MAKING THE GRADE: THE IMPACT OF CLASSROOM BEHAVIOR 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: August 14, 2012

Abstract

Using exogenous variation on course scheduling in Chicago Public Schools, we examine empirical implications of Lazear's (2001) educational production model. Our identification strategy allows us to investigate an underlying mechanism by which class size affects student performance, the behavioral composition of a classroom. Consistent with the Lazear framework, we find that an additional non-disruptive student in attendance increases the probability of passing English I and Algebra I, with larger effects for students in remedial versus regular classes. For regular English I students, we estimate a positive relationship between the number of non-disruptive students in attendance and own reading test score.

JEL Classifications: H75, I21, J13
Keywords: Educational Production, Class Size, Classroom Composition, 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. Heather Hough, Jeffrey Kubik, Lori Taylor, Yisu Zhou, and participants at the Association for Education Finance and Policy conference provided helpful comments and suggestions. 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 for 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 they 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 behavior, 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. 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 ninth graders in Chicago Public Schools (CPS), we study heterogeneity in the impact of class size on student achievement in reading and mathematics. Our identification strategy allows us to analyze an underlying mechanism by which class size affects student performance, the behavioral composition of a classroom. Our classroom composition measure is constructed as the average number of non-disruptive students in attendance on a given school day; we characterize a classroom as being comprised of disruptive and non-disruptive students, where a student is considered disruptive if he or she dropped 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 75 percent of CPS students receive federal lunch subsidies, and the racial/ethnic makeup of the student body is approximately 55 percent black; 30 percent Hispanic; and the remaining 15 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 reading and mathematics, 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 the behavioral composition of a classroom on academic achievement, we use an instrumental variables approach, exploiting exogenous variation in the period of the day a course is offered in CPS. Our analysis focuses on students' course passing and test scores in four ninth grade course subjects: regular English I, remedial English I, regular Algebra I, and remedial Algebra I.

The Lazear theoretical model of educational production suggests that the behavioral composition of a classroom is an important determinant of educational achievement, implying that it should be included as an explanatory variable when estimating an educational production function. Our two-stage least squares (2SLS) results are in line with the Lazear framework, as we find that the behavioral composition of a classroom significantly affects student achievement.

For the full sample of students enrolled in regular English I, our 2SLS estimates indicate that an additional non-disruptive student in attendance on a given school day increases the probability of passing English I by 7.26 percentage points, relative to the mean passing rate of 76.8 percent. Disaggregating this effect by race and ethnicity, we find a larger increase in the probability of passing English I for black versus Hispanic students. We also observe that an additional non-disruptive student in attendance increases a student's own reading test score by 0.0222 student-level standard deviations for the full sample and by 0.0633 student-level standard deviations for Hispanic students; we estimate an insignificant effect on reading scores for black students.

Our full sample and Hispanic subsample effect sizes for reading test scores are similar in magnitude to the effect sizes reported in Finn and Achilles (1990) for first grade students that participated in Tennessee's Project STAR class-size randomized experiment. For the full sample, Finn and Achilles (1990) estimate an effect size for reading test scores of approximately 0.0275 student-level standard deviations for a one-student reduction in class size (based on Table 5, page 566). For minority students, Finn and Achilles (1990) estimate an effect size of approximately 0.0400 student-level standard deviations (based on Table 6, page 567). The full sample effect sizes are close in magnitude across the two studies, while the minority subsample effect size in Finn and Achilles (1990) is approximately two thirds of the magnitude of our Hispanic subsample effect size.

Our estimated impacts of the behavioral composition of a classroom on course passing for students enrolled in regular Algebra I are smaller in magnitude than the corresponding effects for regular English I. For the regular Algebra I sample, we find that an additional non-disruptive student in attendance increases the probability of passing Algebra I by 3.92 percentage points, relative to the mean passing rate of 72.4 percent. Estimates for the black and Hispanic subsamples are of approximately the same magnitude as the full sample estimate.

The Lazear model also suggests that the effects of the behavioral composition of a classroom on student outcomes should be larger for students enrolled in remedial versus regular courses because students in remedial courses have, on average, lower baseline academic performance than students in regular classes. Consistent with this, we find larger overall effects of classroom composition on English I and Algebra I course passing for students enrolled in remedial versus regular classes.

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]

[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

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 sizes 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

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.

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 the behavioral composition of a classroom on course passing and test scores for students in four ninth grade CPS course subjects: regular English I, remedial English I, regular Algebra I, and remedial Algebra 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), excessive absences, and uniform discipline code violation (infringement of the CPS code of conduct). Consider the following linear specification:

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

where Y_{ict} denotes one of two outcome variables for student *i* enrolled in class *c* in the fall or spring semester of academic year *t*: the receipt of a grade of D or better in a particular course or the test score in a subject-relevant (reading or mathematics) standardized examination. The explanatory variable of interest is $Composition_{ct}$, 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 academic year *t* for class *c*. With 90 school days in each semester, $(90 - Absences_{ict})$ is the number of days that student *i* was in attendance in a particular semester in academic year *t* for class *c*. This variable is then summed over every non-disruptive student in the class. This sum is divided by 90 to obtain the daily average number of non-disruptive students in attendance in class *c* in a given semester in academic year *t*, $Composition_{ct}$.

X_{it} is a vector of observable student-specific characteristics, which includes the subject-specific eighth grade Iowa Test of Basic Skills (ITBS) test score measured in student-level standard deviations, demographic variables, and neighborhood variables measured at the level of the Census block group. 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. Three sets of fixed effects are used to capture time-variant, unobserved high school (school attended in ninth grade), middle school (school attended in eighth grade), and/or high school teacher (ninth grade teacher) quality. Specifically, we include, in separate specifications, high school-by-semester fixed effects (a separate fixed effect for each high school in each semester of each academic year), middle school-by-high school-by-semester fixed effects (a separate fixed effect for each combination of middle school and high school in each semester of each academic year), and teacher-by-semester fixed effects (a separate fixed effect for each high school teacher in each semester of each academic year).[2]

The inclusion of these fixed effects ensures that we estimate the effect of the behavioral composition of a classroom on student achievement using only variation in classroom compositions within a given high school in a given semester (first set of fixed effects); within a given middle and high school combination in a given semester (second set of fixed effects), or for a given high school teacher in a given semester (third set of fixed effects). ε_{ict} represents the idiosyncratic error term.

Ordinary least squares (OLS) estimation of equation (1) leads to biased estimates of the effects of the behavioral composition of a classroom on student achievement because our

[2] We do not estimate models with middle school-by-teacher-by-semester fixed effects because of the very large number of singleton groups (i.e., fixed effects with exactly one observation), and thus much smaller effective sample sizes, when using these fixed effects.

explanatory variable of interest, $Composition_{ct}$, is a function of the number of student 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.[3] This implies that the coefficient on classroom composition from an OLS regression has a downward bias.

To eliminate this bias, and hence to estimate the causal effect of classroom composition on student outcomes, we use an instrumental variables regression approach, exploiting exogenous variation in the ordering of classes over the school day. Our set of excluded instrumental variables for $Composition_{ct}$ is $Period_{ct}$ and $Period_{ct}^2$, a quadratic function in the period of the school day in which a class is scheduled. This identification strategy is similar to the one used by Cortes, Bricker, and Rohlfs (2012), in which they use effectively random variation in course scheduling 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. In Section IV, we show graphically that student absences in a particular class vary depending on the period of the day in which the class is offered. Moreover, our measure of the behavioral composition of a classroom generally exhibits an inverted U-shaped pattern when plotted against the period of the school day in which the course is offered, implying that the excluded instrumental variables are highly correlated with classroom composition.

Equation (1) is now the second-stage equation, and the first-stage equation is:

$$(3)\quad Composition_{ct} = \alpha_0 + \alpha_1 Period_{ct} + \alpha_2 Period_{ct}^2 + \theta' X_{it} + \eta_{jkt} + \upsilon_{ict},$$

where the endogenous variable, $Composition_{ct}$, is a function of the excluded instruments, as well as the control variables and fixed effects that appear in equation (1). In Section V, we show that

[3] Our data lend support to this hypothesis. For the regular English I sample, the raw correlation between eighth grade ITBS reading score and the number of ninth grade absences is -0.1569. For the regular Algebra I sample, the raw correlation between eighth grade ITBS mathematics score and the number of ninth grade absences is -0.1918. The corresponding raw correlations for the remedial English I and the remedial Algebra I samples are -0.1325 and -0.1850, respectively.

period of the day and period of the day squared are statistically significant and strong predictors of the behavioral composition of a classroom in almost all of our empirical specifications.

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 ninth 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 titles and numbers, 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 group in which each student resides.

The standardized tests that were administered, and the scores 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, eighth grade reading and mathematics test scores are available from the Iowa Test of Basic Skills (ITBS) for each year of our data. The ninth grade test score data for reading and mathematics are taken from the TAP (Test of Achievement and Proficiency) for the 1993-94 to the 2001-02 school years and from the EXPLORE test for the 2002-03 to the 2005-06 school years. 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 for the analytic samples are given in Table 1. We focus on students enrolled in general (i.e., non-vocational, non-magnet, and non-alternative) high schools.[4] The student-level outcome variables of interest are an indicator for passing regular or remedial English I, an indicator for passing regular or remedial Algebra I, the score on the standardized reading examination, and the score on the standardized mathematics examination. In accordance with CPS policy, we defined a student as having passed a course if he or she received a grade of D or better in that course.[5] Panel A of Table 1 shows that students enrolled in remedial classes have lower course passing rates in both English I and Algebra I, as compared to students in regular classes. The passing rates of students enrolled in regular English I and Algebra I are 77 and 72 percent, respectively, while the passing rates of students enrolled in remedial English I and Algebra I are 75 and 69 percent, respectively.

The lower academic performance of students in remedial classes can also be observed in their ninth grade reading and mathematics test scores. The average ninth grade reading score for the regular English I sample is 0.080 student-level standard deviations versus -0.327 student-level standard deviations for the remedial English I sample. Likewise, the average ninth grade mathematics score for the regular Algebra I sample is 0.076 student-level standard deviations, as compared to -0.424 student-level standard deviations for the remedial Algebra I sample.

This difference in academic performance is also seen in the students' baseline (eighth grade) performance on reading and mathematics examinations, as shown in the last two rows of panel C. The average baseline reading score for the regular English I sample is 0.080 student-level standard deviations, in contrast to -0.352 student-level standard deviations for the remedial English I sample.

---

[4] We also restrict the analysis to first-time ninth grade students; for consistency, if a student repeated ninth grade one or more times, we only use his or her first instance of ninth grade in the data.

[5] Chicago Public Schools Policy Manual Board Report 04-0128-PO1 (January 28, 2004)

Similarly, the average baseline mathematics score for the regular Algebra I sample is 0.042 student-level standard deviations, as opposed to -0.545 student-level standard deviations for the remedial Algebra I sample.

While student achievement is lower for students enrolled in remedial versus 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 ninth graders in all course subjects is 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 34 and 37 percent; and white, American Indian, and Asian students together account for the remaining eight to 12 percent. Eighty-four percent of ninth graders receive free or reduced lunch, and their proportion across course subjects is fairly stable, ranging between 81 and 87 percent. We find that the proportion of students in special education programs is higher in remedial versus regular classes (21 percent compared to 15 percent). Lastly, the neighborhood characteristics of a student's residence (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 the behavioral composition of a classroom on student academic performance, we now make the case for our excluded instrumental variables: period of the day and period of the day squared. 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.[6] Moreover, our analysis focuses on English I and Algebra I, required courses that are offered multiple times during every period of the day. A testable

[6] In private discussions, school administrators have indicated that the process is computerized and essentially random.

implication of the course scheduling process is that student and classroom characteristics should be similar between classes that meet in a particular period of the day and those that meet at other times. In other words, we assert that students enrolled in a given course subject in a given period are otherwise similar to students who take a course in that subject at another time during 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 focus on English and mathematics 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. As Table 2A shows, the breakdown of classes by subject is generally stable over the course of the day; for a particular subject, the percentage of course offerings in that subject differs by, at most, two percentage points across periods. This implies that schools do not appear to systematically schedule academic subjects in certain periods, such as those with lower absence rates (we return to this point later in this section).

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) English I or Algebra I on the student- and neighborhood-level control variables used in our outcome regressions, as well as teacher-by-semester fixed effects. Table 2B reports coefficients from such linear probability models for the sample of students enrolled in regular English I; each column is for a different period of the day.[7] Almost all coefficients are statistically insignificant, and the coefficients that are significant show no apparent pattern across columns.

We next examine the raw, reduced-form (i.e., first-stage) effects of having English I and Algebra I in a particular period on the average number of non-disruptive students in attendance in that period. Panels A and B of Figure 1 show the relationship between period of the day and the endogenous variable of our model, the behavioral composition of a classroom. Panel A gives the average numbers of non-disruptive students in attendance for regular and remedial English I as functions of the period of the day during which these classes meet. The corresponding Algebra I graphs 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 longer dashed lines show the means for black students, and the shorter dashed lines show the means for Hispanic students.

These figures indicate that the average numbers of non-disruptive students in attendance for English I and Algebra I are generally at their lowest levels in first period, gradually rise until approximately fourth period, and then gradually decline over the remainder of the school day. This inverted U-shaped pattern is most pronounced for the regular English I, remedial English I, and regular Algebra I full samples, as well as their corresponding black and Hispanic subsamples, providing compelling evidence of a strong relationship between our excluded instrumental variables and our measure of classroom composition. For the full sample, the average number of non-disruptive students in attendance varies across periods from 16.8 to 18.7 for regular English I, from 16.3 to 18.2 for remedial English I, from 17.2 to 18.9 for regular Algebra I, and from 16.9 to 17.7 for remedial Algebra I.

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

V. Empirical Results and Discussion

In this section, we present our empirical analysis of Lazear's (2001) model, which suggests that the behavioral composition of a classroom is a key contributing factor to educational attainment. Tables 3-6 report the OLS and the two-stage least squares (2SLS) regression results for equation (1), which gives the effects of classroom composition on course passing and test scores. These tables show the estimated coefficient on our measure of classroom composition, as well as the F-statistic for the test of the predictive power of the excluded instruments in first-stage equation (3).

Table 3 presents the results for the regular English I sample; the dependent variable in panel A is an indicator for whether a student received a passing grade in his or her English I course, and the dependent variable in panel B is his or her z-score on the standardized reading examination. Table 4 reports the findings for the regular Algebra I sample; the dependent variable in panel A is an indicator for whether a student received a passing grade in his or her Algebra I class, and the dependent variable in panel B is his or her z-score on the standardized mathematics examination. Tables 5 and 6 display the estimates for the remedial English I and remedial Algebra I samples, respectively. Within each panel, the topmost set of results is for the full (pooled black, Hispanic, white, Asian, and American Indian) sample, the next set is for the black subsample, and the bottommost set is for the Hispanic subsample. We focus on these subgroups because they together comprise 90 percent of our analytic sample (Table 1).

Each column reports the results for a different regression specification. All specifications include student- and neighborhood-level characteristics as baseline controls (listed in the footnotes to Tables 3-6, as well as summarized in Table 1); however, different sets of fixed effects are included in Columns (1)-(3). Specifically, the column layouts are as follows: Column (1) contains high school-by-semester fixed effects, Column (2) replaces the previous set of fixed effects with middle school-by-high school-by-semester fixed effects, and Column (3) includes only teacher-by-semester fixed effects.[8] Since teachers have discretion in determining course grades, for the course passing results we focus on the empirical specification in Column (3) because it controls for teacher fixed effects. For the test score results we focus on the empirical specification in Column (2) because it contains high school fixed effects and also 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. First-Stage Results

Before turning to our estimates of equation (1), we provide further evidence for the validity of our excluded instruments. Table 2C presents estimates of first-stage equation (3) using the full student samples and teacher-by-semester fixed effects. The first two columns of Table 2C display the coefficients on period of the day and period of the day squared using the full regular English I sample from the course passing regression (first column) and from the reading test score regression (second column). These first-stage coefficients correspond to the 2SLS estimates in the top row of Column (3) in each panel of Table 3. The third and fourth columns of Table 2C are for the full regular Algebra I sample (corresponding to Table 4), the fifth and sixth columns are for the full remedial English I sample (corresponding to Table 5), and the last two columns are for the full remedial Algebra I sample (corresponding to Table 6).

In each case, the coefficient on the squared term is negative, implying an estimated inverted-U relationship between period of the day and the behavioral composition of a classroom. The F-statistics for the tests of the joint significance of the excluded instruments are statistically significant at the one percent level and are almost always larger than 10. Consistent with Figure 1, each turning point in the relationship between period of the day and classroom composition (i.e., the particular period of the day at which classroom composition is at its maximum, based on the coefficient

[8] The numbers of observations are different across columns for a given sample because singleton groups (i.e., fixed effects with exactly one observation) were dropped in the estimation, reducing numbers of observations to those shown.

estimates for period and period squared) is approximately fourth period.[9] Overall, Table 2C, along with our discussion in the previous section, provides convincing support for our use of a quadratic function in the period of day as an instrument for classroom composition.

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

We begin with the OLS results in panel A of Table 3, observing in all three columns a very small but mostly statistically significant association between the behavioral composition of a classroom and the probability that a student passes regular English I. In each case, we find that an additional non-disruptive student in attendance on a given school day is associated with a change in the probability of passing English I of less than one half of one percentage point (-0.16 to +0.33 percentage points).

As discussed in Section III, the endogeneity of the behavioral composition of a classroom implies that OLS estimation of equation (1) leads to downwardly biased estimates of the effects of classroom composition on student achievement. As a result, we now turn to our 2SLS estimation results. For each 2SLS regression, we report the F-statistic for the test of the predictive power of the excluded instruments in the first-stage equation. For the full sample and black subsample, we observe large and statistically significant first-stage F-statistics; all are greater than 10. Due to the smaller sample sizes, the first-stage F-statistics for the Hispanic subsample are not as large, but they are still statistically significant at the five percent level or better.

After instrumenting with period of the day and period of the day squared, all of the 2SLS coefficients on classroom composition are larger in magnitude relative to the corresponding OLS coefficients and are statistically significant at the five percent level or better. These findings are

[9] Each turning point is computed as the negative of the coefficient on the linear term divided by twice the coefficient on the squared term.

consistent with the Lazear framework in that the behavioral composition of a classroom is an important determinant of the likelihood that a student passes regular English I.

For the full sample, we estimate that an additional non-disruptive student in attendance increases the probability of passing English I by 6.36 to 7.26 percentage points. As stated earlier, our preferred model specification is Column (3), which includes teacher-by-semester fixed effects. Relative to the mean passing rate of 76.8 percent, the estimated coefficient of 0.0726 in Column (3) translates into an increase of 9.45 percent (0.0726/0.768=0.0945) in the probability of passing English I. When we break down this effect by race and ethnicity, we find a larger increase in the probability of passing English I for black versus Hispanic students: based on the results in Column (3), an extra non-disruptive student in attendance increases the probability of passing English I by 8.53 percentage points for black students (relative to the mean of 75.3 percent) and by 5.58 percentage points for Hispanic students (relative to the mean of 77.6 percent).

Turning next to the OLS results in panel B of Table 3, we observe in all three specifications a positive and at least marginally significant association between the behavioral composition of a classroom and reading test scores for regular English I students. For the full sample, we find that an additional non-disruptive student in attendance is associated with an increase of 0.0060 to 0.0081 student-level standard deviations in the student's reading score, which is less than one percentile point. The OLS coefficients on classroom composition are larger in magnitude for black students and smaller in magnitude for Hispanic students.

Moving to the 2SLS estimation results, for the full sample and black subsample, we again observe large and statistically significant first-stage F-statistics; all are above 10. The F-statistics for the Hispanic subsample are again smaller, but they are still significant at the five percent level or better. All of the 2SLS coefficients on classroom composition for the full sample and Hispanic subsample are larger in magnitude, as compared to the corresponding OLS coefficients, and are statistically significant at the five percent level (except in Column (3) for the full sample).

Focusing on our preferred model specification in Column (2), which controls for middle school-by-high school-by-semester fixed effects, we estimate that an additional non-disruptive student in attendance leads to a 0.0222 student-level standard deviation increase in reading test scores for the full sample (approximately one half to one percentile point) and a 0.0633 student-level standard deviation increase in reading test scores for Hispanic students (approximately two percentile points). The estimated impacts on reading test scores for black students are smaller in magnitude than the estimated effects for Hispanic students and are statistically insignificant in all model specifications.

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

We focus on the 2SLS results in this subsection because the OLS estimates for the regular Algebra I sample have approximately the same magnitude as the OLS estimates for the regular English I sample. The course passing results are reported in panel A of Table 4. All of the 2SLS coefficients on classroom composition are positive and larger in magnitude than the analogous OLS coefficients and are statistically significant at the five percent level or better, again in accordance with the Lazear model. For the full sample, we see that an additional non-disruptive student in attendance increases the probability of passing Algebra I by 3.92 to 5.19 percentage points. The estimated coefficient of 0.0392 in Column (3), our preferred specification, implies an increase of 5.41 percent in the probability of passing Algebra I relative to the mean passing rate of 72.4 percent. This percent increase in the probability of passing regular Algebra I is approximately half the magnitude of the corresponding percent increase in the probability of passing regular English I (9.45 percent). Estimates for the black and Hispanic subsamples in Column (3) are roughly the same as the estimates for the full sample: an additional non-disruptive student in attendance increases the probability of passing Algebra I by 4.21 percentage points for black students (relative to the mean of 70.2 percent) and by 4.32 percentage points for Hispanic students (relative to the mean of 74.4 percent).

The 2SLS results for mathematics test scores are shown in panel B of Table 4. While all first-stage F-statistics are statistically significant at the one percent level and are above 20 for the full sample and black subsample, we estimate small and statistically insignificant effects of the behavioral composition of a classroom on mathematics test scores in all samples and specifications. The one exception is the positive and marginally significant effect of classroom composition on mathematics test scores for black students in Column (3) when including teacher-by-semester fixed effects. This estimated effect size of 0.0114 student-level standard deviations is smaller than the statistically significant effect sizes for reading test scores in panel B of Table 3.

D. Course Passing and Test Score Results for Remedial Classes

Another empirical implication of the Lazear theoretical framework is that the effects of the behavioral composition of a classroom on student achievement should be larger for students enrolled in remedial versus regular courses because students in remedial courses have lower average baseline academic performance. Table 5 presents the results for students enrolled in remedial English I. We again concentrate on the 2SLS estimates.

Starting with the course passing results in panel A of Table 5, we find that all first-stage F-statistics are statistically significant at the one percent level and are greater than 10 for the full sample and black subsample, other than in Column (3) for the black subsample. Each 2SLS coefficient on classroom composition is larger in magnitude than the corresponding OLS coefficient and is always significant at the one percent level. For the full sample in Column (3), we estimate that an additional non-disruptive student in attendance increases the probability of passing English I by 10.50 percentage points. Relative to the mean passing rate of 75.0 percent, the estimated coefficient of 0.1050 translates into an increase of 14.00 percent in the probability of passing English I. This percent increase in the probability of passing remedial English I is approximately 50 percent larger than the magnitude of the analogous percent increase in the probability of passing regular English I (9.45 percent), in line with the Lazear model.

As with the regular English I course passing analysis, we find a larger increase in the probability of passing remedial English I for the black versus Hispanic subsamples: the results in Column (3) imply that an extra non-disruptive student in attendance increases the probability of passing English I by 14.81 percentage points for black students (relative to the mean of 72.4 percent) and by 5.41 percentage points for Hispanic students (relative to the mean of 77.9 percent). For black students, the effects on course passing are larger for remedial versus regular English I, while for Hispanic students the impacts are approximately the same.

The 2SLS results for reading test scores are in panel B of Table 5. While all first-stage F-statistics are statistically significant at the one percent level and are above 10 for the full sample, we estimate small and statistically insignificant effects of classroom composition on reading test scores.

We now discuss the results for students enrolled in remedial Algebra I, which are presented in Table 6. In both panels, the first-stage F-statistics are generally smaller than the corresponding first-stage F-statistics in Tables 3-5, reflecting a weaker relationship between classroom composition and a quadratic function in the period of the day for the remedial Algebra I sample. This is evident from the right graph in panel B of Figure 1. Most first-stage F-statistics are statistically significant at the five percent level or better for the full sample and Hispanic subsample, whereas only the first-stage F-statistics in Column (3) are at least marginally significant for the black subsample.

Focusing on the 2SLS estimates, we begin with the course passing results in panel A of Table 6. For the full sample in Column (3), we estimate that an additional non-disruptive student in attendance increases the probability of passing Algebra I by 6.84 percentage points. Relative to the mean passing rate of 68.5 percent, the estimated coefficient of 0.0684 translates into an increase of 9.99 percent in the probability of passing Algebra I. This percent increase in the probability of passing remedial Algebra I is approximately twice the magnitude of the corresponding percent increase in the probability of passing regular Algebra I (5.41 percent), in agreement with the Lazear model.

We find a larger increase in the probability of passing remedial Algebra I for the black versus Hispanic subsamples: the estimated coefficients in Column (3) imply that an extra non-disruptive student in attendance increases the probability of passing Algebra I by 11.60 percentage points for black students (relative to the mean of 67.3 percent) and by 2.08 percentage points for Hispanic students (relative to the mean of 70.2 percent). The former effect is significant at the five percent level, while the latter effect is significant at the 10 percent level. For black students, the effects on course passing are larger for remedial versus regular Algebra I, while the opposite is true for Hispanic students.

Panel B of Table 6 displays the results for mathematics test scores. In all cases, we estimate statistically insignificant effects of classroom composition on mathematics test scores. The estimated effect sizes are generally small for models with more precisely estimated first-stage relationships.

E. Comparisons of Effect Sizes

It is instructive to compare our estimated effects of classroom composition on test scores with the estimated effects of class size on test scores from Tennessee's Project STAR class-size randomized experiment, as reported in Finn and Achilles (1990). An important difference between our study and that of Finn and Achilles (1990) is the measurement of the classroom-level variable of interest: our classroom composition measure is a specific type of "effective" class size that is based on the observed attendance records of non-disruptive students in a given class, whereas the class sizes studied in Finn and Achilles (1990) are "roster" class sizes based on the number of students officially enrolled in/assigned to a particular class. We generally estimate positive effects of classroom composition on test scores, while Finn and Achilles (1990) find negative effects of class size on test scores, implying that the analysis below will be a comparison of the magnitudes of the effect sizes across the two studies.

We focus on the effect sizes for our regular English I full sample and Hispanic subsample because, for all other samples, we estimate statistically insignificant effects of classroom composition on test scores when using our preferred 2SLS specification. For the regular English I full sample, the estimated 2SLS coefficient on classroom composition for the reading test score regression is 0.0222 (Column (2) in panel B of Table 3), which implies an effect size of 0.0222 student-level standard deviations for a one-student *increase* in the number of non-disruptive students in attendance. For the regular English I Hispanic subsample, the estimated 2SLS coefficient on classroom composition for the reading test score regression is 0.0633 (Column (2) in panel B of Table 3), which indicates an effect size of 0.0633 student-level standard deviations for a one-student *increase* in the number of non-disruptive students in attendance.

Finn and Achilles (1990) report effect sizes for first grade students, disaggregated by examination subject. Because their 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 ours. Finn and Achilles (1990) present results for three standardized examinations in reading: the Basic Skills First (BSF) reading, the Stanford Achievement Test (SAT) word study skills, and the SAT reading examinations. For these examinations, the authors find full sample effect sizes of 0.21, 0.22, and 0.23 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 0.0263, 0.0275, and 0.0288 student-level standard deviations for an approximately one-student *reduction* in class size. These numbers are very similar in magnitude to our full sample effect size of 0.0222 student-level standard deviations for a one-student increase in the number of non-disruptive students in attendance.

Finn and Achilles (1990) also report effect sizes by minority status (Table 6, page 567), finding larger effect sizes for minority students than for white students.[10] For the BSF reading, the SAT word study skills, and the SAT reading examinations, the effect sizes for minority students are 0.35, 0.32, and 0.35 student-level standard deviations, respectively, for an eight-student reduction in class size, translating into effect sizes of 0.0438, 0.0400, and 0.0438 student-level standard deviations for a one-student reduction in class size. These effect sizes are approximately two thirds of the magnitude of our Hispanic subsample effect size of 0.0633 student-level standard deviations for a one-student increase in the number of non-disruptive students in attendance. The effect sizes for white students in Finn and Achilles (1990) are 0.10, 0.16, and 0.15 student-level standard deviations, respectively, for an eight-student reduction in class size, translating into 0.0125, 0.0200, and 0.0188 student-level standard deviations for a one-student reduction in class size.

We also compare our estimated full sample effect size for regular English I to Israeli class size effects reported in Angrist and Lavy (1999). Angrist and Lavy (1999) estimate that a one-student reduction in class size leads to a 0.275 point increase in reading test scores for fifth graders (Table IV, Column (2), page 554), translating into an effect size of about 0.0225 student-level standard deviations for a one-student reduction in class size.[11] The fourth-grade effect size is approximately

[10] Finn and Achilles (1990) do not provide results for more disaggregated racial or ethnic breakdowns.
[11] To obtain the 0.0225 student-level standard deviation effect size for a one-student class size reduction, we divided by eight the 0.18 student-level standard deviation effect size for an eight-student class size reduction reported on page 567.

half this magnitude. We see that the fifth-grade effect size is very similar in magnitude to our full sample effect size of 0.0222 student-level standard deviations for a one-student increase in the number of non-disruptive students in attendance.

VI. Conclusion

Using administrative student transcript files from CPS, we analyze empirical implications of the Lazear educational production model. The Lazear framework suggests that the behavioral composition of a classroom is a central determinant of educational attainment, signifying that it should be included as an explanatory variable when estimating an educational production function. To that end, we exploit exogenous variation on course scheduling in CPS to study heterogeneity in the effect of class size on student achievement in reading and mathematics. Most importantly, our research design permits us to explore an underlying mechanism by which class size affects student achievement, the behavioral composition of a classroom.

In accordance with the theoretical predictions of the Lazear model, we find that, for students enrolled in regular English I, an additional non-disruptive student in attendance increases the probability of passing English I by 7.26 percentage points and raises a student's own reading test score by 0.0222 student-level standard deviations. The estimated impacts of the behavioral composition of a classroom on course passing for students enrolled in regular Algebra I are smaller than the corresponding effects for students enrolled in regular English I. Also consistent with the Lazear framework, we observe larger overall effects of classroom composition on English I and Algebra I course passing for students enrolled in remedial versus regular classes.

References

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. (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 Chris Rohlfs. (2012). 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, 12*(1) (Contributions), Article 27.

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.

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

	Regular English I		Remedial English I		Regular Algebra I		Remedial Algebra I	
Panel A: Outcome Variables								
Course pass rate	0.768	(0.422)	0.750	(0.433)	0.724	(0.447)	0.685	(0.465)
Reading test score (z-score)	0.080	(0.921)	-0.327	(0.718)	-0.008	(0.862)	-0.393	(0.679)
Math test score (z-score)	0.084	(0.912)	-0.274	(0.836)	0.076	(0.880)	-0.424	(0.755)
Panel B: Classroom Characteristic								
Classroom composition	18.111	(6.122)	17.766	(5.873)	18.275	(5.863)	17.565	(5.843)
Panel C: Student Characteristics								
Age	14.225	(0.560)	14.324	(0.554)	14.224	(0.528)	14.319	(0.574)
Male	0.490	(0.500)	0.514	(0.500)	0.496	(0.500)	0.507	(0.500)
White	0.090	(0.286)	0.061	(0.240)	0.083	(0.276)	0.062	(0.242)
Black	0.542	(0.498)	0.550	(0.497)	0.523	(0.499)	0.574	(0.494)
Hispanic	0.342	(0.474)	0.370	(0.483)	0.366	(0.482)	0.350	(0.477)
Asian	0.024	(0.153)	0.017	(0.129)	0.025	(0.157)	0.012	(0.111)
American Indian	0.002	(0.040)	0.001	(0.033)	0.002	(0.039)	0.001	(0.032)
Free or reduced lunch	0.814	(0.390)	0.868	(0.338)	0.841	(0.366)	0.846	(0.361)
Classified as disruptive	0.158	(0.364)	0.104	(0.305)	0.136	(0.343)	0.112	(0.315)
Bilingual education	0.412	(0.585)	0.428	(0.586)	0.450	(0.606)	0.408	(0.600)
Lives with biological parent	0.843	(0.363)	0.802	(0.399)	0.824	(0.381)	0.830	(0.375)
Special education	0.144	(0.351)	0.214	(0.410)	0.148	(0.355)	0.203	(0.402)
8th grade ITBS reading test score (z-score)	0.080	(0.924)	-0.352	(0.783)	0.002	(0.883)	-0.400	(0.765)
8th grade ITBS math test score (z-score)	0.030	(0.928)	-0.314	(0.813)	0.042	(0.888)	-0.545	(0.640)
Panel D: Neighborhood (Census Block Group) Characteristics								
Median family income	31,180	(15,128)	34,119	(16,229)	32,446	(15,781)	32,425	(14,595)
Percent school age (5-18)	0.236	(0.074)	0.242	(0.073)	0.237	(0.074)	0.241	(0.071)
Percent Hispanic	0.259	(0.315)	0.287	(0.333)	0.278	(0.324)	0.270	(0.328)
Percent black	0.487	(0.450)	0.498	(0.448)	0.474	(0.449)	0.514	(0.449)
Mean education	11.916	(1.129)	11.640	(1.357)	11.817	(1.226)	11.734	(1.259)
Percent in poverty	0.249	(0.199)	0.253	(0.189)	0.248	(0.195)	0.251	(0.186)
Observations	237,912		292,136		279,050		134,336	

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

Notes: Regular English I (regular Algebra I) represents the sample of students enrolled in only one regular Algebra I (regular English I) course per semester. Remedial English I (remedial Algebra I) represents the sample of students enrolled in at least one remedial English I (remedial Algebra I) course per semester; some students take regular English I (regular Algebra 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	0.22	0.21	0.21	0.20	0.20	0.20	0.20
Mathematics	0.16	0.16	0.17	0.16	0.16	0.17	0.16
Social Studies	0.14	0.14	0.14	0.14	0.14	0.14	0.14
Science	0.13	0.13	0.13	0.14	0.13	0.14	0.14
Foreign Language	0.03	0.03	0.03	0.03	0.03	0.03	0.03
Shop	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Business	0.05	0.06	0.05	0.06	0.05	0.05	0.05
Vocational	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Art, Music, and Physical Education	0.24	0.25	0.25	0.25	0.26	0.25	0.24
Other	0.00	0.00	0.01	0.00	0.00	0.00	0.00

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

Note: The fractions in each column sum to one.

Table 2B: Determinants of Period of the Day (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	-0.0000	-0.0040***	-0.0023*	0.0010	-0.0004	0.0022*	0.0035***
	(0.0011)	(0.0013)	(0.0012)	(0.0011)	(0.0011)	(0.0012)	(0.0013)
Male	0.0004	0.0024	-0.0000	0.0022*	-0.0031**	-0.0033**	0.0021
	(0.0013)	(0.0016)	(0.0014)	(0.0013)	(0.0012)	(0.0014)	(0.0015)
White	-0.0260*	0.0114	0.0187	0.0215	0.0110	-0.0211*	0.0039
	(0.0139)	(0.0146)	(0.0125)	(0.0140)	(0.0121)	(0.0125)	(0.0144)
Black	-0.0221	0.0078	0.0149	0.0229*	0.0139	-0.0182	0.0041
	(0.0138)	(0.0145)	(0.0121)	(0.0138)	(0.0123)	(0.0127)	(0.0141)
Hispanic	-0.0230*	0.0143	0.0168	0.0172	0.0139	-0.0207*	0.0054
	(0.0137)	(0.0147)	(0.0124)	(0.0140)	(0.0120)	(0.0126)	(0.0144)
Asian	-0.0262*	0.0156	0.0176	0.0154	0.0167	-0.0148	-0.0011
	(0.0144)	(0.0151)	(0.0131)	(0.0144)	(0.0128)	(0.0131)	(0.0153)
Free or reduced lunch	-0.0023*	0.0056***	-0.0035**	-0.0023	-0.0020	-0.0019	0.0042***
	(0.0014)	(0.0018)	(0.0016)	(0.0015)	(0.0014)	(0.0015)	(0.0016)
Classified as disruptive	0.0038**	0.0012	-0.0004	0.0034**	-0.0017	-0.0017	-0.0021
	(0.0016)	(0.0017)	(0.0017)	(0.0017)	(0.0016)	(0.0015)	(0.0016)
Bilingual education	-0.0001	-0.0028*	0.0003	0.0026	-0.0012	0.0010	-0.0035**
	(0.0015)	(0.0016)	(0.0018)	(0.0018)	(0.0015)	(0.0014)	(0.0017)
Lives with biological parent	-0.0016	-0.0025	0.0021	-0.0002	-0.0023	0.0025*	0.0021
	(0.0014)	(0.0016)	(0.0018)	(0.0016)	(0.0015)	(0.0014)	(0.0015)
Special education	0.0043	-0.0065	0.0030	0.0026	0.0006	-0.0134*	0.0044
	(0.0058)	(0.0079)	(0.0080)	(0.0080)	(0.0069)	(0.0069)	(0.0063)
Special education x year trend	-0.0009	0.0003	0.0013	-0.0010	0.0012	0.0015	-0.0014
	(0.0009)	(0.0012)	(0.0012)	(0.0011)	(0.0010)	(0.0010)	(0.0010)
8th grade ITBS reading z-score	0.0052***	0.0004	0.0014	0.0024	-0.0039**	0.0013	-0.0059***
	(0.0015)	(0.0018)	(0.0018)	(0.0016)	(0.0016)	(0.0017)	(0.0015)
Neighborhood median family income	-0.0010	0.0009	-0.0003	-0.0000	0.0004	-0.0002	-0.0002
	(0.0007)	(0.0009)	(0.0008)	(0.0007)	(0.0007)	(0.0007)	(0.0007)
Neighborhood percent school age (5-18)	-0.0254***	0.0044	0.0189	-0.0101	-0.0054	-0.0083	0.0250**
	(0.0096)	(0.0124)	(0.0118)	(0.0100)	(0.0099)	(0.0110)	(0.0121)
Neighborhood percent Hispanic	0.0074*	0.0003	0.0059	-0.0019	-0.0126***	-0.0036	-0.0014
	(0.0041)	(0.0047)	(0.0049)	(0.0047)	(0.0043)	(0.0045)	(0.0051)
Neighborhood percent black	0.0038	0.0008	0.0049	-0.0082**	-0.0075**	0.0016	-0.0020
	(0.0036)	(0.0039)	(0.0040)	(0.0036)	(0.0033)	(0.0038)	(0.0042)
Neighborhood mean education	0.0006	-0.0013	0.0017*	-0.0005	0.0001	-0.0009	-0.0003
	(0.0009)	(0.0010)	(0.0010)	(0.0009)	(0.0009)	(0.0009)	(0.0010)
Neighborhood percent in poverty	-0.0000	0.0070	-0.0011	0.0036	0.0009	-0.0046	-0.0050
	(0.0051)	(0.0061)	(0.0057)	(0.0050)	(0.0048)	(0.0052)	(0.0052)
Observations	235,853	235,853	235,853	235,853	235,853	235,853	235,853

Notes: Each specification includes teacher-by-semester fixed effects. Neighborhood median family income is measured in $10,000s. Singleton groups (i.e., fixed effects with exactly one observation) were dropped in the estimation, reducing the sample size to that shown. Numbers in parentheses represent standard errors clustered at the high school-by-semester level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table 2C: First-Stage Results for Full Sample Column (3) Regression Specifications of Tables 3-6

	Table 3: Regular English I		Table 4: Regular Algebra I		Table 5: Remedial English I		Table 6: Remedial Algebra I	
	Panel A: Course Passing	Panel B: Reading Test Score	Panel A: Course Passing	Panel B: Math Test Score	Panel A: Course Passing	Panel B: Reading Test Score	Panel A: Course Passing	Panel B: Math Test Score
Period	0.5315***	0.5089***	0.7497***	0.9204***	0.4250***	0.4389***	0.5410***	0.4874***
	(0.0977)	(0.1004)	(0.0985)	(0.1109)	(0.0907)	(0.0945)	(0.1339)	(0.1393)
$Period^2$	-0.0678***	-0.0646***	-0.0911***	-0.1121***	-0.0497***	-0.0514***	-0.0661***	-0.0591***
	(0.0116)	(0.0119)	(0.0118)	(0.0132)	(0.0106)	(0.0110)	(0.0157)	(0.0164)
F-Statistic	18.18***	15.45***	29.80***	35.96***	11.08***	10.92***	8.86***	6.53***
Turning Point	3.92	3.94	4.11	4.11	4.28	4.27	4.10	4.12
Observations	235,853	202,158	278,624	195,980	291,037	250,663	133,957	112,120

Notes: Singleton groups (i.e., fixed effects with exactly one observation) were dropped in the estimation, reducing the sample sizes to those shown. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. The first-stage F-statistic is the F-statistic on the excluded instruments for classroom composition. The turning point is the period of the day at which classroom composition is at its maximum, based on the coefficient estimates for Period and $Period^2$.

Table 3: OLS and 2SLS Results for Course Passing and Reading Test Scores for Regular English I Sample

	Sample Mean	(1) OLS	(1) 2SLS	(2) OLS	(2) 2SLS	(3) OLS	(3) 2SLS
		Panel A: English I Course Passing					
Full Student Sample							
Classroom Composition	0.768	0.0001	0.0657***	-0.0005	0.0636***	0.0028***	0.0726***
		(0.0003)	(0.0113)	(0.0004)	(0.0117)	(0.0004)	(0.0129)
First-Stage F-Statistic		21.55***		19.89***		18.18***	
Observations		237,912		204,976		235,853	
Black Student Subsample							
Classroom Composition	0.753	0.0009**	0.0760***	0.0003	0.0747***	0.0033***	0.0853***
		(0.0004)	(0.0145)	(0.0005)	(0.0147)	(0.0005)	(0.0181)
First-Stage F-Statistic		18.25***		18.17***		12.50***	
Observations		129,042		102,867		127,096	
Hispanic Student Subsample							
Classroom Composition	0.776	-0.0013***	0.0554***	-0.0016***	0.0320**	0.0020***	0.0558***
		(0.0005)	(0.0202)	(0.0005)	(0.0163)	(0.0006)	(0.0171)
First-Stage F-Statistic		5.00***		4.19**		6.92***	
Observations		81,260		71,404		79,900	
		Panel B: Reading Test Score					
Full Student Sample							
Classroom Composition		0.0081***	0.0246**	0.0075***	0.0222**	0.0060***	0.0142
		(0.0007)	(0.0106)	(0.0007)	(0.0112)	(0.0010)	(0.0111)
First-Stage F-Statistic		16.59***		15.44***		15.45***	
Observations		203,926		174,200		202,158	
Black Student Subsample							
Classroom Composition		0.0099***	0.0097	0.0097***	0.0089	0.0083***	0.0159
		(0.0008)	(0.0128)	(0.0009)	(0.0130)	(0.0012)	(0.0141)
First-Stage F-Statistic		13.75***		13.43***		10.15***	
Observations		108,152		84,758		106,482	
Hispanic Student Subsample							
Classroom Composition		0.0052***	0.0670**	0.0052***	0.0633**	0.0025*	0.0413**
		(0.0010)	(0.0281)	(0.0012)	(0.0290)	(0.0015)	(0.0205)
First-Stage F-Statistic		4.81***		4.45**		6.66***	
Observations		71,556		62,423		70,415	
Fixed Effects:							
High School-by-Semester		Yes	Yes	-	-	-	-
Middle School-by-High School-by-Semester		-	-	Yes	Yes	-	-
Teacher-by-Semester		-	-	-	-	Yes	Yes

Notes: The samples include students enrolled in only one regular English I course per semester. The sample mean is the mean course pass rate for the sample in Column (1). Singleton groups (i.e., fixed effects with exactly one observation) were dropped in the estimation, reducing the sample sizes to those shown. All regressions include student controls (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) and neighborhood controls (median family income, percent school age (5-18), percent Hispanic, percent black, mean education, and percent in poverty). Numbers in parentheses represent standard errors clustered at the high school-by-semester level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. The first-stage F-statistic is the F-statistic on the excluded instruments for classroom composition.

Table 4: OLS and 2SLS Results for Course Passing and Math Test Scores for Regular Algebra I Sample

	Sample Mean	(1) OLS	(1) 2SLS	(2) OLS	(2) 2SLS	(3) OLS	(3) 2SLS
		Panel A: Algebra I Course Passing					
Full Student Sample							
Classroom Composition	0.724	-0.0020***	0.0518***	-0.0023***	0.0519***	0.0009**	0.0392***
		(0.0003)	(0.0097)	(0.0004)	(0.0103)	(0.0004)	(0.0064)
First-Stage F-Statistic		20.49***		18.99***		29.80***	
Observations		279,050		242,419		278,624	
Black Student Subsample							
Classroom Composition	0.702	-0.0020***	0.0502***	-0.0027***	0.0502***	0.0009*	0.0421***
		(0.0005)	(0.0110)	(0.0006)	(0.0119)	(0.0005)	(0.0075)
First-Stage F-Statistic		16.81***		15.95***		25.62***	
Observations		146,064		116,921		145,302	
Hispanic Student Subsample							
Classroom Composition	0.744	-0.0022***	0.0672***	-0.0023***	0.0609***	0.0002	0.0432**
		(0.0005)	(0.0225)	(0.0005)	(0.0227)	(0.0006)	(0.0174)
First-Stage F-Statistic		5.88***		4.98***		5.61***	
Observations		102,093		91,120		101,495	
		Panel B: Math Test Score					
Full Student Sample							
Classroom Composition		0.0054***	0.0037	0.0044***	0.0053	0.0040***	0.0068
		(0.0005)	(0.0059)	(0.0005)	(0.0071)	(0.0007)	(0.0054)
First-Stage F-Statistic		25.33***		21.37***		35.96***	
Observations		196,369		168,071		195,980	
Black Student Subsample							
Classroom Composition		0.0057***	0.0073	0.0048***	0.0076	0.0042***	0.0114*
		(0.0007)	(0.0063)	(0.0008)	(0.0077)	(0.0009)	(0.0060)
First-Stage F-Statistic		26.69***		22.74***		35.44***	
Observations		100,741		78,481		100,078	
Hispanic Student Subsample							
Classroom Composition		0.0047***	0.0010	0.0036***	0.0110	0.0031***	-0.0050
		(0.0006)	(0.0127)	(0.0007)	(0.0151)	(0.0010)	(0.0146)
First-Stage F-Statistic		6.60***		5.66***		6.15***	
Observations		73,238		64,640		72,749	
Fixed Effects:							
High School-by-Semester		Yes	Yes	-	-	-	-
Middle School-by-High School-by-Semester		-	-	Yes	Yes	-	-
Teacher-by-Semester		-	-	-	-	Yes	Yes

Notes: The samples include students enrolled in only one regular Algebra I course per semester. The sample mean is the mean course pass rate for the sample in Column (1). Singleton groups (i.e., fixed effects with exactly one observation) were dropped in the estimation, reducing the sample sizes to those shown. All regressions include student controls (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) and neighborhood controls (median family income, percent school age (5-18), percent Hispanic, percent black, mean education, and percent in poverty). Numbers in parentheses represent standard errors clustered at the high school-by-semester level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. The first-stage F-statistic is the F-statistic on the excluded instruments for classroom composition.

Table 5: OLS and 2SLS Results for Course Passing and Reading Test Scores for Remedial English I Sample

	Sample Mean	(1) OLS	(1) 2SLS	(2) OLS	(2) 2SLS	(3) OLS	(3) 2SLS
Panel A: English I Course Passing							
Full Student Sample							
Classroom Composition	0.750	-0.0004	0.0819***	-0.0011***	0.0828***	0.0031***	0.1050***
		(0.0004)	(0.0151)	(0.0004)	(0.0153)	(0.0004)	(0.0226)
First-Stage F-Statistic		16.65***		16.78***		11.08***	
Observations		292,136		283,765		291,037	
Black Student Subsample							
Classroom Composition	0.724	0.0009*	0.1086***	0.0002	0.1069***	0.0038***	0.1481***
		(0.0004)	(0.0250)	(0.0005)	(0.0246)	(0.0005)	(0.0434)
First-Stage F-Statistic		10.25***		10.57***		5.84***	
Observations		160,726		154,352		159,494	
Hispanic Student Subsample							
Classroom Composition	0.779	-0.0025***	0.0375***	-0.0029***	0.0427***	0.0014**	0.0541***
		(0.0005)	(0.0110)	(0.0005)	(0.0123)	(0.0007)	(0.0167)
First-Stage F-Statistic		9.76***		8.49***		6.84***	
Observations		108,173		105,560		107,200	
Panel B: Reading Test Score							
Full Student Sample							
Classroom Composition		0.0038***	-0.0061	0.0025***	-0.0057	0.0024***	0.0013
		(0.0005)	(0.0073)	(0.0004)	(0.0068)	(0.0006)	(0.0088)
First-Stage F-Statistic		14.82***		14.29***		10.92***	
Observations		251,668		244,362		250,663	
Black Student Subsample							
Classroom Composition		0.0037***	-0.0054	0.0017***	-0.0065	0.0023***	0.0107
		(0.0006)	(0.0089)	(0.0005)	(0.0080)	(0.0008)	(0.0125)
First-Stage F-Statistic		8.38***		8.54***		5.21***	
Observations		134,674		129,195		133,541	
Hispanic Student Subsample							
Classroom Composition		0.0030***	0.0020	0.0026***	0.0052	0.0012	-0.0037
		(0.0007)	(0.0096)	(0.0007)	(0.0100)	(0.0010)	(0.0134)
First-Stage F-Statistic		9.45***		7.40***		7.14***	
Observations		96,814		94,526		95,908	
Fixed Effects:							
High School-by-Semester		Yes	Yes	-	-	-	-
Middle School-by-High School-by-Semester		-	-	Yes	Yes	-	-
Teacher-by-Semester		-	-	-	-	Yes	Yes

Notes: The samples include students enrolled in at least one remedial English I course per semester; some students take regular English I in addition to the remedial class. The sample mean is the mean course pass rate for the sample in Column (1). Singleton groups (i.e., fixed effects with exactly one observation) were dropped in the estimation, reducing the sample sizes to those shown. All regressions include student controls (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) and neighborhood controls (median family income, percent school age (5-18), percent Hispanic, percent black, mean education, and percent in poverty). Numbers in parentheses represent standard errors clustered at the high school-by-semester level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. The first-stage F-statistic is the F-statistic on the excluded instruments for classroom composition.

Table 6: OLS and 2SLS Results for Course Passing and Math Test Scores for Remedial Algebra I Sample

	Sample Mean	(1) OLS	(1) 2SLS	(2) OLS	(2) 2SLS	(3) OLS	(3) 2SLS
Panel A: Algebra I Course Passing							
Full Student Sample							
Classroom Composition	0.685	-0.0003	0.1158***	-0.0007	0.1056***	0.0014**	0.0684***
		(0.0007)	(0.0423)	(0.0007)	(0.0390)	(0.0007)	(0.0175)
First-Stage F-Statistic		4.29**		4.42**		8.86***	
Observations		134,336		124,070		133,957	
Black Student Subsample							
Classroom Composition	0.673	0.0005	0.2553	0.0006	0.2669	0.0018**	0.1160**
		(0.0009)	(0.1983)	(0.0009)	(0.2491)	(0.0008)	(0.0457)
First-Stage F-Statistic		0.84		0.58		3.19**	
Observations		77,123		68,885		76,658	
Hispanic Student Subsample							
Classroom Composition	0.702	-0.0012	0.0424**	-0.0012	0.0421***	0.0010	0.0208*
		(0.0009)	(0.0172)	(0.0009)	(0.0163)	(0.0011)	(0.0113)
First-Stage F-Statistic		7.38***		8.65***		9.64***	
Observations		46,948		43,880		46,646	
Panel B: Math Test Score							
Full Student Sample							
Classroom Composition		0.0033***	0.0320	0.0027***	0.0029	0.0010	-0.0123
		(0.0007)	(0.0254)	(0.0007)	(0.0201)	(0.0008)	(0.0138)
First-Stage F-Statistic		2.53*		2.97*		6.53***	
Observations		112,447		103,586		112,120	
Black Student Subsample							
Classroom Composition		0.0041***	0.0491	0.0028***	-0.0326	0.0013	-0.0149
		(0.0008)	(0.0721)	(0.0009)	(0.0596)	(0.0010)	(0.0232)
First-Stage F-Statistic		0.44		0.58		2.51*	
Observations		64,619		57,541		64,207	
Hispanic Student Subsample							
Classroom Composition		0.0022**	0.0183	0.0018	0.0185	-0.0006	-0.0050
		(0.0011)	(0.0197)	(0.0011)	(0.0168)	(0.0014)	(0.0149)
First-Stage F-Statistic		5.25***		6.45***		7.19***	
Observations		39,420		36,726		39,117	
Fixed Effects:							
High School-by-Semester		Yes	Yes	-	-	-	-
Middle School-by-High School-by-Semester		-	-	Yes	Yes	-	-
Teacher-by-Semester		-	-	-	-	Yes	Yes

Notes: The samples include students enrolled in at least one remedial Algebra I course per semester; some students take regular Algebra I in addition to the remedial class. The sample mean is the mean course pass rate for the sample in Column (1). Singleton groups (i.e., fixed effects with exactly one observation) were dropped in the estimation, reducing the sample sizes to those shown. All regressions include student controls (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) and neighborhood controls (median family income, percent school age (5-18), percent Hispanic, percent black, mean education, and percent in poverty). Numbers in parentheses represent standard errors clustered at the high school-by-semester level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. The first-stage F-statistic is the F-statistic on the excluded instruments for classroom composition.

Appendix A: Additional Regressions for Testing Instrument Validity

Table A1: Determinants of Period of the Day (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	-0.0022**	0.0030**	0.0016	-0.0020**	-0.0004	-0.0003	0.0005
	(0.0009)	(0.0012)	(0.0011)	(0.0010)	(0.0009)	(0.0010)	(0.0011)
Male	0.0011	0.0031**	-0.0010	0.0015	-0.0008	-0.0020	-0.0020
	(0.0012)	(0.0014)	(0.0012)	(0.0013)	(0.0011)	(0.0012)	(0.0013)
White	0.0040	-0.0246*	0.0118	-0.0042	-0.0029	-0.0074	0.0215
	(0.0095)	(0.0141)	(0.0124)	(0.0125)	(0.0117)	(0.0123)	(0.0136)
Black	0.0042	-0.0271*	0.0120	-0.0109	-0.0004	-0.0078	0.0247*
	(0.0096)	(0.0141)	(0.0122)	(0.0123)	(0.0116)	(0.0119)	(0.0134)
Hispanic	0.0021	-0.0253*	0.0098	-0.0061	0.0004	-0.0094	0.0230*
	(0.0096)	(0.0140)	(0.0123)	(0.0124)	(0.0118)	(0.0123)	(0.0134)
Asian	0.0008	-0.0165	0.0042	-0.0104	-0.0014	-0.0054	0.0237*
	(0.0098)	(0.0150)	(0.0125)	(0.0126)	(0.0120)	(0.0128)	(0.0144)
Free or reduced lunch	0.0063***	-0.0018	0.0035**	-0.0001	-0.0011	-0.0035**	-0.0025*
	(0.0014)	(0.0017)	(0.0016)	(0.0015)	(0.0014)	(0.0015)	(0.0014)
Classified as disruptive	0.0005	-0.0029	-0.0006	-0.0043***	0.0037**	-0.0013	0.0024
	(0.0016)	(0.0018)	(0.0017)	(0.0016)	(0.0016)	(0.0016)	(0.0017)
Bilingual education	-0.0005	0.0004	-0.0003	-0.0007	0.0012	0.0010	0.0022
	(0.0015)	(0.0016)	(0.0015)	(0.0017)	(0.0015)	(0.0015)	(0.0016)
Lives with biological parent	0.0021*	0.0003	0.0001	0.0016	0.0000	-0.0007	-0.0024*
	(0.0011)	(0.0015)	(0.0014)	(0.0014)	(0.0012)	(0.0013)	(0.0014)
Special education	-0.0044	-0.0023	0.0137*	-0.0097	0.0028	0.0171**	-0.0077
	(0.0062)	(0.0078)	(0.0080)	(0.0072)	(0.0066)	(0.0077)	(0.0075)
Special education x year trend	0.0003	0.0009	-0.0017*	0.0017*	0.0005	-0.0017	-0.0001
	(0.0009)	(0.0010)	(0.0010)	(0.0009)	(0.0009)	(0.0010)	(0.0010)
8th grade ITBS math z-score	-0.0046***	-0.0013	0.0017	-0.0026*	0.0033**	0.0031**	-0.0017
	(0.0014)	(0.0017)	(0.0018)	(0.0016)	(0.0015)	(0.0015)	(0.0017)
Neighborhood median family income	-0.0002	0.0005	0.0008	-0.0015**	-0.0009	0.0001	0.0009
	(0.0006)	(0.0006)	(0.0007)	(0.0006)	(0.0006)	(0.0006)	(0.0006)
Neighborhood percent school age (5-18)	-0.0087	-0.0184*	-0.0163	0.0156	-0.0126	0.0341***	-0.0059
	(0.0095)	(0.0111)	(0.0105)	(0.0108)	(0.0095)	(0.0096)	(0.0103)
Neighborhood percent Hispanic	-0.0016	0.0043	0.0097**	0.0001	0.0075*	-0.0043	-0.0100**
	(0.0035)	(0.0047)	(0.0047)	(0.0040)	(0.0041)	(0.0042)	(0.0044)
Neighborhood percent black	-0.0003	0.0057	0.0087**	0.0042	0.0026	-0.0099***	-0.0036
	(0.0032)	(0.0038)	(0.0038)	(0.0033)	(0.0037)	(0.0033)	(0.0036)
Neighborhood mean education	0.0003	0.0003	-0.0001	0.0004	0.0008	0.0011	-0.0028***
	(0.0007)	(0.0010)	(0.0009)	(0.0008)	(0.0008)	(0.0008)	(0.0008)
Neighborhood percent in poverty	0.0028	0.0052	0.0056	-0.0121**	-0.0059	-0.0050	0.0023
	(0.0048)	(0.0053)	(0.0054)	(0.0051)	(0.0043)	(0.0044)	(0.0050)
Observations	278,624	278,624	278,624	278,624	278,624	278,624	278,624

Notes: Each specification includes teacher-by-semester fixed effects. Neighborhood median family income is measured in $10,000s. Singleton groups (i.e., fixed effects with exactly one observation) were dropped in the estimation, reducing the sample size to that shown. Numbers in parentheses represent standard errors clustered at the high school-by-semester level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Table A2: Determinants of Period of the Day (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	0.0005	0.0015	0.0004	-0.0014	-0.0017*	0.0011	0.0030***
	(0.0008)	(0.0011)	(0.0009)	(0.0009)	(0.0009)	(0.0010)	(0.0009)
Male	-0.0007	-0.0010	0.0019	-0.0019	0.0010	0.0004	0.0018
	(0.0013)	(0.0013)	(0.0013)	(0.0012)	(0.0011)	(0.0012)	(0.0013)
White	0.0126	-0.0178	0.0399***	-0.0074	-0.0130	-0.0053	0.0254*
	(0.0121)	(0.0130)	(0.0137)	(0.0158)	(0.0129)	(0.0158)	(0.0138)
Black	0.0117	-0.0190	0.0304**	-0.0041	-0.0127	-0.0029	0.0271*
	(0.0124)	(0.0130)	(0.0141)	(0.0156)	(0.0129)	(0.0156)	(0.0140)
Hispanic	0.0095	-0.0162	0.0405***	-0.0057	-0.0104	-0.0068	0.0229
	(0.0120)	(0.0129)	(0.0139)	(0.0156)	(0.0128)	(0.0156)	(0.0140)
Asian	0.0127	-0.0119	0.0473***	-0.0007	-0.0156	-0.0135	0.0192
	(0.0123)	(0.0137)	(0.0155)	(0.0160)	(0.0135)	(0.0159)	(0.0141)
Free or reduced lunch	0.0011	0.0009	0.0006	-0.0002	-0.0005	0.0016	-0.0024
	(0.0014)	(0.0017)	(0.0015)	(0.0014)	(0.0014)	(0.0014)	(0.0016)
Classified as disruptive	-0.0007	0.0022	-0.0005	-0.0001	-0.0001	0.0006	-0.0004
	(0.0017)	(0.0018)	(0.0019)	(0.0016)	(0.0017)	(0.0016)	(0.0018)
Bilingual education	0.0017	0.0016	-0.0043***	-0.0001	-0.0022	-0.0009	0.0029*
	(0.0012)	(0.0014)	(0.0015)	(0.0014)	(0.0016)	(0.0015)	(0.0016)
Lives with biological parent	0.0007	0.0006	-0.0024*	0.0011	-0.0050***	0.0007	0.0023*
	(0.0011)	(0.0013)	(0.0013)	(0.0012)	(0.0013)	(0.0012)	(0.0012)
Special education	0.0001	-0.0004	-0.0045	-0.0017	-0.0081	0.0090	-0.0009
	(0.0051)	(0.0070)	(0.0062)	(0.0061)	(0.0052)	(0.0066)	(0.0054)
Special education x year trend	0.0003	-0.0001	0.0003	0.0005	0.0003	-0.0006	-0.0002
	(0.0005)	(0.0007)	(0.0006)	(0.0006)	(0.0006)	(0.0007)	(0.0006)
8th grade ITBS reading z-score	0.0004	-0.0006	0.0009	-0.0004	0.0015	-0.0001	-0.0022*
	(0.0010)	(0.0012)	(0.0011)	(0.0012)	(0.0010)	(0.0011)	(0.0012)
Neighborhood median family income	0.0004	0.0010	-0.0005	0.0016***	-0.0012**	-0.0005	-0.0002
	(0.0005)	(0.0007)	(0.0006)	(0.0005)	(0.0005)	(0.0005)	(0.0005)
Neighborhood percent school age (5-18)	-0.0144	-0.0059	0.0234**	0.0064	0.0050	0.0095	0.0265***
	(0.0088)	(0.0103)	(0.0101)	(0.0094)	(0.0090)	(0.0102)	(0.0096)
Neighborhood percent Hispanic	0.0037	0.0082*	-0.0041	0.0056	0.0007	-0.0126***	-0.0036
	(0.0032)	(0.0045)	(0.0046)	(0.0040)	(0.0043)	(0.0043)	(0.0043)
Neighborhood percent black	0.0023	0.0118***	0.0053	0.0005	-0.0009	-0.0130***	-0.0022
	(0.0029)	(0.0037)	(0.0041)	(0.0035)	(0.0036)	(0.0037)	(0.0035)
Neighborhood mean education	-0.0010	-0.0005	0.0008	0.0005	0.0007	-0.0006	-0.0001
	(0.0006)	(0.0007)	(0.0008)	(0.0007)	(0.0007)	(0.0007)	(0.0007)
Neighborhood percent in poverty	0.0082*	0.0040	-0.0009	-0.0012	0.0073*	-0.0040	-0.0131***
	(0.0045)	(0.0057)	(0.0052)	(0.0046)	(0.0044)	(0.0046)	(0.0044)
Observations	291,037	291,037	291,037	291,037	291,037	291,037	291,037

Notes: Each specification includes teacher-by-semester fixed effects. Neighborhood median family income is measured in $10,000s. Singleton groups (i.e., fixed effects with exactly one observation) were dropped in the estimation, reducing the sample size to that shown. Numbers in parentheses represent standard errors clustered at the high school-by-semester level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

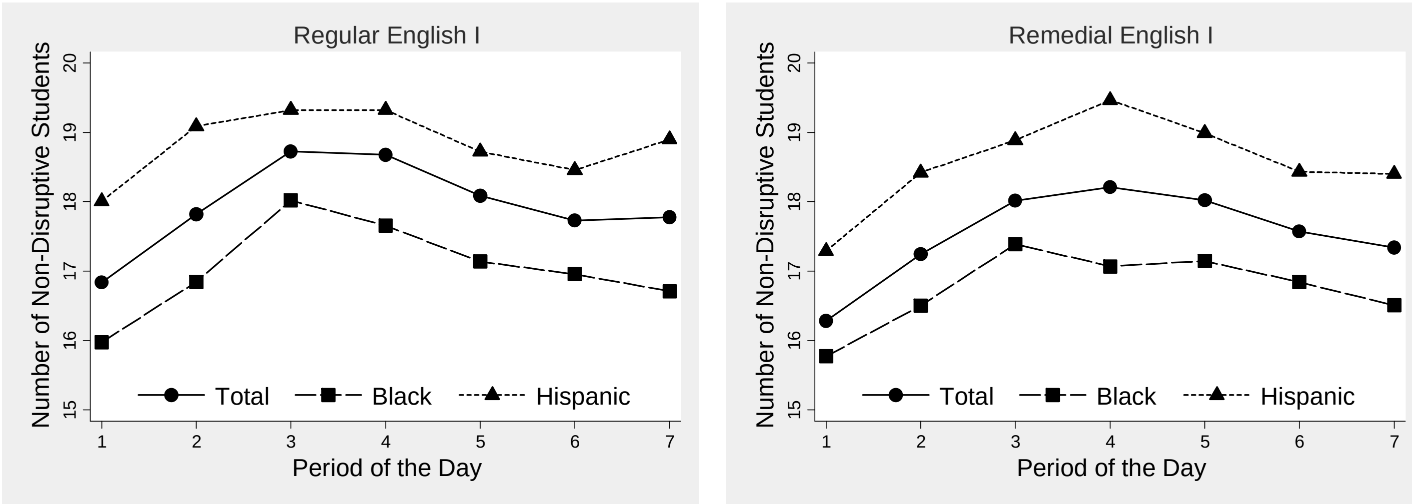
Table A3: Determinants of Period of the Day (Remedial Algebra I Sample)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Dependent Variable: Period of the day is …						
	1st	2nd	3rd	4th	5th	6th	7th
Age	0.0017	-0.0026*	-0.0022*	0.0005	-0.0001	-0.0022	0.0029**
	(0.0012)	(0.0014)	(0.0013)	(0.0012)	(0.0014)	(0.0013)	(0.0012)
Male	-0.0010	0.0013	-0.0003	-0.0035**	-0.0012	0.0018	0.0001
	(0.0017)	(0.0015)	(0.0016)	(0.0017)	(0.0018)	(0.0016)	(0.0019)
White	0.0236	-0.0028	-0.0301	0.0102	-0.0173	0.0197	-0.0256
	(0.0184)	(0.0217)	(0.0245)	(0.0146)	(0.0242)	(0.0153)	(0.0195)
Black	0.0228	-0.0067	-0.0227	0.0046	-0.0149	0.0130	-0.0226
	(0.0185)	(0.0222)	(0.0240)	(0.0139)	(0.0238)	(0.0156)	(0.0192)
Hispanic	0.0221	-0.0032	-0.0223	0.0058	-0.0160	0.0203	-0.0293
	(0.0186)	(0.0215)	(0.0243)	(0.0141)	(0.0235)	(0.0154)	(0.0192)
Asian	0.0235	-0.0034	-0.0318	0.0172	-0.0155	0.0099	-0.0245
	(0.0188)	(0.0220)	(0.0240)	(0.0150)	(0.0242)	(0.0152)	(0.0207)
Free or reduced lunch	-0.0033	0.0007	0.0019	-0.0058***	0.0022	-0.0022	0.0032
	(0.0020)	(0.0023)	(0.0021)	(0.0019)	(0.0021)	(0.0020)	(0.0022)
Classified as disruptive	-0.0037	0.0050**	0.0021	-0.0017	-0.0003	0.0041*	-0.0046*
	(0.0028)	(0.0026)	(0.0028)	(0.0024)	(0.0029)	(0.0024)	(0.0027)
Bilingual education	-0.0024	-0.0001	0.0015	0.0012	0.0011	-0.0015	-0.0022
	(0.0018)	(0.0020)	(0.0021)	(0.0017)	(0.0023)	(0.0021)	(0.0019)
Lives with biological parent	-0.0019	0.0014	0.0044**	0.0003	0.0011	0.0052***	-0.0054***
	(0.0016)	(0.0018)	(0.0017)	(0.0017)	(0.0018)	(0.0016)	(0.0019)
Special education	-0.0027	0.0119*	-0.0159**	0.0070	0.0001	0.0023	-0.0121*
	(0.0057)	(0.0063)	(0.0068)	(0.0061)	(0.0050)	(0.0044)	(0.0066)
Special education x year trend	0.0004	-0.0007	0.0021***	-0.0003	0.0004	-0.0007	0.0008
	(0.0006)	(0.0007)	(0.0007)	(0.0007)	(0.0006)	(0.0006)	(0.0007)
8th grade ITBS math z-score	-0.0023	-0.0014	0.0012	-0.0011	-0.0007	0.0007	0.0022
	(0.0017)	(0.0019)	(0.0018)	(0.0015)	(0.0018)	(0.0015)	(0.0019)
Neighborhood median family income	0.0006	-0.0006	0.0008	-0.0008	-0.0004	0.0004	-0.0004
	(0.0008)	(0.0010)	(0.0009)	(0.0008)	(0.0008)	(0.0008)	(0.0008)
Neighborhood percent school age (5-18)	0.0214	-0.0169	0.0045	0.0254*	-0.0325**	-0.0234*	0.0098
	(0.0133)	(0.0156)	(0.0140)	(0.0134)	(0.0133)	(0.0123)	(0.0141)
Neighborhood percent Hispanic	0.0011	-0.0058	-0.0074	-0.0098**	0.0131**	-0.0002	0.0105*
	(0.0050)	(0.0064)	(0.0062)	(0.0050)	(0.0065)	(0.0055)	(0.0061)
Neighborhood percent black	0.0001	0.0079	-0.0076	-0.0053	0.0033	-0.0041	0.0043
	(0.0046)	(0.0049)	(0.0051)	(0.0043)	(0.0051)	(0.0053)	(0.0050)
Neighborhood mean education	0.0013	-0.0011	-0.0008	-0.0012	0.0008	-0.0006	0.0010
	(0.0011)	(0.0012)	(0.0011)	(0.0009)	(0.0011)	(0.0010)	(0.0010)
Neighborhood percent in poverty	-0.0039	-0.0054	0.0045	-0.0077	0.0020	0.0179***	-0.0020
	(0.0078)	(0.0079)	(0.0071)	(0.0073)	(0.0070)	(0.0064)	(0.0071)
Observations	133,957	133,957	133,957	133,957	133,957	133,957	133,957

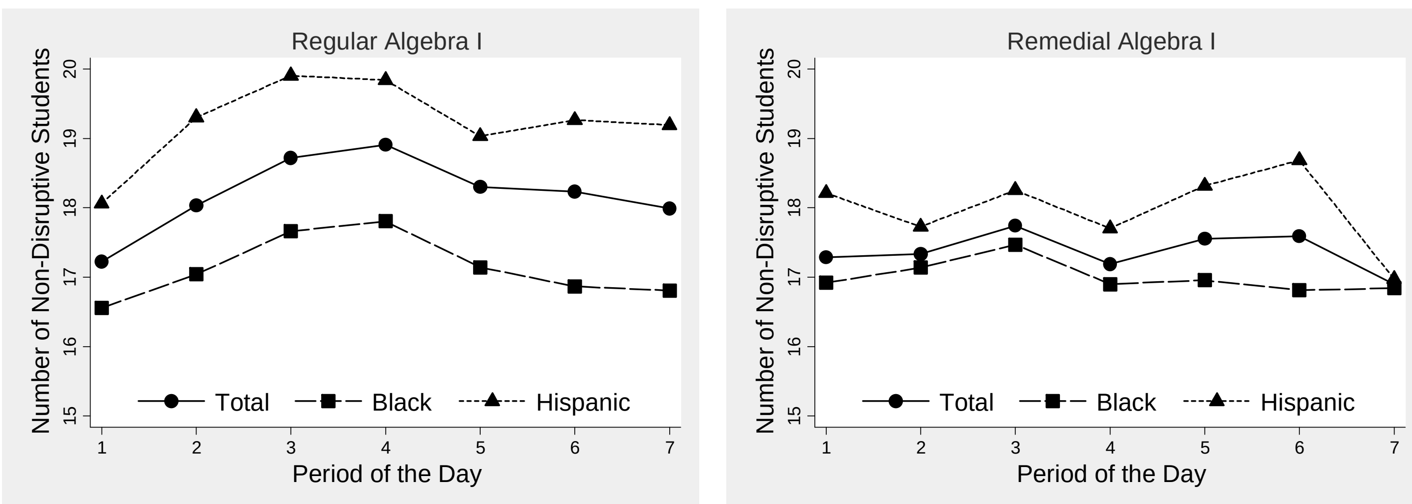
Notes: Each specification includes teacher-by-semester fixed effects. Neighborhood median family income is measured in $10,000s. Singleton groups (i.e., fixed effects with exactly one observation) were dropped in the estimation, reducing the sample size to that shown. Numbers in parentheses represent standard errors clustered at the high school-by-semester level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

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

Panel A: Regular English I and Remedial English I



Panel B: Regular Algebra I and Remedial Algebra 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 English I, regular Algebra I, and their respective remedial courses.