COMBINATION CLASSES AND EDUCATIONAL ACHIEVEMENT

Jaime Lynn Thomas*

jlthomas@ucsd.edu

University of California, San Diego

Abstract: This paper determines the effect of membership in a K-1 or 1-2 combination class in first grade on student achievement. I address the selection that occurs at all three levels of implementing a combination class. In order to control for any systematic differences between schools that offer combination classes and those that do not, I conduct a within-school analysis using school fixed effects. I find little evidence of meaningful nonrandom assignment of teachers to combination classes. There is, however, evidence that 1-2 students are positively selected based on ability. Using a rich set of covariates, I am able to control for the variables influencing selection. Estimates of the effect of combination class membership in first grade on reading and general knowledge test scores are not significantly different from zero. The estimate of the effect on math scores for first graders in 1-2 combination classes is positive and significant, indicating that 1-2 students can be expected to outperform single-grade students by one-seventh of a standard deviation. This result is not sensitive to functional-form assumptions. In addition, I find no evidence that first-graders in schools offering combination classes perform worse than first-graders in schools that do not offer such classes. Therefore, I conclude that combination classes may be a Pareto-improving option for school administrators.

JEL Codes: I21

* I thank Julie Cullen, Nora Gordon, Julian Betts, and Kate Antonovics for their valuable advice and helpful suggestions.

I. Introduction

The combination class, in which students from two adjacent grades are grouped within one classroom under one teacher, is an increasingly common method of classroom organization, yet has received little attention in the literature. The nationwide trend toward class-size reduction suggests that combination classes will only become more prevalent, since they can be used to attain class-size goals by smoothing enrollment across grades. They are a cost-saving option, allowing schools to use fewer teachers and classrooms. If combination-class membership has a nonnegative effect on student outcomes, offering such classes is an attractive strategy for schools looking to save money without sacrificing educational quality.

Combination classes also offer another avenue besides age at school entry to assess the effect of relative age on student performance. The age-at-school entry literature focuses on students in similar learning environments and assesses the effect of relative and absolute age on student achievement and other outcomes. Relatively older students are consistently shown to perform better on reading and math tests.[1] Rather than being a relative age effect, more recent research has established that this is likely to be an absolute age effect, and that being younger might actually lead to higher test scores.[2] This paper compare students of the same absolute age who are placed in different learning environments—single-grade and combination classes—in which curricula and teaching methods differ along with students' relative ages. I am asking a different question, but my results support the recent findings in the age-at-school-entry literature.

[1] See Stipek (2002) for a detailed literature review.
[2] See, for example, Bedard and Dhuey (2006), Datar (2006), Black et al. (2008), and Elder and Lubotsky (2006).

This paper also contributes to the small body of literature that directly addresses the effect of combination classes on student achievement. Within this literature, there is little consensus.[3] In addition to differential success in dealing with nonrandom selection, prior studies have not distinguished between relatively older students (those in the higher grade of the combination class) and relatively younger students (those in the lower grade), even though treatment systematically differs along this dimension.

I seek to determine the effects of membership in K-1 and 1-2 combination classes on student achievement in first grade, as measured by test scores from the spring of the first-grade year. I address the selection that occurs at all three levels of implementing combination classes. In order to control for any systematic differences between schools that offer combination classes and those that do not, I conduct a within-school analysis of schools offering combination classes using school fixed effects. I find little evidence of meaningful nonrandom assignment of teachers to combination classes, indicating that differences in outcomes are not due to differences in teacher quality. There is some evidence that first-graders in 1-2 classes are positively selected based on ability. Using a rich set of covariates, however, I am able to effectively control for the variables influencing selection. Estimates of the effect of combination class membership in first grade on reading and general knowledge test scores are not significantly different from zero. The estimate of the effect on math scores for first-graders in 1-2 combination classes is positive and significant, indicating that 1-2 students—that is, students who are young relative to their classmates—can be expected to outperform their single-grade peers by one-seventh of a standard deviation. In addition, I find no evidence that first-

[3] See Veenman (1995) for a meta-analysis, and Hill and Rowe (1998) and Sims (2008) for more recent studies.

graders in schools offering combination classes perform worse than first-graders in schools that do not offer such classes, indicating that offering combination classes may be a Pareto-improving option for school administrators.

This paper proceeds as follows. Section II discusses the conceptual framework underlying the estimation of combination-class effects. Section III describes the rich data set used in the analysis. Section IV documents the selection issues that arise at the school, classroom, and student level when a school chooses to offer combination classes. Section V presents the main results of the paper and discusses some robustness checks. Section VI demonstrates that first-graders in schools offering combination classes do not seem to perform worse than first-graders in schools that do not offer combination classes, and Section VII concludes.

II. Conceptual Framework

II.A. Class Type as an Input in an Education Production Function

Student achievement as measured by test scores is a function of many variables. A child's performance in elementary school depends on the characteristics of the child's home, such as socioeconomic status (SES) and parental involvement. School performance also depends on the child's own characteristics, such as scholastic ability, past educational experience, behavior, and motivation. In addition, performance hinges on attributes of the student's school, such as demographics, school resources, and calendar type (year-round or traditional nine-month calendar). Finally, class characteristics such as teacher quality, curriculum, and classroom organization influence student achievement.

In this paper, I am interested in the effect of classroom organization in first grade on test scores—specifically, whether the class is K-1 combination, a 1-2 combination, or a single-grade first-grade class. If schools randomly chose to offer combination classes, and if teachers and students were randomly assigned to combination classes, a simple linear regression of first-grade test scores on dummy variables for K-1 and 1-2 combination-class membership would yield estimates of combination-class treatment effects that could be interpreted causally.

In the following two subsections, I discuss what exactly the combination-class treatment entails and the obstacles to the causal interpretation of the coefficients on combination-class dummy variables that emerge under nonrandom assignment.

II.B. The Combination-Class Treatment

Combination classes differ from single-grade classes on several dimensions. Some of these differences are inherent to combination classes and would exist even if schools randomly decided to offer combination classes. It is the effect of these inherent characteristics that I would like to isolate.

First, the age span within a combination class is wider than within a single-grade class. For instance, if a kindergarten class contains five- and six-year-olds, and a first grade class contains six- and seven-year-olds, a K-1 combination would contain children aged five to seven. Evidence on the effect of age diversity within a classroom is inconclusive.[4]

A related but separate characteristic of the combination-class treatment is that students are systematically placed within this wider age range so that they end up as relatively older or relatively younger than their classmates. First-graders in K-1 classes

[4] See, for example, Miller (1995).

are relatively old, and first-graders in 1-2 classes are relatively young. Because of the wider age span in combination classes, these relative age differences are more pronounced than in single-grade classes. Elder and Lubotsky (2006) show that having older classmates tends to raise reading and math achievement, conditional on the student's own age. This is in contrast to earlier findings within the age-at-school entry literature that relatively older students do better. This literature considers students within one type of class whose ages are different because of school entry cutoff dates. Instead, I am looking at students who are the same age in different types of classrooms, where relative age depends on the type of combination class in which the student is placed.

In addition to their relative age differences, first-graders in K-1 and in 1-2 combination classes are likely to experience different teaching methods and curricula than students in single-grade classes. In a survey of 35 combination-class teachers in California, Mason, Burns, and Armesto (1993) find that teachers tend to use a mixed approach in combination classes, in which the teacher separates students by grade level for certain subjects such as math and reading and uses large-group instruction for subjects such as science and social studies. We can assume that the large-group curriculum in a K-1 combination class will be aimed at a lower level than the large-group curriculum in a 1-2 combination class. In this way, the combination class effect will differ depending on a student's relative grade level within the class. At first glance, the mixed approach would seem to have a positive effect on first-graders in a 1-2 combination and a negative effect on those in a K-1 combination, relative to the performance of single-grade first-graders. It is not inconceivable, however, that first-graders in a K-1 combination would benefit from the review of kindergarten concepts and do better in a K-1 combination class than they would have in a single-grade class.

II.C. Confounding Factors Resulting from Nonrandom Assignment of Students and Teachers to Combination Classes

Combination classes differ from single-grade classes in other ways if there is nonrandom assignment to combination classes. First, schools that decide to offer combination classes may be systematically different from those that do not. For example, year-round schools may have a small number of students per grade level and choose to offer combination classes in order to use fewer classrooms. Calendar type may have an effect on student achievement apart from its association with combination classes. If a year-round calendar has a negative effect on student achievement, as shown in Graves (2007), the combination-class effect would be biased downwards.

Second, the resources available to students in combination classes may be different from those available to single-grade students. If combination classes are systematically larger, for example, students may be adversely affected. Teaching quality may also differ by class type. In a survey of 72 school principals in California, Burns, Mason, and Demiranda (1993) find that many select only the best teachers for combination classes. If this is indeed the general selection criterion, the positive effect of the teachers' skill will bias estimates of the combination-class effect upwards.

Finally, students are assigned to combination classes. The main reason for nonrandom assignment is to make these classes more attractive to teachers (Mason et al., 1993). Generally, the goal is to make student ability more homogeneous than it would be under random assignment, or to populate the class with independent workers. First-graders placed in combination classes are likely to be positively selected on behavior in all cases. Selection on ability is likely to be positive for 1-2 placement and negative for K-1 placement. I document the selection that occurs on all three levels and discuss how I deal with selection at each level in Section IV, below.

III. Data

The Early Childhood Longitudinal Study, Kindergarten Class of 1998-1999 (ECLS-K) Restricted-Use Data Set is an ongoing study focusing on children's early school experiences. It has a rich set of student-, classroom-, and school-level variables, allowing me to determine what factors influence a school's decision to offer combination classes and to analyze the teacher and student characteristics that influence assignment, as well as measure the effect of combination class membership on test scores.

In this study, I use data collected in the spring of the children's kindergarten year and the spring of their first-grade year. Spring first-grade reading, math, and general knowledge standardized test scores are the outcome variables of interest. Spring kindergarten scores are prior test score controls. I use a variety of child-, classroom-, and school-level controls: child characteristics (gender, race and ethnicity, age), family background variables (SES, home language), teacher characteristics (race and ethnicity, education, and experience), classroom characteristics (demographics, student performance, classroom activities, age distribution, class size), and school characteristics (location, calendar type, percent minority students, percent of students eligible for free lunch). In addition, the ECLS-K contains behavior measures that are typically unobservable to the econometrician. Finally, I match schools to the National Center for

Education Statistics (NCES) Common Core of Data (CCD) to obtain data on enrollment by grade level for the 1999-2000 school year.

I restrict the sample to public-school first-graders, and only include students who were first-time kindergarteners in the 1998-1999 school year and remain in the dataset as first-graders in the 1999-2000 school year, resulting in a sample of 10,640 students. I further restrict the sample to students whose first-grade class type can be accurately identified, resulting in a final sample of 9,339 individuals.

IV. Selection Issues

In order to identify the causal effect of combination class membership, one must address the selection that occurs when combination classes are offered. First, I discuss school-level selection, then selection at the classroom level, and finally student-level selection.

IV.A. School-level Selection

Are schools that offer combination classes systematically different from those that do not? 17 percent of the public elementary schools sampled in the ECLS-K offer some type of combination (K-1 or 1-2) or multi-grade (K-1-2, 1-2-3, K-1-2-3, etc.) class. Table 1 contains a breakdown of schools by the types of first-grade classes they offer. Schools that offer some type of combination or multi-grade class fall into two broad categories. 92 schools offer single grade first-grade classes and one or both combination classes (K-1, 1-2, or K-1 and 1-2), or the two combination classes only. I will call these "combination-class schools."

76 schools offer K-1 and 1-2 classes only (that is, no single-grade classes), or offer first and K-1 or 1-2 along with some other type of multi-grade class. These "multi-

grade schools" appear to have so few students per grade level that their only option is to combine grades, as in a one-room schoolhouse. In the analysis that follows, I drop students from multi-grade schools and consider the sample of students attending combination-class schools only. In doing so, I am able to divide students cleanly into three groups of first-graders: those in a single-grade class, those in a K-1 combination, and those in a 1-2 combination. In addition, I am able to focus on schools in which the only first-grade class options come from this set of possible class types.

Table 2 compares combination-class schools to those offering only single-grade classes. Combination-class schools seem to base the decision to offer these classes on classroom constraints (i.e., crowding) and school calendar type. Indeed, these 92 schools are more likely to have a year-round calendar than the schools offering only single-grade classes. Classroom constraints may be a function of school calendar type, especially if the school operates on a multi-track year-round calendar. In this type of school, the student body and staff are divided into three to five tracks. At any one time, all but one of the tracks is attending school and the last track is on vacation.

The ECLS-K does not reveal if a school is a single- or a multi-track year-round school. However, multi-track year-round schools are fairly common among year-round schools. In California, for example, which is home to 44 percent of year-round schools nationwide (National Association for Year-Round Education, 2007), 48 percent of year-round schools use a multi-track calendar (California Department of Education, Statistical Summary of Year-Round Programs, 2005-2006). Burns, Mason, and Demiranda (1993) find that multi-track principals are constrained in their assignment of students to different types of classes since there are relatively low numbers of students in each grade level per track, and principals may have little choice but to combine adjacent grades into a combination class.

Combination-class schools differ on other dimensions as well. They are more likely to be in the West. They have a higher percentage of minorities and larger average enrollments in grades K through two. They also have a significantly lower number of full-time equivalent (FTE) teachers per student. This could be both a cause and an effect of combination classes. An overcrowded school is more likely to switch to a multi-track year-round calendar, which in turn may lead to the adoption of combination classes. On the other hand, one of the intended results of combination classes is that students from two grade levels are combined into one class, necessitating one teacher instead of two and lowering the teacher-pupil ratio.

Combination-class schools appear to be more disadvantaged than single-grade schools, which could bias estimates of the combination-class effect downwards if these differences are not addressed. Sims (2008) finds that second- and third-graders in schools with a higher percentage of students in combination classes perform worse than second- and third-graders in schools with fewer combination-class students. He uses an instrumental variables technique to account for the school's decision to offer combination classes, but shows that his instrument is correlated with observable school characteristics. Because it may also be correlated with unobservable school characteristics, the estimates in Sims (2008) may be biased downwards. In order to avoid any school-level bias, I make my sample of schools as homogeneous as possible by considering only combination-class schools. I address any additional systematic, school-level differences by focusing on within-school differences between combination- and single-grade classes using school fixed effects.

IV.B. Classroom-level Selection

Within combination-class schools, are K-1 or 1-2 teachers systematically different from single-grade teachers? Do K-1 or 1-2 classes systematically differ from single-grade classes? This section answers these questions.

The teacher-level variables I analyze are as follows: gender, race and ethnicity, experience, education, job satisfaction (enjoys present teaching job, believes teacher makes a difference in children's lives, and would choose teaching again),[5] and paid and unpaid preparation hours per week. Table 3 contains the means of each of these variables by class type. The sample is restricted to teachers within combination-class schools. Of these teachers, 293 teach single-grade first, 47 teach K-1 combinations, and 99 teach 1-2 combinations.

Comparing K-1 teachers to single-grade teachers, K-1 teachers are slightly less likely to be white but are similar to single-grade teachers in other respects. Comparing 1-2 teachers to single-grade teachers, 1-2 teachers are less likely to be male, more likely to be white, less likely to be Hispanic and more likely to be Asian. 1-2 teachers also appear to be happier with their career choice than single-grade teachers, answering more positively to the question of whether they would choose teaching again. This could be either a result of their experience teaching a 1-2 combination, or a reason for their assignment to such a class. The latter would result in an upward bias in estimating the effect of 1-2 membership on test scores. Because of the direction-of-causality problem

[5] The job satisfaction variables contain teachers' responses on a five-point Likert scale in which one = "strongly disagree" and 5 = "strongly agree."

and since the difference in means is only significant at the ten percent level, I ignore this possibility in the analysis that follows.

A more complete indication of nonrandom selection of teachers is to see if the teacher-level variables, taken together, influence assignment to combination classes within schools. I model the selection of teachers using a simple linear model that includes school fixed effects.[6] All of the job satisfaction variables have a direction-of-causality problem, so I do not include them in the following model:

$$\begin{aligned} class\_type_i = {} & \beta_1 male_i + \beta_2 black_i + \beta_3 hispanic_i + \beta_4 asian_i + \beta_5 other_i \\ & + \beta_6 yrs\_teach_i + \beta_7 some\_grad_i + \beta_8 grad\_dgr_i + \sum_{j=1}^{J} \delta_j s_j + \varepsilon_i \end{aligned} \quad (1)$$

Teachers are indexed by *i*, schools by *j*. I run two separate regressions. In the first, $class\_type_i$ equals one if teacher *i* teaches a K-1 class. In this regression, I restrict my sample to teachers within combination-class schools offering only single-grade and K-1 combination classes. Recall from Table 1 that this is the second most common type of combination-class school. Thus, $class\_type_i$ equals zero if teacher *i* teaches a single-grade class within this type of school.

In the second regression, $class\_type_i$ equals one if teacher *i* teachers a 1-2 class, and the sample is restricted to teachers within combination-class schools offering only single-grade and 1-2 combination classes. This is the most common type of combination-class school. The independent variables in both regressions are the observed teacher characteristics: gender, ethnicity, experience, and education, as well as school fixed effects.

Table 4 contains the results. Though K-1 teachers are less likely to be male and more likely to be Hispanic or Black, and 1-2 teachers are more likely to be Asian, there

[6] Probit results produce similar marginal effects.

appears to be no evidence of meaningful selection on observables—the coefficients on years of teaching experience and the education dummies are not significant individually or jointly in either regression.[7] This lack of evidence on selection based on experience and education suggests that nonrandom assignment of teachers is not a source of bias in the outcome regressions in Section V.

In addition to comparing teachers by class type, I compare the following classroom characteristics: size, percent boys, percent minority, percent gifted, percent limited English proficiency, percent below grade level in reading and math, age distribution, and teaching methods (use of whole-class, small-group, or individual activities).

Table 5 contains the means of classroom-level variables obtained from regressions on dummies for K-1 class and 1-2 class and school fixed effects. Single-grade classes form the base case. We observe the obvious differences in age distribution: K-1 classes are younger and 1-2 classes are older than single-grade classes. There is debate about the effect of class size on student achievement,[8] but in any case, combination classes do not differ from single-grade classes along this dimension. This lack of variation could be due to the fact that a plurality of the students in the sample (31%) lives in California, which implemented its Class Size Reduction Act in the 1996-1997 school year, giving financial rewards to schools that reduced class size in grades K-3 to 20 students or fewer. By the 1999-2000 school year, 99% of first-graders were in classes of 20 or fewer students (California Department of Education, 2009).

[7] $F(3,82) = 0.90$, p-value = 0.446 in the K-1 regression; $F(3,176) = 0.67$, p-value = 0.574 in the 1-2 regression.
[8] See, for example, Hoxby (2000) or Mishel and Rothstein (2002).

Teaching methods also differ according to Table 5. Teacher-directed whole-class and individual activities are less common in K-1 than in single-grade classes. Child-selected activities are more common in K-1 classes than they are in single-grade or 1-2 classes. Differences in teaching methods are part of the combination-class treatment effect that I want to estimate. Class composition, however, also differs, and this is a result of nonrandom selection—a confounding factor that could bias estimates of the combination-class effect. 1-2 classes contain more gifted students, which points to the possibility of positive peer selection (though these classroom-level data do not specify if the gifted students are first- or second-graders). Positive selection of peers into combination classes will bias estimates upwards. In order to address this source of bias, I run outcome regressions in Section V including average peer ability as measured by kindergarten test scores as a partial control.

IV.C. Student-level Selection

In this section, I analyze student-level variables to determine if there is positive or negative selection into combination classes. The student-level variables are as follows: sex, age, ethnicity, home language, SES, kindergarten behavior measures, and kindergarten and first-grade math, reading, and general knowledge standardized test scores.

Behavior is a typically unobservable determinant of student achievement, but the ECLS-K contains several behavior measures. Students' kindergarten teachers rated their behavior along five dimensions. The Approaches to Learning Scale measures behaviors that affect the ease with which children can benefit from the learning environment. The Self-Control Scale has four items that indicate the child's ability to control behavior. The five Interpersonal Skills items rate the child's ability to get along with others. The Externalizing Problem Behaviors scale rate the frequency with which a child acts out, and the Internalizing Problem Behavior Scale asks about the apparent presence of anxiety, loneliness, low self-esteem, and sadness.

Table 6 contains means of student-level variables by class type. K-1 students are more likely to be Hispanic than single-grade students, more likely to internalize problem behaviors, and have lower kindergarten and first-grade reading scores. 1-2 students are more likely to be white and less likely to be black or speak a language other than English than single-grade students. In addition, they appear to be positively selected on behavior and prior test scores—they are better behaved and have higher kindergarten test scores than K-1 or single-grade students. They also have higher first-grade test scores, which could be a result of a 1-2 treatment effect or of positive selection. Note that Table 6 does not take school fixed effects into account—there may be systematic differences across schools that offer 1-2 classes and schools that offer K-1 classes that are being picked up in these average child characteristics.

In order to consider the joint effect of these variables on the assignment to combination classes within schools, I model student selection using school fixed effects. The model is as follows:

$$\begin{aligned} class\_type1_i = {} & \beta_0 + \beta_1 male_i + \beta_2 black_i + \beta_3 hispanic_i + \beta_4 asian_i + \beta_5 other_i \\ & + \beta_6 non\_eng_i + \beta_7 ses_i + \beta_8 learnK_i + \beta_9 controlK_i + \beta_{10} personalK_i \quad (2) \\ & + \beta_{11} externK_i + \beta_{12} internK_i + \beta_{13} readK_i + \beta_{14} mathK_i + \beta_{15} genK_i \\ & + \sum_{j=1}^{J} \delta_j s_j + \varepsilon_i \, . \end{aligned}$$

Students are indexed by i, schools by j; s_j is a school fixed effect. Background characteristics and kindergarten test scores and behavior measures are used as predictors of first-grade class type I run two separate regressions—one for schools offering only single-grade and K-1 classes, and one for schools offering only single-grade and 1-2 classes, as in the previous subsection. In the first, $class\_type1_i$ equals one if the student is in a K-1 combination; in the second, $class\_type1_i$ equals one if the student is in a 1-2 combination.

Regression results are contained in Table 7. There is little evidence for selection into K-1 classes. K-1 students are more likely to internalize problem behaviors than their single-grade counterparts, but F-tests of the joint significance of kindergarten test scores and behavior measures fail to reject the null hypothesis.[9]

The table gives mixed evidence for selection into 1-2 classes. 1-2 students have significantly higher kindergarten math scores but appear to be less well behaved than single grade students—are they placed into 1-2 classes because they would be bored in a single-grade class? Considering the results of F-tests of joint significance of the test score and behavior measures, however, we see strong evidence that high-achieving first-graders are assigned to 1-2 classes. An F-test of kindergarten behavior measures alone fails to reject the null hypothesis, but F-tests of kindergarten test scores alone and with the behavior measures show that these variables are jointly significant.[10]

This positive selection will bias estimates of the combination-class effect upwards unless I can control for the variables influencing class assignment. Including prior-year

[9] Testing the joint significance of kindergarten behavior measures, I obtain $F(5, 235) = 1.20$, p-value = 0.312. Testing kindergarten test scores, I obtain $F(3, 235) = 0.42$, p-value = 0.70. Testing behavior measures and test scores, I obtain $F(8, 235) = 1.05$, p-value = 0.402.

[10] Testing the joint significance of kindergarten social rating scores, I obtain $F(5, 521) = 1.30$, p-value = 0.261. Testing kindergarten test scores, I obtain $F(3, 521) = 3.76$, p-value = 0.011. Testing social rating and test scores, I obtain $F(8, 521) = 2.26$, p-value = 0.022.

test scores and behavior measures in the outcome regressions, discussed below, seems to accomplish this and allows me to estimate a coefficient that can be interpreted causally.

V. Results

In this section, I discuss the results from four outcome-regression models. The dependent variables are first-grade reading, math, and general knowledge test scores. I run one regression per test score for a total of three regressions per model. The independent variables differ by model, but all include school fixed effects. Model 1 contains only dummies for class type, with single-grade classes being the omitted category. Model 2 contains class-type dummies as well as the student background characteristics sex, age, ethnicity, home language, and SES. Model 3 contains combination-class dummies, background characteristics, and kindergarten test scores. Finally, Model 4 contains class-type dummies, background characteristics, kindergarten test scores, and kindergarten behavior measures. This information is summarized in Table 8.

Table 9 contains the coefficients on the K-1 and 1-2 dummies from each of the three regressions in each of the four models. As controls are added, the coefficient on K-1 membership moves from negative to positive but is insignificant at the 5% level in all cases (it is significant at the 10% level in the Model 2 regression of first-grade reading test scores on combination-class dummies and student characteristics). The coefficient on 1-2 membership shrinks as controls are added, but retains significance for math scores when the full set of controls is used.

The 1-2 coefficient of 1.3 in the Model 4 regression of first-grade math scores on combination class dummies, student characteristics, kindergarten test scores, and kindergarten behavior measures indicates that 1-2 membership is associated with nearly a two-percentile-point gain in first-grade math test scores relative to single-grade students. This is approximately one-seventh of a standard deviation. Interpreting this causally could be problematic due to the positive selection of 1-2 students documented in the previous section. Note, however, that the signs, magnitudes, and significance levels of the class-type dummies do not change much between Model 3 and Model 4 in the math test score regressions. Nor do the adjusted R-squared values change substantially between these two models. The difference between them is that the Model 4 contains child social rating scores that proxy for qualities such as behavior and motivation that are usually unobservable to the econometrician. That signs, magnitudes, significance levels, and adjusted R-squared values do not change substantially between Models 3 and 4 is one indication that kindergarten test scores are a good proxy for ability and other usually unobservable characteristics such as behavior and motivation, and allows me to conclude that I have adequately controlled for selection bias.[11]

In order to address the possible upward bias from possible peer selection, I run the Model 4 regression including average peer test scores as independent variables. The signs, magnitudes, and significance levels of the coefficients are nearly identical,[12] indicating that peer effects are not a source of bias.

In the models discussed above, I have imposed a linear relationship on student characteristics and test scores. In order to determine whether my results are sensitive to

[11] In addition, I tried several different specifications of these models in which I included behavior measures first, then added test scores. (Please see Appendix A, at http://econ.ucsd.edu/~jlthomas/research, for detailed results). In the model with background characteristics and behavior measures but no test scores, the coefficient on the 1-2 dummy in the reading regression was 1.372 and significant at the 10 percent level; otherwise, estimates were qualitatively similar to Model 3, above. The adjusted R-squared values were approximately 0.4, however—much smaller than in Model 3, indicating that including test scores without behavior measures adds more information than including behavior measures without test scores.
[12] Appendix A contains these results.

this linear structure, I re-estimate the combination-class treatment effect using propensity score matching. I find that none of these estimates is significant. The coefficient of interest, however, (the effect of being in a 1-2 combination class on first-grade math scores) is 1.7, slightly larger than the OLS estimate,[13] indicating that the linear model does not overstate the effect of being in a 1-2 class on math scores.

VI. Overall Impact

Sims (2008) finds that children in schools with a higher percentage of students in combination classes perform worse than children in schools with fewer combination-class students. This could be because, once combination classes are implemented, single-grade students do worse than they would have if the school had not implemented combination classes, perhaps because resources are diverted to the combination classes and away from single-grade classes. In this case, my finding that first-graders in 1-2 classes outperform their single-grade peers could be explained by single-grade students doing worse than they would have had the school not implemented combination classes.

Addressing the question of whether 1-2 students benefit at the expense of other first-graders is difficult, however, because schools that offer combination classes are quite different from schools that do not. One way to address this would be to regress test score outcomes on a dummy for whether the school offers combination classes, kindergarten test scores, kindergarten behavior measures, student background characteristics, as well as school-level controls. If the school-level controls accounted for all the relevant differences between schools that choose to offer combination classes and those that do not, the coefficient on the school-type dummy could be interpreted as the causal effect of offering combination classes on first-grade test scores.

[13] These results are contained in Appendix A.

As a rudimentary check that, overall, first-graders are not harmed by a school's decision to offer combination classes, I compare single-grade schools to schools offering first grade and a 1-2 combination by regressing first-grade test scores on a dummy indicating that the school offers single-grade and 1-2 classes. As other independent variables, I include the student-level variables from Model 4, as well as the following school-level covariates: indicators for region, community size and year-round school, average grade-level enrollment, standard deviation of enrollment across grades, full-time equivalent teachers per student, percent minority, and percent eligible for free lunch. Table 10 contains the coefficients of interest (please see Appendix A for full regression results).[14]

None of the coefficients is significant, though the point estimates are negative for reading and math test scores. As discussed above, schools that offer combination classes tend to be larger, have a higher percentage of minority students, and have fewer teachers per student than schools that do not offer combination classes. That is, combination-class schools tend to be more disadvantaged than single-grade schools. To the extent that this is true for unobservable school characteristics influencing the choice to offer combination classes and student outcomes, we can assume that these coefficients are biased downwards. This reinforces the conclusion that, overall, the decision to offer combination classes, at least at the 1-2 level, does not harm first-graders overall. Thus, combination classes may be a Pareto-improving option for school administrators.

[14] Appendix A can be found at http://econ.ucsd.edu/~jlthomas/research.

VII. Conclusion

In this paper, I address the selection issues involved in all three levels of implementing a combination class. To address school-level selection, I limit the sample to combination-class schools and use school fixed effects in the outcome regressions. To address teacher-level selection, I model teacher assignment to combination classes and find little evidence for meaningful nonrandom selection. To address student-level selection, I model student assignment to combination classes and find evidence for positive selection into 1-2 classes. I therefore use a rich set of control variables, including behavior measures that are usually unavailable to the econometrician, to more plausibly assume ignorability of treatment and estimate the causal effect of combination-class membership in first-grade on first-grade test scores.

I find that there is no effect on reading or general knowledge scores for students in either type of class, but that 1-2 combination-class membership is associated with an increase of one-seventh of a standard deviation on math test scores relative to single-grade students. This result is not sensitive to functional form assumptions. In addition, I find little evidence that 1-2 students benefit at the expense of other first-graders.

These results indicate that combination-class membership in first grade has at worst, no effect and at best, a small positive effect on student achievement as measured by test scores. I conclude that combination classes may be a Pareto-improving option for school administrators. Given that more and more states are implementing class size reduction initiatives, and that combination classes conserve scarce resources by allowing schools to use fewer teachers and classrooms, it is more important than ever for school administrators to find ways to reduce class size in the least costly manner. Combination classes allow school administrators to reduce class size within one grade while smoothing class size across grades, and should be considered as a viable means of classroom organization.

It should be acknowledged, however, that implementing combination classes is problematic for other reasons. Teachers do not like them (Mason, Burns, and Armesto, 1993). In addition, parents may not want their children to be placed in the higher grade of a combination class because they perceive this as a signal that their children are low achievers, even though the data I have presented here indicate that K-1 students are statistically indistinguishable from their single-grade counterparts.

This paper shows that 1-2 (i.e., lower-grade) students benefit from combination-class membership. These students are relatively young compared to their classmates, and this result supports recent findings in the age-at-school entry literature that relatively younger students benefit from having older peers. An interesting direction for future research would be to determine if lower-grade students in other combination classes (e.g., third graders in a 3-4 class) also benefit, and if these benefits persist over time.

References

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber (2005). "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy* 113, pp. 151-184.

Bedard, Kelly and Elizabeth Dhuey (2006). "The Persistence of Early Childhood Maturity: International Evidence of Long-run Age Effects," *Quarterly Journal of Economics* 121, pp. 1437-1472.

Betts, Julian (1996). "Is There a Link Between School Inputs and Earnings? Fresh Scrutiny of an Old Literature," University of California at San Diego, Economics Working Paper Series 96-09, Department of Economics, UC San Diego.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes (2008). "Too Young to Leave the Nest? The Effects of School Starting Age," NBER Working Paper 13969, National Bureau of Economic Research, Cambridge, MA.

Burns, Robert, DeWayne Mason, and Michael A. Demiranda (1993). "How Elementary Principals Assign Teachers and Students to Combination Classes," California Educational Research Cooperative, University of California, Riverside.

California Department of Education (2006). "Statistical Summary of Year-round Programs, 2005-2006," <http://www.cde.ca.gov/ls/fa/yr/guide.asp>

California Department of Education (2009). "Fingertip Facts: K-3 Class Size Reduction Program History and Summarized Data," <http://www.cde.ca.gov/ls/cs/k3/facts.asp>.

Datar, Ashlesha (2006). "Does Delaying Kindergarten Entrance Give Children a Head Start?" *Economics of Education Review* 25, pp. 43-62.

Elder, Todd E. and Darren H. Lubotsky (2006). "Kindergarten Entrance Age and Children's Achievement: Impacts of State Policies, Family Background, and Peers," *Journal of Human Resources*.

Graves, Jennifer (2007). "The Effect of Year Round School Calendars on Academic Performance," Unpublished Manuscript, University of California, Irvine.

Hanushek, Eric A. (1986). "The Economics of Schooling: Production and Efficiency in Public Schools," *Journal of Economic Literature* 24, pp. 1141-1177.

Hill, Peter W. and Kenneth J. Rowe (1998). "Modelling Student Progress in Studies of Educational Effectiveness," *Social Effectiveness and School Improvement* 9, pp. 310-333.

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

Mason, DeWayne A. and Robert B. Burns (1994). "A Review of the Literature on Combination Classes," California Educational Research Cooperative, University of California, Riverside.

______ (1995a). "Organizational Constraints on the Formation of Elementary School Classes," *American Journal of Education* 103:2, pp. 185-212.

______ (1995b). "Teachers' views of combination classes," *Journal of Educational Research*, 89:1, pp. 36–45.

______ (1996). "'Simply No Worse and Simply No Better' May Simply Be Wrong: A Critique of Veenman's Conclusion about Multigrade Classes," *Review of Educational Research* 66:3, pp. 307-322.

Mason, DeWayne A. and Roland Doepner (1998). "Principals' Views of Combination Classes," *Journal of Educational Research*, 91:3, pp. 160–172.

Mason, DeWayne A., Robert Burns, and Jorge Armesto (1993). "Teachers' Views about Combination Classes," California Educational Research Cooperative, University of California, Riverside.

Mason, DeWayne A. and Janet Stimson (1994). "A National Survey of Combination and Nongraded Classes," California Educational Research Cooperative, University of California, Riverside.

Miller, William (1995). "Are Multi-Age Grouping Practices a Missing Link in the Educational Reform Debate?" *NAASP Bulletin 1995*, National Association of Secondary School Principals, Reston, VA.

Mishel, Lawrence and Richard Rothstein, eds. (2002). *The Class Size Debate*, Economic Policy Institute, Washington, D.C.

National Association for Year-round Education (2007). "Statistical Summaries of Year-round Education Programs, 2006-2007."

Veenman, Simon (1995). "Cognitive and Noncognitive Effects of Multigrade and Multi-Age Classes: A Best-Evidence Synthesis," *Review of Educational Research* 65:4, pp. 319-381.

______ (1996). "Effects of Multigrade and Multi-Age Classes Reconsidered," *Review of Educational Research* 66:3, pp. 323-340.

Wooldridge, Jeffrey M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Massachusetts: The MIT Press.

Stipek, Deborah (2002). "At What Age Should Children Enter Kindergarten? A Question for Policy Makers and Parents," *Social Policy Report: a Publication of the Society for Research in Child Development*, 16, pp. 3-16.

Sims, David (2008). "A Strategic Response to Class Size Reduction: Combination Classes and Student Achievement in California," *Journal of Policy Analysis and Management* 27, pp. 457-478.

Table 1: Types of First-grade Classes Offered

Combination and/or multi-grade class offering	*No. of schools*	*Pct. of schools*
Single-grade first grade class only	845	0.834
Any type of combination or multi-grade class	168	0.166
Combination-class schools	*92*	*0.091*
First, 1-2	55	0.054
First, K-1	26	0.026
First, K-1, 1-2	9	0.009
K-1, 1-2	2	0.002
Multi-grade schools	*76*	*0.075*
First, other	39	0.038
K-1 only	11	0.011
1-2 only	8	0.008
Other only	8	0.008
First, 1-2, other	6	0.006
1-2, other	2	0.002
First, K-1, 1-2, other	1	0.001
First, K-1, other	1	0.001
K-1, 1-2, other	0	0.000
K-1, other	0	0.000

Table 2: Comparison of Combination Class Schools to Schools Offering Single-Grade Classes Only

School characteristic	*Schools offering single-grade classes only*	*Combination-class schools*
West	0.206 (0.014)	0.446 (0.052)***
FTE teachers per student	0.061 (0.000)	0.057 (0.001)***
Northeast	0.187 (0.013)	0.076 (0.028)***
Midwest	0.239 (0.015)	0.130 (0.035)**
Percent minority	41.722 (1.209)	50.244 (3.406)**
Year-round	0.041 (0.009)	0.095 (0.032)**
Average grade-level enrollment over grades K-2	90.111 (1.566)	98.417 (4.661)*
Std. dev. of grade-level enrollment over grades K-2	10.871 (0.327)	9.210 (0.581)*
Total enrollment	558.743 (8.929)	598.761 (27.487)
Average grade-level enrollment over all grades	87.356 (1.457)	93.739 (4.173)
Suburb	0.398 (0.017)	0.337 (0.050)
City	0.396 (0.017)	0.457 (0.052)
Town	0.086 (0.010)	0.109 (0.033)
Rural	0.120 (0.011)	0.098 (0.031)
South	0.368 (0.017)	0.348 (0.050)
Standard deviation of grade-level enrollment over all grades	18.626 (0.568)	19.005 (1.877)
Percent of students eligible for free lunch	33.705 (1.246)	33.623 (4.042)

Note: This table contains the results of a two-sample Student's t-test assuming equal variances. * denotes that the means are significantly different at the 10% level, ** at the 5% level, and *** at the 10% level.

Table 3: Means of Teacher Characteristics by Class Type

Teacher characteristic	*Single-grade 1st mean*	*K-1 mean*	*1-2 mean*
Male	0.049 (0.013)	0 (0)	0.010 (0.010)*
White	0.761 (0.025)	0.638 (0.071)*	0.835 (0.038)†††
Black	0.035 (0.011)	0.085 (0.041)	0.010 (0.010)††
Hispanic	0.165 (0.022)	0.213 (0.060)	0.082 (0.028)**,††
Asian	0.021 (0.009)	0.043 (0.030)	0.062 (0.025)**
Other	0.018 (0.008)	0.021 (0.021)	0.010 (0.010)
Years teaching	11.846 (0.564)	13.553 (1.431)	12.402 (0.953)
B.A. or less	0.226 (0.025)	0.25 (0.066)	0.245 (0.045)
Some graduate school	0.373 (0.029)	0.295 (0.070)	0.394 (0.051)
Graduate degree	0.401 (0.029)	0.455 (0.076)	0.362 (0.050)
Enjoys present teaching job	4.358 (0.044)	4.447 (0.109)	4.371 (0.094)
Makes a difference	4.503 (0.036)	4.574 (0.073)	4.618 (0.056)
Would choose teaching again	4.292 (0.057)	4.319 (0.140)	4.484 (0.090)*
Paid prep hours per week	1.906 (0.049)	1.804 (0.115)	1.889 (0.070)
Unpaid prep hours per week	3.613 (0.060)	3.362 (0.123)	3.793 (0.098)†††

Notes: I consider only teachers in combination-class schools. Of these, 293 teach single-grade first, 47 teach K-1 combinations, and 99 teach 1-2 combinations. Standard errors are in parentheses. * denotes that the K-1 or the 1-2 mean is different from the single-grade mean at the 10% level, ** at the 5% level, and *** at the 1% level. † denotes that the 1-2 mean is different from the K-1 mean at the 10% level, †† at the 5% level, and ††† at the 1% level.

Table 4: Modeling Teacher Selection

Teacher characteristic	Regression 1: K-1 combination dummy as dependent variable	Regression 2: 1-2 combination dummy as dependent variable
Male	-0.365 (0.158)**	-0.303 (0.189)
Black	0.536 (0.186)***	-0.153 (0.216)
Hispanic	0.232 (0.138)*	-0.094 (0.135)
Asian	0.458 (0.360)	0.482 (0.260)*
Other	0.451 (0.395)	0.237 (0.341)
Years teaching	0.006 (0.005)	0.005 (0.197)
Some graduate school	-0.151 (0.155)	-0.049 (0.676)
Graduate degree	-0.107 (0.148)	-0.116 (0.335)
Constant	0.202 (0.137)	0.285 (0.088)**
Regression statistics	*Regression 1*	*Regression 2*
Number of obs.	116	239
p-value of F statistic	0.025	0.264
Adj. R-squared	-0.061	-0.132

Notes: Table 4 contains the results of two linear regressions of class-type dummies on teacher characteristics. Both regressions include school fixed effects. In Regression 1, the sample is restricted to the 26 schools offering only single-grade first and K-1 classes. In Regression 2, the sample is restricted to the 55 schools offering only single-grade first and 1-2 classes. Robust standard errors are in parentheses.
a) F(8, 82) for the K-1 regression; F(8, 176) for the 1-2 regression.

Table 5: Within-school Means of Classroom Characteristics by Class Type

Classroom characteristic	*Single-grade 1st mean*	*K-1 mean*	*1-2 mean*
Class size	20.800 (0.169)	20.288 (0.733)	21.338 (0.669)
Percent boys	0.512 (0.006)	0.513 (0.026)	0.509 (0.016)
Percent minority	53.949 (0.822)	52.996 (2.596)	51.663 (2.149)
Percent gifted	0.016 (0.006)	0.053 (0.024)	0.082 (0.026)**
Percent limited English proficiency	0.395 (0.029)	0.315 (0.056)	0.283 (0.070)
Percent reading below grade level	0.269 (0.011)	0.315 (0.046)	0.252 (0.025)
Percent math below grade level	0.185 (0.009)	0.238 (0.040)	0.193 (0.020)
Teacher-directed whole class activity	3.966 (0.044)	3.561 (0.151)***	3.795 (0.112)
Teacher-directed small group activities	3.550 (0.050)	3.670 (0.140)	3.654 (0.108)
Teacher-directed individual activities	2.794 (0.052)	2.458 (0.139)**	2.689 (0.120)
Child-selected activities	2.483 (0.042)	2.873 (0.114)***	2.589 (0.091)†
Percent 5 years or younger	0.001 (0.003)	0.150 (0.025)***	0.0001 (0.009)†††
Percent 6 years old	0.406 (0.011)	0.495 (0.027)***	0.201 (0.025)***,†††
Percent 7 years old	0.558 (0.011)	0.352 (0.033)***	0.475 (0.027)***,†††
Percent 8 years old	0.035 (0.005)	0.002 (0.013)**	0.299 (0.022)***,†††
Percent 9 years old	0.0001 (0.0001)	-0.0001 (0.004)	0.024 (0.007)***,†††
Percent 10 years or older	0.0001 (0.0001)	0.0001 (0.0002)	0.001 (0.001)

Table 5 contains the results of regressions of each of the classroom-level variables on dummies for K-1 class and 1-2 class and school fixed effects. Single-grade first grade classes form the base case. Standard errors are in parentheses. * denotes that the K-1 or the 1-2 mean is different from the single-grade mean at the 10% level, ** at the 5% level, and *** at the 1% level. † denotes that the 1-2 mean is different from the K-1 mean at the 10% level, †† at the 5% level, and ††† at the 1% level.

Table 6: Means of Student Characteristics by Class Type

Student characteristic	*Single-grade 1st mean*	*K-1 mean*	*1-2 mean*
Male	0.513 (0.016)	0.458 (0.051)	0.489 (0.033)
Age in months (Spring 1st)	85.874 (0.135)	86.066 (0.419)	86.103 (0.264)
White	0.460 (0.016)	0.417 (0.051)	0.529 (0.033)*,†
Black	0.130 (0.011)	0.094 (0.030)	0.08 (0.018)**
Hispanic	0.290 (0.015)	0.375 (0.050)*	0.258 (0.029)††
Asian	0.063 (0.008)	0.031 (0.018)	0.076 (0.018)
Other	0.057 (0.008)	0.083 (0.028)	0.058 (0.016)
Language other than English spoken at home	0.222 (0.014)	0.161 (0.038)	0.146 (0.024)**
SES	-0.119 (0.027)	-0.189 (0.079)	-0.048 (0.051)
Approaches to learning (Spring K)	3.097 (0.022)	3.058 (0.068)	3.215 (0.043)**,†
Self-control (Spring K)	3.165 (0.021)	3.152 (0.072)	3.271 (0.039)**
Interpersonal (Spring K)	3.108 (0.022)	3.097 (0.071)	3.228 (0.039)**,†
Externalizing problem behaviors (Spring K)	1.684 (0.022)	1.615 (0.079)	1.535 (0.039)***
Internalizing problem behaviors (Spring K)	1.534 (0.016)	1.656 (0.064)**	1.524 (0.032)††
Reading test score (Spring K)	50.862 (0.315)	48.629 (1.251)**	52.694 (0.695)**,†††
Math test score (Spring K)	49.763 (0.308)	48.424 (1.222)	51.520 (0.651)**,††
General Knowledge test score (Spring K)	49.724 (0.335)	48.085 (1.166)	51.322 (0.644)** ,††
Reading test score (Spring 1st)	50.388 (0.283)	48.643 (1.034)*	51.737 (0.669)** ,††
Math test score (Spring 1st)	49.908 (0.297)	48.685 (1.313)	52.052 (0.596)*** ,†††
General Knowledge test score (Spring 1st)	49.588 (0.321)	48.403 (1.130)	51.043 (0.618)** ,††

Notes: 931 students in single-grade first; 96 in K-1; 225 in 1-2. Standard errors in parentheses. * denotes that the K-1 or the 1-2 mean is different from the single-grade mean at the 10% level; **, the 5% level; ***, the 1% level. † denotes that the 1-2 mean is different from the K-1 mean at the 10% level; ††, the 5% level, †††, the 1% level.

Table 7: Modeling Student Selection

Student characteristic	Regression 1: K-1 combination dummy as dependent variable	Regression 2: 1-2 combination dummy as dependent variable
Male	-0.005 (0.052)	-0.023 (0.037)
Age in months (Spring 1st)	0.009 (0.007)	0.004 (0.005)
Black	-0.243 (0.090)***	-0.125 (0.069)*
Hispanic	0.080 (0.104)	0.021 (0.061)
Asian	0.015 (0.111)	-0.025 (0.082)
Other	0.071 (0.115)	0.005 (0.089)
Language other than English spoken at home	-0.013 (0.107)	0.023 (0.071)
SES	-0.008 (0.040)	0.041 (0.030)
Approaches to learning	0.015 (0.064)	-0.086 (0.050)*
Self-control	0.017 (0.078)	-0.027 (0.066)
Interpersonal	-0.030 (0.073)	0.079 (0.053)
Externalizing problem behaviors	-0.058 (0.055)	-0.058 (0.044)
Internalizing problem behaviors	0.105 (0.054)	0.007 (0.044)
Reading test score (Spring K)	-0.0002 (0.004)	0.007 (0.003)**
Math test score (Spring K)	-0.005 (0.005)	0.002 (0.003)
General Knowledge test score (Spring K)	0.001 (0.004)	0.0004 (0.003)
Constant	-0.486 (0.647)	-0.350 (0.471)
Regression statistics	*Regression 1*	*Regression 2*
Number of obs.	277	591
p-value of F statistic	0.205	0.001
Adj. R-squared	0.085	0.103

Notes: Table 7 contains the results of two linear regressions of class-type dummies on student characteristics. Both regressions include school fixed effects. In Regression 1, the sample is restricted to the 26 schools offering only single-grade first and K-1 classes. In Regression 2, the sample is restricted to the 55 schools offering only single-grade first and 1-2 classes. Robust standard errors are in parentheses.

Table 8: Four Outcome-regression Models

Independent variables	*Model 1*	*Model 2*	*Model 3*	*Model 4*
K-1 and 1-2 dummies	X	X	X	X
Student characteristics		X	X	X
Kindergarten test scores			X	X
Kindergarten social rating scores				X
School fixed effects	X	X	X	X

Table 9: Coefficients of Interest from Four Outcome Regressions

	Dependent variable: first-grade reading test score		
	K-1 combination	*1-2 combination*	
	Coeff.	Coeff.	Adjusted R-squared
Model 1	-1.270 (1.326)	2.319 (0.866)***	0.177
Model 2	-2.224 (1.294)*	1.331 (0.865)	0.265
Model 3	-0.186 (0.768)	-0.071 (0.466)	0.707
Model 4	0.211 (0.784)	0.186 (0.478)	0.708
	Dependent variable: first-grade math test score		
	K-1 combination	*1-2 combination*	
	Coeff.	Coeff.	Adjusted R-squared
Model 1	-2.642 (1.553)*	3.700 (0.814)***	0.183
Model 2	-3.205 (1.477)**	2.637 (0.779)***	0.275
Model 3	0.442 (1.007)	1.286 (0.525)**	0.642
Model 4	0.640 (1.051)	1.333 (0.551)**	0.646
	Dependent variable: first-grade general knowledge test score		
	K-1 combination	*1-2 combination*	
	Coeff.	Coeff.	Adjusted R-squared
Model 1	-1.008 (1.396)	1.975 (0.813)**	0.208
Model 2	-2.100 (1.302)	0.555 (0.720)	0.393
Model 3	0.232 (0.831)	-0.316 (0.471)	0.712
Model 4	0.253 (0.891)	-0.381 (0.481)	0.714

Note: Robust standard errors are in parentheses. Longitudinal weights are used; results are not sensitive to the inclusion of weights or to clustering at the class level. Please see Appendix A, at http://econ.ucsd.edu/~jlthomas/research, for full regression results.

Table 10: Do 1-2 Students Benefit at the Expense of Other First Graders? Results from a Regression of Test Scores on a School-type Dummy and School Characteristics

	Independent variable: dummy for school offering first and 1-2 classes
Dependent variable	*Coefficient*
Reading test score (Spring 1st)	-0.342 (0.414)
Math test score (Spring 1st)	-0.111 (0.812)
General Knowledge test score (Spring 1st)	0.088 (0.786)

Note: Table 10 contains the results of regressing first grade test scores on a dummy indicating that the school offers single-grade and 1-2 classes, as well as student-level variables from Model 4 and following school-level covariates: indicators for region, community size and year-round school, average class size, standard deviation of enrollment across grades, full-time equivalent teachers per student, percent minority, and percent eligible for free lunch. Robust standard errors are in parentheses.