**	
First-stage F-Statistic (Size)		3.16***		3.07***		2.30**		2.18**	
Observations (student-by-year)		101,506		101,506		74,925		74,925	
Controls:									
Student and Neighborhood Characteristics		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Middle School-by-High School-by-Year F.E.		Yes	Yes	-	-	Yes	Yes	-	-
Middle School-by-Teacher-by-Year F.E.		-	-	Yes	Yes	-	-	Yes	Yes

Notes: The samples include students enrolled in at least one remedial Algebra I (or remedial English I) course per semester; some students take regular Algebra I (or regular English I) in addition to the remedial class. Student controls include age, male, race/ethnicity, free or reduced lunch, classified as disruptive, bilingual education, guardian status, special education, special education x year trend, 8th grade ITBS score, and enrolled in 8th+ period. Neighborhood characteristics include median family income, percent school age (5-18), percent Hispanic, percent black, mean education, and percent in poverty. Also included is a quadratic function in course subject enrollment by school and semester, as well as an indicator for the fall semester. Numbers in parentheses represent standard errors clustered at the high school-by-year level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. First-stage F-statistics represent the F-statistics on the excluded instruments for both first-stages (Composition and Size) using standard errors clustered at the high school-by-year level.

Appendix A: Additional Regressions for Testing Instrument Validity

Table A1: Instrument Validity Regressions for Period of the Day Instruments (Regular English I Sample)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Dependent Variable: Period of the day is …						
	1st	2nd	3rd	4th	5th	6th	7th
Age	0.0004	-0.0039	-0.0012	0.0014	0.0003	0.0015	0.0017
	(0.0016)	(0.0018)	(0.0017)	(0.0017)	(0.0016)	(0.0017)	(0.0018)
Male	0.0011	0.0019	-0.0008	0.0020	-0.0030	-0.0027	0.0027
	(0.0018)	(0.0022)	(0.0020)	(0.0019)	(0.0018)	(0.0020)	(0.0021)
White	-0.0025	-0.0009	0.0025	0.0043	-0.0061	-0.0000	-0.0008
	(0.0029)	(0.0037)	(0.0038)	(0.0037)	(0.0034)	(0.0037)	(0.0037)
Black	0.0005	-0.0046	-0.0013	0.0024*	-0.0016	-0.0007	0.0003
	(0.0036)	(0.0048)	(0.0045)	(0.0039)	(0.0041)	(0.0038)	(0.0042)
American Indian	0.0110	-0.0074	-0.0130	-0.0214	-0.0007	0.0090	0.0053**
	(0.0164)	(0.0214)	(0.0180)	(0.0197)	(0.0175)	(0.0179)	(0.0208)
Asian	-0.0027	0.0016	-0.0021	-0.0021	-0.0034	0.0072	-0.0028
	(0.0060)	(0.0061)	(0.0060)	(0.0056)	(0.0057)	(0.0062)	(0.0055)
Free or reduced lunch	-0.0020***	0.0054	-0.0031	-0.0010	-0.0021	-0.0019	0.0033
	(0.0020)	(0.0026)	(0.0023)	(0.0021)	(0.0020)	(0.0022)	(0.0022)
Bilingual education	-0.0011	-0.0037	0.0011	0.0012	0.0001	0.0011	-0.0018*
	(0.0022)	(0.0024)	(0.0026)	(0.0024)	(0.0021)	(0.0022)	(0.0025)
Lives with biological parent	-0.0016*	-0.0023	0.0007	-0.0005	-0.0009	0.0011	0.0020
	(0.0020)	(0.0025)	(0.0026)	(0.0024)	(0.0021)	(0.0021)	(0.0024)
Special education	0.0041	-0.0097	-0.0072	0.0059	0.0008	-0.0092*	0.0043
	(0.0083)	(0.0110)	(0.0106)	(0.0108)	(0.0100)	(0.0103)	(0.0088)
Special education x year trend	-0.0009	0.0010	0.0026	-0.0014*	0.0011	0.0006	-0.0015
	(0.0013)	(0.0016)	(0.0015)	(0.0015)	(0.0014)	(0.0015)	(0.0014)
8th grade ITBS math z-score	0.0022	0.0037	0.0022	0.0006	-0.0042*	-0.0027	-0.0017*
	(0.0015)	(0.0019)	(0.0017)	(0.0017)	(0.0014)	(0.0017)	(0.0017)
8th grade ITBS reading z-score	0.0036	-0.0004	-0.0009	0.0008	-0.0012	0.0024	-0.0035*
	(0.0020)	(0.0022)	(0.0021)	(0.0020)	(0.0019)	(0.0020)	(0.0019)
Classified as disruptive	0.0055	0.0004	0.0012	0.0019*	-0.0018***	-0.0034	-0.0020
	(0.0023)	(0.0023)	(0.0022)	(0.0022)	(0.0022)	(0.0021)	(0.0023)
Neighborhood median family income	0.0000	-0.0000	-0.0000**	0.0000	0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Neighborhood percent school age (5-18)	-0.0157	0.0181	0.0260	-0.0106	-0.0017	-0.0163**	0.0196
	(0.0143)	(0.0180)	(0.0180)	(0.0154)	(0.0156)	(0.0165)	(0.0186)
Neighborhood percent Hispanic	0.0067	-0.0098	0.0063**	-0.0003	0.0004	-0.0024	-0.0008*
	(0.0063)	(0.0077)	(0.0085)	(0.0079)	(0.0072)	(0.0079)	(0.0076)
Neighborhood percent black	0.0040	-0.0009	0.0054*	-0.0057	-0.0062	-0.0024	-0.0026
	(0.0058)	(0.0063)	(0.0067)	(0.0063)	(0.0054)	(0.0062)	(0.0058)
Neighborhood mean education	0.0007	-0.0005	0.0023	-0.0021	0.0001	-0.0014	0.0005**
	(0.0013)	(0.0016)	(0.0017)	(0.0015)	(0.0013)	(0.0014)	(0.0016)
Neighborhood percent in poverty	-0.0047	-0.0001	0.0005	0.0104	-0.0041	-0.0021	-0.0076
	(0.0077)	(0.0092)	(0.0084)	(0.0073)	(0.0072)	(0.0083)	(0.0091)
Observations (student-by-year)	230,585	230,585	230,585	230,585	230,585	230,585	230,585

Notes: Each regression specification corresponds to the OLS specification in column (4) of Table 4, which includes middle school-by-teacher-by-year fixed effects. Also included is a quadratic function in course subject enrollment by school and semester, as well as an indicator for the fall semester. Numbers in parentheses represent standard errors clustered at the high school-by-year level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table A2: Instrument Validity Regressions for Period of the Day Instruments (Remedial Algebra I Sample)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Dependent Variable: Period of the day is …						
	1st	2nd	3rd	4th	5th	6th	7th
Age	0.0014	-0.0022	-0.0021	-0.0009	-0.0006	-0.0013	0.0030
	(0.0016)	(0.0019)	(0.0017)	(0.0017)	(0.0020)	(0.0019)	(0.0018)
Male	-0.0010	0.0018	-0.0004	-0.0039	-0.0023	0.0011	0.0004
	(0.0024)	(0.0022)	(0.0022)	(0.0024)	(0.0025)	(0.0023)	(0.0026)
White	0.0013	-0.0023	-0.0101	0.0029	0.0021	0.0004	0.0058
	(0.0042)	(0.0051)	(0.0054)	(0.0040)	(0.0051)	(0.0041)	(0.0049)
Black	0.0029	-0.0013	-0.0000	0.0022*	0.0009	-0.0074	0.0044
	(0.0047)	(0.0056)	(0.0046)	(0.0048)	(0.0049)	(0.0051)	(0.0050)
American Indian	-0.0218	0.0053	0.0311	-0.0032	0.0063	-0.0175	0.0152**
	(0.0288)	(0.0347)	(0.0346)	(0.0201)	(0.0342)	(0.0165)	(0.0285)
Asian	0.0026	-0.0062	-0.0136	0.0122	0.0053	-0.0049	0.0076
	(0.0071)	(0.0077)	(0.0080)	(0.0088)	(0.0077)	(0.0083)	(0.0082)
Free or reduced lunch	-0.0040***	0.0010	0.0002	-0.0055	0.0023	-0.0020	0.0045
	(0.0028)	(0.0031)	(0.0029)	(0.0026)	(0.0028)	(0.0028)	(0.0030)
Bilingual education	-0.0018	0.0001	0.0011	0.0003	0.0009	0.0006	-0.0030*
	(0.0026)	(0.0027)	(0.0029)	(0.0026)	(0.0032)	(0.0029)	(0.0027)
Lives with biological parent	-0.0012*	-0.0000	0.0046	0.0012	0.0016	0.0052	-0.0063
	(0.0023)	(0.0025)	(0.0024)	(0.0023)	(0.0025)	(0.0023)	(0.0026)
Special education	-0.0025	0.0097	-0.0182	0.0101	0.0011	0.0036*	-0.0122
	(0.0079)	(0.0091)	(0.0100)	(0.0090)	(0.0071)	(0.0063)	(0.0093)
Special education x year trend	0.0004	-0.0006	0.0024	-0.0003*	0.0001	-0.0007	0.0009
	(0.0009)	(0.0009)	(0.0011)	(0.0010)	(0.0008)	(0.0008)	(0.0010)
8th grade ITBS math z-score	-0.0046	0.0001	0.0025	-0.0002	-0.0010*	0.0013	0.0020*
	(0.0021)	(0.0024)	(0.0027)	(0.0021)	(0.0022)	(0.0020)	(0.0025)
8th grade ITBS reading z-score	0.0019	-0.0041	-0.0017	-0.0004	0.0012	0.0016	0.0002*
	(0.0018)	(0.0018)	(0.0016)	(0.0016)	(0.0017)	(0.0016)	(0.0020)
Classified as disruptive	-0.0027	0.0061	0.0012	0.0002*	-0.0012***	0.0009	-0.0029
	(0.0033)	(0.0033)	(0.0035)	(0.0029)	(0.0034)	(0.0030)	(0.0032)
Neighborhood median family income	0.0000	-0.0000	0.0000**	-0.0000	-0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Neighborhood percent school age (5-18)	0.0023	-0.0248	0.0210	0.0179	-0.0088	-0.0200**	0.0023
	(0.0201)	(0.0216)	(0.0205)	(0.0194)	(0.0182)	(0.0181)	(0.0188)
Neighborhood percent Hispanic	0.0035	0.0021	-0.0079**	-0.0100	0.0038	-0.0061	0.0075*
	(0.0085)	(0.0095)	(0.0108)	(0.0077)	(0.0099)	(0.0089)	(0.0100)
Neighborhood percent black	0.0001	0.0056	-0.0092*	-0.0021	0.0023	-0.0029	0.0028
	(0.0076)	(0.0081)	(0.0082)	(0.0059)	(0.0081)	(0.0076)	(0.0083)
Neighborhood mean education	0.0008	-0.0008	0.0011	-0.0012	0.0017	0.0004	-0.0011**
	(0.0014)	(0.0016)	(0.0016)	(0.0013)	(0.0016)	(0.0014)	(0.0013)
Neighborhood percent in poverty	0.0004	-0.0152	0.0126	-0.0014	-0.0074	0.0167	-0.0041
	(0.0107)	(0.0118)	(0.0103)	(0.0104)	(0.0104)	(0.0091)	(0.0116)
Observations (student-by-year)	138,386	138,386	138,386	138,386	138,386	138,386	138,386

Notes: Each regression specification corresponds to the OLS specification in column (2) of Table 5, panels A and B, which includes middle school-by-teacher-by-year fixed effects. Also included is a quadratic function in course subject enrollment by school and semester, as well as an indicator for the fall semester. Numbers in parentheses represent standard errors clustered at the high school-by-year level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table A3: Instrument Validity Regressions for Period of the Day Instruments (Remedial English I Sample)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Dependent Variable: Period of the day is …						
	1st	2nd	3rd	4th	5th	6th	7th
Age	0.0015	0.0015	-0.0002	-0.0012	-0.0014	0.0008	0.0028
	(0.0011)	(0.0014)	(0.0013)	(0.0013)	(0.0012)	(0.0013)	(0.0012)
Male	-0.0005	-0.0000	0.0013	-0.0015	0.0004	0.0005	0.0012
	(0.0018)	(0.0018)	(0.0018)	(0.0017)	(0.0016)	(0.0016)	(0.0018)
White	0.0025	0.0005	0.0000	-0.0008	-0.0043	-0.0004	0.0020
	(0.0022)	(0.0035)	(0.0032)	(0.0035)	(0.0033)	(0.0032)	(0.0033)
Black	-0.0006	-0.0001	-0.0087	0.0031*	-0.0020	0.0020	0.0036
	(0.0029)	(0.0034)	(0.0035)	(0.0035)	(0.0035)	(0.0037)	(0.0037)
American Indian	-0.0146	0.0074	-0.0304	0.0046	0.0130	0.0006	-0.0188**
	(0.0166)	(0.0190)	(0.0194)	(0.0209)	(0.0181)	(0.0212)	(0.0194)
Asian	0.0049	0.0046	0.0104	0.0038	0.0028	-0.0080	-0.0074
	(0.0040)	(0.0069)	(0.0075)	(0.0059)	(0.0056)	(0.0061)	(0.0049)
Free or reduced lunch	0.0007***	0.0007	-0.0012	0.0005	0.0002	0.0007	-0.0025
	(0.0018)	(0.0022)	(0.0020)	(0.0018)	(0.0019)	(0.0019)	(0.0022)
Bilingual education	0.0007	0.0026	-0.0043	-0.0014	-0.0012	-0.0018	0.0028*
	(0.0017)	(0.0019)	(0.0021)	(0.0020)	(0.0022)	(0.0022)	(0.0021)
Lives with biological parent	0.0008*	-0.0002	-0.0033	0.0018	-0.0044	0.0007	0.0026
	(0.0016)	(0.0019)	(0.0018)	(0.0017)	(0.0018)	(0.0017)	(0.0017)
Special education	-0.0035	-0.0038	-0.0019	-0.0051	-0.0091	0.0133*	-0.0053
	(0.0075)	(0.0101)	(0.0097)	(0.0089)	(0.0080)	(0.0090)	(0.0081)
Special education x year trend	0.0005	0.0001	0.0002	0.0008*	0.0007	-0.0009	0.0003
	(0.0008)	(0.0010)	(0.0010)	(0.0009)	(0.0009)	(0.0009)	(0.0009)
8th grade ITBS math z-score	0.0010	-0.0013	-0.0012	0.0002	0.0000*	0.0014	0.0001*
	(0.0013)	(0.0014)	(0.0014)	(0.0013)	(0.0013)	(0.0013)	(0.0014)
8th grade ITBS reading z-score	-0.0006	-0.0008	0.0006	-0.0001	0.0024	0.0013	-0.0025*
	(0.0013)	(0.0016)	(0.0014)	(0.0016)	(0.0014)	(0.0014)	(0.0016)
Classified as disruptive	-0.0019	0.0049	-0.0017	0.0017*	-0.0006***	-0.0015	-0.0006
	(0.0020)	(0.0022)	(0.0023)	(0.0019)	(0.0020)	(0.0021)	(0.0021)
Neighborhood median family income	0.0000	0.0000	-0.0000**	0.0000	-0.0000	-0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Neighborhood percent school age (5-18)	-0.0100	-0.0121	0.0142	0.0116	0.0171	-0.0006**	0.0311
	(0.0122)	(0.0141)	(0.0147)	(0.0139)	(0.0127)	(0.0129)	(0.0138)
Neighborhood percent Hispanic	0.0046	-0.0013	-0.0066**	0.0035	-0.0001	-0.0067	-0.0073*
	(0.0055)	(0.0071)	(0.0076)	(0.0072)	(0.0066)	(0.0072)	(0.0072)
Neighborhood percent black	0.0025	0.0112	-0.0033*	-0.0007	-0.0065	-0.0056	-0.0069
	(0.0049)	(0.0064)	(0.0063)	(0.0059)	(0.0055)	(0.0058)	(0.0054)
Neighborhood mean education	0.0006	-0.0009	0.0001	0.0010	0.0011	-0.0010	-0.0005**
	(0.0010)	(0.0011)	(0.0012)	(0.0011)	(0.0011)	(0.0010)	(0.0011)
Neighborhood percent in poverty	0.0098	0.0072	-0.0005	-0.0015	0.0038	-0.0024	-0.0088
	(0.0065)	(0.0078)	(0.0075)	(0.0062)	(0.0068)	(0.0070)	(0.0067)
Observations (student-by-year)	300,962	300,962	300,962	300,962	300,962	300,962	300,962

Notes: Each regression specification corresponds to the OLS specification in column (2) of Table 5, panels C and D, which includes middle school-by-teacher-by-year fixed effects. Also included is a quadratic function in course subject enrollment by school and semester, as well as an indicator for the fall semester. Numbers in parentheses represent standard errors clustered at the high school-by-year level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Appendix B: Selected First-Stage Results

Table B1: First-Stage Results for Full Sample Column (4) Regression Specifications of Tables 3 and 4

	Table 3: Regular Algebra I				Table 4: Regular English I			
	Panel A: Course Passing		Panel B: Math Test Scores		Panel A: Course Passing		Panel B: Reading Test Scores	
	Composition	Size	Composition	Size	Composition	Size	Composition	Size
Predicted Class Size	-0.222**	-0.108*	-0.171	-0.064	-0.209	-0.057	-0.266*	-0.071
	(0.101)	(0.057)	(0.117)	(0.055)	(0.131)	(0.064)	(0.152)	(0.070)
Period 2	0.545***	0.088	0.674***	0.112	0.445**	-0.216	0.341	-0.229
	(0.204)	(0.232)	(0.231)	(0.260)	(0.205)	(0.208)	(0.228)	(0.226)
Period 3	1.038***	0.389*	1.352***	0.583**	0.930***	0.165	0.936***	0.219
	(0.221)	(0.231)	(0.254)	(0.257)	(0.218)	(0.217)	(0.247)	(0.247)
Period 4	0.805***	0.020	1.012***	0.112	0.253	-0.345	0.236	-0.226
	(0.222)	(0.238)	(0.254)	(0.271)	(0.222)	(0.217)	(0.247)	(0.236)
Period 5	0.642***	0.251	0.755**	0.231	0.358	-0.514**	0.335	-0.446*
	(0.247)	(0.275)	(0.300)	(0.331)	(0.225)	(0.225)	(0.250)	(0.256)
Period 6	0.462**	-0.257	0.746***	-0.021	0.313	-0.602***	0.236	-0.662***
	(0.225)	(0.244)	(0.268)	(0.284)	(0.214)	(0.210)	(0.235)	(0.231)
Period 7	0.209	-0.237	0.260	-0.122	0.087	-0.450**	0.008	-0.490**
	(0.209)	(0.247)	(0.252)	(0.295)	(0.208)	(0.205)	(0.232)	(0.224)
Obs (student-by-year)	252,771		179,734		216,715		160,837	
F-statistic	4.94***	2.56**	5.88***	1.90*	3.87***	3.06***	3.29***	2.86***

Notes: Period 1 is the omitted variable in all regression specifications. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. First-stage F-statistics represent the F-statistics on the excluded instruments for both first-stages (Composition and Size) using standard errors clustered at the high school-by-year level.