Enrollment Manipulation, Class Size Caps and Educational Segregation*

Michael Gilraine and Robert McMillan

August 31, 2018

ABSTRACT

This paper examines public school responses to financial incentives, showing evidence of enrollment manipulation that leads to increased segregation across schools. We do so in the context of California's statewide class size reduction program – up to that time, the most costly state education reform ever implemented in the United States. The program used financial incentives to encourage schools to keep class sizes at twenty (or below) in grades K-3. As a response to those incentives, we document extensive enrollment manipulation at enrollment multiples of twenty, particularly concentrated in schools likely to be over-subscribed under California's open enrollment system – those with high test scores and high shares of advantaged students. That system allows schools to refuse admission to students choosing schools from within the district but outside the relevant school attendance zone. Given such discretion, we hypothesize that some schools were able to manipulate enrollment by restricting access to students living outside the school attendance zone, especially minority students from disadvantaged backgrounds. Using the end of the twenty-student cap in the 2008-09 school year, we employ a transparent triple-differences strategy to show that this shift led to a decline in both racial and income segregation within districts. In general, our analysis draws attention to the importance of financial incentives in determining public school behavior, sometimes in ways policy makers did not intend.

Keywords: Class Size Cap, Enrollment Manipulation, Bunching, Financial Incentives, Open Enrollment, Segregation, Difference-in-Differences, Triple-Differences

JEL Classifications: H40, I21, I22

*We would like to thank workshop participants at the University of Toronto for comments. Financial support from SSHRC is gratefully acknowledged. All remaining errors are our own. Contact information: Gilraine, New York University (mike.gilraine@nyu.edu); McMillan, University of Toronto and NBER (mcmillan@chass.utoronto.ca).

1 Introduction

In light of chronic concerns about the underperformance of many public schools, policy makers in nearly every state have implemented open enrollment policies that allow students to apply to attend schools outside their school attendance zone.[1] Such policies have the potential to increase household choice and stimulate competition, with their attendant beneifts. Yet the likely potency of open enrollment policies in practice depends in a key way on the extent to which school capacity is allowed to adjust to reflect shifts in demand. In many state education systems, where teacher unions have played an important role in shaping public education provision,[2] such adjustments are typically muted, with policies making a fixed (and often limited) number of 'open enrollment' seats available at a particular school.

In this paper, we shed new light on how flexible public school capacity actually is, and how public schools seek to manipulate it. The bunching patterns we document in school enrollments provide particularly clear evidence of the extent of manipulation.[3] In turn, we are able to draw out consequences of that manipulation in the context of a heavily regulated public school system, showing segregation-related outcomes likely at odds with the intent of education policy makers.

Our analysis focuses on a class size cap reform in California that provided strong financial incentives to schools when they maintained class sizes at twenty students or below. We show that many public schools responded to this financial incentive by restricting capacity. Coupled with the operation of California's intra-district open enrollment system, those schools with excess demand were likely to be the better schools, at least in multi-school districts, and were able to limit access to minority students from disadvantaged

[1] Open enrollment policies vary significantly both across- and within-state, with some being mandatory while others are voluntary. Only three states lack any form of open enrollment policy. For an up-to-date listing by state, see `http://ecs.force.com/mbdata/MBQuestNB2n?rep=OE1701`.

[2] See Hoxby (1996) for a compelling empirical analysis of the impact of unions on educational inputs and productivity.

[3] Urquiola and Verhoogen (2009) find similar manipulation with respect to class size caps in the Chilean private school system.

backgrounds living outside the relevant attendance zone. Further, using the end of the twenty-student cap in the 2008-09 school year, we employ a transparent triple-differences strategy to show that this shift led to a decline in both racial and income segregation within districts. The class size caps thus had an unintended consequence of restricting student access to the highest quality schools, further exacerbating educational segregation. In turn, such segregation is likely to have broader adverse consequences for the inequality of educational outcomes.

2 California's Class Reduction Program

The financial incentives we focus on were associated with California's class size reduction (CSR) program of the late-1990s.

2.1 Background

California implemented the largest state-sponsored class size reduction program in the United States in July 1996.[4] The law's introduction was part of a series of education reforms undertaken by the state in response to the 1992 release of California's dismal national test rankings. Inspired by Project STAR, the class size reduction program provided financial incentives for schools to reduce class sizes in grades K-3 to twenty student or less, relative to a pre-CSR statewide average of 28.5 students in elementary schools.

The program was enacted in July 1996 and went into effect for grade 1 that September. Participation in the program was voluntary at both the district- and school-level. At the district-level, participation was virtually universal due to the high level of financial incentives, while at the school-level participation levels were high but not universal.[5] The program paid $650 for each pupil enrolled in a class of 20 or fewer students in grade 1

[4]See Schrag (2006) for a more detailed description of CSR in California.

[5]For districts to participate in CSR, they only had to opt in to the program. For its schools to receive CSR funding, however, they had to reduce class sizes in the relevant grade.

which, relative to average 1995-96 per-pupil expenditures of $6,068, provided a sizeable incentive for schools to reduce class sizes to 20 students or less. After reducing their grade 1 class sizes, schools could then receive funding for decreasing their grade 2 class sizes the following year. Afterwards, schools would receive funding for decreasing class sizes in either kindergarten or grade 3. By the 1999-2000 school year, all grades K-3 would be subject to CSR.

The broad support from parents and teachers for policies reducing class size notwithstanding, California's class size reduction reform was completely unanticipated. Leading up to the reform, the Republican Governor Peter Wilson dismissed class size reduction as a waste of money.

His change of heart was precipitated by a requirement to spend a large portion of California's tax revenue on education due to a narrowly-passed constitutional initiative in 1988. Sensing the Governor's reluctance to give the tax revenues to schools, the California Teachers' Association (CTA) began attacking Governor Wilson over the state's crowded classrooms. The Governor viewed the union's class size campaign as a cover, thinking that the real agenda of the CTA was to funnel unrestricted funding to school districts, whose boards were dominated by local unions, to use for additional teacher pay. To prevent funds going directly to boost teacher salaries, the Governor created the incentive framework to reduce class sizes.

The law that emerged was written with little regard to optimal policy. Initially proposed for only grades 1 and 2, it was expanded to kindergarten and grade 3 when the Democratic Senate balked at tax cuts embedded in the class size reduction bill. Additionally, the class size goal of 20 was set, according to an unidentified Wilson staffer in Schrag (2006), because it "was what we could afford."

Despite its scale, California's CSR remains a relatively unstudied reform. Researchers analyzing CSR face challenges due to its near-universal adoption and data issues during the time period, with California lacking student-level microdata, teacher assignment data

and student achievement measures.[6] Data constraints to one side, several researchers have studied the implementation of CSR, including the hiring of uncertified teachers and the migration of higher quality teachers to wealthier districts (Bohrnstedt and Stecher, 2002; Jepsen and Rivkin, 2009). Others have looked at the effects of CSR on student achievement, though delivering conflicting results.[7]

2009-10 CSR rule change: In time for the beginning of the 2009-10 school year, the rule was changed so that the school would instead lose 5 percent of its CSR funding in a class for each student above 20 in that class. Thus the 'hard' cap was eliminated, under which a school would lose all of its CSR funding for a class if its class size went above 20 students.

2.2 Incentives

This subsection highlights the financial incentives that schools faced under the program. As a representative illustration, we consider the incentives faced by schools in the 1998-99 school year, when per-pupil funding was $6,025 per student and CSR funding was $832 per student.[8] Average teacher salaries were $46,129.[9] I define the 'net funding' of a school as the total funding that it receives from the state, minus the cost of teacher salaries.

Financial Incentives to Reduce Class Size: Figure 1(a) represents the financial incentives faced by a school-grade in the 1998-99 school year. The figure plots total funding minus teacher salaries by school-grade enrollment for a hypothetical school that would, in the absence of CSR, only lower class sizes once average class sizes exceeded 35

[6]Standardized testing in California started in the 1997-98 school year – one year after CSR was implemented.

[7]Bohrnstedt and Stecher (2002), Stecher et al. (2003) and Funkhouser (2009) find that CSR has, at best, minimal effects on student achievement. Unlu (2005) and Jepsen and Rivkin (2009), on the other hand, find that CSR improved student test scores by 0.2-0.3 and 0.06-0.1 of a standard deviation, respectively.

[8]In 1996-97, this payment was $650 per student. The payment increased each year: by the 2002-03 school year, it was $906.

[9]Average teacher salaries are used for the exercises below, though schools could hire less experienced teachers at lower salaries: the minimum salary for a newly-hired teacher in California was $31,129.

students. As we can see, if the school-grade does not follow CSR and therefore keeps only one class once enrollment exceeds 20 students, the school loses $17,472 ($832*21) in CSR funding because its class size now exceeds 20 students. If the school followed CSR, however, it would retain the $17,472 in CSR funding, but would be required to hire a new teacher at a cost of $46,129. Therefore, the marginal cost of adding a class for a school with grade-level enrollment of twenty-one students goes is $28,657 with CSR, implying that schools have a strong financial incentive to not follow CSR, at least when their school-grade enrollment just exceeds a multiple of 20.

Financial Incentives to Create Combination Classes: Figure 1(b) displays the underlying financial incentives in the 1998-99 school year for a hypothetical two-grade school with identical enrollments in each grade. The line with circles shows the net funding that a school with a given enrollment level receives when it keeps its class sizes below twenty students. The line with triangles shows the net funding such a school receives if it keeps class sizes at twenty students or below, but uses 'combination' classes – those that mix students from different grades (most likely, adjacent grades) in the same classroom. We can see a clear discontinuity at grade-level enrollments of thirty (and fifty), where a school with an enrollment of thirty in grades two and three can create three classes of twenty students each: one for grade two, one for grade three, and one combination class. If the enrollment in each grade is thirty-one, however, the school must have four classes.

Financial Incentives to Manipulate Enrollment: This discussion makes clear that for schools following the maximum class size rule, a school with a grade-level enrollment of twenty students has more net funding than a school with a grade-level enrollment of twenty-one students, even neglecting the marginal cost of educating another student. Schools therefore have an incentive to limit their enrollment to twenty students, rather than adding a twenty-first student. This intuition holds true for each enrollment multiple of twenty.

Similarly, schools that use combination classes receive more money if they keep the combined grade 2 and grade 3 enrollment at 60 (thirty in each grade) rather than adding the sixty-first student. The school therefore has a strong incentive to limit its enrollment to sixty students, rather than accepting the sixty-first student. This intuition holds true for each combined enrollment multiple of $60, 100, 140, 180, ...$

3 Data

Our data is derived from publicly available data provided by the California Department of Education (CDE). First, we use the public school enrollment data files to calculate student enrollment and racial composition at the school-grade level for the 1990-91 through the 2014-15 school years. We then augment these enrollment data with CDE datasets that provides 'English as a Second Language' (ESL) status at the school-grade level and Free or Reduced-Price Meal status at the school level for this time period.[10]

We also use test score data from California's Standardized Testing and Reporting Program (STAR) for second through eleventh grade for school years 1997-98 through 2012-13.[11] All students in second through eleventh grade (with some minor exceptions[12]) took the Stanford Achievement Test in both mathematics and English near the end of the academic year. The Stanford Achievement Test was a national norm-referenced multiple-choice test introduced in the 1997-98 school year. To allow comparisons across grades and over time, we standardize test scores to have mean zero and standard deviation for each grade-year.

Table 1 displays mean K-3 school-grade characteristics for high- and low-bunching

[10]This serves as a measure of the poverty rate of the entire student body since only students whose household income is below a threshold based on a percentage of the poverty line are eligible. It is not available at the grade level, unlike our other public school demographic variables.

[11]California moved to a new testing system for the 2013-14 school year, which did not test students in either math or English in that year.

[12]Students were exempted if they were special education students or if a parent or guardian submitted a written request for exemption. Test taking rates were high nonetheless. For example, in 1998-99, over ninety-three percent of students in grades 2-11 took the relevant test.

schools before and after the hard class size cap was lifted. We define high-bunching schools as those whose school-grade enrollment falls on an enrollment multiple of 20 (or one below that multiple) twice as often as expected by chance.[13] The first row shows the rate of bunching for high- and low-bunching schools. Column (1) shows that during the years the hard cap was in place a high-bunching school ended up at the enrollment multiple of 20 (or one below) 32 percent of time; this declined to 12.4 percent (column (2)) after the hard cap was eliminated. Contrasting this change to the change in low-bunching schools in column (5) indicates that high-bunching schools experienced a 19.4 percent decline in their probability of locating at an enrollment multiple of 20 (or one below) relative to low-bunching schools after the hard cap was lifted. Looking at the other rows, we see that these high-bunching schools also experienced a 3.6 student increase in enrollment relative to low-bunching schools once the hard class size cap was lifted, a decline in the proportion of their student body that is white with a corresponding increase in proportion hispanic, and a relative decline in test scores. All these changes align with what we would expect if bunching schools, which are predominantly high-performing and advantaged, raised their enrollment after the hard class size cap was eliminated, allowing for more students residing elsewhere in the district to attend.

4 Descriptives

Given the strong incentives, we first document that schools responded in a manner consistent with the financial incentives in place. We highlight three measures of adjustment by schools: class size reduction, enrollment manipulation, and the creation of combination classes.

[13]Under random chance, we expect the school to fall on an enrollment multiple of 20 (or one below that multiple) about ten percent of the time. The probability that a school will exceed our twenty percent threshold by chance in our sample is 2.6%.

4.1 Class Size Reduction

Over time: Figure 2 shows student-to-teacher ratios for elementary and middle schools in the 1990-91 through 2014-15 school years.[14] As we can see, there is a sharp drop in the student-to-teacher ratio in elementary schools following the CSR reform in 1996-97. Additionally, there is a sharp increase in student-to-teacher ratio in elementary schools following the removal of the CSR 'hard cap' in the 2009-10 school year. No such sharp changes are observed for middle schools.

Class Size Rule: Given the presence of hard caps during CSR years, we may *a priori* expect to see the sawtooth pattern in class size as a function of enrollment as in Angrist and Lavy (1999) or Hoxby (2000). Figure 3(a) plots average class size and the class size that would be predicted if the school followed CSR class size caps in a passive way. The pattern in the data does not, for the most part, resemble the predicted class size function. While schools do seem to follow the class size rule leading up to (and near) enrollment multiples of 20, once the multiple is exceeded (with, say, an enrollment of 21 students), we do not observe the steep drop in class size that would be predicted if schools followed the CSR class size rule passively. Figure 3(b) shows why: there are steep increases in the proportion of schools with average class sizes above 20 students once the enrollment multiple of 20 is exceeded. Therefore, in line with financial incentives, some schools forego CSR funding in order to avoid hiring an additional teacher.

Figures 4(a) and 4(b) plot mathematics and English test scores on the y-axis rather than average class size. A curious pattern emerges for test scores, with large jumps in test scores at enrollment multiples of twenty. Indeed, when we focus on these enrollment multiples of twenty in Figure 5, we see extreme jumps in test scores at enrollment multiples of twenty during the years the CSR 'hard cap' was in place. These jumps are not observed once the CSR 'hard cap' is removed in the 2009-10 school year.

[14]The student-to-teacher ratio rather than class size is used as this variable is available over the entire time horizon. In our setting, student-to-teacher ratios are only available at the school level.

4.2 Combination Classes

Over time: Data constraints do not allow a graph to be constructed over time.

Strategic Use of Combination Classes: Figure 6 plots the fraction of students that are in combination classes by the distance to the enrollment multiple of 20 for CSR and non-CSR grades. We observe a steep decline in the proportion of students in CSR grades at enrollment multiples of 20 that is not observed for non-CSR grades. This type of behaviour aligns with the financial incentives: there is no financial incentive for schools to create combination classes at enrollment multiples of 20 since all the students can be accommodated into single-grade classes. This suggests that schools are strategically using combination classes to reduce their class sizes below the CSR limit as first noted by Sims (2008).

4.3 Enrollment Manipulation

In the first set of descriptive results, we show that schools in California manipulate their enrollment with respect to the class size cap. Figure 7(a) shows the distribution of school-grade level enrollment, normalized so that a value of zero is a multiple of twenty.[15] A visual analysis is clear: there is a substantial excess distribution of school-grades with enrollment at multiples of twenty. Schools in California are therefore able to manipulate their enrollment, consistent the financial incentives for school to limit their enrollment at a multiple of twenty, which was described in Section 2.2. Figure 7(b) shows that this bunching does not occur in non-CSR grades and Figure 8 shows a similar result for CSR grades in the pre- and post-period.[16]

Next, Figure 9 shows that this sorting leads to clear discontinuities in the observable characteristics of schools. Schools to the left of the discontinuity have a higher proportion

[15]Any multiple of twenty is itself a class size cap. Figure 17 shows the distribution of school-grade enrollment for grades K-3 during CSR 'hard cap' years.

[16]There remains some minor bunching in the post period in CSR grades, which is driven by the fact while schools eliminated their bunching once the CSR 'hard cap' was eliminated quite quickly, they did not do so immediately.

of students who are white or Asian, and fewer students who are Hispanic, eligible for free or reduced price meals, or are an English learner. In light of these discontinuities in observable characteristics, it is unlikely that the identifying assumption of the RD design that outcomes would be continuous through the discontinuity in the absence of the policy is satisfied.

With Respect to Combination Classes: In Sims (2008), the key incentive leading schools to create combination classes arose when adjacent grades had combined enrollments were multiples of 60 (so that the combined 'left over' students in the two grades was 20). Figure 10 shows combined adjacent grade enrollment for enrollment multiples of 60, omitting the adjacent combination of 120 and 240 students as there may be bunching at those numbers caused solely through the aforementioned bunching at enrollment multiples of 20.[17] It is clear in the figure that bunching occurs at these combined adjacent grade enrollment multiples of 60, though not nearly to the same extent as bunching at enrollment multiples of 20.[18]

How Schools Manipulate Enrollment: Given that public schools in California must, by law, accept all students within their school attendance zone, it is surprising that schools are able to manipulate enrollment to such an extent. The ability for this to occur is driven by California's intradistrict policy, which allows some parents to enroll their children at a school other than their designated attendance zone school. Specifically, California Education Code Section 35160.5(b) permits parents to request transfers to other schools within their school district, and those schools must take them if there is space. If there are too many students applying to a specific school, a lottery is held. Since schools only must take these students if there is space, schools can *choose* their

[17]For instance, an adjacent combination of 120 can be caused by both school-grades have enrollments of 60, which occurs when both school-grades are on an enrollment multiple of 20. We rule out possible bunching at adjacent multiples of 60 stemming from say enrollments of 20 and 40 in adjacent grades by imposing a restriction that adjacent school-grade enrollments must be within ten students of one another.

[18]Figure 10 graphs combined enrollment for grade g and $g-1$. A very similar figure arises if one uses combined enrollment for grade g and $g+1$. The slight bunching observed for grades 4-6 occurs because there is some manipulation of grade 4 enrollment to ensure combined grade 3 and 4 enrollment are at a multiple of 60.

capacity, conditional on the constraint that every student *within* their school attendance zone who applies is admitted.

To provide evidence supporting this, Figure 11 shows bunching around enrollment multiples of twenty for CSR grades for both *one-school* districts and multi-school districts. Since one-school districts, by definition, have no intra-district transfer opportunities, schools should be much more restricted in their ability to choose enrollment. This is what we find: while the bunching the one school districts is statistically significant, the bunching in the many-school districts is sixfold that of the bunch in the one school districts.[19]

Types of School Able to Manipulate Enrollment: Given that schools must be over-subscribed in order to choose their enrollment, it seems likely that schools with high demand from parents are more likely to be oversubscribed and thus be able to manipulate their enrollment. Given parents value school composition, we divide school into poverty deciles based on their *within* district poverty rank.[20] We proxy a school's poverty measure by using its proportion of students on Free or Reduced Price Meals. Figure 12 plots the bunching for the highest and lowest decile poverty schools based on their within-district rank. While both low- and high-poverty schools experience bunching, bunching in low-poverty schools (relative to their district) is over fifty percent higher than the bunching in high-poverty schools (relative to their district).

[19]One-school districts are likely still able to manipulate their enrollment somewhat due to the inter-district transfer law (California Education Code sections 46600 – 46610). This law permits parents to request an inter-district transfer, though this transfer must be approved by both the student's original district and the district to which the student seeks to transfer to. Therefore, it is more difficult for inter-district transfer to be approved relative to intra-district transfer, generating the differences in the ability for schools to manipulate enrollment.

[20]Qualitatively similar results can be obtained by ranking schools based on their within-district test score percentile.

5 The Consequences of Enrollment Manipulation

In this section, we use the ending of the CSR 'hard cap' policy to show that the manipulation of enrollment by schools came at the expense of minority students. The ending of the policy led to a decline in the proportion of white students in schools that were more likely to manipulate enrollment, with an associated rise in the proportion of Hispanic students. Since schools that manipulated enrollment were likely to have higher proportions of white students and high-socioeconomic status (SES) students, we show that the end of the CSR 'hard cap' policy, in turn, led to a decline in both racial and income segregation within districts.

5.1 School-Level Responses

Since high-demand schools manipulated their enrollment by restricting the number of seats available to students outside the attendance zone, we would expect that the end of the CSR 'hard cap' should induce these school to open up new seats. Figure 13 shows enrollment trends for high- and low-bunching schools,[21] and we observe an increase in enrollment for the high-bunching schools in the 2009-10 school year not observed for the low-bunching schools.

Comparing high- to low-bunching schools directly may be problematic, however, as a school's ability to bunch is directly related to the parental demand for that school. We therefore use the fact that the financial incentives make it so that small schools are far more likely to bunch than larger schools.[22] To avoid any enrollment changes that are

[21] High-bunching schools are defined as schools that hit an enrollment multiple of 20 at least twice over the 10 year period 1999-00 through 2008-09. (Without bunching, the probability a school would hit an enrollment multiple of 20 at least twice is about one percent.)

[22] This arises due to the nature of the financial incentives that CSR gives a per-student reward to schools for *every* student in a CSR-compliant class. Therefore, a school with enrollment – say – of 21 students would lose their entire CSR funding ($17,472 in 1998-99) if they do not add another class. In contrast, while a school-grade with 121 students that does not add another class still loses the CSR funding for that class ($17,472 in 1998-99), that only represents twenty-one percent of the school's total CSR funding, since they would still get CSR funding for the 100 students in CSR-compliant classes ($83,200 in 1998-99).

endogenous to CSR, we use enrollment well before the introduction of CSR (specifically, 1990-91 school year enrollment) to indicate the size of the school.

Figure 14 shows the probability that a K-3 school-grade bunches in a given year based on 1990-91 school-grade enrollment. A clear pattern emerges: school-grades with smaller 1990-91 enrollment bunch at much higher rates than school-grades with higher 1990-91 enrollment. In fact, large school-grades with 1990-91 enrollment of nearly 200 students show almost no enrollment manipulation since the probability they are at an enrollment multiple of 20 is nearly five percent; the probability a school would reach that multiple without any bunching.

More formally, we can investigate these jumps in enrollment combined with possible changes to student demographics (which will occur if students outside the attendance zone differ from those within) by comparing small and large schools based on pre-CSR enrollment levels after versus before the elimination of the CSR hard cap. In addition, since we expect these changes in student demographics to occur only for schools serving CSR grades, we run a difference-in-differences design by grade, which will form the basis of an additional layer of differencing later. Formally, we run the following specification for every grade K-12:

$$y_{sgt} = \beta_0 + \beta_1 post_t + \beta_2 enroll_{sg,1990} + \beta_3 (post_t * enroll_{sg,1990}) + \eta_s + \delta_t + \phi X_{sgt} + \epsilon_{sgt}\,, \tag{5.1}$$

where y_{sgt} is the outcome of interest (enrollment or demographics) at school s in grade g at time t, $post_{gt}$ is a post-CSR hard cap dummy variable, $enroll_{sg,1990}$ is a school's average K-3 1990-91 enrollment level (measured in hundreds), X_{sgt} is a vector of school-grade-year controls, and η_s and δ_t are school and year fixed effects, respectively. The coefficient of interest is given by β_3.

Given that we can compare high- to low-enrollment (and thus bunching) schools by grade, we can use grades (CSR versus non-CSR) as an additional layer of differencing. A

concern arises, however, that a school serving grades K-5 will see demographic changes to every grade over time as students that entered in CSR grades (K-3) progress through the school.[23] Appendix Table 4 reports the distribution of Californian schools by grade span. It is apparent from this table that the majority of schools that serve a K-3 grade are either K-5 or K-6 schools, so we should expect the end of the CSR 'hard cap' to almost entirely effect grades K-5. Therefore, we can add an additional layer of differencing by comparing grades K-5 to grades 7-12.[24] To do so, we run the following triple-differences specification:[25]

$$y_{sgt} = \beta_0 + \beta_1 post_t + \beta_2 enroll_{sg,1990} + \beta_3 (post_t * enroll_{sg,1990}) + \beta_4 CSR_g + \beta_5 (post_t * CSR_g) + \beta_5 (enroll_{sg,1990} * CSR_g) + \beta_7 (post_t * enroll_{sg,1990} * CSR_g) + \eta_s + \lambda_{gt} + \phi X_{sgt} + \epsilon_{sgt} \,, \tag{5.2}$$

where CSR_g is indicator for grades K-5 and λ_{gt} are grade-by-year fixed effects. All other variables are defined in Equation 5.1. The coefficient of interest is given by β_7.

School-Level Results

Figure 15 displays the difference-in-differences estimates from Equation 5.1 by grade. For schools with grade spans K-5, the majority of which serve the CSR grades, we see that larger schools in 1990 (who bunch at low rates) see an increase in the proportion of white students in their schools, with an associated decline in percent Hispanic, English learners, and students on free or reduced price lunch relative to smaller schools in 1990 (who bunch at high rates). Around half of these differences are observed in grade 6, given

[23] This would be a good opportunity to cite GMM – students stay in schools until they are forced to leave.

[24] Since schools are roughly divided into K-5 and K-6 configurations, grade 6 is omitted from the triple-differences framework since about half of grade 6 students are in a school serving CSR grades, while the other half are not.

[25] We used grade level differencing to map our triple-differences results directly to the visual difference-in-differences evidence in Figure 15. An alternative triple-difference specification would compare schools serving CSR grades versus not as the third layer of differencing. The results we present are very similar to this alternative specification.

that approximately half of grade 6 schools serve CSR grades, while for grades 7-12, whose schools do not serve CSR grades, these differences are not observed.

Table 2 combines these difference-in-differences estimates from Equation 5.1 into grades where a majority of schools serve CSR grades (grades K-5) in Column (1) and into grades where a majority of schools do not serve CSR grades (grades 7-12). Column (3) then takes the difference between these two columns through the triple-differences specification given by Equation 5.2. Focusing on column (3), we first see that the end of the CSR 'hard cap' led a large decline in school bunching at enrollment multiples of 20 for small school-grades relative to large school-grades (based on 1990-91 enrollment). This decline in bunching, in turn, led to a steep rise in enrollment of about 8 students (off a mean of 84) for these small school-grades relative to larger school-grades. This increase in enrollment comes from schools admitting more Hispanic students from outside their school attendance zone, causing a 1.6 percent decrease in the proportion white in their student body, with an associated 0.75 percent increase in the proportion Hispanic and a 3.43 percent increase in the proportion of students who are English learners. Given these numbers, of the 8 extra students enrolled by the public school, approximately 1.5 of them were white and about 5 were Hispanic (with the remaining 1.5 students being of a different race), and 5.5 students were English learners.

5.2 District-Level Analysis

The school-level results suggest that CSR caused schools to manipulate their enrollment, which came at the expense of denying minority students entry at these (generally) high-SES and predominately white schools. We therefore expect that the end of the CSR 'hard cap' would allow these bunching schools to open more enrollment slots for low-SES and minority students, which in turn should lower school segregation.

To do so, we construct dissimilarity indices at the district-grade level and run a version of Equation 5.1. We use pre- versus post- CSR 'hard cap' and high- versus low-bunching

school district as our layer of differencing.[26] Specifically, for every grade K-12, we regress:

$$D_{dgt} = \beta_0 + \beta_1 post_t + \beta_2 bunch_{dg} + \beta_3 (post_t * bunch_{dg}) + \psi_d + \delta_t + \phi X_{dgt} + \epsilon_{sgt}\,, \quad (5.3)$$

where D_{dgt} is a dissimilarity index (race or socioeconomic) at district d in grade g at time t, $bunch_{dg}$ is the degree of K-3 bunching at enrollment multiples of 20 during CSR years, $post_t$ is a post-CSR dummy, X_{dgt} is a vector of district-grade-year controls, and ψ_d and δ_t are district and year fixed effects, respectively. The coefficient of interest is given by β_3.

District Results

Figure 16 displays the difference-in-differences estimates from Equation 5.3 by grade. For grades K-5, whose majority of schools serve the CSR grades, we see that district with more schools bunching at enrollment multiples of 20 see a decline in their dissimilarity indices relative to school districts who have less bunching. For grades 7-12, whose schools do not serve CSR grades, these differences are not observed for the racial dissimilarity index, while for the sociodemographic index we see a large increase in the index for grades 7-12.

Table 3 combines these difference-in-differences estimates from Equation 5.3 into grades where a majority of schools serve CSR grades (grades K-5) in Column (1) and into grades where a majority of schools do not serve CSR grades (grades 7-12). Column (3) then takes the difference between these two columns through the triple-differences specification (similar to Equation 5.2, but rather at the district level). Focusing on column (3), we see a decline in the dissimilarity index for race and socioeconomic status of 0.079 and 0.136, respectively. These represent a 0.38 and 0.71 standard deviation decline in the dissimilarity index for race and socioeconomic status, respectively.

[26] Since dissimilarity indices are by definition zero for one-school districts, these districts are eliminated from our analysis.

6 Conclusion

In this paper, we have documented clear evidence of enrollment manipulation in public schools. We link this manipulation to the strong financial incentives associated with the introduction of California's vast class size reduction program of the late-1990s, coupled with schools having discretion to set 'capacity' under the state's intra-district open enrollment system. Consistent with the precise incentives under the CSR program, we showed that school enrollments are bunched at multiples of twenty, especially in schools likely to be over-subscribed. Of note, we also showed how the removal of the 'hard cap' in the 2008-09 academic year led to a reduction in segregation within-district, based on a triple differences design.

From a policy perspective, our analysis underlines an undesirable consequence of open enrollment programs that do not allow school capacity to adjust in a non-marginal way: local discretion can lead schools to manipulate capacity in a way that serves to exclude under-privileged students elsewhere in the district. The study also has implications for open enrollment policies where schools are forced to randomize students when seats are over-subscribed. Although they may appear superior from a fairness perspective, schools may actually restrict capacity under such a system, giving rise to the sorts of adverse segregation effects we have documented.

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, 533–575.

Bohrnstedt, George W. and Brian M. Stecher (2002), "What we have learned about class size reduction in California. Capstone report." URL `http://www.classize.org/techreport/index-02.htm`.

Funkhouser, Edward (2009), "The effect of kindergarten classroom size reduction on second grade student achievement: Evidence from California." *Economics of Education Review*, 28, 403–414.

Hoxby, Caroline M. (1996), "How teachers' unions affect education production." *Quarterly Journal of Economics*, 111, 671–718.

Hoxby, Caroline M. (2000), "The effects of class size on student achievement: New evidence from population variation." *Quarterly Journal of Economics*, 115, 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, 223–250.

Schrag, Peter (2006), "Policy from the Hip: Class size reduction in California." *Brookings Papers on Education Policy*, 2006, 229–243.

Sims, David (2008), "A strategic response to class size reduction: Combination classes and student achievement in California." *Journal of Policy Analysis and Management*, 27, 457–478.

Stecher, Brian M., Daniel McCaffrey, and Delia Bugliari (2003), "The relationship between

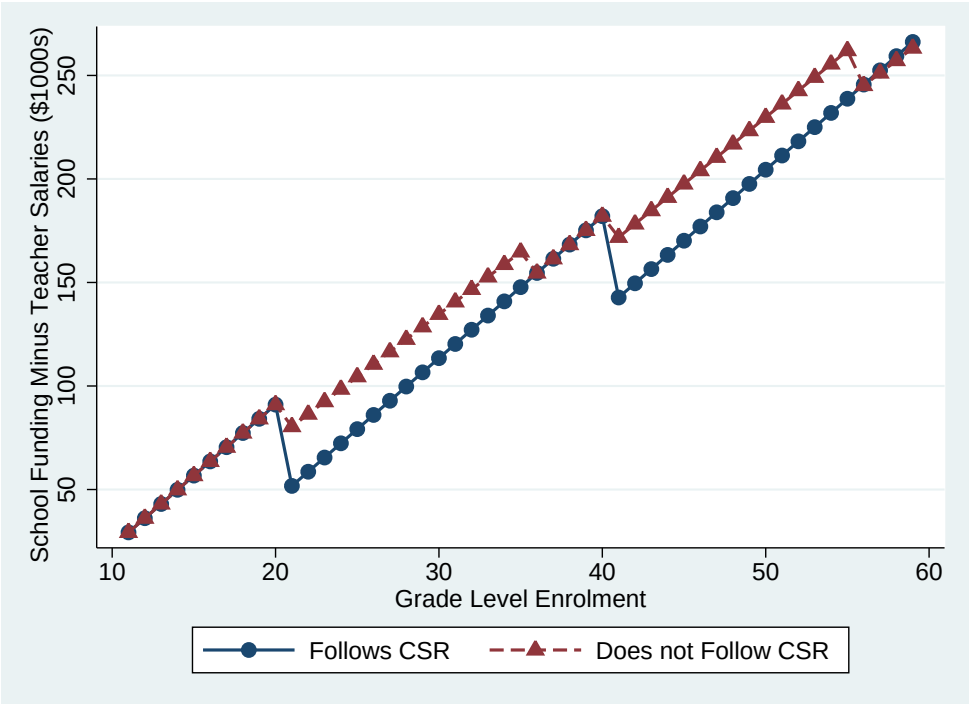
exposure to class size reduction and student achievement in California." *Education Policy Analysis Archives*, 11, 1–27.

Unlu, Fatih (2005), "California class size reduction reform: New findings from the NAEP." URL `http://www.classsizematters.org/wp-content/uploads/2011/04/California\_CSR\_Fatih\_Unlu.pdf`.

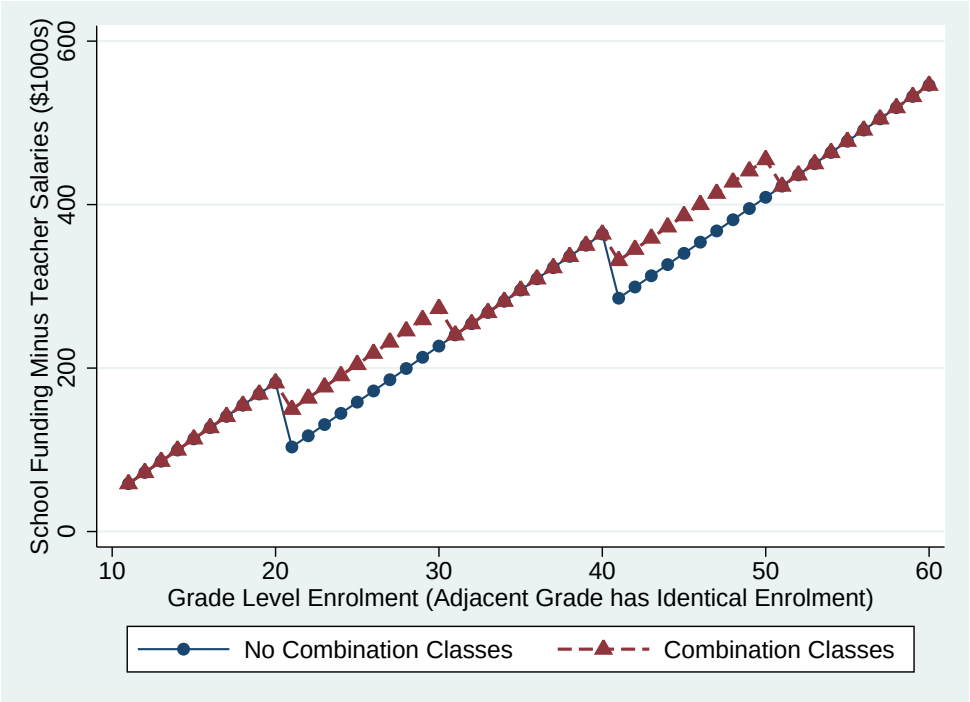
Urquiola, Miguel and Eric Verhoogen (2009), "Class-size caps, sorting, and the regression-discontinuity design." *American Economic Review*, 99, 179–215.

Figure 1: CSR Financial Incentives

(a) Incentives Under CSR Rules

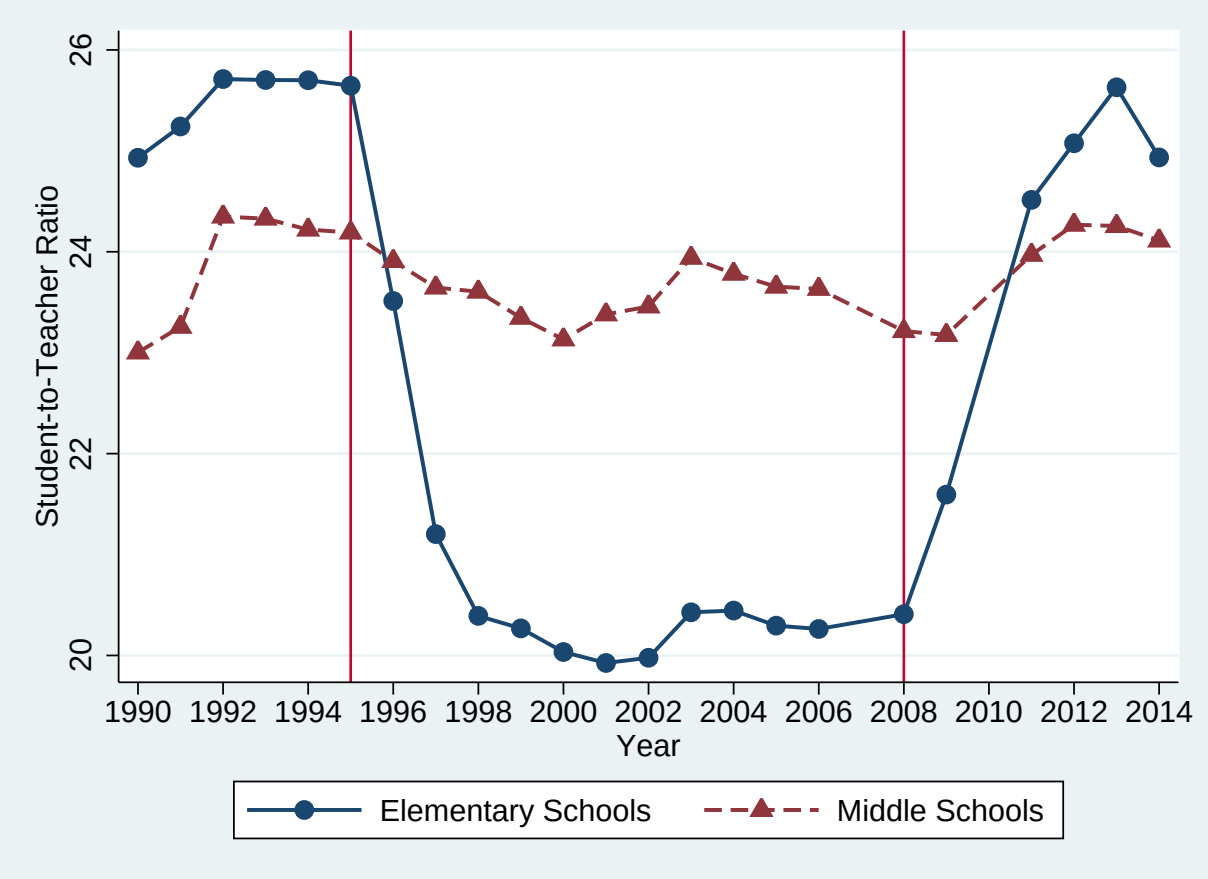


(b) Incentives to Form Combination Classes



Notes: Figure 1(a) depicts the financial incentives for a school that in the absence of CSR, adds another teacher whenever average class sizes exceed thirty-five students. Figure 1(b) depicts the financial incentives for a school with two grades with identical enrollment, with each grade's enrollment on the x-axis. The dollar amounts correspond to 1998-99 school year funding allotments. The 'no combination class' case considers a school that always keeps its class sizes below 20, without resorting to combination classes. The 'combination class' case allows the school to utilize a combination to maximize its total level of funding minus teacher salaries.

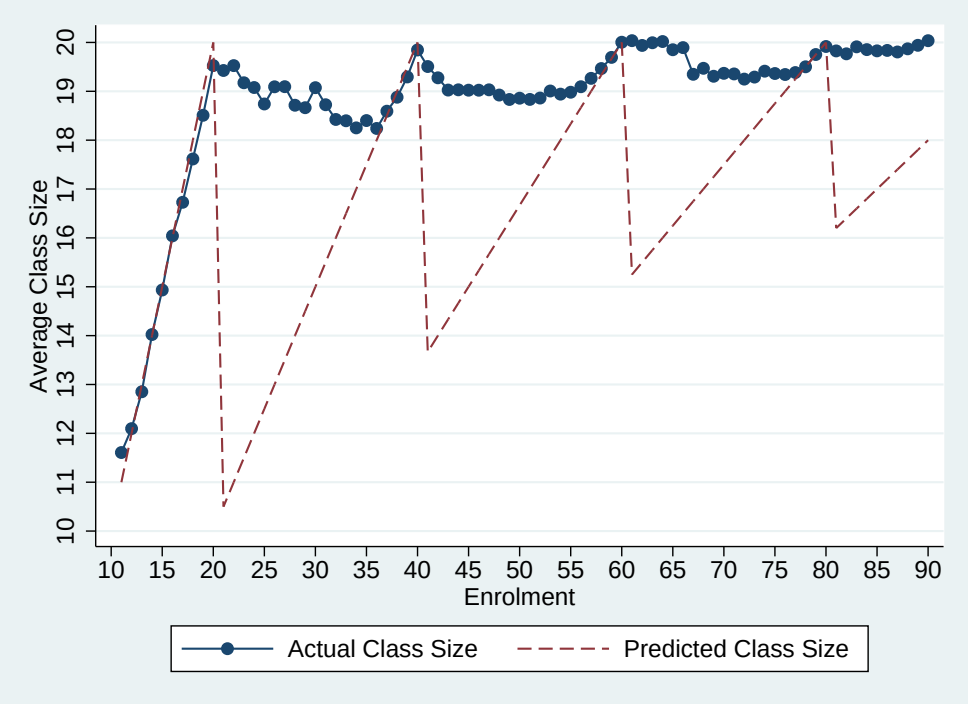
Figure 2: Teacher-to-Student Ratio by Year



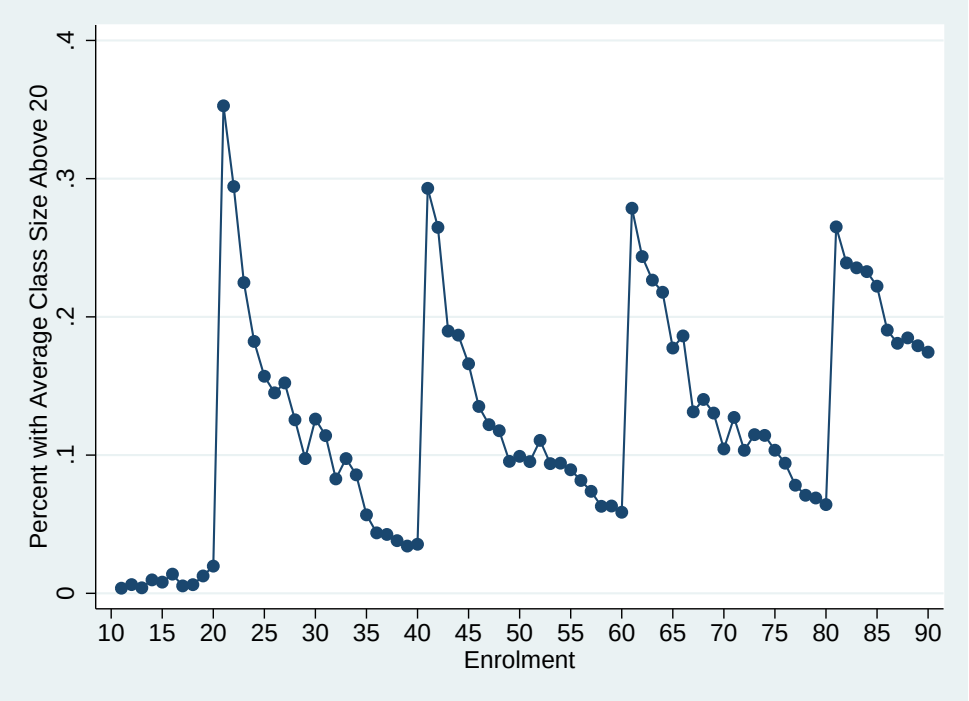
Notes: The vertical lines correspond to the last 'pre-CSR' year and the last CSR year, respectively. The 2007-08 school year is omitted as there is an anomoly in the data for that year.

Figure 3: Class Sizes by Enrollment in CSR grades

(a) Average Class Size by enrollment in CSR grades



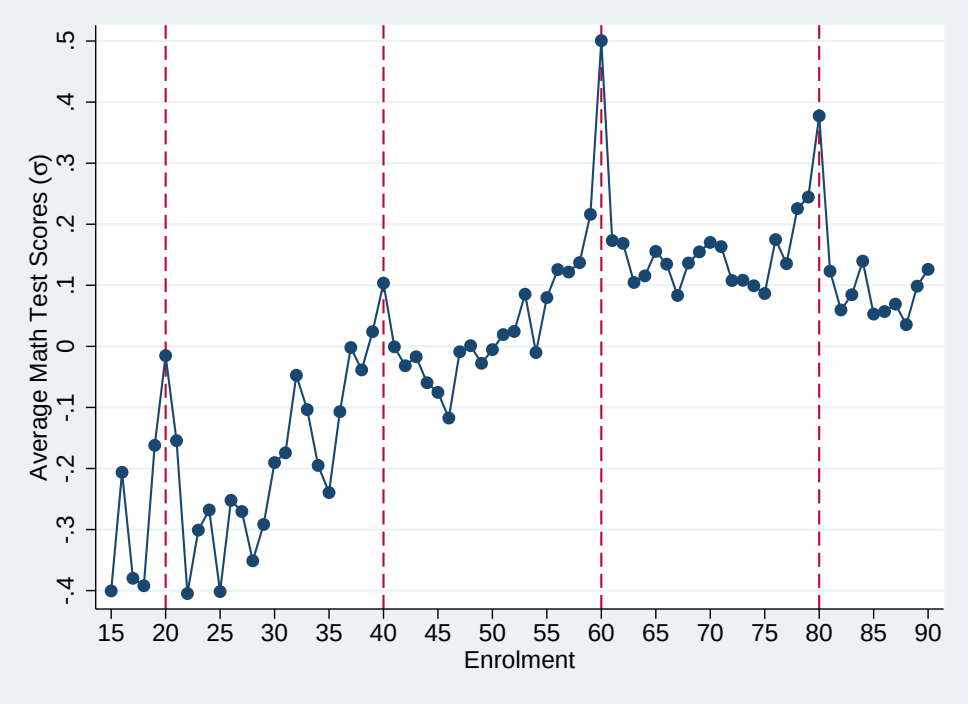
(b) Percent with Average Class Size over 20 by enrollment in CSR grades



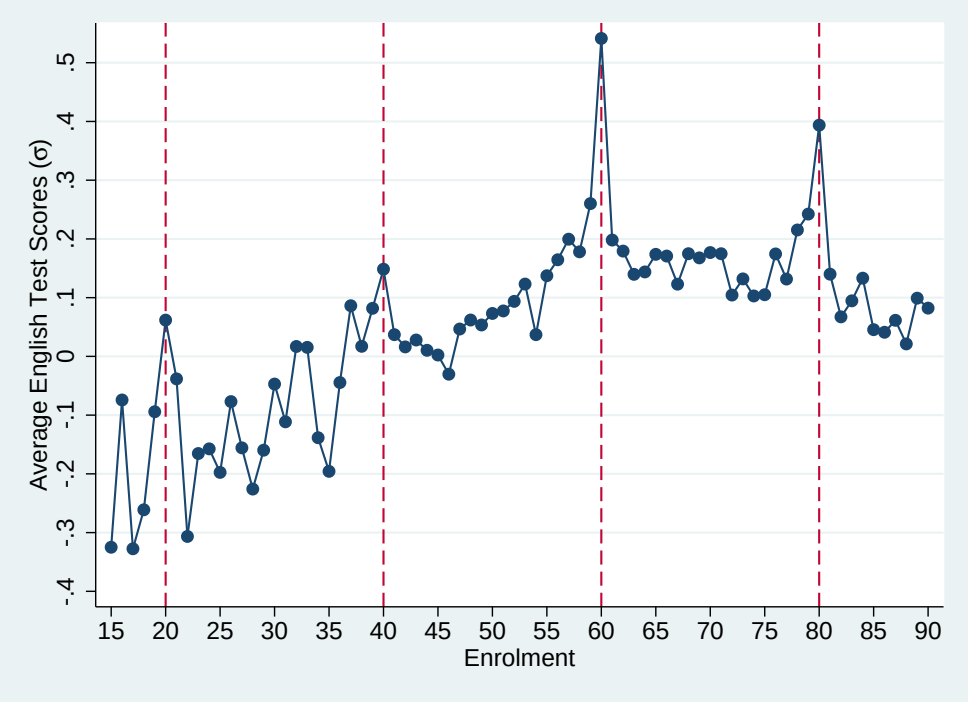
Notes: Figures 3(a) and 3(b) cover grades K-3 over the 1999-00 through 2008-09 school years and consist of 119,143 school-grade-year observations.

Figure 4: Test Scores by Enrollment in CSR grades

(a) Math Test Scores



(b) English Test Scores



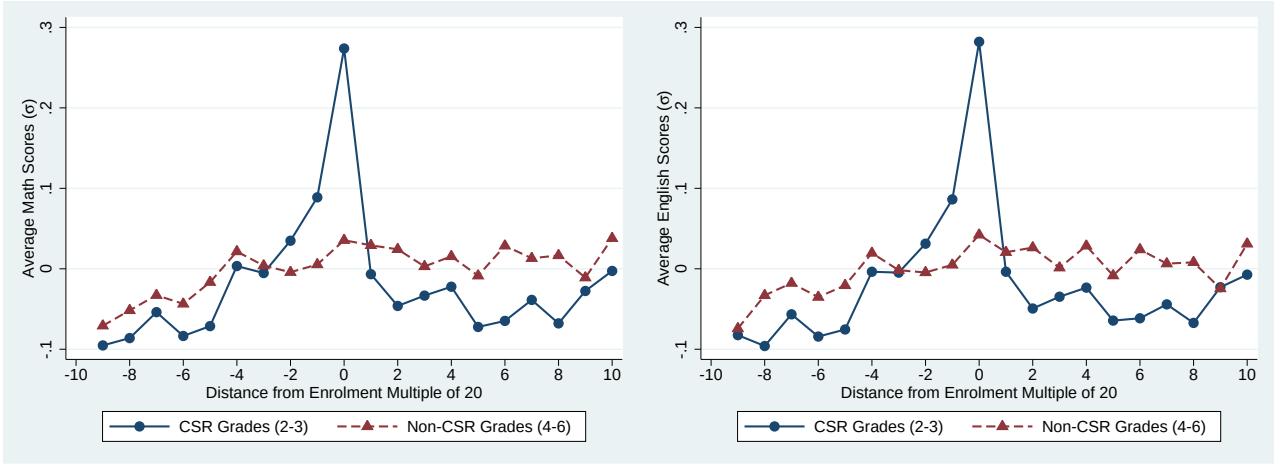
Notes: Figures 4(a) and 4(b) cover grades 2 and 3 over the 1999-00 through 2008-09 school years and consist of 56,091 and 56,076 school-grade-year observations, respectively. The vertical dashed lines represent enrollment multiples of 20.

Figure 5: Test Scores by Enrollment Multiples

(CSR 'Hard Cap' Years)
(School-years 1999-00 through 2008-09)

(a) Math

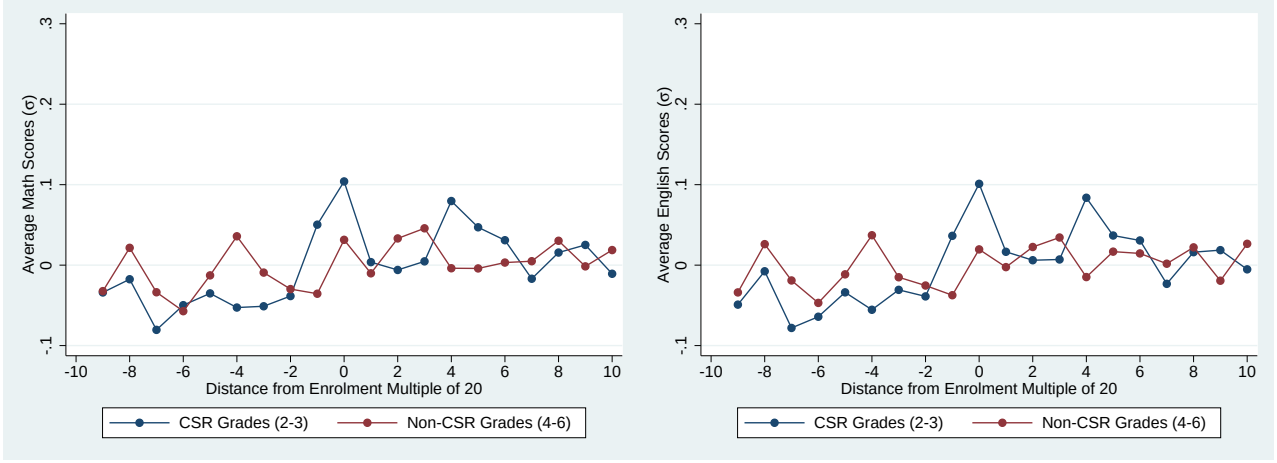
(b) English



Post-CSR 'Hard Cap' Years
(School-years 2009-10 through 2014-15)

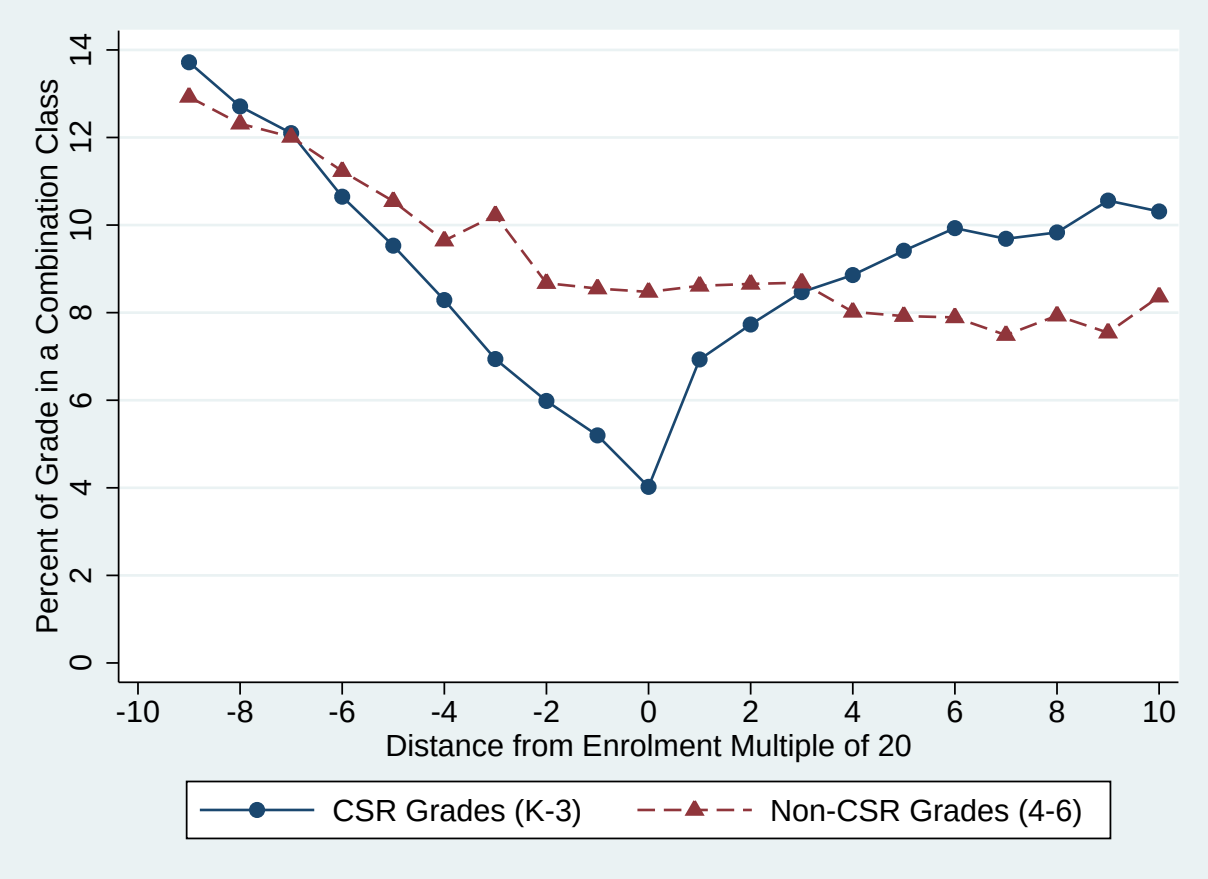
(c) Math

(d) English



Notes: CSR grades only refers to grades 2 and 3, as test scores are unavailable for grades K-1. CSR 'Hard Cap' years refer to years when a school would lose all of its CSR funding for a class if its class size went above 20 students. Post-CSR 'Hard Cap' Years refer to school years where the school would lose 5 percent of its CSR funding in a class for each student above 20 in that class.

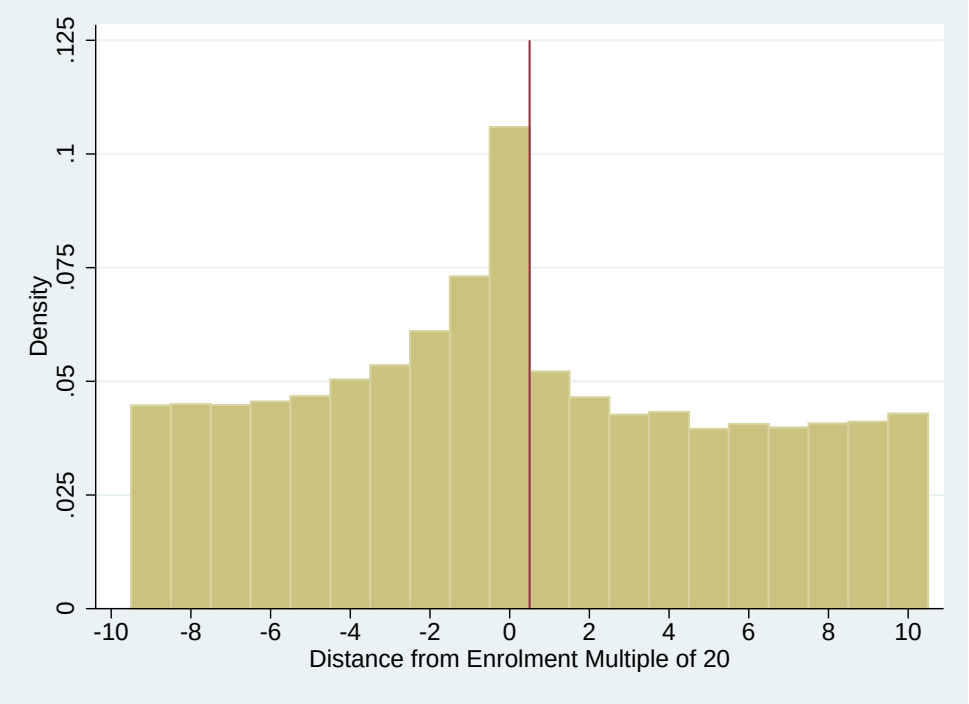
Figure 6: Combination Classes by enrollment Multiples



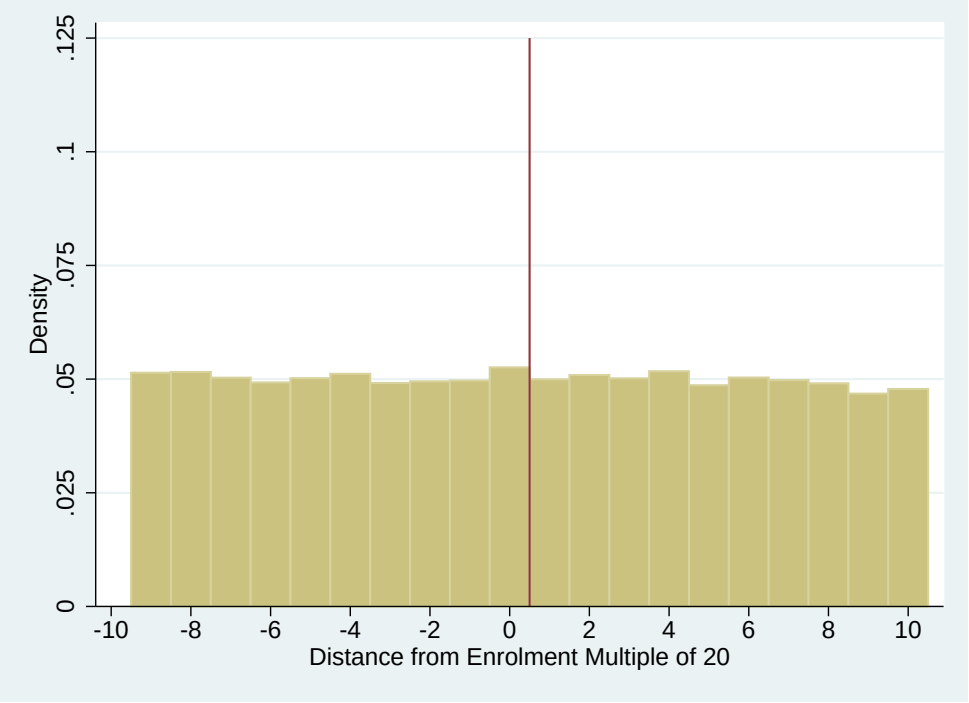
Notes: Figure 6 covers the 1999-00 through 2008-09 school years and consists of 338,984 school-grade-year observations.

Figure 7: Enrollment Bunching in Public Schools (CSR Years)

(a) CSR grades (K-3) – b=1.52*** (0.09)



(b) Non-CSR Grades (4-6) – b=-0.02 (0.04)



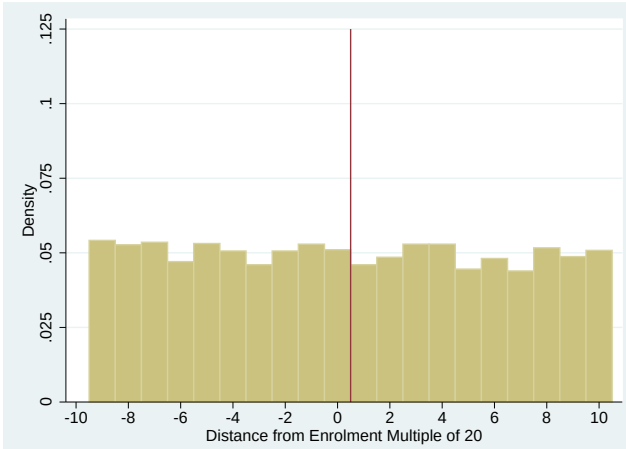
Notes: Observations are at the school-grade-year level. Figures 7(a) and 7(b) are based on 211,270 and 137,465 observations, respectively. CSR grades refer to grade K-3 and non-CSR grades refer to grades 4-6. CSR years cover the 1999-00 through 2008-09 school years.

Figure 8: Enrollment Bunching in Public Schools (CSR Years)

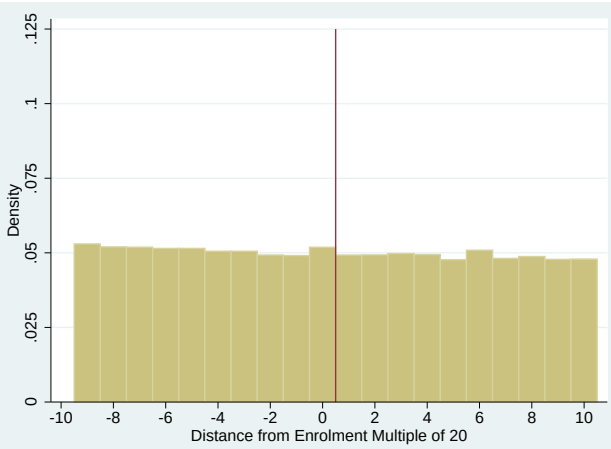
Pre-CSR Years

(School-years 1990-91 through 1995-96)

(a) CSR grades (K-3) – b=0.05 (0.05)



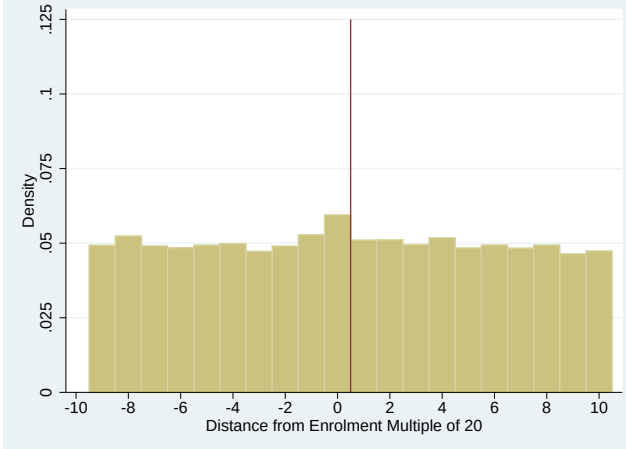
(b) Non-CSR grades (4-6) – b=-0.01 (0.04)



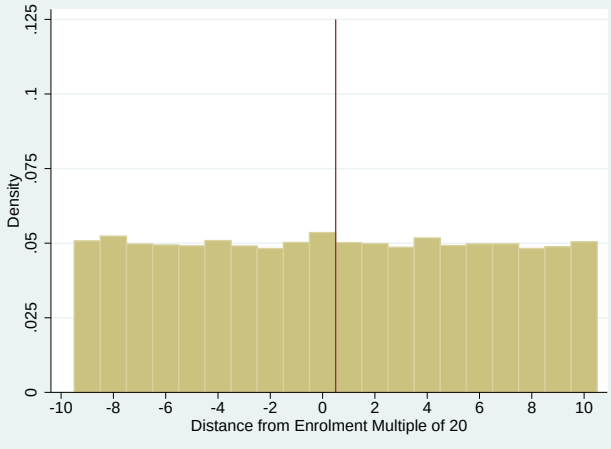
Post-CSR 'Hard Cap' Years

(School-years 2009-10 through 2014-15)

(c) CSR grades (K-3) – b=0.14** (0.06)



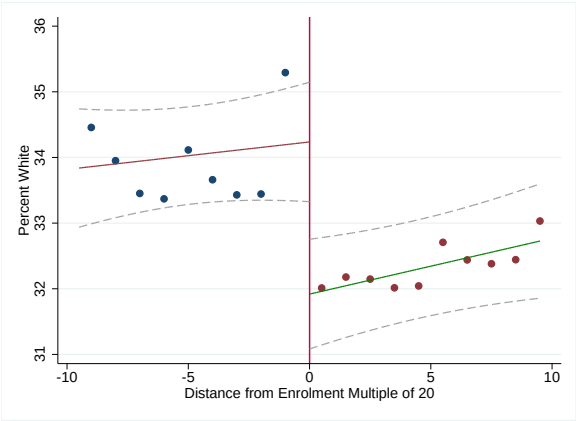
(d) Non-CSR Grades (4-6) – b=0.01 (0.05)



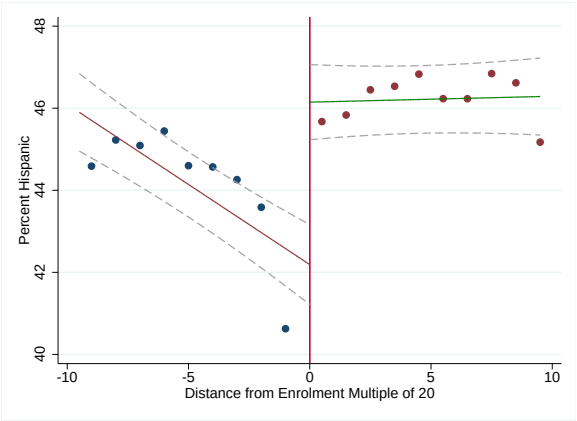
Notes: Observations are at the school-grade-year level. Figures 8a, 8b, 8c, 8d are based on 111,758, 76,085, 134,435, and 89,076 observations, respectively. CSR grades refer to grade K-3 and non-CSR grades refer to grades 4-6. CSR 'Hard Cap' years refer to years when a school would lose all of its CSR funding for a class if its class size went above 20 students. Post-CSR 'Hard Cap' Years refer to school years where the school would lose 5 percent of its CSR funding in a class for each student above 20 in that class.

Figure 9: Observable Covariates by Enrollment Multiple (CSR Grades/Years)

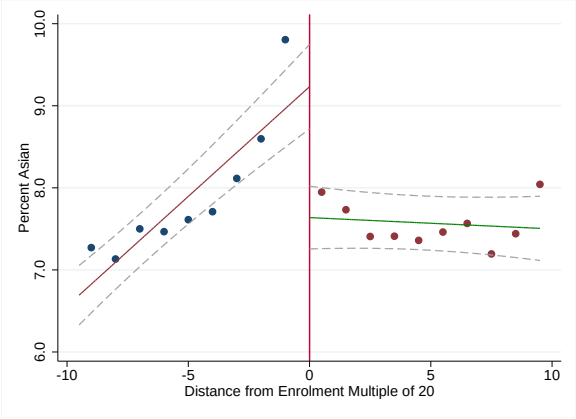
(a) Percent White
RD Estimate: -2.32*** (0.40)



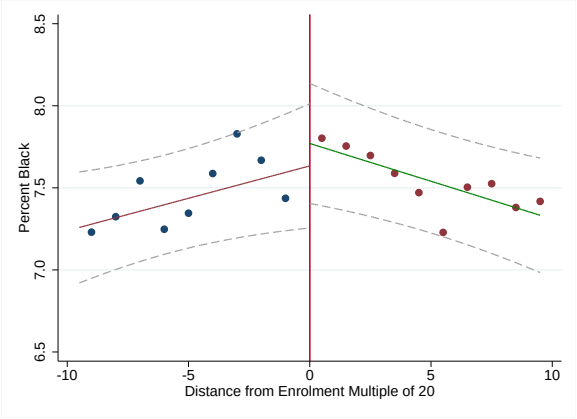
(b) Percent Hispanic
RD Estimate: 3.96*** (0.43)



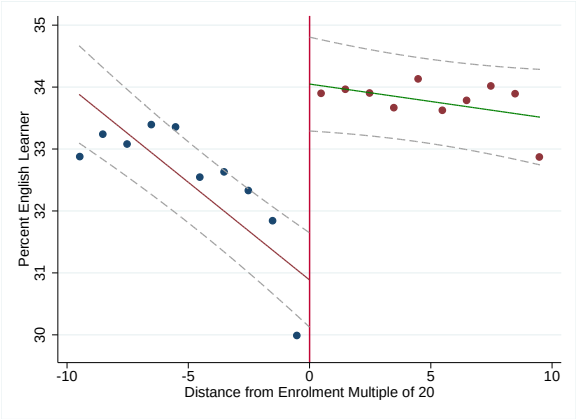
(c) Percent Asian
RD Estimate: -1.60*** (0.23)



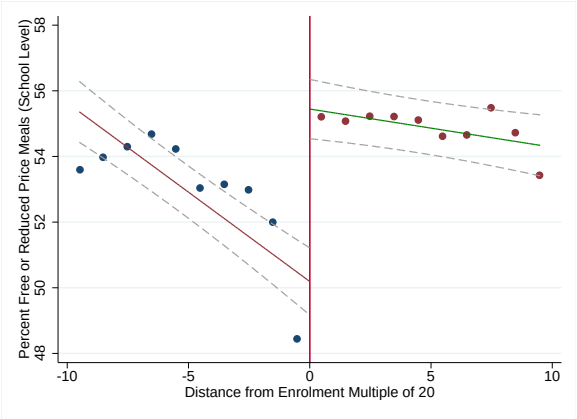
(d) Percent Black
RD Estimate: 0.14 (0.16)



(e) English Learners
RD Estimate: 3.16*** (0.34)



(f) Percent Free or Reduced Price Meal (School Level)
RD Estimate: 5.25*** (0.45)

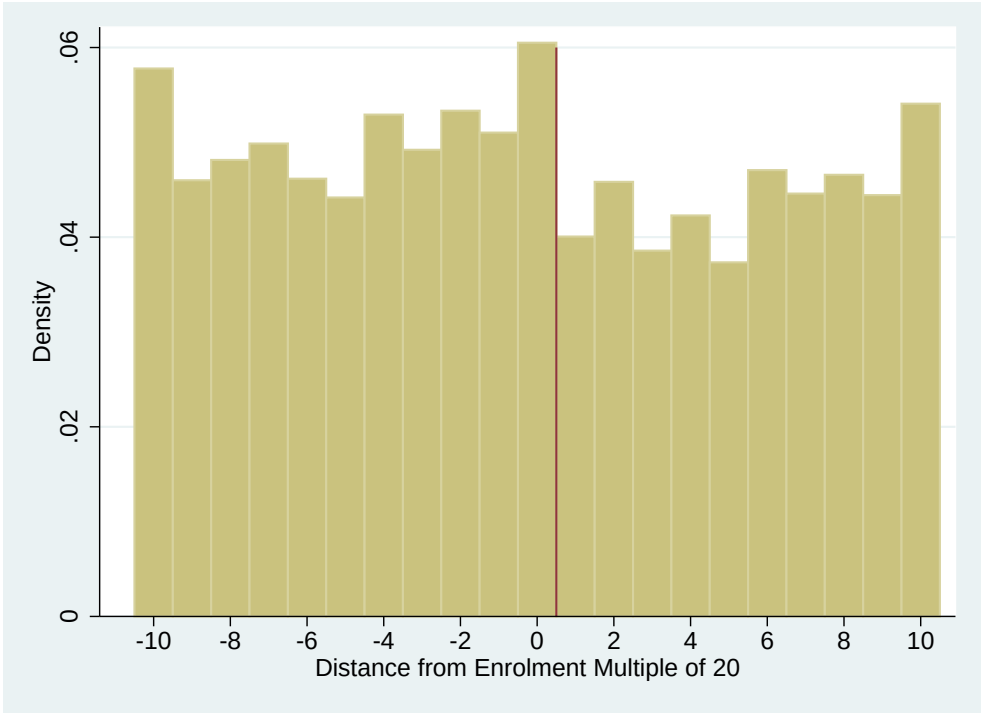


Notes: Figures are based off 211,270 school-grade-year level observations. Free and Reduced Price Meal data is only available at the school-year level. RD estimates use a bandwidth of ten and standard errors are clustered at the school level.

Figure 10: Enrollment Bunching for Combination Classes (CSR Years)

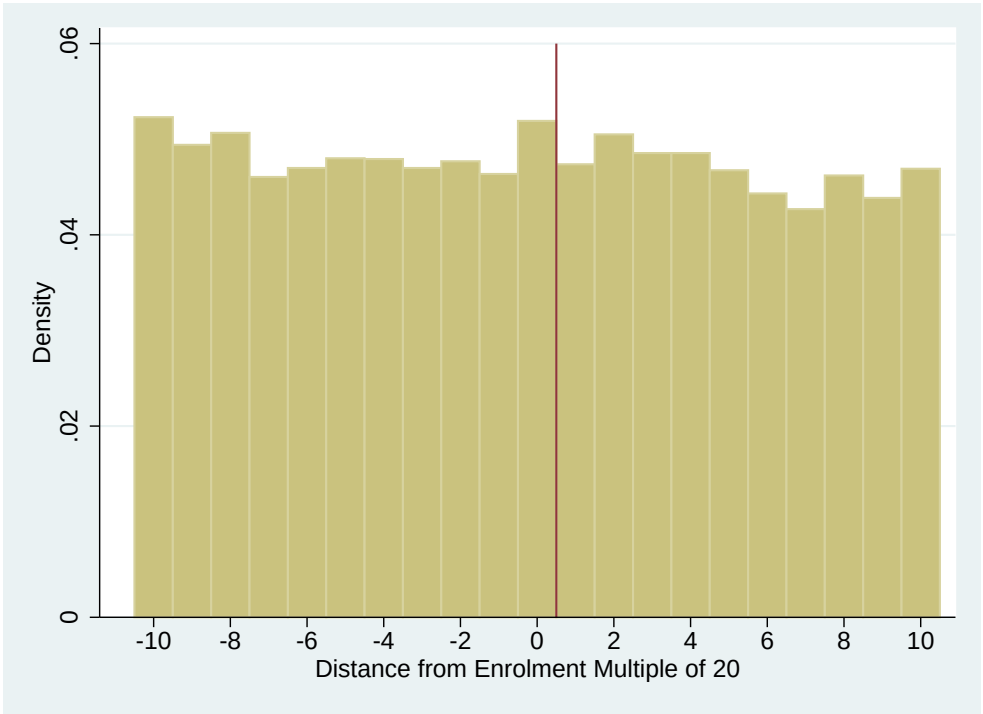
(a) CSR Grades (K-3)

b= 0.60** (0.27)



(b) Non-CSR Grades (4-6)

b= 0.26* (0.14)

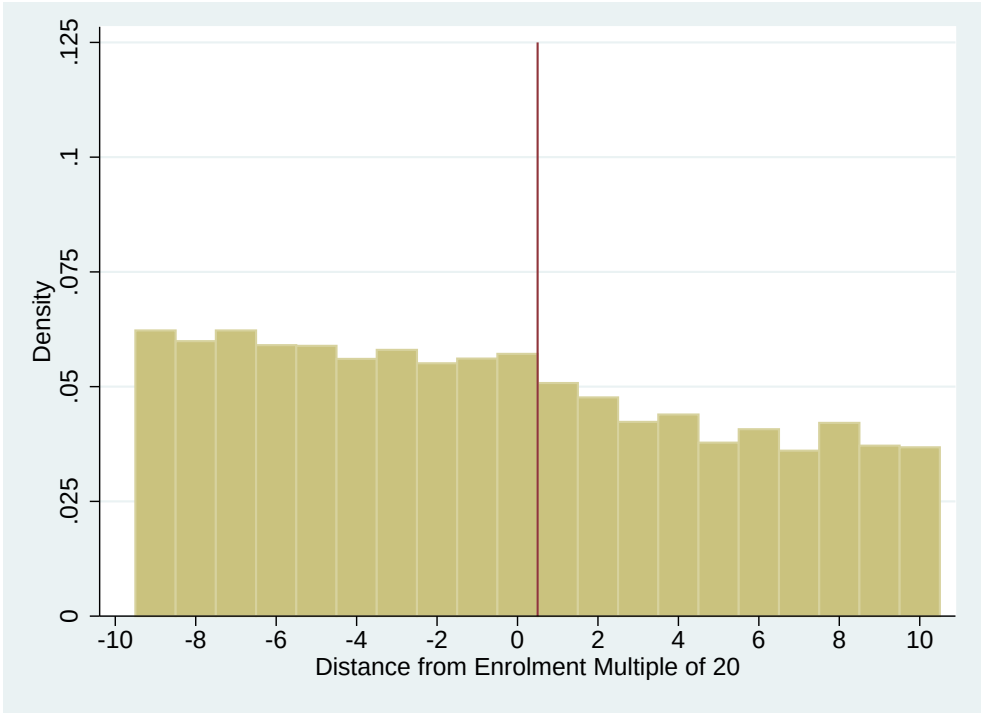


Notes: Figures 10(a) and 10(b) are based off 11,430 and 12,122 school-grade-year level observations, respectively. The enrollment multiple of 60 is based off enrollment in a given K-3 grade and the enrollment of its adjacent grade that is one grade below it.

Figure 11: Enrollment Bunching in One- versus Many-School Districts

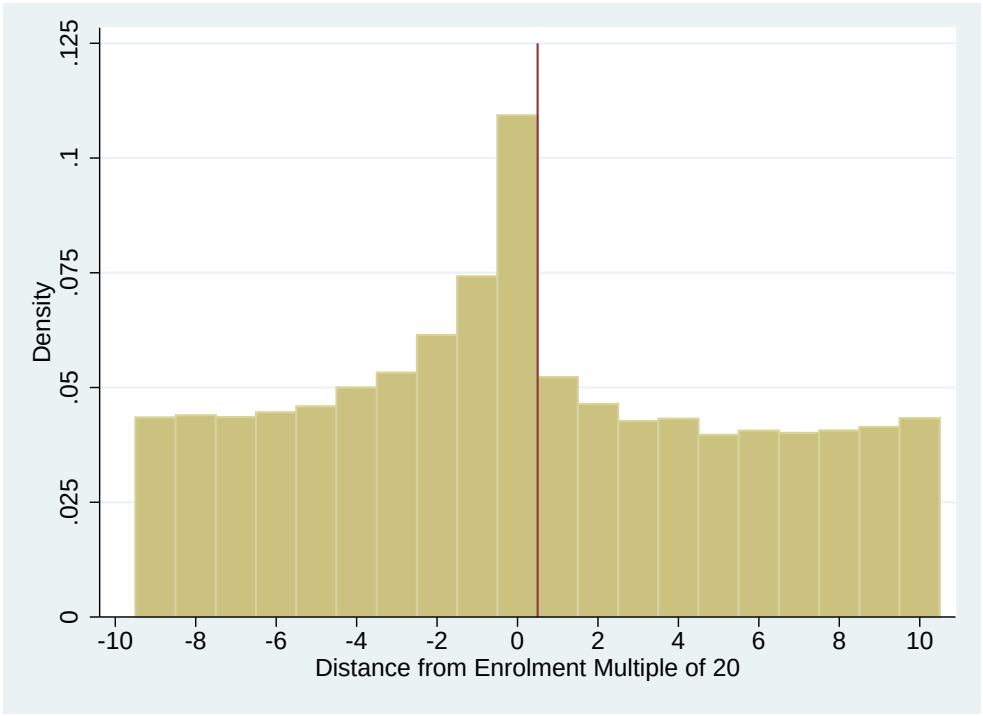
(a) One-School Districts

b= 0.26*** (0.10)



(b) Many-School Districts

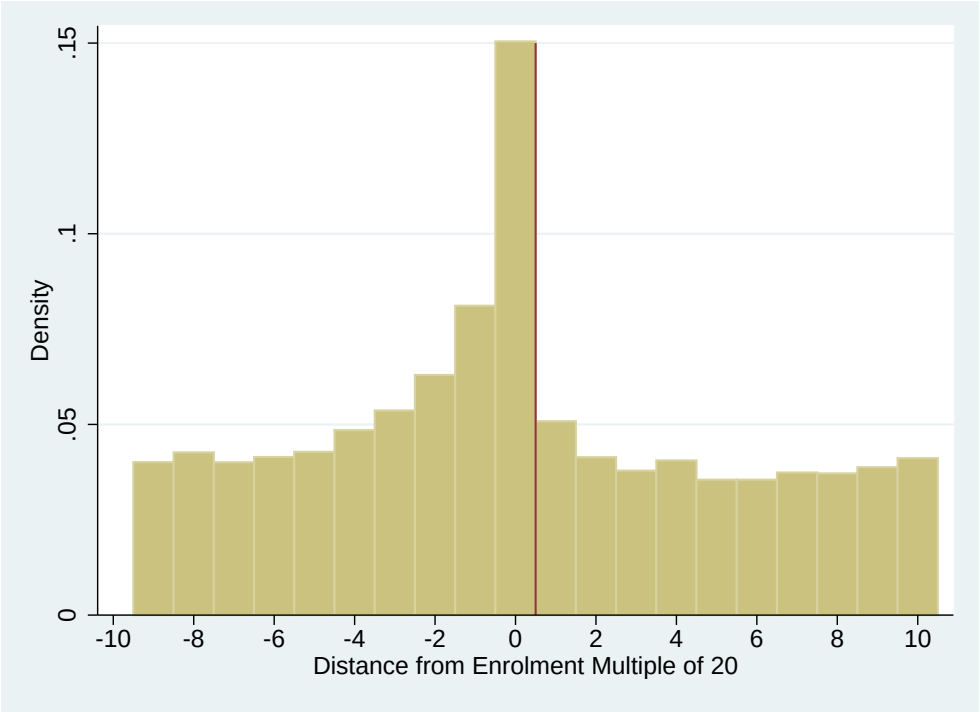
b= 1.61*** (0.10)



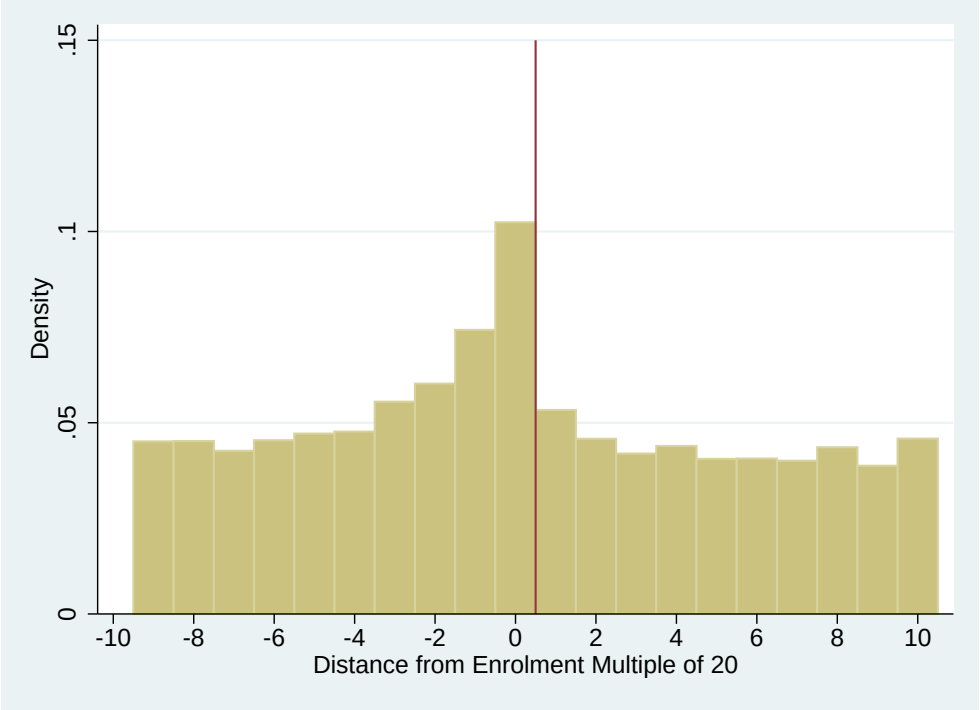
Notes: Figures 11(a) and 11 are based off 13,704 and 197,566 school-grade-year level observations, respectively.

Figure 12: Enrollment Bunching by Top and Bottom Within-District Poverty Decile

(a) Lowest Within-District Poverty Rank Decile
b= 2.421*** (0.19)

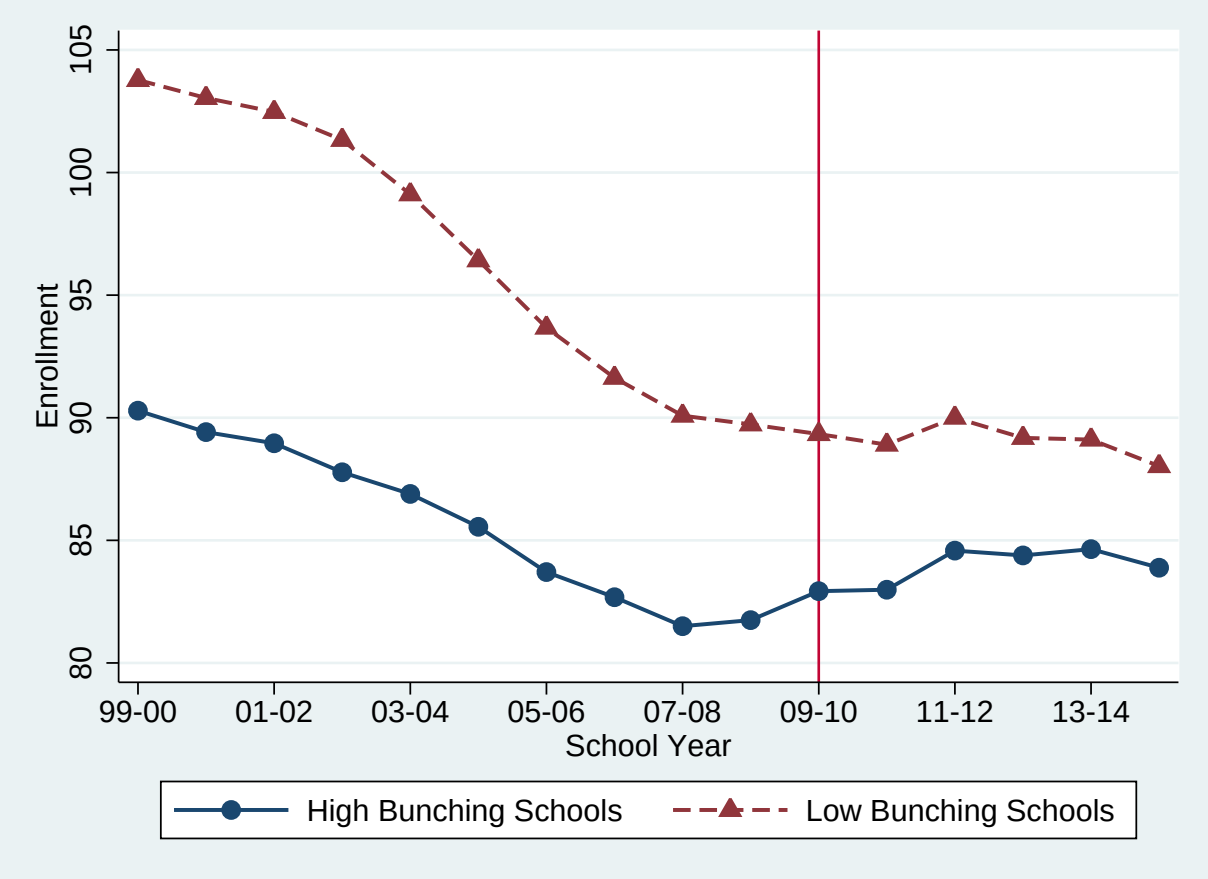


(b) Highest Within-District Poverty Rank Decile
b= 1.59*** (0.15)



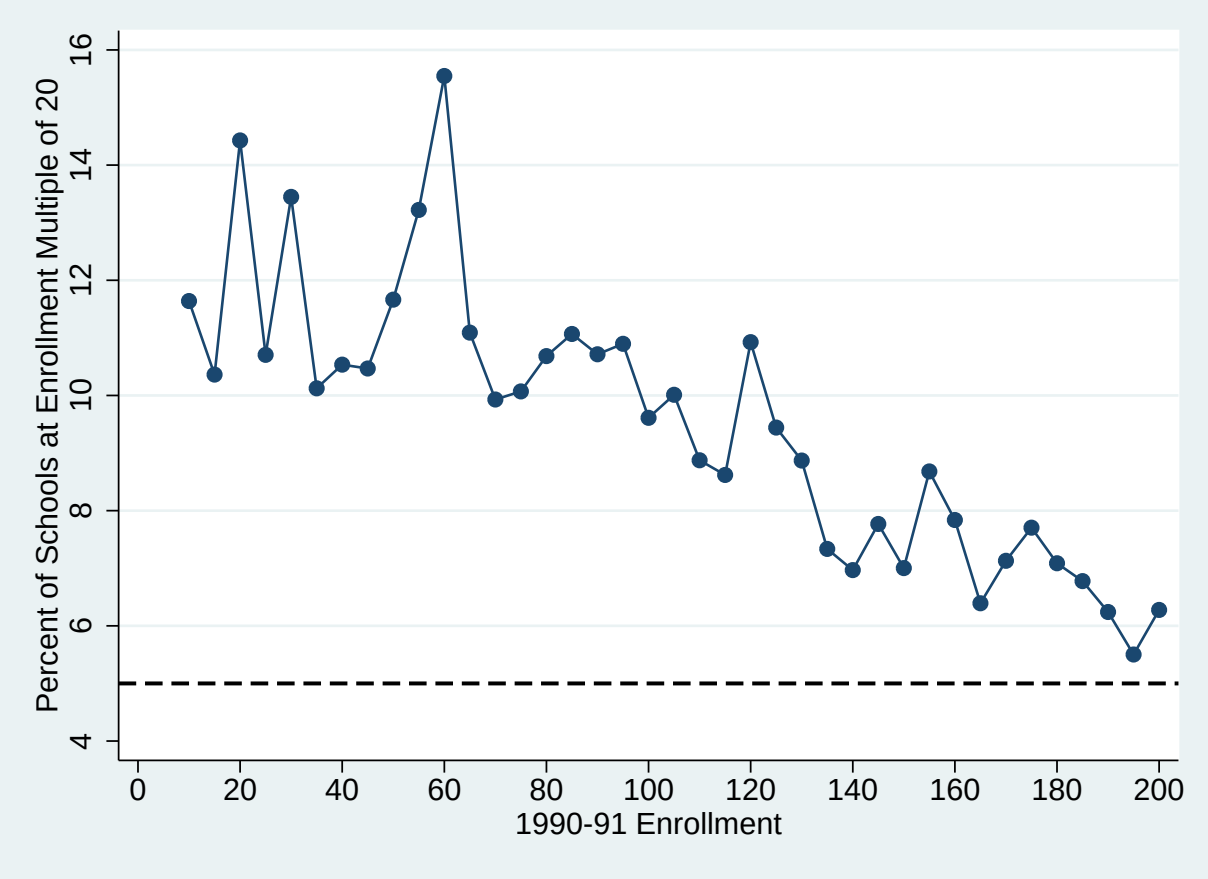
Notes: Figures 12a(a) and 12b are based off 22,361 and 20,722 school-grade-year level observations, respectively. Only districts with five or more schools in a given grade-year combination are included.

Figure 13: Average Enrollment by Year



Notes: Figure 13 covers grades K-3 over the 1999-00 through 2008-09 school years and consists of 197,566 school-grade-year observations. Only schools in districts with more than one school are included.

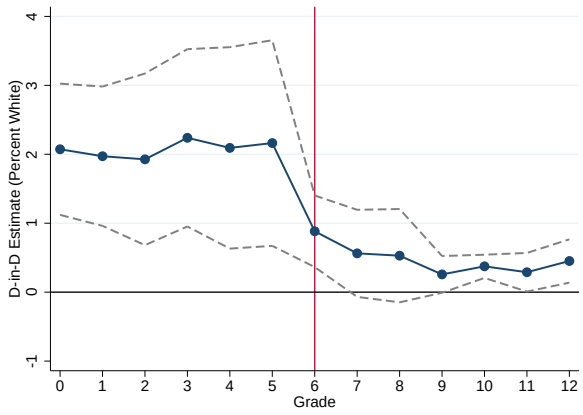
Figure 14: Bunching by Enrollment in CSR Grades



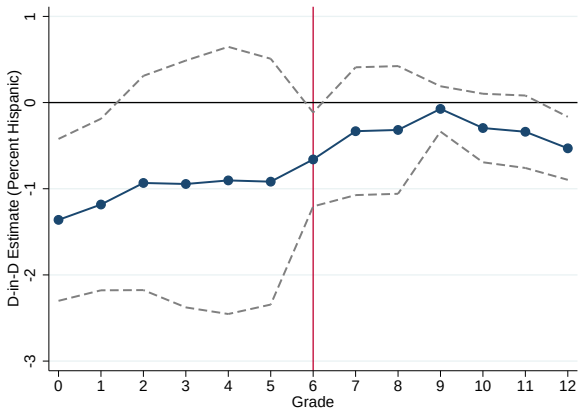
Notes: Figure 14 covers grades K-3 over the 1999-00 through 2008-09 school years and consists of 197,566 school-grade-year observations. Only schools in districts with more than one school are included. The horizontal dashed line represents 5 percent, which is the fraction of schools at an enrollment multiple of 20 that would be expected under no bunching.

Figure 15: Difference-in-Differences Estimates of Student Composition (School-grade Level)

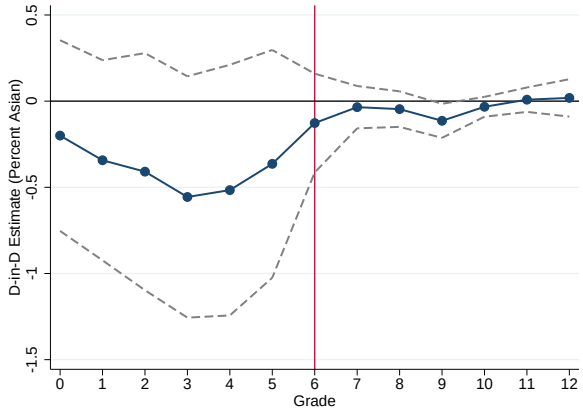
(a) Percent White



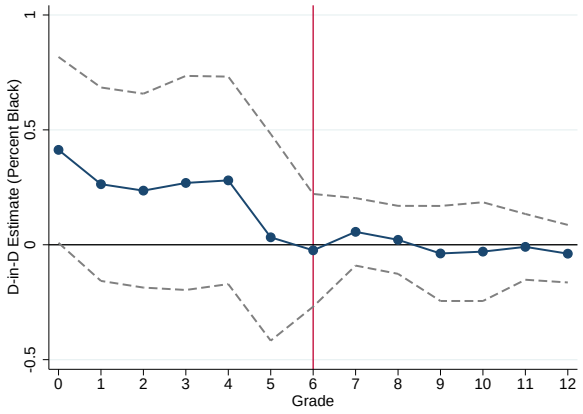
(b) Percent Hispanic



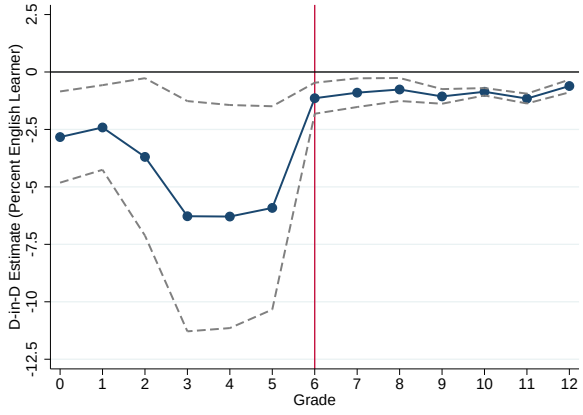
(c) Percent Asian



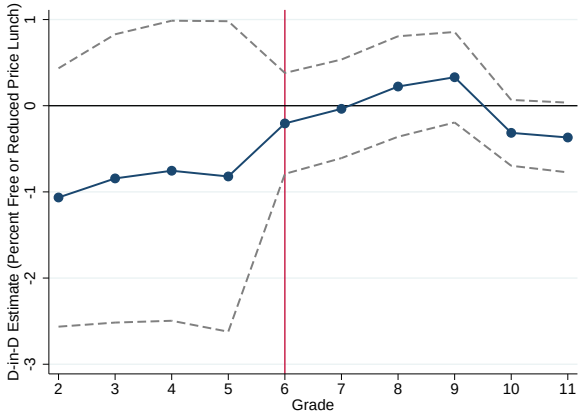
(d) Percent Black



(e) English Learners



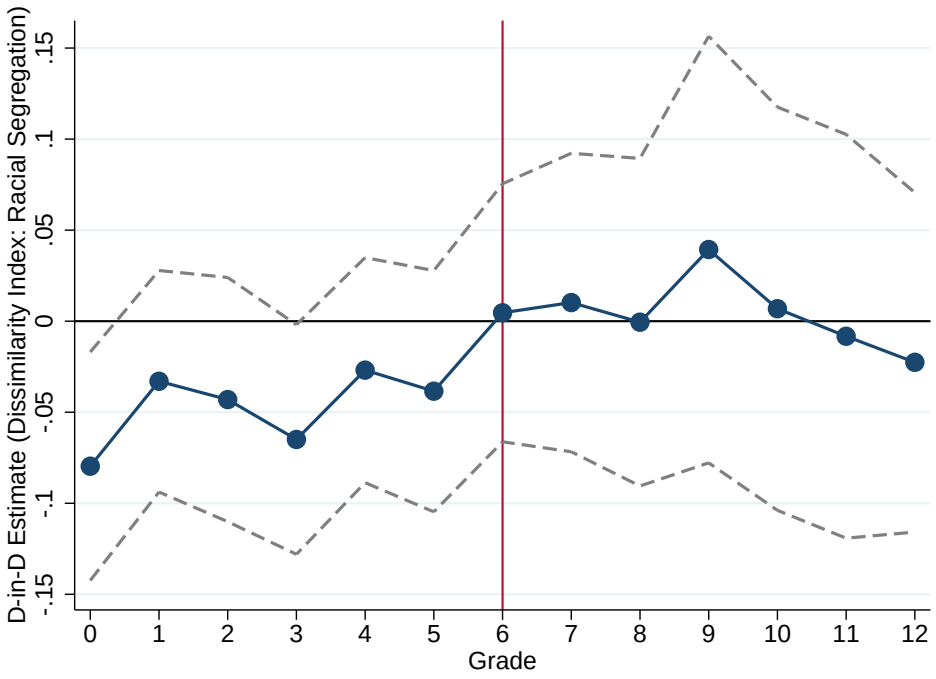
(f) Percent Free or Reduced Price Meal



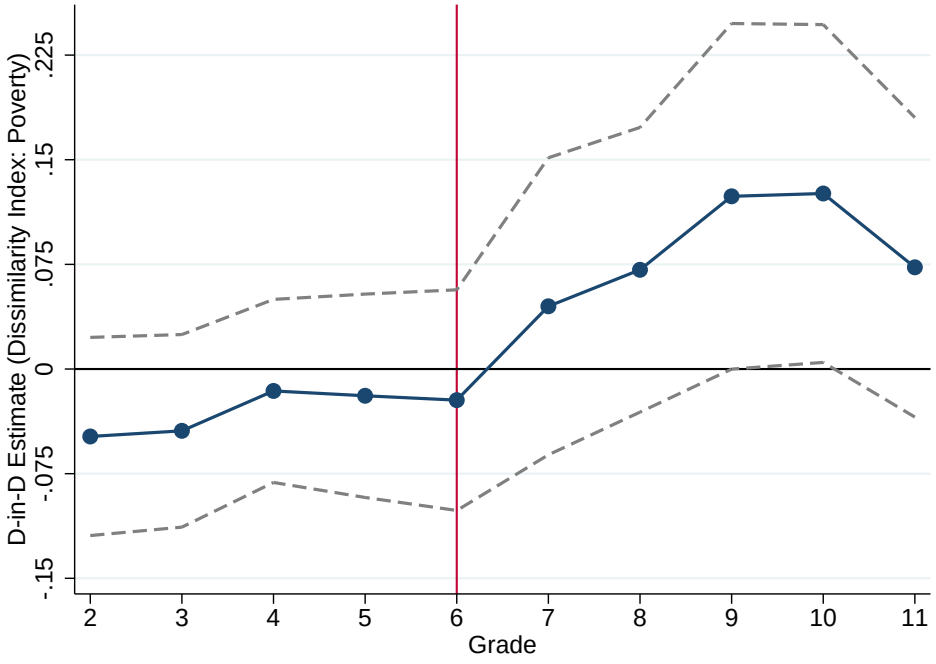
Notes: All figures, with the exception of Figure 15(f), cover grades K-12 for the 1999-00 through 2014-15 school years. (Free and Reduced Price Meal data is only available for grades 2-11 for the 1999-00 through 2012-13 school years.)

Figure 16: Difference-in-Differences Estimates of School Segregation (District-grade Level)

(a) Racial (White-Hispanic) Dissimilarity Index



(b) Socioeconomic (FRPM-non FRPM) Dissimilarity Index



Notes: Figures 12a and 12b are based off 22,361 and 20,722 school-grade-year level observations, respectively. Only districts with five or more schools in a given grade-year combination are included.

Table 1: Descriptive Changes in School Characteristics (CSR Grades Only)

	High-Bunching Schools		Low-Bunching Schools		High_{Post} - High_{Pre}
	Hard Cap (Pre)[1]	Post-Hard Cap[2]	Hard Cap (Pre)[1]	Post-Hard Cap[2]	- (Low_{Post} - Low_{Pre})
	(1)	(2)	(3)	(4)	(5)
Bunching[3]	32.0	12.4	10.8	10.5	-19.4
Enrollment	82.4	80.7	92.7	87.3	3.6
% White	35.8	28.3	32.0	25.9	-1.4
% Hispanic	39.3	47.1	47.8	54.4	1.2
% Black	7.7	6.4	7.4	5.9	0.2
% Asian	10.2	10.0	6.7	6.8	-0.3
% English Learner	28.7	31.9	35.1	36.7	1.7
% FRPM[4]	44.1	56.4	52.2	64.5	-0.1
Math Scores[5]	0.22	0.16	-0.11	-0.09	-0.07
Reading Scores[5]	0.23	0.18	-0.12	-0.10	-0.07
Observations (School-Grade-Year)	67,527	46,014	135,242	80,024	328,807

Notes: This table shows various summary statistics for CSR grades high- and low-bunching schools before and after the hard cap was lifted. Column (5) then takes the difference-in-differences (High-Bunching$_{Post\text{-}Hard\ Cap}$ - High$_{Hard\ Cap}$ - ($_{Post\text{-}Hard\ Cap}$ - Low$_{Hard\ Cap}$)) between these to show the change in high- vs. low-bunching schools after the hard cap was lifted. High-bunching schools are defined as schools whose K-3 school-grade enrollment falls on an enrollment multiple of 20 or one below more than twice as often as would happen by chance during the 1999-00 through 2008-09 school years. All other schools are defined as low-bunching schools.

[1] 'Hard cap (Pre)' refers to school years the 20 student class size hard cap was in place. This time period covers the 1999-00 through 2008-09 school years.

[2] 'Post-Hard cap' refers to years after the 20 student class size hard cap was lifted. This time period covers the 2009-10 through 2013-14 school years.

[3] 'Bunching' is defined as a school-grade with an enrollment multiple of 20 or one below.

[4] 'FRPM' is an acronym for 'Free or Reduced Price Meal Eligible' and is a measure of socioeconomic disadvantage. % FRPM is only available at the school-year level.

[5] Test scores are standardized to have mean zero and standard deviation one for each grade-year. Observations for test scores are 31,927, 15,923, 62,903 and 28,126 for columns (1), (2), (3) and (4), respectively.

Table 2: Triple-Differences Estimates of CSR 'Hard Caps' on School enrollment and Demographics

Outcome Variable	D-in-D: Most Schools Serve CSR grades (K-6)	D-in-D: Most Schools *do not* Serve CSR grades (7-12)	Triple-differences
	(1)	(2)	(3)
School Bunching	3.45***	0.14*	3.24***
	(0.57)	(0.08)	(0.56)
Enrollment	-26.53***	-18.06***	-8.62***
	(4.93)	(3.78)	(1.76)
Percent White	2.06***	0.40***	1.67***
	(0.57)	(0.13)	(0.50)
Percent Hispanic	-1.05*	-0.32	-0.75*
	(0.60)	(0.20)	(0.44)
Percent Black	0.25	-0.01	0.27
	(0.20)	(0.08)	(0.19)
Percent Asian	-0.38	-0.03	-0.36
	(0.31)	(0.03)	(0.30)
Percent English Learner	-4.37***	-0.92***	-3.43***
	(1.65)	(0.08)	(1.61)
Percent Free or Reduced Price Lunch (FRPM)[1]	-0.86	-0.15	-0.74
	(0.82)	(0.18)	(0.70)
Observations	362,576	84,171	446,747
School/Grade-by-Year FE	Yes	Yes	Yes

Notes: Column (1) provides estimates from the difference-in-differences specification given by Equation (5.2) for grades K-6 in which most schools serve CSR grades, while Column (2) gives these estimates for grades 7-12 in which most schools do not serve CSR grades. Column (3) provides estimates from the triple-differences specification given by Equation (5.2), which roughly calculates the difference between Columns (1) and (2). Observations are at the school-grade-year level and cover 1999-00 through 2014-15 school years. Only schools in many-school districts are included. Standard errors are clustered at the district level. ***,** and * denote significance at the 1%, 5% and 10% levels, respectively.

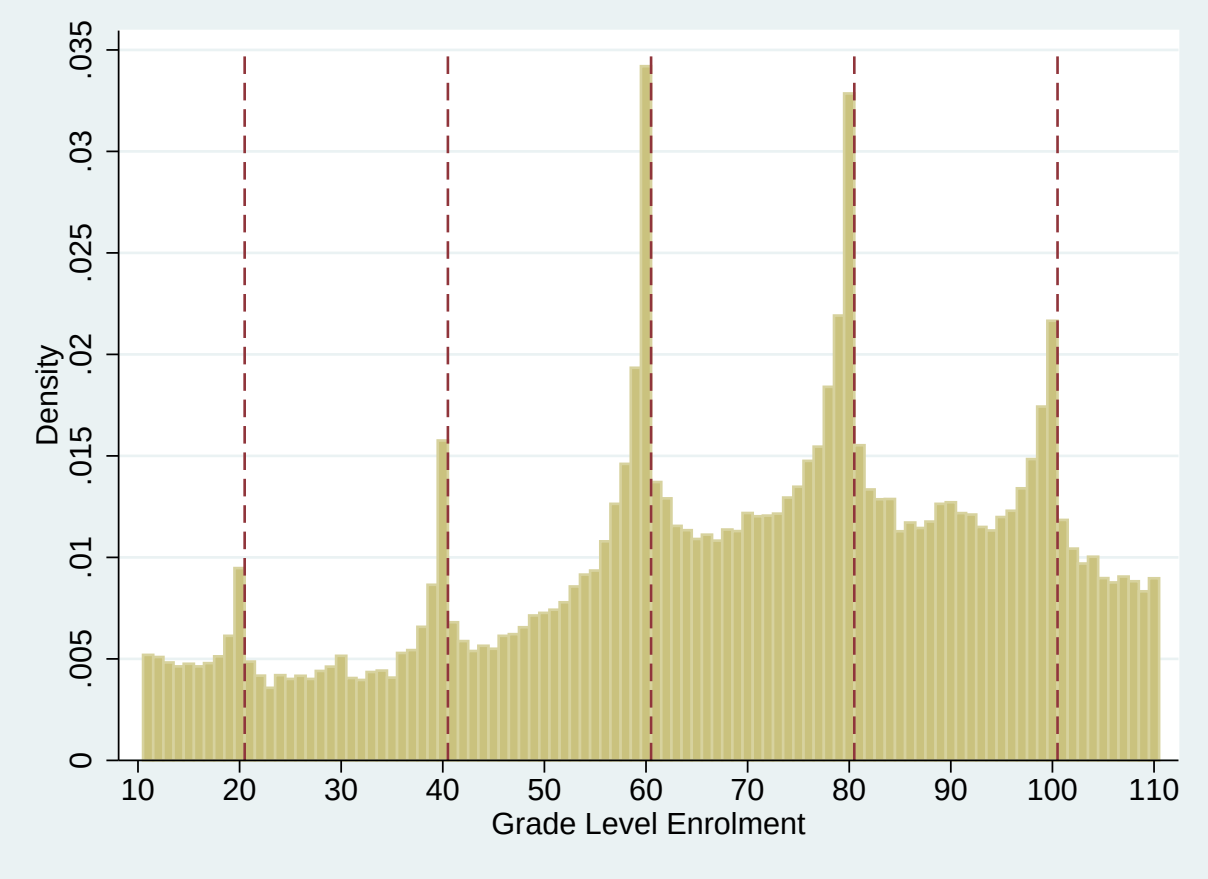
[1] Percent FRPM has 212,225, 61,071 and 273,296 observations in Column (1), (2), and (3), respectively. This is because FRPM data are only available for grades 2-11 until the 2012-13 school year.

Table 3: Triple-Differences Estimates of CSR 'Hard Caps' on District Level Segregation

Outcome Variable	D-in-D: Most Schools Serve CSR grades (K-6)	D-in-D: Most Schools *do not* Serve CSR grades (7-12)	Triple-differences
	(1)	(2)	(3)
Dissimilarity Index: Racial	-0.046*	-0.008	-0.079*
	(0.27)	(0.038)	(0.043)
Dissimilarity Index: Socio-economic	-0.031	0.081*	-0.136***
	(0.031)	(0.042)	(0.050)
Observations (Dissimilarity Index: Racial)	46,756	27,928	74,684
Observations (Dissimilarity Index: Socio-economic)	26,534	19,954	46,488
District/Grade-by-Year FE	Yes	Yes	Yes

Notes: Column (1) provides estimates from the difference-in-differences specification given by Equation 5.3 for grades K-6 in which most schools serve CSR grades, while Column (2) gives these estimates for grades 7-12 in which most schools do not serve CSR grades. Column (3) provides estimates from the triple-differences specification, which roughly calculates the difference between Columns (1) and (2). Observations are at the district-grade-year level and cover grades K-12 for 1999-00 through 2014-15 school years for the racial dissimilarity index and grades 2-11 for 1999-00 to 2012-13 for the socioeconomic dissimilarity index. Covariates including enrollment, race, English learner status, and free and reduced price meal status are included. Only school districts with many schools and whose relevant demographic characteristics make up none or one-hundred percent of the student body are not are included. Standard errors are clustered at the district level. ***,** and * denote significance at the 1%, 5% and 10% levels, respectively.

Figure 17: Distribution of School-grade Enrollment in CSR grades



Notes: Figure 17 covers grades K-3 over the 1999-00 through 2008-09 school years.The vertical dashed lines represent enrollment multiples of 20.

Table 4: School Statistics by Grade Span Configuration

Grade Span	Number of Schools	% of 2008-09 Total enrollment
	(1)	(2)
K-5	2,468	21.9
K-6	1,976	18.3
K-8	798	5.7
K-12	194	1.6
6-8	843	12.1
7-8	324	4.4
9-12	1,289	28.2
Other	1,371	7.8

Notes: Number of Schools refers to number of schools of that grade span in California in the 2008-09 school year.