

Procedure and Results

1. Empty strings ("") treated as NA. Categorical variable created for treatment type for each school year k-3. Each academic year treatment variable, student's treatment status takes a value of 1 for placement in a large class, 2 for regular class, and 3 for small sized class. Large class is taken as 30 student plus a teacher's aide and regular as 30 students with no aide. 1900 students were in small classes in kindergarten. 70 students who were in small classes in kindergarten were in large classes in 3rd grade.
2. Balance testing was done using the *ietoolkit* package offered by the World Bank, which allows for balance testing across multiple treatment arms. Balance testing was done between kindergarten class sizes: Large – Regular; Large – Small; Regular – Small. Variables for balance testing were selected according to the pre-treatment student and teacher characteristics requirement outlined in the guidelines. A list of selected variables is available in the footnotes¹. It is worth noting there are significant imbalances across multiple variables raising concern. Figure 1 provides a table of t-test results, where column 1 represents the test between Large and Regular classes, column 2 compares Large to Small, and column 3 represent Regular and Small.

Columns 1 and 3 indicate differences exist in school location (urban/suburban/rural/inner-city), and the race of the teachers. The race composition between Large and Small classes is significantly different. Columns 1 and 2 show differences in the school's offering of free lunches and the number of years instructors had been teaching. Also, it is worth noting that there are no recorded male teachers of any kindergarten class size, indicated by the N/A for all 3 columns for the gktgen variable. We should consider these imbalances when estimating and interpreting results of class size on academic performance.
3. Figure 2 shows results from regressions comparing the 4th grade math and reading test scores of students in different class sizes in kindergarten. I control for the pre-treatment characteristics from the balance test. To facilitate comparison between all groups, I omit the small class treatment group. Compared to students assigned to Regular or Large classes in kindergarten, the 4th grade reading test scores of those placed in Small classes were 8.4 and 6.1 points higher, respectively, significant below the 1% level. For math scores, compared to Regular classes, students in Small kindergarten classes scored 5.24 points higher, significant at the 5% level.

¹ Gender of student (gender), race of student (race), urban/rural/suburban/inner-city status of kindergarten (gksurban), gender of kindergarten teacher (gktgen), race of kindergarten teacher (gktrace), highest degree of kindergarten teacher (gkthighdegree), free lunch offered at kindergarten (gkfreelunch), number of years kindergarten instructor has been teaching (gktyears)

Looking strictly at those student in Small classes and Regular classes, we can say that students in smaller classes perform better.

4. Using the treatment indicator created in exercise 1, I can assess student treatment by academic year k-3. Accordingly, 51 students who were in Small kindergarten classes were in Regular classes in 3rd grade.

To identify the number of years a student was in a small class, I create dummies for each academic year that are equal to 1 if the student was in a small class and 0 otherwise. I sum these dummies for all years k-3rd grade to create the *num_yearssmall* variable. Table 1 presents a cross tabulation of the number of years a student was in small classes with their treatment status in kindergarten.

5. To estimate the treatment intensity effect, I regress the number of years a student was in a Small class against their reading and math scores. Figure 3 displays the results. Again I control for the balance test variables, which drastically improves the precisions (adjusted R^2) of both estimations compared to no controls. Being in Small classes for an additional year is associated with an increase in reading scores by 3.7 points and math scores by 1.4 points. Both are significant below the 1% level.

While the included control variables account for more variation in test scores, our model overall is still imprecise. Additional covariates should be considered if available. Of particular interest are household characteristics such as income, education, and marital status of the student's family.