CAPSTONE REPORT

CSR Research Consortium

a partnership researching California's class size reduction reform

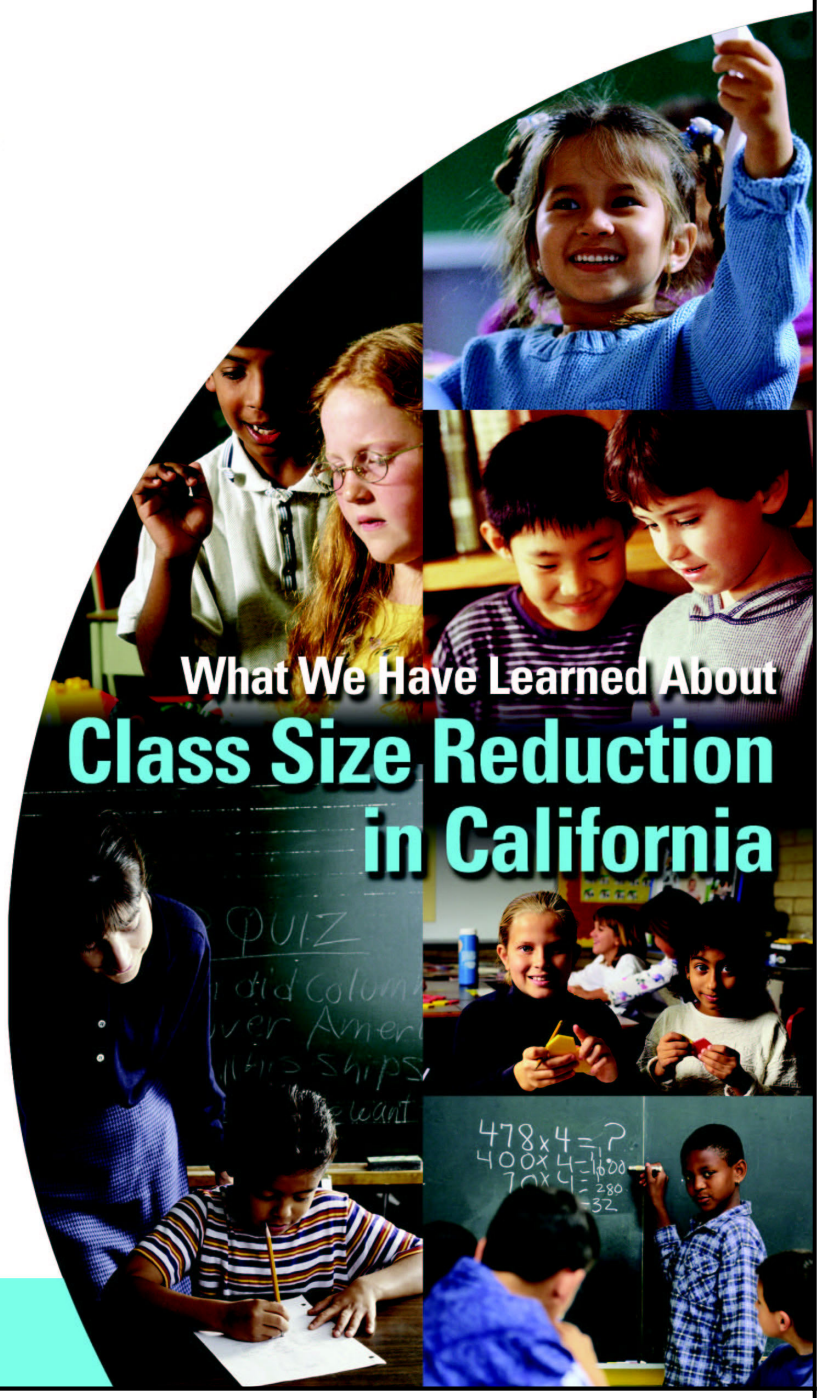
American Institutes for Research

RAND

PACE

EdSource

WestEd



What We Have Learned About Class Size Reduction in California

Editors

George W. Bohrnstedt

Brian M. Stecher

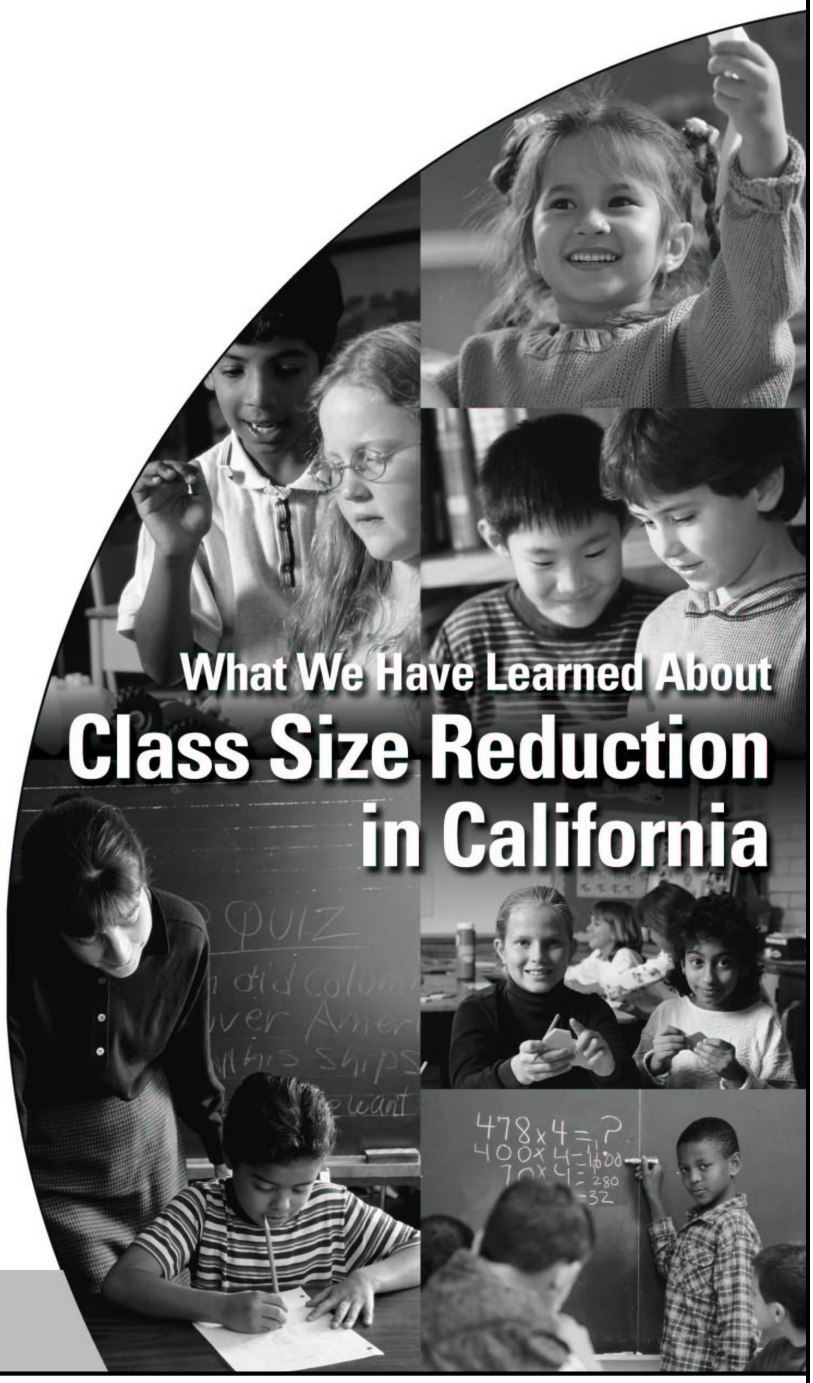
SEPTEMBER 2002

Technical Appendix

CAPSTONE REPORT

CSR Research Consortium

a partnership researching California's class size reduction reform



What We Have Learned About Class Size Reduction in California

SEPTEMBER 2002

Technical Appendix

Table of Contents

APPENDIX A: Diminishing Resources and the CSR Program A-1
Adequacy of Funding A-1
Prospects for the 2002–03 School Year A-2

APPENDIX B: The Relationship Between CSR Exposure and Achievement at the School Level B-1
Introduction B-1
Methods B-2
Results B-11

APPENDIX C: Class Size Reduction and Teacher Migration C-1
Methods and Data Sources C-3
New Hires C-4
Cross-School Migration C-6
Cross-Grade Migration C-20
Teacher Attrition C-26
Conclusions C-30
Supplement to Appendic C: Detailed Grade Migration Tables C-34

APPENDIX D: The Relationship Between Teacher Characteristics and Student Achievement in Reduced-size Classes: A Study of 6 California Districts D-1
Study Objectives D-1
Sample D-2
Data D-2
Characteristics of the Data, by District D-3
Model D-11
Results D-12
Teacher Education and Credential Status D-15
Student Characteristics D-15
Conclusions D-18
Discussion D-18
Supplement to Appendix D D-19

APPENDIX E: State Class Size Reduction Measures E-1

Appendix A

Diminishing Resources and the CSR Program

Georges Vernez

In the winter of 2002, declines in state and local revenues fueled speculation that funding for education might be reduced in spite of expectations that enrollment would continue to increase in most of the state's districts. Newspapers began reporting that some districts were considering curtailing the CSR program, if not dropping it entirely, in an attempt to balance their budgets.

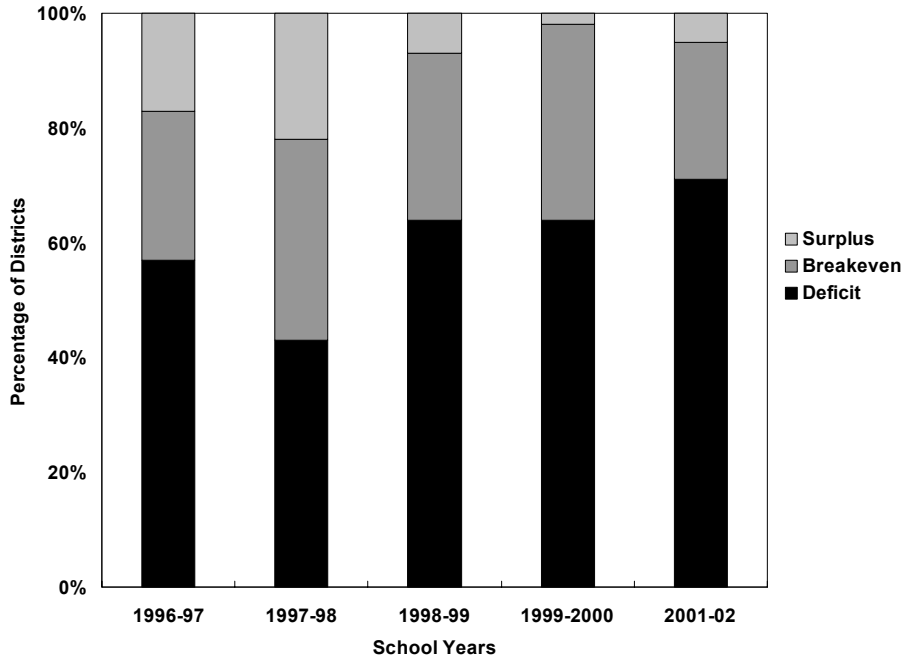
To assess the extent of this problem, we conducted a telephone survey in April 2002 of 38 district superintendents randomly selected from the original 84 districts that were surveyed by mail in 2000.[1] As in our 1998 and 2000 surveys, we asked districts about the adequacy of state reimbursement to cover the cost of implementing the CSR program and about the size of their deficit or surplus. In addition, we asked about their anticipated budget for school year 2002–03 and, if they expected funding to be inadequate to maintain all academic and other programs, about how they planned to balance their budget.

Adequacy of Funding

Two-thirds of the districts contacted reported that their operating costs of implementing the CSR program exceeded state reimbursement in school year 2001–02—about the same proportion as in the previous school years (Figure A1). All but two of the remaining districts broke even. These latter districts are generally smaller and have declining enrollments.

[1] Bohrnstedt, G.W. and Stecher, B.M. (1999). *Class size reduction in California: Early evaluation findings, 1996–1998*. Palo Alto, CA: American Institutes for Research.

Figure A1–
Adequacy of CSR Funding



Sources: CSR Consortium 1998 and 2000 Surveys of Superintendents and 2002 telephone survey of superintendents. Based on a random subsample of 38 districts drawn from the Consortium's larger 84-district sample.

In districts where the costs of operating the CSR program exceeded state reimbursement, the estimated deficit averaged from $57 per pupil to $350 per pupil. Larger districts generally faced larger average deficits.

Prospects for the 2002–03 School Year

Nearly all districts expected that their respective financial situations would not change for the better in school year 2002–03 and might even worsen depending on actual state reimbursement for CSR, the outcome of collective bargaining—increases in teacher salaries have exceeded the COLA adjustment applied to state reimbursement for the CSR program—and increases or decreases in school enrollment.

Sixteen percent of the districts contacted indicated they were considering cutting back their CSR program by discontinuing it in Kindergarten and/or third grade. None of these districts had made a final decision, however. They were considering other options as well, assessing the potential community and internal (teachers, unions, and other staff) responses to curtailing the CSR program before making a decision.

The majority of districts that were not considering a cutback in the CSR program felt that, although it was expensive, its benefits outweighed its costs. And even if there were

no documented educational benefits, they anticipated that cutting back would be politically difficult given the popularity of the program among both parents and teachers.

Finally, half of the districts anticipated budgets for school year 2002–03 that would be inadequate to maintain the level of academic and support programs offered in the previous year. Reasons given included increased enrollments, state budget cuts, salary increases (some dependent on bargaining agreements reached during healthy economic times) and increases in health and utility costs. To balance their budgets, some districts were considering cutting after school and community programs. But most districts were considering cutting back rather than eliminating programs. Candidate programs cited most frequently for cutting back included nursing, tutoring, counseling, music, library, drama, visual arts and physical education. These are the very same programs that have been continuing targets for cutbacks since implementation of the CSR program began.[2]

In summary, the Consortium's survey of 38 districts indicates that the vast majority of districts are not considering a reduction in class size in California's K–3 classrooms, as various newspaper articles have suggested. The districts did indicate, however, that they were increasingly facing budget shortfalls, but that the funds would more likely come from programs other than CSR to make up the deficits.

[2] Ibid

Appendix B

The Relationship Between CSR Exposure and Achievement at the School-Level

Brian M. Stecher, Daniel F. McCaffrey, and Delia Bugliari

Introduction

In the first two Class Size Reduction (CSR) evaluation reports (see Bohrnstedt and Stecher 1999; Stecher and Bohrnstedt, 2000), we estimated the impact of CSR on student achievement by comparing the Stanford Achievement Test, 9th Edition (SAT-9) test scores of third-grade students taught in reduced size classes with those of third-grade students taught in non-reduced size classes.[1] Pre-existing differences between the CSR and non-CSR students were adjusted for statistically using student and teacher background characteristics as well as scores from fourth- and fifth-grade students who had little or no exposure to CSR.

We could not use a similar, comparative approach in subsequent evaluation reports because CSR had been implemented in over 95 percent of the third-grade classes in California by 2000–01. The rapid implementation left too few untreated students to serve as a comparison group. Furthermore, by 2000-01 some or all of the upper-grade (i.e., fourth- and fifth-grade) students in most schools had participated in reduced size classes in earlier years, so we could not use their test results to control for pre-existing differences.

As a result, we approached the problem differently in the third CSR evaluation report, focusing on differences in cumulative exposure to CSR and their relationship to achievement (Stecher and Bohrnstedt, 2002). The uneven growth in participation in CSR from 1996-97 to 2000-01 provided an opportunity to examine trends in achievement among cohorts of students who had *different* levels of exposure to CSR.

The analysis focused on statewide average achievement scores during the period 1997–98 to 2000–01. We compared the average achievement of successive cohorts of students as they moved through the system with their average exposure to CSR.

[1] The Consortium's analyses were limited by the fact that there were no achievement data for kindergarten students or first grade students in any year, and there were no achievement data for any students prior to 1998.

Trends in achievement that correspond to patterns of exposure provide evidence in support of the hypothesis that CSR improves achievement; trends that bear no relationship to CSR participation offer no such support.

Successive cohorts of students had higher achievement during this period, which suggests that one or more of the state educational reforms (which include CSR, new curriculum standards, a statewide standardized testing program, the end of bilingual education, and high stakes accountability) had a positive effect. However, the trend in test scores over this period was unrelated to the trend in CSR exposure, so we could not make a strong case that CSR was chiefly responsible for achievement gains. Yet, aggregate analyses do not tell the whole story. For example, the state level analysis could not control for external effects, such as student mobility. Neither did it permit us to examine the influence of student or teacher background characteristics. As a result, we conducted additional analyses of trends in exposure and achievement at the school level.

Methods

Achievement Data

Beginning in 1998, California students in grades 2–11 have been required to complete the SAT-9 annually in the spring. The test results are reported in the summer and fall, and they are made available for research purposes in the public release California Standardized Testing and Reporting (STAR) data files. All analyses reported below use the public release STAR data.

As part of STAR testing, students complete standardized multiple-choice tests in mathematics, reading, language and spelling. We focus here on mathematics, reading and language. We use SAT-9 scale scores (rather than raw scores, percentile ranks, or normal curve equivalents) as measures of achievement in these analyses because scale scores are designed so that score differences are comparable for the entire range of scores. In addition, the scales are equated across grade levels, facilitating cross grade comparisons.

School Sample

The initial school sample included 4,961 elementary schools in the STAR data files from school years 1997-98 through 2000-01. We excluded those schools for which the STAR file in any year contained scores for 10 or fewer students and those schools for which the STAR files were missing basic demographic data (gender, ethnicity, English language fluency status) on all students. These criteria excluded 2,069 school (42%), leaving 2,892 schools in our analysis file.

Despite the exclusions, the schools in our sample closely resemble the schools in the state as a whole in terms of student demographic characteristics. Table B1 shows the comparison between the sample schools and the whole state in terms of participation in CALWORKS, eligibility for free or reduced priced lunches, race/ethnicity, and

language status for the 1999-2000 school year. The mean values for sample schools are within one to three percentage points of the mean values for the state as a whole on all variables, so the generalizability of the results from our analyses are not limited by the populations served by sampled schools.

Table B1–
Demographic Characteristics of Sample Schools and All Elementary Schools

Demographic Feature	All Elementary Schools*	Analysis Sample Schools
Percent CALWORKS Participants	13.59 (12.88)	14.64 (12.97)
Percent Free or Reduced Price Lunch Eligible	51.99 (30.27)	53.76 (30.17)
Percent White	38.39 (29.38)	34.86 (28.66)
Percent Hispanic	40.98 (29.31)	43.19 (29.60)
Percent African American	8.13 (12.60)	9.28 (14.00)
Percent Asian	7.76 (12.01)	7.88 (11.40)
Percent Minority	61.61 (29.38)	65.15 (28.66)
Percent ELL	27.14 (24.10)	29.29 (24.30)
Total Enrollment	609.94 (282.39)	660.40 (276.38)

*State sample includes 4761 elementary schools open since 1996 with CDS codes

Class Size Reduction Participation

Class size reduction began with the 1996-97 school year, one year prior to STAR testing. By the 1999-2000 school year over 90 percent of all students in kindergarten through third grade were participating in CSR. However, for earlier cohorts, CSR participation varied across schools. This variation provided an opportunity to compare achievement with CSR exposure. The first step in our analysis, therefore, was to determine CSR participation by grade and school year for each of the 2,892 schools in the analysis file. We focused on CSR participation for three cohorts of students—those who entered kindergarten in 1995-96 (K95), 1996-97 (K96) or 1997-98 (K97). These three cohorts of students reached the third grade in 1999, 2000, and 2001 and they are the only cohorts with exposure to CSR for whom we have SAT-9 scores in both second and third grade.

For each elementary school in California we developed an indicator of CSR participation by grade level by year. Unfortunately, the state did not collect comparable information about CSR participation every year, so we had to use multiple data sources to infer CSR status. The primary data for assessing CSR status were the individual student SAT-9 answer files, which included indicator variables for CSR participation for every student. We also used teacher reports of classroom enrollment from the CBEDS Professional Assignment Information Form (PAIF). A third source of information was the district level J-7 CSR report, which describes district participation in CSR for the 1996-97 and 1997-98 school years. The J-7 information was only useful when participation was uniform across the district.

Finally, the CBEDS School Information File (SIF) data contain school and grade level CSR indicators for the 1998-99, 1999-2000, and 2000-01 school years.

The CSR indicator development process began with the student-level STAR data file. If 10% or fewer students within a grade at a school were coded as participating in the CSR program (either option 1 or 2), we classified that grade as not reduced. If 90% or more students within a grade at a school were indicated as in the CSR program, we classified that grade as reduced. We classified a grade as undetermined by STAR if between 10% and 90% of students were indicated as CSR. Let $C_{gjt,STAR}$ denote the CSR status for grade g = kindergarten, 1, 2 or 3, in school j and school-year t = 1996-97, 1997-98, 1998-99, 1999-00 or 2000-01. $C_{gjt,STAR}$ equals "R" if we determine the school had reduced classes for grade g in year t; $C_{gjt,STAR}$ equals "N" if not reduced and "U" if undetermined.

Because the STAR data did not permit clear classification for every school, grade level, or school year, i.e., in some instances $C_{gjt,STAR}$ equals "U," we turned to other sources to make our final determination of CSR participation. The PAIF data provide the number of students in each teacher's classroom and the number of teaching assignments. The distribution of students across classrooms for teachers with multiple assignments cannot be determined from the PAIF. Therefore, for determining CSR participation we used only teachers with a single teaching assignment. Also, some teachers report over 50 students or fewer than 14 students in their classroom. We excluded these teachers from the classification process, arguing that they represented data errors or nontraditional education assignments.

A school was judged to have reduced size classes for a given grade in a given year if over 65% of included teachers in that grade reported 21 or fewer students. If fewer than 35% of included teachers in a grade reported 21 or fewer students, we classified that grade as not reduced. We classified a grade as undetermined by PAIF if between 35% and 65% of the classes were reported as having 21 or fewer students. We let $C_{gjt,PAIF}$ denote the CSR status as determined by the PAIF where the variable again takes on the values of "R", "N" and "U" for reduced, not reduced or undetermined.

We also created variables for the CSR participation as determined by the SIF ($C_{gjt,SIF}$) and the J-7 data ($C_{gjt,J7}$). $C_{gjt,SIF}$ equals "U" for the 1996-97 and 1997-98 school years for all grades and schools because grade-level CSR indicators were not added to SIF until 1998-99. Finally, $C_{gjt,J7}$ takes on values "R" and "N" only if the district had uniform CSR practices at a grade level across all schools.

For final CSR classification, we compared the CSR indicators based on STAR, PAIF, SIF and J-7. In the majority of cases, all determinable sources agreed, $C_{gjt,STAR} = C_{gjt,PAIF} = C_{gjt,SIF} = C_{gjt,J7}$ or some variables equaled "U" and the remaining variables agreed. In these cases we assigned the common value to the CSR indicator. In the cases of disagreement, we examined the longitudinal trend in CSR indicators before making a final determination. For example, if $C_{gjt,STAR}$ = R and $C_{gjt,PAIF}$ = N for year t we checked the data for the previous year ($t - 1$). If $C_{gjt-1,STAR} = C_{gjt-1,PAIF}$ = R then we

decided that the school probably had reduced class size in year *t* as well. Schools for which we were unable to resolve data conflicts confidently were excluded from the final analytic file. We excluded 543 schools because of indeterminate CSR status, leaving a sample of 2,349 schools. The excluded schools constituted 19% of the 2,892 schools that met the data and size conditions described above. The remaining schools constituted 47% of the original sample.

CSR Exposure by Cohort

For each of the three focal cohorts, K95, K96 and K97, Table B2 presents the distribution of CSR exposure across the final sample of schools. Table B2a shows that nearly 90 percent of the schools in the sample had one of two patterns of CSR exposure for the K95 student cohorts: CSR in grades 2 and 3 only (22.3%) or CSR in grades 1, 2 and 3 (66.8%). For the K96 cohort there was even less variation in CSR exposure. Table B2b shows that these students participated in CSR for grades 1, 2 and 3 in almost every school (89.9%). By the K97 cohort, Table B2c shows that more schools introduced CSR in kindergarten, and the schools fell, almost exclusively, into one of two patterns of CSR exposure: kindergarten through grade 3 (38.8%) or grades 1, 2 and 3 (59.9%).

Table B2a–
Distribution of CSR Exposure for Cohort K95

Exposure Pattern	Number of Schools	Percent of Sample
Indeterminate	20	0.9
None	25	1.1
Grade 3 Only	7	0.3
Grade 2 Only	66	2.8
Grades 2 and 3	525	22.3
Grade 1 Only	10	0.4
Grades 1 and 3	5	0.2
Grades 1 and 2	105	4.5
Grades, 1, 2 and 3	1,569	66.8
Kindergarten and Grade 3	1	0.0
Kindergarten, Grades 2 and 3	1	0.0
Kindergarten, Grades 1, 2 and 3	15	0.6

Table B2b–
Distribution of CSR Exposure for Cohort K96

Exposure Pattern	Number of Schools	Percent of Sample
Indeterminate	12	0.5
None	1	0.0
Grade 3 Only	1	0.0
Grades 2 and 3	12	0.5
Grade 1 Only	4	0.2
Grades 1 and 3	1	0.0
Grades 1 and 2	50	2.1
Grades, 1, 2 and 3	2,112	89.9
Kindergarten, Grades 2 and 3	1	0.0
Kindergarten, Grades 1, 2 and 3	155	6.6

Table B2c–
Distribution of CSR Exposure for Cohort K97

Exposure Pattern	Number of Schools	Percent of Sample
Indeterminate	7	0.3
Grades 2 and 3	1	0.0
Grades 1 and 2	20	0.9
Grades, 1, 2 and 3	1,406	59.9
Kindergarten, Grades 2 and 3	1	0.0
Kindergarten, Grades 1 and 2	3	0.1
Kindergarten, Grades 1, 2 and 3	911	38.8

Grouping Schools by CSR Exposure

We focused our analyses on four groups of schools with distinctive patterns of CSR exposure. These 1,918 schools constitute 82% of the schools in the final analysis sample and 40% of the schools in the original sample. Table B3 shows these four

patterns. Because few schools had any of the remaining exposure patterns, we restrict the study to schools in these four groups.

Table B3–
Distribution of CSR Exposure for All Three Cohorts

Group	K95	K96	K97	Number of Schools
A	1, 2, 3	1, 2, 3	1, 2, 3	877
B	2, 3	1, 2, 3	1, 2, 3	348
C	2, 3	1, 2, 3	K, 1, 2, 3	152
D	1, 2, 3	1, 2, 3	K, 1, 2, 3	541

Demographic differences across groups (described below) led us to focus our primary comparisons of outcomes on Group A and Group B. These two groups contain 1,225 schools. In Group A, students who entered kindergarten in 1995-96, 1996-97 or 1997-98 had reduced-size classes in grades 1, 2 and 3 (but not kindergarten). Group B schools serve a similar population of students, but the three cohorts had different exposure to CSR. Students entering kindergarten in 1995-96 had two years of exposure to CSR in second and third grade, those entering in subsequent years had an additional year of CSR in first grade.

Student Sample

As noted above, our analyses are restricted to students in the K95, K96 and K97 cohorts. From these cohorts we included only those students who: 1) attended the same school for kindergarten through second or third grade, depending on the grade of the outcome used in the analysis; 2) did not have a test identified as "Out of Level"; and 3) were not identified as receiving Special Education services. We also excluded students when their STAR data CSR flag was inconsistent with the data from the vast majority (over 90 percent) of their fellow students in the same grade and school. For example if the STAR student data file indicated that for a particular school over 90 percent of third graders in a cohort were in reduced size classes, then we excluded any third graders from that school and cohort for whom the STAR data indicated they were not in reduced size classes.

Table B4 contains summaries of the student demographic characteristics and teacher qualifications of the identified cohorts of students in schools in the four groups. Groups A and B are similar in terms of students and teacher characteristics, while Groups C and D are distinctly different. Schools in Groups A and B have greater percentages of minority students, EL students, and students from families receiving public assistance than schools in Groups C and D. Groups A and B also are similar in terms of teacher characteristics, and they have fewer teachers who are fully-credentialed than schools in Groups C and D. These differences make comparisons between Groups C and D and the other groups difficult because such comparisons would confound student demographics and teacher qualifications with CSR effects. Therefore we focus only on Groups A and B.

There is one instance in which schools in Groups A and B differ with respect to teacher credentials that only is apparent when the data are disaggregated by cohort. Group B schools have more uncredentialed first-grade teachers than Group A schools for cohorts K96 and K97. This difference appeared when Group B introduced CSR at first grade, and it probably is a result of these schools hiring new teachers in the tight teacher labor market that followed the introduction of CSR. (See Tables B7 to B12 for student and teacher characteristics disaggregated by cohort and grade level.)

Table B4–
Average Student and Teacher Characteristics for Cohorts K95, K96 and K97, by Group

Group	Student Characteristics*			Teacher Characteristics**	
	Minority %	EL %	AFDC %	Experience	Credential
A	66.84	33.23	20.40	13.30	89.13
B	69.23	32.06	21.09	13.25	88.51
C	57.66	25.91	18.38	13.46	93.10
D	51.99	20.67	18.26	13.52	94.71

Note: *Average for the three cohorts during their kindergarten, first, second, and third grades.
**Average years of experience for teachers of the identified cohorts of students; percentage of teachers of the identified cohorts of students with full credentials.

Group A schools had between 53,000 and 59,000 students per cohort when the cohorts reached grade 2 and between 46,000 and 48,000 students per cohort when the cohorts reached third grade. For Group B, the numbers of second graders per cohort ranged from 23,000 to 25,000 and the number of third graders per cohort ranged from 19,000 to 21,000. The samples are smaller in third grade than in second grade because they are restricted to students who attended the same school for one additional year.

Analysis

Our goal is to determine if cohort-to-cohort variation in CSR exposure predicts cohort-to-cohort variation in test scores. On the basis of the exposure patterns presented in Table B3, we note that a comparison of schools across years, groups and cohorts can only provide data on the effects of a one-year variation in exposure to CSR. Larger differences in exposure do not exist among comparable groups of schools. In addition, other reforms and changes were taking place during this period that might have affected test scores. As a result, a simple comparison of scores for students in the K95 cohort with scores for students in the K96 or K97 cohorts might confound CSR effects with these other changes. More complex comparisons however can isolate the effects of CSR with less confounding of alternative effects. For example, because the exposure to CSR was the same for all three cohorts in Group A, these schools provide a measure of the effects of factors unrelated to CSR on the trend in scores over these three years. Similarly, differences between K96 and K97 scores in Group B schools also are unrelated to CSR because exposure was the same for these two cohorts (but not for the K95 cohort). Thus, differences among

these five cohorts in Groups A and B can be used to estimate the effects of other programs and the effects of cohort-to-cohort variation.

On the other hand, the students in the Group B-K95 cohort had one year less CSR exposure during first grade than the students in the two later Group B cohorts and than students in all three cohorts in Group A. By comparing scores for the Group B-K95 students to those of other students, we can observe differences between groups with varying exposure to CSR. However, we must make judicious use of the data from the other students to limit the confounding effects of other programs and cohort-to-cohort variation in scores. The following list of comparisons with Group B-K95 highlights the assumptions about groups and time trends that are required for the comparisons to provide unconfounded estimates of the CSR effect. It also points out the comparisons that we believe provide the best estimates of the CSR effect.

Comparison 1: Compare Group B-K95 scores to Group B-K96 or Group B-K97 scores. The comparison yields unconfounded estimates of the CSR effect if we assume that, in the absence of CSR, scores do not change systematically over time. However, research has consistently shown that score gains occur in the years following the introduction of a new, high-stakes testing program even in the absence of other initiatives. Thus, this assumption seems unwarranted, i.e., scores are likely to change over time even in the absence of CSR. In fact, this change is evident in Group A where CSR exposure is constant. As a result, we will not use these within-Group B comparisons as an estimate of the CSR effect.

Comparison 2: Compare Group B-K95 scores to Group A-K95 scores. This comparison yields unconfounded estimates of the CSR effect if we assume that, in the absence of CSR, the groups would have the same scores on average. At first this assumption seems reasonable because the schools in the two groups are very similar on student demographic and teacher characteristics. However, the schools in Group A implemented CSR more quickly than schools in Group B, and the factors that led to this alternative behavior might be related to average scores. Thus, we do not think this assumption is warranted. (Alternatively, comparison of Group B-K95 to Group A-K96 or K97 would be affected both by time trends and cross group differences. The required assumptions for unconfounded estimation are not tenable in these comparisons either.)

Comparison 3: Compare the difference between Group B-K96 and Group B-K95 to the difference between Group B-K97 and Group B-K96. This comparison attempts to remove the time trend by using the difference between Group B-K97 and Group B-K96 scores as an estimate of the time trend between K95 and K96. The comparison yields unconfounded estimates if we assume that the time trend in scores is linear across the three cohorts. This is one of the estimates that will be presented in the results section. (In Table B5, Comparison 3 is found in the row labeled Difference and the column labeled Group B.)

Comparison 4: Compare the difference between Group B-K95 and Group A-K95 to the difference between Group B-K96 and Group A-K96. (This is equivalent to comparing the difference between K96 and K95 for Group B to the difference between K96 and K95 in Group A.) This comparison uses differences across Groups in K96 to estimate differences across groups in K95. Alternatively, we can view this estimate as using Group A to estimate the time trend from K95 to K96. This estimate is unconfounded if we assume that, in the absence of CSR, group differences would be constant over time. (We could also include the K97 cohorts in these comparisons.) We also present this comparison in the results section. (In Table B5, Comparison 4 is found in the row labeled K96 less K95 and the column labeled Difference.)

Comparison 5: Compare the difference in differences for Group B (i.e., compare the difference between K96 and K95 and the difference between K97 and K96) to the difference in differences for Group A. This model uses Group A to estimate the size of cohort-to-cohort deviations from a linear time trend in Group B. This model produces unconfounded estimates of the CSR effect if we assume that no interactions would exist in between groups and deviations from time trends in the absence of CSR. (In Table B5, Comparison 5 is found in the row labeled Difference and column labeled Difference.)

Because scores for students within the same school might be positively correlated and because schools vary in size, the simple average estimators described above might not be efficient. Therefore, we also fit a hierarchical linear model to estimate Comparison 5 while allowing for possible intra-school correlation. Model 1 for a score for the kth student in cohort t (t = 1 for K95, 2 for K96 and 3 for K97), school j of group i, y_{ijtk}, is given by

$$y_{ijtk} = \alpha_{ij} + \beta_{ij} t + \varphi_1 I(t=1) + \varphi_2 I(t=2) + \varepsilon_{ijtk}$$

$$\alpha_{ij} = \alpha_i + \eta_{ij}$$

$$\beta_{ij} = \beta_i + \zeta_{ij}$$

$$\varepsilon_{ijtk} \; i.i.d. \; N(0,\sigma^2), \eta_{ij} \; i.i.d. \; N(0,\tau^2), \zeta_{ij} \; i.i.d. \; N(0,\nu^2), Corr(\eta_{ij},\zeta_{ij}) = \rho$$

The functions I(t = 1) and I(t = 2) equal one if t = 1 or 2 respectively and zero otherwise. SAS Proc Mixed provided estimates of the coefficients of the random effects model. We also used fixed school effects models and the results were nearly identical. Sensitivity analyses were conducted to explore the effects of teacher credentials, and the results were essentially unchanged.

We fit Model 1 separately for grade 2 and 3. Individual student scores are not linkable over time in the STAR data, so growth modeling was not possible. Models of change for cohorts within school were feasible but because we had no hypotheses

on the effects of a year's delay in CSR for growth in the following two years we looked only at the effects within each grade.

Results

CSR Effects on Math, Reading and Language Test Score

There is an upward trend in scores across cohorts K95, K96, and K97 in both Group A and Group B schools (see Figure B1). The top panel of the figure shows the box and whisker plots of the distribution of school mean math scores for the three cohorts of students from Group A schools. The dot corresponds to the median score, the upper and lower sides of the rectangle correspond to the 25th and 75th percentiles of the distribution, and the brackets at the ends of the whiskers correspond to the 5th and 95th percentiles of the distribution of scores. Dots beyond the whiskers are extreme outliers.

There is an obvious upward trend in scores across cohorts over time, as the distribution shifts to the right for each successive cohort. However, in Group A schools, all three cohorts experienced exactly the same pattern of CSR exposure (grades 1 through 3). Thus, the trend in scores is not related to changes in the level of CSR exposure.[2] During the time period that our three study cohorts were in kindergarten through grade 3, California enacted several other statewide education initiatives including the introduction of demanding new curriculum standards, a statewide standardized testing program with high-stakes accountability, and the end of bilingual education. All of these programs might contribute to rising test scores across cohorts, even if differences in CSR have no effect.

The lower panel in Figure B1 shows box and whisker plots for the cohorts in the Group B schools. The plots for Group B show a nearly identical trend to the plots for Group A, even though students in cohort K95 in Group B had one year less exposure to CSR than students in the other two cohorts in Group B. Figures for reading and language scores show similar patterns (see Figures B2 and B3). On the basis of this figure, it seems clear that the additional year of CSR in first grade did not have large effects on mathematics scores.

[2] Although the trend in scores is not related to level of CSR exposure, the size of gains might be sensitive to class size reduction overall. For example, the achievement gains for primary grades were larger than for upper elementary where classes remained large. Small classes might allow teachers to better implement reforms or to respond more quickly to the incentives of the accountability system. However, we do not have adequate data to test for effects between grades; we can only compare differential amounts of CSR among students in the same grades.

Figure B1–
Third Grade SAT-9 Score Distributions in Mathematics for Successive Cohorts of Students with Constant vs. Increasing CSR Exposure

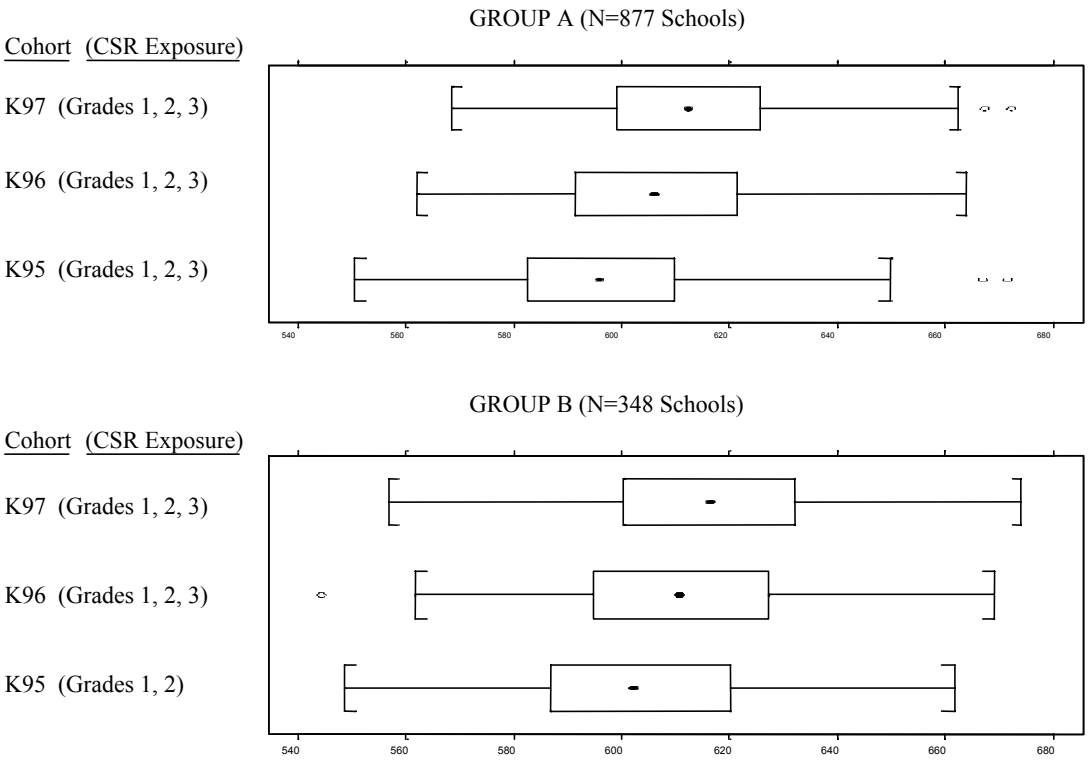


Table B5 provides further evidence that, for students in these cohorts and schools, the effects of an additional year of CSR were small. In Table B5a, the first row presents the differences between mean second-grade math scores for K96 and K95 for Groups A and B, and the difference between these differences. The second row contains the differences between mean second-grade math scores for K97 and K96 for the two groups and the difference between the differences. In the third row we have the difference of these two cohort-to-cohort differences in each Group and between the groups. Tables B5b-4f contain similar differences for grade 3 mathematics scores and for grades 2 and 3 reading and language scores.

The table contains the results of Comparisons 3, 4 and 5 among cohort means by group, grade, and subject. For example, for Group B, the difference in mean scores for K96 and K95 is the difference between a cohort of students that participated in CSR in grades 1, 2 and 3 and a cohort that participated only in grade 2 and 3. Thus, the value of 6.49 from Table B5a represents in part an effect of one additional year of CSR when students were tested in second grade. It also includes other effects occurring during this time. Comparison 3 attempts to remove the time trend in this comparison by using the difference between K97 and K96 in Group B, which is

found in Table B5a to be 8.05. Under the assumptions listed above, the difference between these two values produces an unconfounded estimate of the CSR effect as –1.57 (the last row of Table B5a in the Group B column).

Comparison 4 uses the difference between K96 and K95 in Group A to estimate the natural trend in scores, and adjusts the Group B differences accordingly. This produces an estimate of the CSR effect as 1.15 (the last column in the first row of Table B5a). As noted above, each estimate makes different assumptions about what has remained constant across time or groups. The estimate in the Group B column assumes that changes from cohort-to-cohort in Group B are constant except for CSR. The estimate in the K96 less K95 row assumes that changes from K95 to K96 are constant across Groups A and B except for CSR.

Comparison 5 assumes that, except for the effects of CSR, cohort specific deviations from a linear trend are constant across groups. This difference of differences approach provides an estimate of the CSR effect equal to –0.52. This value is computed as the difference of the values for Groups B and A in the last row of Table B5a. (The estimate is given in the Difference column of the Difference row of Table B4a.)

Table B5a– Second Grade Math

	Group A	Group B	Dif-ference
K96 less K95	5.34	6.49	1.15
K97 less K96	6.39	8.05	1.67
Dif-ference	-1.05	-1.57	-0.52

Table B5b– Third Grade Math

	Group A	Group B	Dif-ference
K96 less K95	6.79	8.17	1.38
K97 less K96	6.53	7.29	0.71
Dif-ference	0.26	0.93	0.67

Table B5c– Second Grade Reading

	Group A	Group B	Dif-ference
K96 less K95	2.05	3.66	1.61
K97 less K96	6.26	6.05	-0.21
Dif-ference	-4.21	-2.39	1.82

Table B5d– Third Grade Reading

	Group A	Group B	Dif-ference
K96 less K95	4.63	4.04	-0.59
K97 less K96	6.23	6.77	0.54
Dif-ference	-1.59	-2.72	-1.13

Table B5e– Second Grade Language

	Group A	Group B	Dif-ference
K96 less K95	2.00	3.35	1.35
K97 less K96	5.25	5.54	0.29
Dif-ference	-3.26	-2.20	1.06

Table B5f– Third Grade Language

	Group A	Group B	Dif-ference
K96 less K95	6.03	5.83	-0.20
K97 less K96	5.78	6.55	0.76
Dif-ference	0.24	-0.71	-0.96

The estimates in Table B5 ignore random school effects that are included in Model 1 to produce efficient estimates and test the null hypothesis that the effect is zero. The results of this model are reported in Table B6, and the full model estimates are included in Table B12. The estimated effects are uniformly small in absolute value ranging from –1.1 to 1.7; these estimates are also small relative to the standard deviation in SAT-9 scores (about 40 scale score points). In addition, the effects across grades are offset--the negative estimate for math in grade 2 is followed by a positive estimate at grade 3, and the positive estimates for reading and language at grade 2 are followed negative estimates at grade 3. Overall, the estimates from Table B5 and Table B6 are very similar and suggest little CSR effect. We also explored school fixed effects models and the results were nearly identical to those in Table B6.

Table B6–
Estimates of 95% Confidence Intervals of CSR Effects from Model 1

	Grade 2	Grade 3
Math	-0.9 (-2.3, 0.5)	0.7 (-0.7, 2.2)
Reading	1.7 (0.3, 3.1)	-1.1 (-2.6, 0.3)
Language	0.9 (-0.4, 2.2)	-0.8 (-2.3, 0.6)

We also conducted some sensitivity analyses to see whether these results were consistent for across student and teacher characteristics. We found similar results when we restricted the analysis to schools with more than 65% minority students, suggesting that the CSR effect was not larger for minority students. (This analysis included about one-half of the schools.) To address the possible bias introduced by the difference between Groups A and B in the change in the percentage of fully-credentialed first grade teachers, we restricted the analysis to schools with no change in the percentage of fully-credentialed teachers during this time period. The results of this analysis were similar, as well. Finally, we ran the analyses with both restrictions, and although the sample of schools was small, we saw no substantial differences in the results.

Caveats

These school-level analyses were less susceptible to confounding from external sources than the statewide analyses presented in our third evaluation report. For

example, we were able to control for student mobility by only including students who attended the same school from kindergarten through second or third grade. Yet, there are still limitations in these analyses. The greatest limitation comes from the lack of variation that existed in exposure to CSR. Our comparisons were limited to a one-year difference in exposure to reduced size classes among students whose total exposure was two or three years. The one-year difference occurred in first grade, and all students subsequently participated in reduced size classes in second and third grade—the points at which their achievement was measured. The Tennessee STAR experiment compared students who attended reduced size classes for four consecutive years with students who attended normal size classes for four consecutive years. They found that at least two years of exposure were needed to produce lasting differences. Those conditions for comparison did not exist in California.

There have also been modest changes in the demographic characteristics of students during this period that might have affected achievement trends. Table B6 shows selected demographic characteristics of California public school students during this time period. There has been a modest increase in the percentage of Hispanic students during this time period, but our differencing approach should have minimized the impact of this gradual change. Yet, our models were simple and did not adjust for demographic or other student background variables. Given the small size of effects and the general similarity of the comparison groups we used a simple analysis rather than complex models. However, small differences among the groups might have affected our results, and more complex models might have removed some of these differences.

Table B6–
Demographic Characteristics of California Students, 1995-2000 (percentages)

School Year	Total Enrollment	Limited English Proficient (LEP)	Race/Ethnicity				
			Asian	Hispanic or Latino	African American	White (not Hispanic)	Other
1995–96	5,467,224	23.6	8.2	38.7	8.8	40.4	3.9
1996–97	5,612,965	24.2	8.2	39.7	8.7	39.5	3.9
1997–98	5,727,303	24.6	8.1	40.5	8.8	38.8	3.9
1998–99	5,844,111	24.6	8.1	41.3	8.7	37.8	4.2
1999–00	5,951,612	24.7	8.0	42.2	8.6	36.9	4.3
2000–01	6,050,895	24.9	8.0	43.2	8.4	35.9	4.5

Note: Starting in 1998–99, all figures include California Youth Authority (CYA) schools. "Other" includes American Indian or Alaskan Native, Pacific Islander, Filipino, and, beginning in 1998, Multiple or No Response.

Source: California Department of Education, Education Demographics Unit.

There have been significant policy and program changes during this period that also affected student achievement. These changes include new state standards and curricula, revised grade-level promotion policies, a new test-based school-level accountability system with large rewards for increases in scores, and the elimination

of traditional bilingual education programs. Because they occurred simultaneously, we used various forms of differencing to disentangle their separate effects and to isolate the unique contribution of CSR to score improvement during this period. However, the differencing requires many assumptions about the equivalence of groups and cohorts in the absence of CSR and the large of number of changes in other programs calls into question the validity of those assumptions.

In addition, there is some reason to doubt the validity of the score gains we used as the basis for these analyses. The California school accountability system has created a high-stakes atmosphere that may lead to changes in test scores that are independent of actual changes in achievement. The gains in SAT-9 scores observed in California are well within the range that might be associated with such score inflation. Again, differencing removes general trends due to score inflation but cannot account for differential inflation.

Another limitation is the restricted sample of the schools and students used in our study. Many schools did not have complete student demographic data, and they were eliminated from our sample. Others had too few valid test scores and were eliminated for this reason. Still other schools were dropped because of indeterminacy in CSR exposure. In addition our analyses focus on students who did not change schools during the K-3 years. The effects of CSR might be different for the schools and students we excluded from our analysis, but we do not have the data to determine the effects of these restrictions on our results. We do not have any good hypotheses about the likely direction of differences between the CSR effects in our sample and those for the entire state.

Finally, the available data do not allow us to judge the impact of the entire CSR program and its effects on students for the last five years. Rather we look for evidence that reduced size classes can make a difference by testing whether additional exposure yields greater achievement. A positive result would be encouraging evidence that small classes are beneficial and that offering them to students in California could have positive effects. Our null finding, however, cannot be interpreted as evidence that the CSR program is not effective. Our results are consistent with at least two possible inferences: a.) reduced size classes have no effect, or b.) two, three or four years of exposure to reduced size classes do have a positive effect compared to no exposure but the difference between two years of exposure and three years of exposure is negligible. One should not make the most pessimistic interpretation of our results (e.g., that reduced size classes have no effect and therefore the entire CSR program is a failure). Rather we should make the most cautious interpretation that, in the context of a K-3 program of reduced size classes, a one-year incremental difference in exposure has no effect. K-3 CSR might have large positive effects on students but differential gains among students with small differences in exposure cannot be used as evidence of those larger effects.

Conclusions

The goal of this investigation was to determine the extent to which changes in achievement correspond to the implementation of the CSR program. The analyses show that scores at the elementary level have been rising at the same time that increasing percentages of students have been taught in reduced size classes. However, many other educational reforms were enacted during this period that might have contributed to the achievement gains, and it is impossible for us to determine how much the various factors may have influenced trends in overall student achievement. Our analyses that used differences in group means to control for the other factors showed that a one-year difference in exposure occurring in first grade is not associated with greater gains in achievement. Due to the rapidity of CSR implementation, we could not test the cumulative effects of two or three years of exposure. Thus while the analyses presented in this chapter find no association between one year's difference in exposure and differences in achievement, we cannot draw any conclusions about the effects of CSR in larger doses.

Supplement to Appendix B

Table B7–
Percentage of Students in Cohort Whose Families Receive AFDC During Four Years, by Group

Group	Cohort	Kindergarten	First Grade	Second Grade	Third Grade
A	K956	24.84	23.71	20.60	19.12
	K967	23.71	22.06	19.12	17.43
	K978	22.06	19.12	17.43	15.62
B	K956	25.36	24.34	22.01	20.06
	K967	24.34	22.99	20.06	17.74
	K978	22.99	20.06	17.74	15.39
C	K956	23.11	22.67	17.82	17.12
	K967	22.67	20.12	17.12	15.00
	K978	20.12	17.12	15.00	12.64
D	K956	21.39	20.97	21.70	17.44
	K967	20.97	19.16	17.44	15.13
	K978	19.16	17.44	15.127	13.16

Table B8–
Percentage of Minority Students in Cohort During Four Years, by Group

Group	Cohort	Kindergarten	First Grade	Second Grade	Third Grade
A	K956	64.79	65.69	63.78	67.55
	K967	65.69	66.57	67.55	68.42
	K978	66.57	67.55	68.42	69.57
B	K956	66.57	67.98	62.68	70.36
	K967	67.98	69.20	70.36	71.57
	K978	69.20	70.36	71.57	72.92
C	K956	55.10	56.45	52.88	58.64
	K967	56.45	57.08	58.64	59.86
	K978	57.08	58.64	59.86	61.21
D	K956	49.11	49.84	56.78	52.05
	K967	49.84	50.80	52.05	53.14
	K978	50.80	52.05	53.14	54.32

Table B9–
Percentage of EL Students in Cohort During Four Years, by Group

Group	Cohort	Kindergarten	First Grade	Second Grade	Third Grade
A	K956	32.51	32.92	33.26	33.37
	K967	32.92	33.33	33.37	33.48
	K978	33.33	33.37	33.48	33.40
B	K956	30.79	31.81	29.59	32.50
	K967	31.81	32.24	32.50	32.86
	K978	32.24	32.50	32.86	32.97
C	K956	24.50	25.45	23.58	26.26
	K967	25.45	26.49	26.26	26.52
	K978	26.49	26.26	26.52	27.15
D	K956	19.00	19.72	24.88	20.35
	K967	19.72	20.50	20.35	20.83
	K978	20.50	20.35	20.83	20.96

Table B10–
Average Years of Teaching Experience for Teachers of Cohort During Four Years, by Group

Group	Cohort	Kindergarten	First Grade	Second Grade	Third Grade
A	K956	16.08	13.24	12.85	12.51
	K967	16.87	11.15	12.61	12.94
	K978	14.58	10.72	12.59	13.45
B	K956	15.54	13.33	13.11	12.18
	K967	16.45	11.25	12.69	12.70
	K978	15.14	10.75	12.85	12.99
C	K956	16.30	13.04	12.38	13.86
	K967	15.58	11.17	12.62	14.44
	K978	13.24	11.53	12.96	14.42
D	K956	16.42	11.99	13.04	13.75
	K967	15.33	11.93	12.98	14.03
	K978	12.88	12.27	13.34	14.24

Table B11–
Percentage of Teachers of Cohort with Full Credentials During Four Years, by Group

Group	Cohort	Kindergarten	First Grade	Second Grade	Third Grade
A	K956	98.06	95.21	87.34	85.21
	K967	96.22	85.56	87.10	87.05
	K978	88.09	85.11	86.65	88.01
B	K956	98.78	95.74	86.56	84.18
	K967	96.26	83.11	86.09	86.28
	K978	88.96	82.29	85.75	88.11
C	K956	98.99	97.31	92.66	90.27
	K967	97.85	92.24	89.47	92.27
	K978	91.98	89.21	91.67	93.27
D	K956	98.45	96.58	93.17	94.05
	K967	97.01	93.63	94.41	94.28
	K978	92.39	94.02	94.19	94.34

Table B12–
Parameter Estimates (Standard Errors) for Model 1

	Grade 2			Grade 3		
	Math	**Reading**	**Language**	**Math**	**Reading**	**Language**
Mean Group A, K95	569.5 (0.9)	571.5 (1)	583.3 (0.9)	603.3 (1)	608.8 (1.1)	607.6 (1)
Difference, Group A K96 less K95	7.3 (0.3)	5 (0.3)	4.4 (0.3)	6.9 (0.3)	4.6 (0.3)	5.8 (0.3)
Difference, Group A K97 less K95	13.2 (0.4)	10.9 (0.4)	9.1 (0.4)	12.3 (0.4)	9.4 (0.3)	10.4 (0.4)
Difference between Groups K95	-8.6 (1.7)	-4.5 (1.8)	-5.4 (1.7)	-5.8 (1.8)	-6.7 (2)	-7.3 (1.8)
Group B Linear Trend	1.9 (0.5)	-0.1 (0.5)	0.5 (0.5)	0.7 (0.5)	0.7 (0.5)	0.8 (0.5)
Effect of Additional Year CSR at Grade 1	-0.9 (0.7)	1.7 (0.7)	0.9 (0.7)	0.7 (0.7)	-1.1 (0.7)	-0.8 (0.7)

Note: The difference parameter estimates of the Difference, Group A K96 less K95 and the Difference, Group A K97 less K95 contain the Group A linear trend and the common (across Groups) cohort deviations from the linear trend.

Figure B2–
Third Grade SAT-9 Score Distributions in Reading for Successive Cohorts of Students with Constant vs. Increasing CSR Exposure

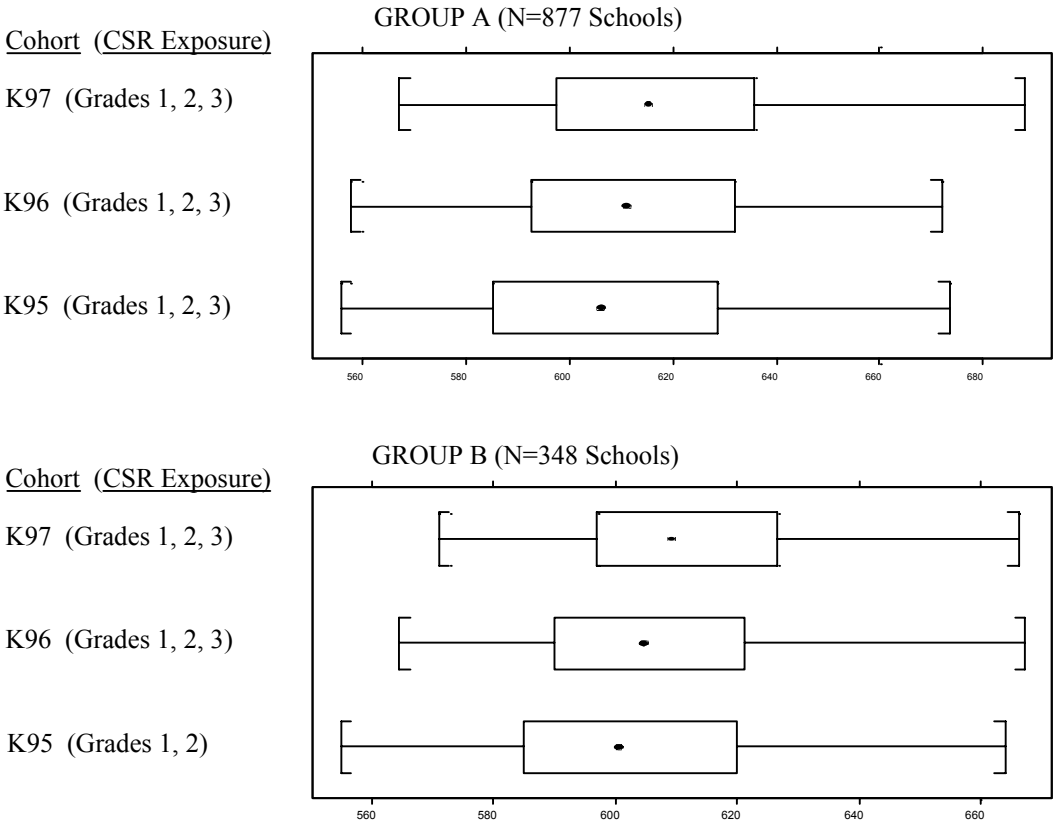
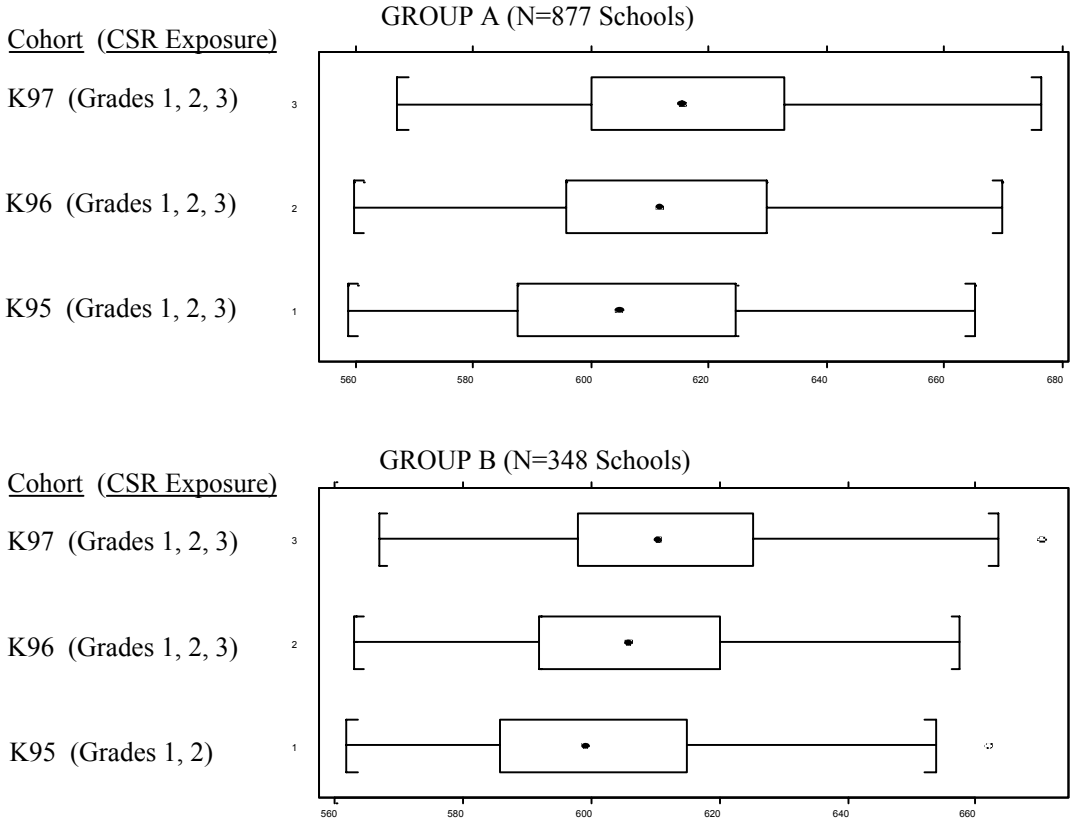


Figure B3–
Third Grade SAT-9 Score Distributions in Language for Successive Cohorts of Students with Constant vs. Increasing CSR Exposure



Appendix C

Class Size Reduction and Teacher Migration 1995–2000

Lawrence P. Gallagher

The advent of CSR in 1996–97 created an unprecedented demand for K–3 teachers. Between the years 1995 and 1999, the size of the California K–3 teaching force increased by 43 percent. This report addresses a central question: where did those new K–3 teachers come from?

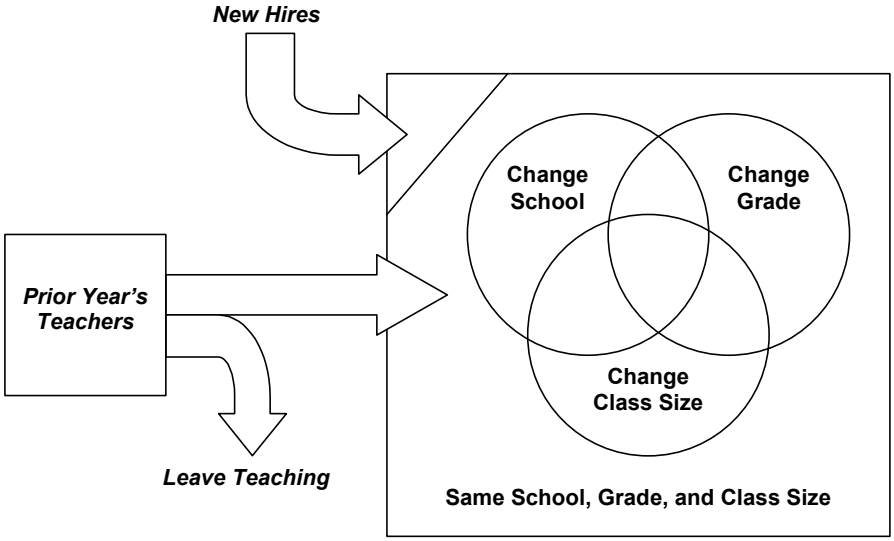
One possibility, of course, is that these new classrooms were filled through hiring of more teachers. Yet, it is not immediately clear that novice teachers were hired directly into vacancies opened up in reduced-size classes. Another possibility is that the vacancies were primarily filled by veteran teachers, while the vacancies created by these migrating veterans were subsequently filled by new hires.

Furthermore, the introduction of CSR was not uniform across the state. In prior reports, we note that higher-SES districts implemented CSR earlier and more fully than their lower-SES counterparts. Perhaps this created a stronger demand in higher-SES districts, drawing teachers away from lower-SES schools with the promise of a smaller class.

The terms of a teacher's labor contract often grant first choice of assignment to the most senior teachers. Even if newly created CSR classrooms were initially staffed by new hires, perhaps their more senior colleagues displaced these teachers in subsequent years, forcing the new hires to teach in non-CSR targeted grade levels.

Figure C1 illustrates the year-to-year possibilities for changes in teaching assignment. Every year some number of new teachers is hired into the profession. Similarly, of last year's teaching cohort, a certain number may leave teaching. Teachers from the prior year who remain in the profession may retain their identical teaching assignment. On the other hand, they may change any combination of three critical factors: their assigned school, their assigned grade, and (especially with the introduction of CSR in 1996–97) the size of their class.

Figure C1–
Teacher Migration Possibilities



The possibility of significant changes in teaching assignment suggests the need for an in-depth analysis of teacher migration patterns over the course of CSR implementation. Using a longitudinal database provided by the CDE, this report tracks teachers over six years of employment. In particular, we address the following questions.

- What percentage of newly created K–3 classrooms are being filled by new hires?
- To what degree are teachers changing schools with the onset of CSR?
- To what degree are teachers changing grade level assignments with the onset of CSR?
- Does teaching in a reduced-size class predict teacher attrition?

For each of these questions, where appropriate, we examine key school, teacher, and classroom characteristics. When cross-school migration is considered, we pay attention to the socio-economic status of that school, as measured by the percent of students receiving free/reduced price lunch. The primary teacher characteristic considered is the number of years of experience, particularly differentiating between novices (0–3 years) and experienced teachers (4 or more years.) The classroom characteristic most salient to our analysis is, of course, class size.

As with other aspects of our report, it is difficult to make a causal claim in a non-experimental situation. Because we have teacher employment data going back to 1994—two years prior to the implementation of CSR—we can examine sudden changes in teacher employment trends coincident with the onset of CSR at particular grade levels. Of course, we still cannot rule out the possible effects of other changes that occurred about the same time that have nothing to do with CSR.

Methods and Data Sources

The primary data source is based on the California state Professional Assignment Information Form (PAIF). The California Department of Education (CDE) compiled a longitudinal database of teacher assignments spanning six years. Included in these records are demographic and credential characteristics of the teachers, their length of service, and of particular interest, their current school and grade level assignments. Information about particular schools was linked from a variety of state and federal databases, most notably the Common Core of Data (CCD). This includes information on ethnic composition and school lunch eligibility for each school and grade.

The CSR legislation specifically targets the K-3 grade levels. However, it has been hypothesized that grades 4 and 5 could be inadvertently impacted by CSR, in that some grade 4 and 5 teachers may desire to move into a reduced-size K-3 classroom. We therefore focus our data analysis on the K-5 grade level teachers. It turns out that migration from grades 6 through 12, as well as other positions (e.g., administration, special education, etc.), is negligible and does not vary significantly with the onset of CSR – analyses of these grade levels and positions are not included in the body of this report[1].

The linked PAIF file contained a significant percentage of records without linking identifiers. Because these teachers could not be tracked over time, they have been dropped from subsequent analysis. Still other records had duplicate identifiers — these were cross-checked with the sex and birth year fields of the records. Records with the same identifier but different sex and/or birth year fields across time were dropped. Finally, there was a particular problem with data from the Los Angeles Unified School District. A significant proportion of the grade assignments was miscoded. After follow-up inquiries to LAUSD, we decided to drop all LAUSD teacher records from this data set.

Table C1 shows the number of records in the original data (excluding LAUSD), the number of records after filtering, and the overall percentage of records dropped. Of particular note is the increase of 12 percentage points between the 1996–97 and 1997–98 school years. Beginning in 1997–98, we are forced to drop at least 25 percent of the teacher records.[2].

Table C1–
Number of K–5 PAIF records used in analysis after filtering

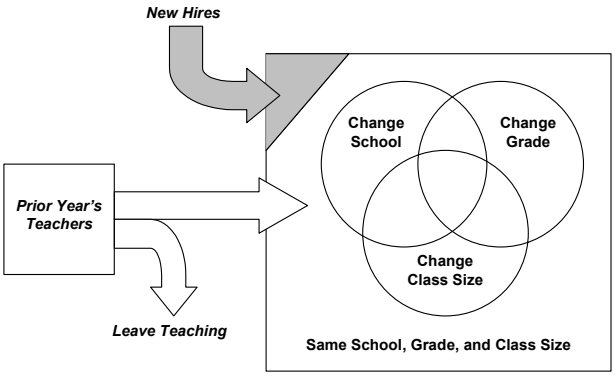
	1995–96	1996–97	1997–98	1998–99	1999–2000
Original K–5 PAIF records	77,254	88,309	102,293	107,596	109,958
Records after filtering	65,722	76,585	76,609	78,589	78,539
Percent records dropped	15%	13%	25%	27%	29%

[1] Grades 6-12, special education, and administrative positions were collapsed into a category of "Other" in the detailed migration tables presented in the supplement to this report.

[2] A separate analysis of the records with missing identifiers suggests that this population did not differ significantly from the overall population, but that analysis was not included in this report.

Using the filtered data described above, simple descriptive statistics were used to examine each of the research questions. In particular we report percentages of various subpopulations of teachers who migrate between schools and/or grade levels. Detailed tables showing the absolute frequencies of grade assignments and cross-grade movement are shown in a supplement to this report.

New Hires



We have documented in the main report the relative increase in novice teachers with the onset of CSR. In prior reports, we define novice teachers as those with three or fewer years of experience. Of particular interest are those teachers newly hired into the profession in a given year. Here we consider teachers who report having one year or less of experience.[3]

Figure C2 shows the percentage of new hires at each K–5 grade level and year. While all grade levels experience a sharp increase in new hiring with the onset of CSR, grades 1 and 2 show the most dramatic change. After the first year of CSR implementation, the new hiring in grades 1 and 2 begins to taper off, declining to pre-CSR levels by the 1999–2000 school year. Similarly, grades K and 3 experience hiring growth for the first two years of CSR implementation, and show a similar tendency to taper off to pre-CSR levels by the last year studied. This pattern is to be expected, given that the law authorizing CSR gave funding priority to grades 1 and 2 in the first year of implementation

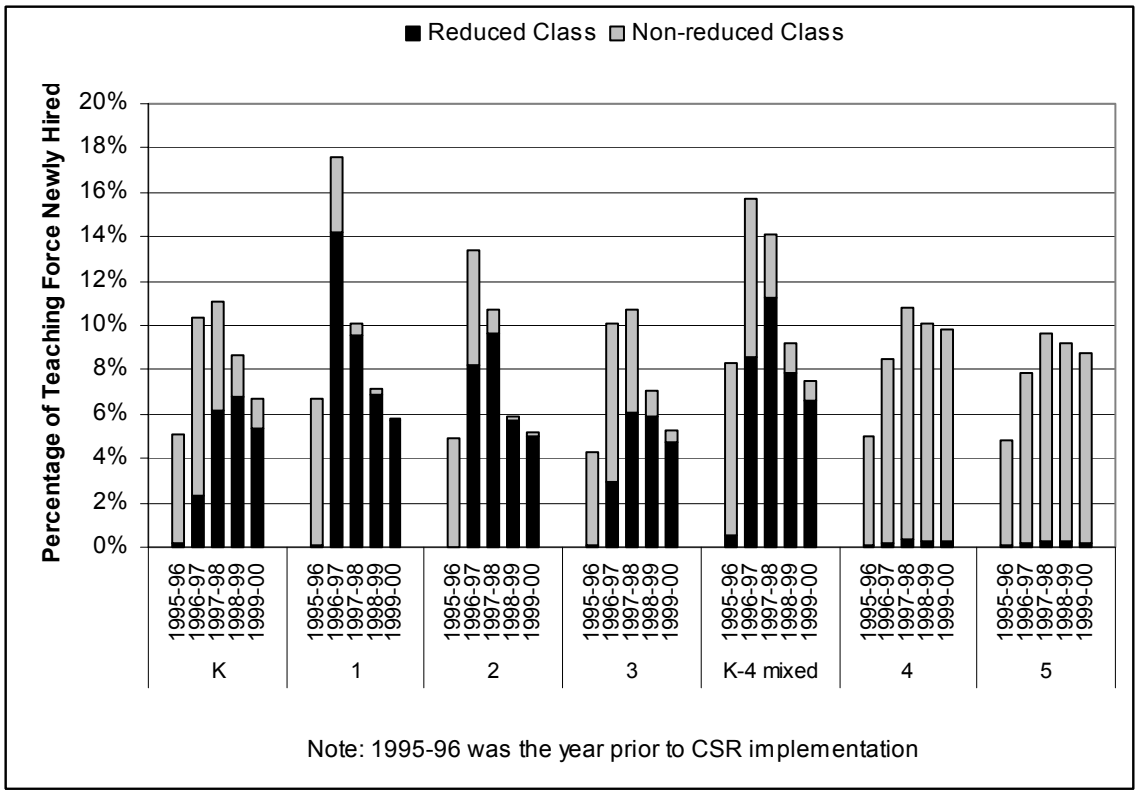
The K–4 mixed grade category[4] mirrors the grade 1 and 2 pattern of a sharp increase followed by decline to pre-CSR levels. It is possible that many of these new hires are filling mixed grade 1 and 2 classes (thereby eligible for CSR remuneration), but we lack information about the specific grade levels combined in mixed-grade classrooms.

[3] The PAIF survey form is administered every autumn. The instructions for the years of experience question instruct the teacher to fill out a "1" if this is their first year teaching. However, the wording of the question can easily be misconstrued, and there were a significant number of teachers responding "0" for years of prior experience. In this section we interpret these respondents as new hires.

[4] The CDE data source designates teachers as teaching in a K-3 mixed-grade class, as well as a mixed grade 3 and 4 class. Because the latter category includes third-grade students, these categories were combined into a single K-4 mixed group.

Grades 4 and 5 are the most puzzling: they increase their rate of new hires over the first two years of CSR implementation, but continue to sustain that rate beyond the first two years. New hiring rates taper off only slightly for the years following CSR implementation. It may be that these new hires are filling vacancies as teachers move from grades 4 and 5 to K–3 — we examine that hypothesis in a later section of this appendix.

Figure C2–
Percentage of teachers newly hired, by grade assignment and class size



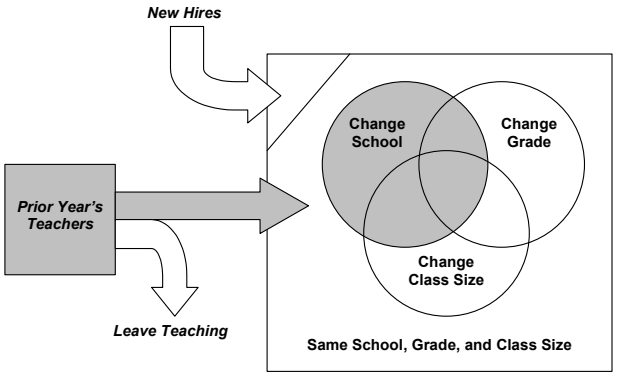
Each bar in Figure C2 is divided into two sections, indicating the class size assigned to the new hires. It is clear that once CSR is implemented in a grade level, the majority of new hires are being placed in reduced-size classes. That is, in this chart we see little evidence that experienced teachers are vacating their large classes to take advantage of a CSR classroom, while their large-class vacancies are being filled by new hires. In a later section we corroborate this finding with an examination of veteran teacher movements.

The number of new hires by year and grade assignment is depicted in Table C2.

Table C2–
Number of new K–5 hires by year and grade assignment

	1995–96		1996–97		1997–98		1998–99		1999–00	
Assignment	Reduced	Non-reduced	Reduced	Non-reduced	Reduced	Non-reduced	Reduced	Non-reduced	Reduced	Non-reduced
K	19	528	269	923	667	530	826	228	666	164
1	12	644	2,008	469	1,261	80	940	31	766	15
2	4	465	939	591	1,222	143	763	27	670	22
3	9	382	298	718	649	494	768	149	634	69
K–4 mixed	39	568	580	486	668	171	447	76	357	51
4	8	412	15	703	24	769	24	758	22	781
5	8	384	11	629	16	644	21	643	13	637
Total	99	3,383	4,120	4,519	4,507	2,831	3,789	1,912	3,128	1,739

Cross-School Migration



In any single year of this study, fewer than 12 percent of teachers in a grade level had migrated from another school. In most cases, this rate is less than 8 percent. Figure C3 shows the average school migration rate by grade and year.

Across all grades K through 5 there is an overall increase in teacher mobility with the onset of CSR. As expected given the implementation schedule of CSR legislation, the cross-school migration rates rise most quickly for grades 1 and 2, followed by a more delayed rise in grades K and 3. This suggests that the rapid expansion of classes due to CSR may have had an impact on cross-school migration. By the last year studied, we see the school migration rate drop below that of the pre-CSR years.

Figure C3–
Percentage of teachers changing schools in a given grade level and year

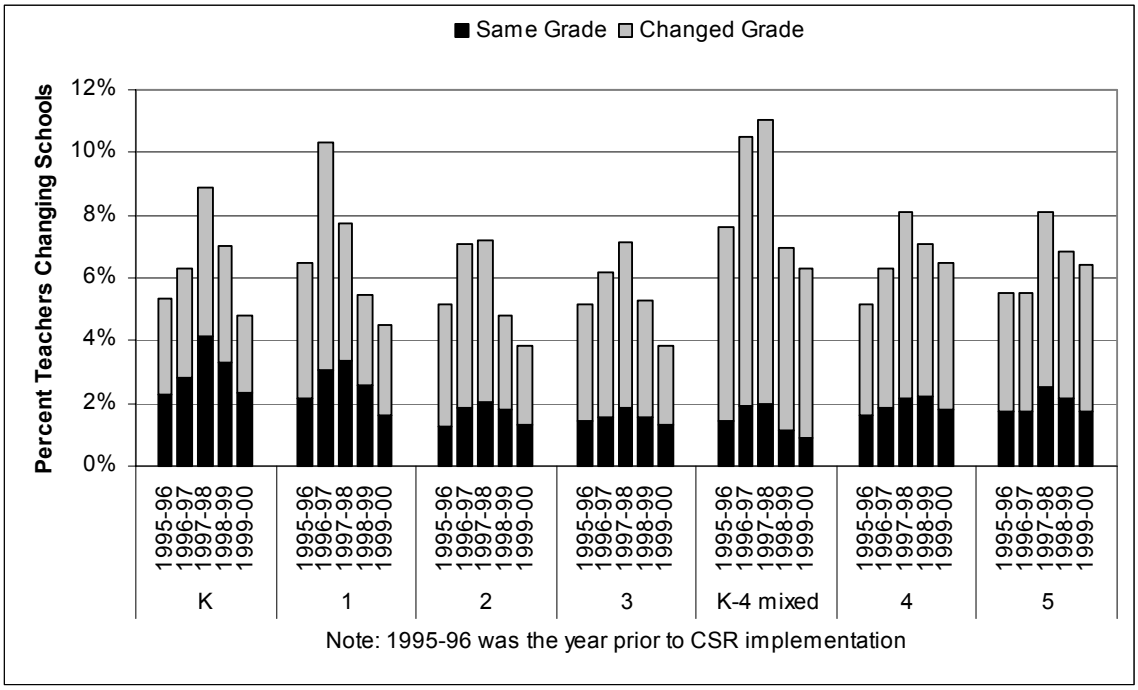


Table C3–
Number of K–5 teachers changing schools in a given grade level and year

Grade assignment	1995–96		1996–97		1997–98		1998–99		1999–00	
	School Movers	School Stayers	School Movers	School Stayers	School Movers	School Stayers	School Movers	School Stayers	School Movers	School Stayers
K	544	10,712	650	11,772	854	11,408	777	12,387	558	12,769
1	595	9,723	1,199	14,408	919	14,118	687	13,951	570	13,720
2	463	9,379	702	11,788	819	13,670	606	13,740	485	13,631
3	451	8,992	561	10,293	678	11,316	636	13,303	487	13,654
K–4 mixed	512	7,203	599	7,015	563	6,276	361	5,904	314	5,528
4	411	8,304	486	8,534	530	7,752	495	7,908	480	8,331
5	427	8,006	411	8,167	500	7,206	447	7,387	436	7,576

The highest rates of school migration occur for the K–4 mixed grade teachers. It is possible that mixed-grade classes are being created from students who otherwise would have tipped a CSR classroom over the average daily attendance (ADA) of 20.4. For example, if a school has a first and second grade classroom of 30 students each, they could split each class in two, resulting in two classes at each grade level of 15 students. This requires the hiring of two additional teachers. An alternative would be to create one

combined first and second grade classroom, and assign 10 students from each grade to this mixed grade class. In this case, only one additional teacher need be hired.[5]

There also was a noticeable increase in the percentage of grade 4 and 5 teachers changing schools. These grade levels are not directly targeted by the CSR legislation — it remains to be seen why we observe an increase in cross-school teacher mobility after the onset of CSR. This increase in teacher mobility mirrors the increase in new hiring observed for grades 4 and 5 in the above section.

In all years an grade levels the majority of teachers changing schools also are changing grades. This holds true whether the teachers are changing into a CSR-targeted grade level or not. In particular, teachers changing schools into K–4 mixed classes also are much more likely to have come from a single-grade classroom.

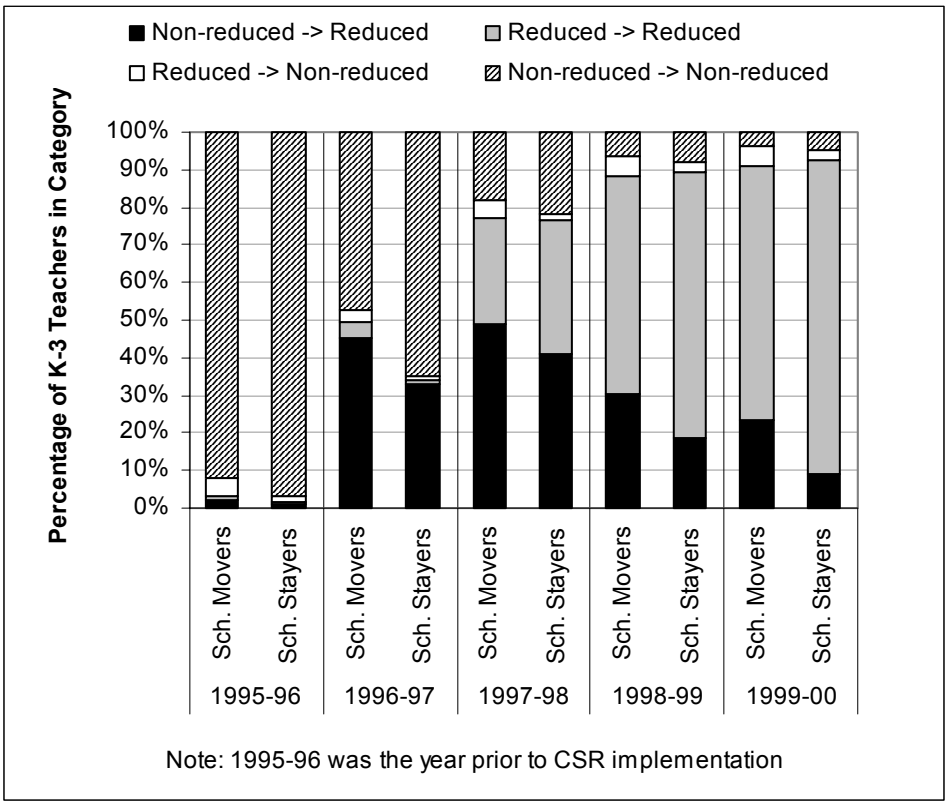
Cross-School Teacher Migration and Class Size

The rise in cross-school teacher migration rates corresponding to the onset of CSR suggests that class size may have been a motivating factor in teacher movement. In this section we examine the correspondence of cross-school migration to change in class size.

A teacher may change schools and/or class sizes through two routes. The most direct route is when a teacher changes schools in order to fill an opening in a reduced-size class. Another possibility, though, might be the displacement of junior teachers by more senior teachers coveting a CSR-eligible classroom. We examine both hypotheses in this section.

[5] According to CDE, mixed-grade classrooms are eligible for CSR funding. See http://www.cde.ca.gov/classsize/facts.htm

Figure C4–
Proportion of K–3 teachers changing class sizes, by cross-school migration status



We show the four possible class size transitions (remain large, small to large, remain small, and large to small) by school migration status and year in Figure C4. This allows one to directly compare the trends in class size change by school change.

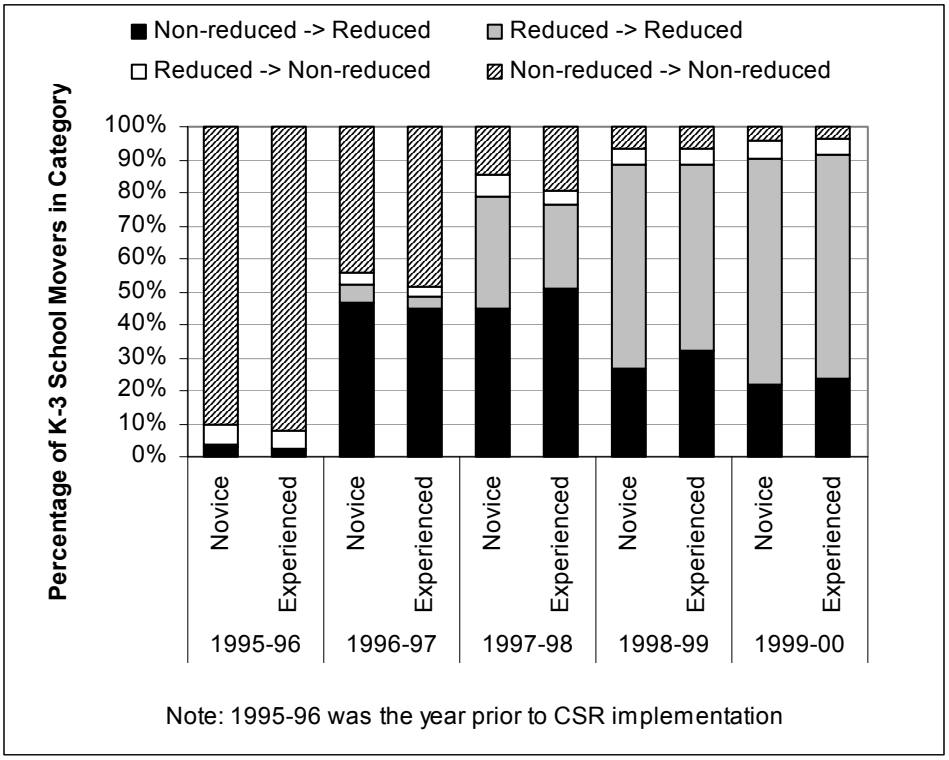
Beginning in 1996–97 we see that school movers are indeed more likely to change from a large to a small class than are their school-staying peers. As CSR becomes more fully implemented over the years, the overall percentage of teachers moving from large to small classes drops for both groups—there simply are fewer large classes in the K–3 grade levels to be moving out of.

Note that from 1997–98 onward, the percentage of teachers in small classes (the combination of transitions from large to small, as well as those remaining small) is essentially equal for both school movers and school stayers. This can be seen by comparing the heights of the first two stacked bar segments.

Are the class-size changing patterns significantly different for novice and experienced teachers? Surprisingly, perhaps, the answer is no. In Figure C5 we show the distribution of class size changes for school movers only, broken out by two levels of teacher experience. In every year of CSR, novices are slightly more likely to remain in small

classes, while experienced teachers are slightly more likely to move from a large to small class. In the end, however, combining both the "remaining small" and "large to small" categories (i.e., those teachers who are teaching in a small class in a given year), experience seems to bear little relationship to the proportion of school movers in small classes.

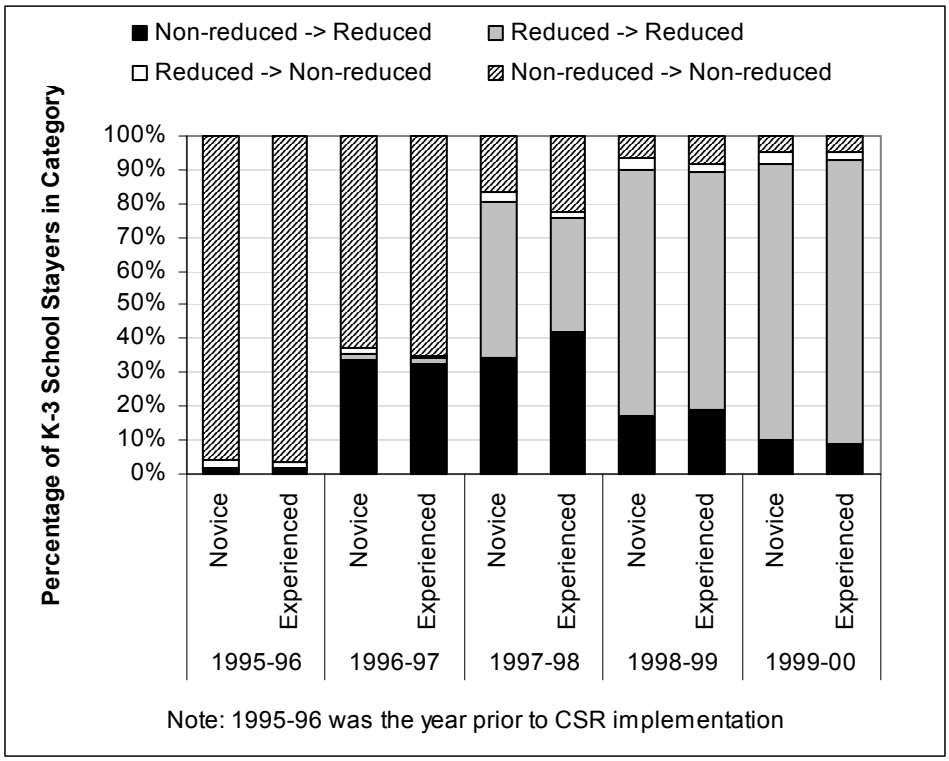
Figure C5–
School Movers change in class size, by experience level



We have already established that no more than 12 percent of teachers change schools in a given year. For these school movers, there is no evidence that experienced teachers are acquiring smaller classes any more rapidly than novice teachers.

However, in most years more than 90 percent of the teaching force remains at the same school. We should examine these school stayers for evidence of displacement of junior faculty by more senior teachers. In Figure C6 we see that, in fact, school stayers exhibit similar change patterns to school movers. There is very little difference in the proportion of teachers ending up in small classes, when level of experience is examined. Novice teachers do exhibit a slightly greater tendency to move from large to small classes than do their more experienced peers, although the experienced teachers are more likely to have already been in a small class and remain in such.

Figure C6–
School Stayers change in class size, by experience level



In particular, in neither case do we see any significant difference in the proportion of teachers moving from a small class to a large class. If there were significant displacement of junior K–3 teachers by senior teachers, we would expect to see a greater percentage of novice teachers moving from small to large classes. We see very little evidence of such displacement — movement from small to large classes is a relatively rare event for teachers of all experience levels.

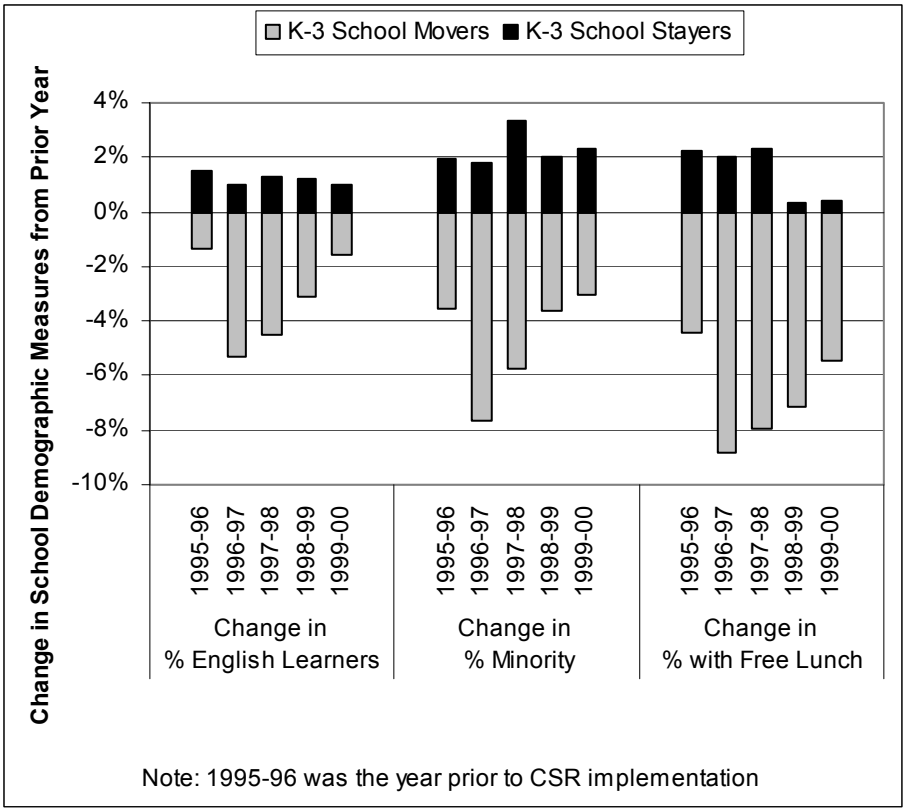
We have evidence, then, that school movement might be somewhat influenced by the desire to teach in a reduced-size class. There seems to be little support for the theory that more junior teachers are being displaced out of small classes. For any given year and experience category, the proportion of teachers moving from a small class to a large class is never greater than 4 percent

Cross-School Teacher Migration and School Characteristics

We would like to know more about why teachers are changing schools. As shown in the main report, wealthier districts implemented CSR earlier and more fully. It is widely argued that veteran teachers from lower SES districts are filling these new teaching positions. We examine this hypothesis by comparing the characteristics of schools connected to teacher migration.

When teachers change schools, the overall demographic composition of their classes is likely to change. We examine three school-level demographic variables: the percentage of English Language Learners, the percentage of ethnic minority students, and the percentage of students with free/reduced price lunch. Figure C7 shows the change in demographic measures from the prior year for both school movers and stayers. For school movers (shown in light gray), this quantity represents the difference between their new school and their prior school assignment. For those who stay in the same school (shown in black), this figure represents the normal year-to-year demographic change at the school site.

Figure C7–
School demographic changes for both school movers and school stayers



The results in Figure C7 corroborate the hypothesis that school movers are generally moving to schools with a different demographic makeup. School movers tend to migrate to schools with lower percentages of English learners, minorities, and students with free/reduced lunch. Those who remain at the same school experience a year to year upward shift in each of these measures, corresponding to the changing makeup of California elementary school students over a five-year span.

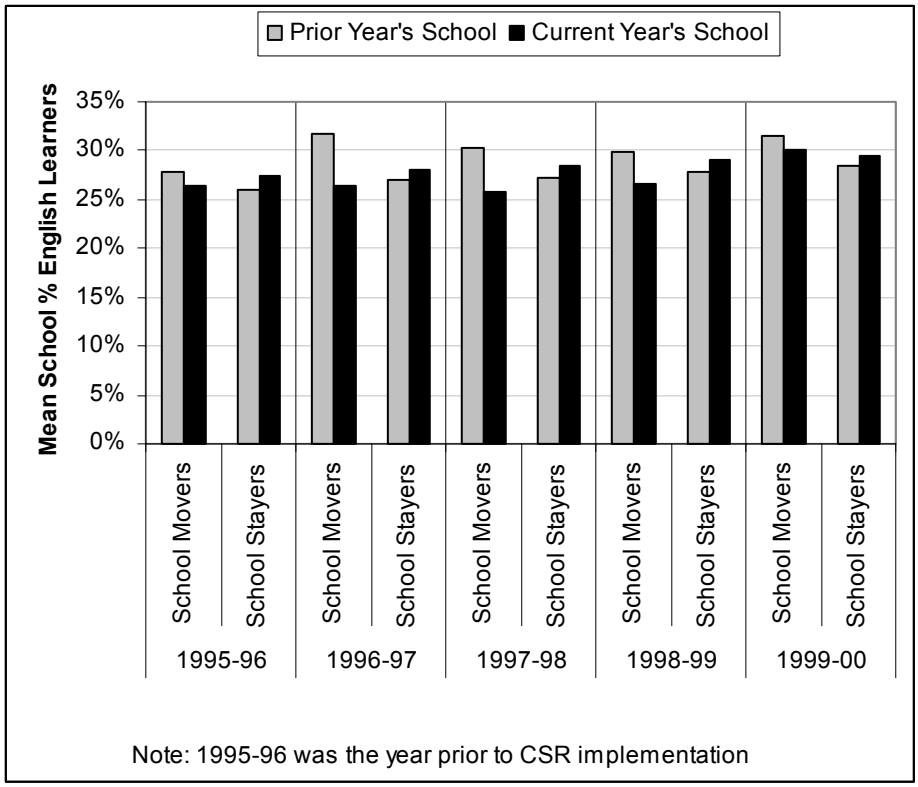
While the data in Figure C7 indicate that teachers who change schools generally move to a higher SES situation, of interest to us is the impact that CSR may have had on this phenomenon. While school movers on average changed to a higher SES school in the

year prior to CSR, with the onset of CSR in 1996–97 we see that the demographic shift experienced by school movers increased markedly. This trend continues for the next two years. By the last year studied, we see a return to pre-CSR levels in two out of three measures.

It appears, then, that not only was there an increase in teacher mobility with the onset of CSR (see Figure C3), but also that these migrating teachers had greater opportunity to move to higher SES schools. However, one must put these numbers in perspective. Figure C7 only shows the changes in demographic characteristics between prior and current schools for school movers—it says nothing about the absolute level of the SES measures. We cannot tell, for instance, whether school movers tend to come from much lower SES schools than do the school stayers.

In Figure C8 we show the mean school percentage of English Language Learners for both school movers and school stayers. In each year, we show the percentage of English Language Learners for the prior year and the current year. For school movers, the prior year's value represents the percentage of English Language Learners at their old school. For school stayers, the prior year's value represents the level at their current school.

Figure C8–
Mean school percentage English learners for both school movers and school stayers



In any given year, school movers came from schools with a higher percentage of English learners than did their school staying peers. This can be seen by comparing the gray bars (prior year's school level) within a single year. Similarly, school movers moved to schools with a slightly lower percentage of English learners than school stayers, which can be seen by comparing the black bars within a given year. Overall, there is no strong change in level either for prior schools or current schools coinciding with the onset of CSR. Similar trends are found for the percent of minority students (Figure C9) and percent of students with free or reduced lunch (Figure C10).

Figure C9–
Mean school percentage minority for both school movers and school stayers

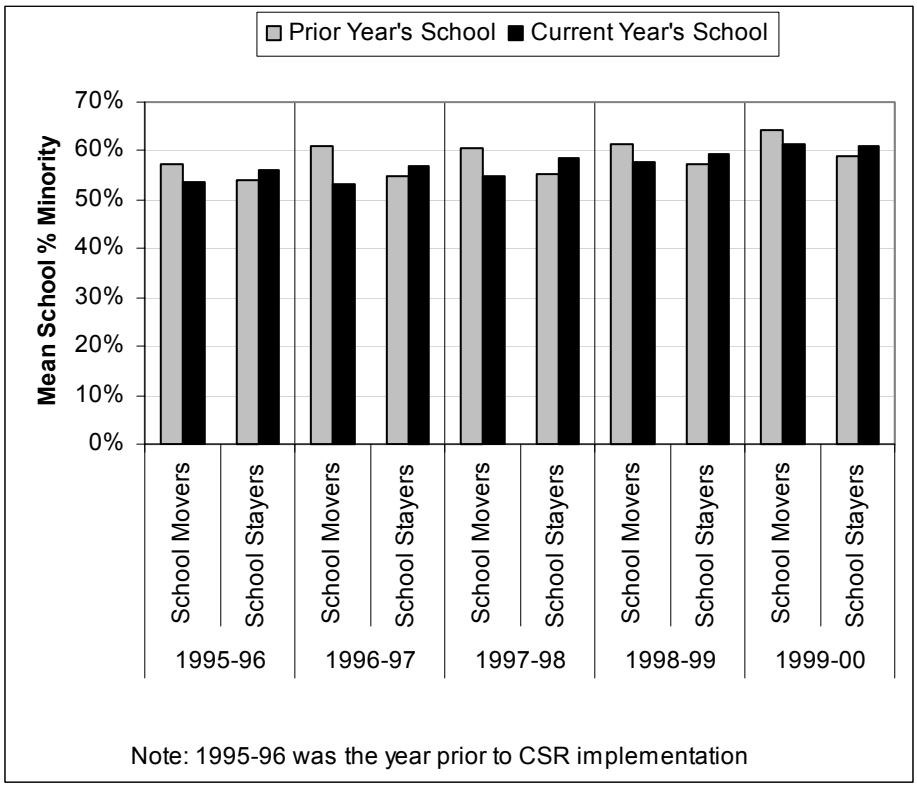
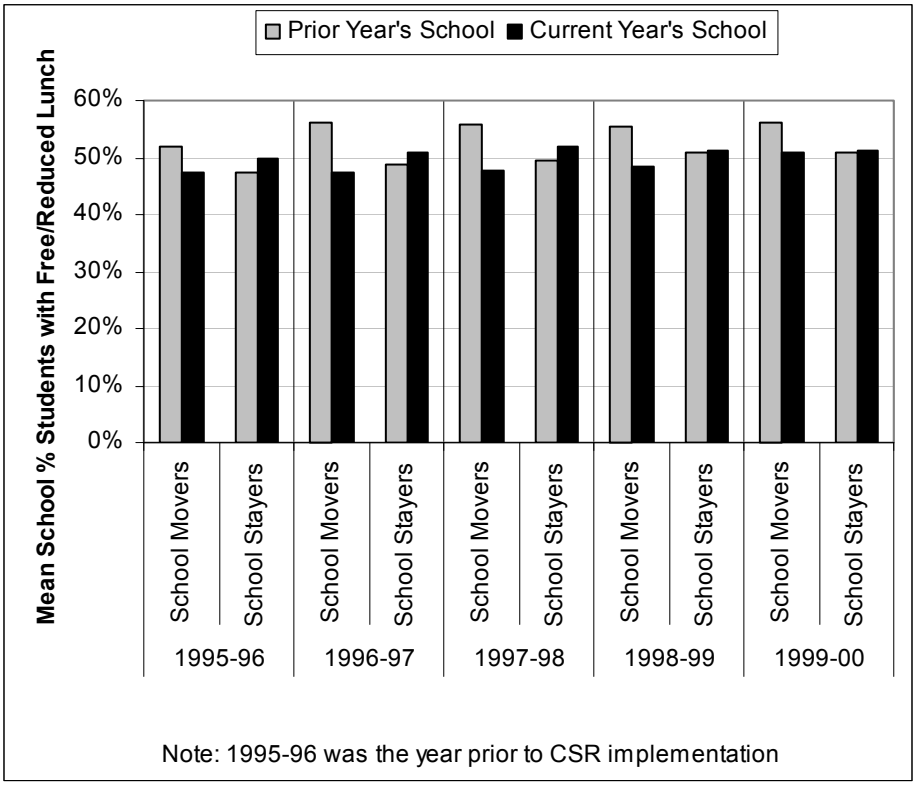


Figure C10–
Mean school percentage students with free/reduced lunch for both school movers and school stayers



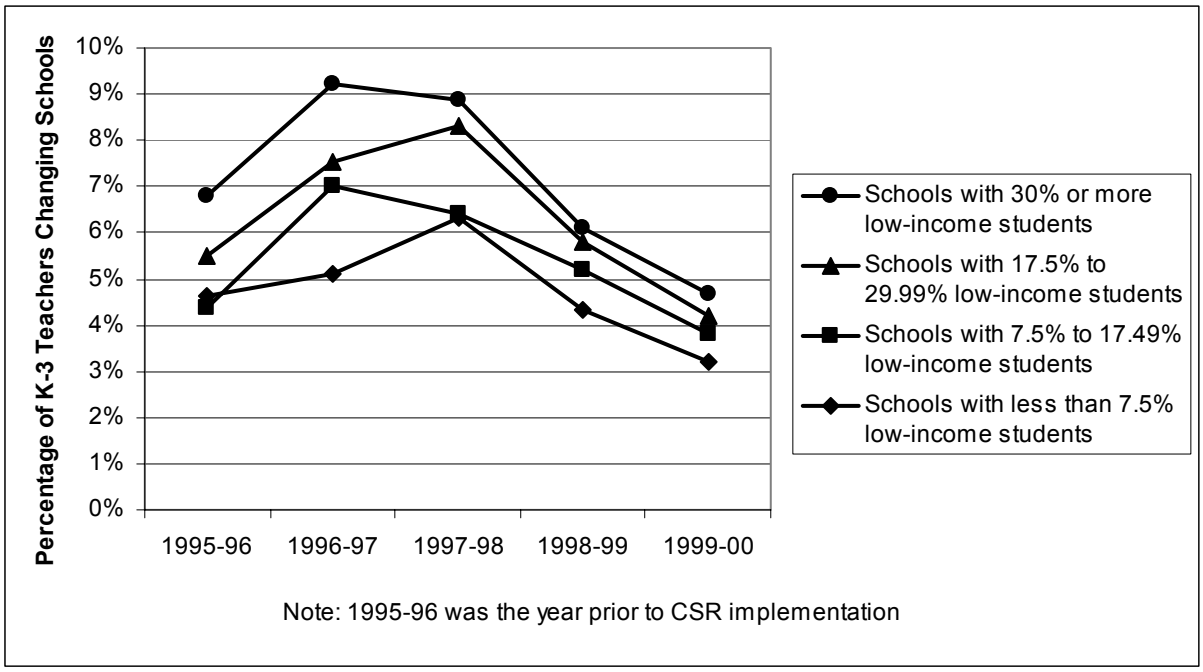
Overall, there are detectable differences in teacher migration rates when schools are broken out into SES quartiles.[6] Figure C11 shows the school migration rates by school SES quartile. In general, the rates of migration track the SES quartiles: schools with higher rates of free/reduced lunch eligibility show higher rates of teacher mobility. During the first two years of CSR implementation, however, there is considerable variation in teacher migration rates that is not necessarily proportional to the SES quartile of the school.

In the first year of CSR (1996–97), all schools shows a rise in teacher migration, although the schools with the lowest percentage of low-income students only show a slight increase when compared to the other three quartiles. It is during the second year of CSR that we see marked differences among the quartiles. The quartile with fewest low-income students continues to see an increase in migration, while the next highest quartile (7.5%–17.5% low-income students) shows a decrease — the two quartiles converge to similar rates. Likewise, the quartile depicting 17.5 percent to 30 percent low-income students shows an increase in migration, while the quartile with the most low-income students shows a slight decrease.

[6] These are the same SES quartiles—based on the percent of students on free/reduced lunch—used in prior CSR reports. Schools are categorized according to their 1996–97 SES level.

By the year 1998–99 all SES quartiles are migrating less frequently than during the first two years of CSR, and have converged to more similar values. These quartiles continue to decrease their migration rates in parallel through the last year tracked (1999–00).

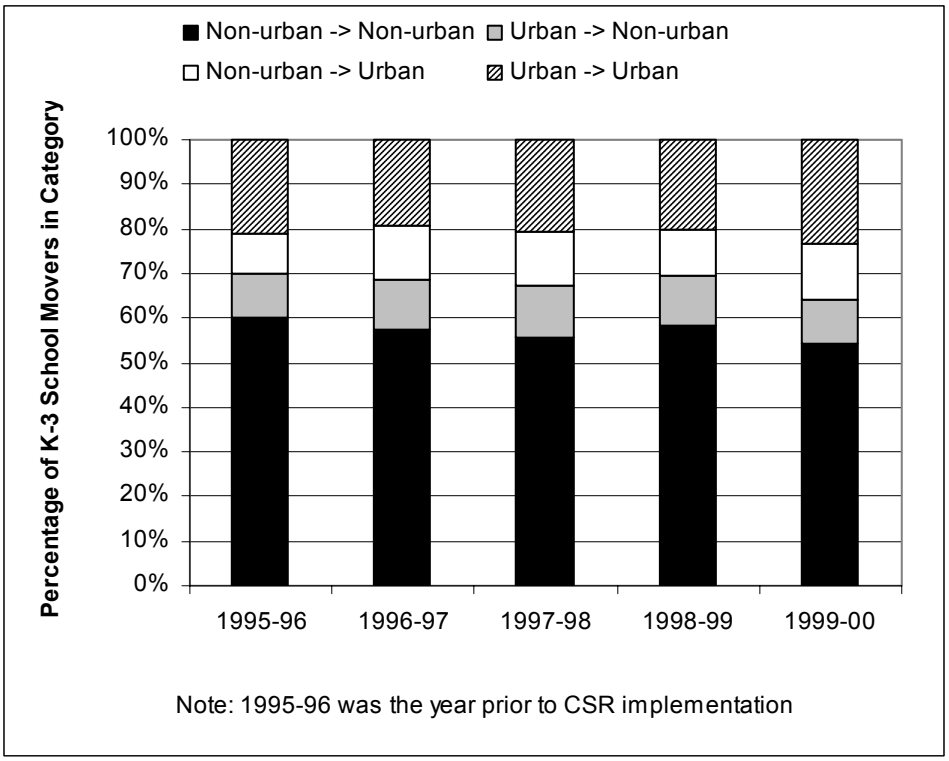
Figure C11–
Percentage of teachers changing schools, by SES quartile of prior school assignment



In addition to the SES of a school, policymakers have raised concerns about "urban flight"—teachers moving from urban schools to suburban schools as CSR is introduced. However, we find virtually no differences in teacher migration rates between urban and non-urban[7] districts. The data shown in Figure C12 indicate that for all K–3 school movers, the proportion moving between urban and non-urban schools (as well as remaining in schools of the same type) holds remarkably constant over the years studied. There is a consistent, minor difference in teacher mobility between novice and experienced teachers, with novice teachers slightly more likely to switch urban status in their move (again attesting to the overall increased mobility of younger teachers.) Significantly, however, we do not see an increase in teacher mobility from urban to non-urban schools with the onset of CSR, regardless of teacher experience. Rather, as shown above in Figure C7, specific demographic markers are a better predictor of cross-school migration.

[7] In this section we collapse rural and suburban districts into the single category of "non-urban."

Figure C12–
Distribution of urban migration patterns among K–3 school movers

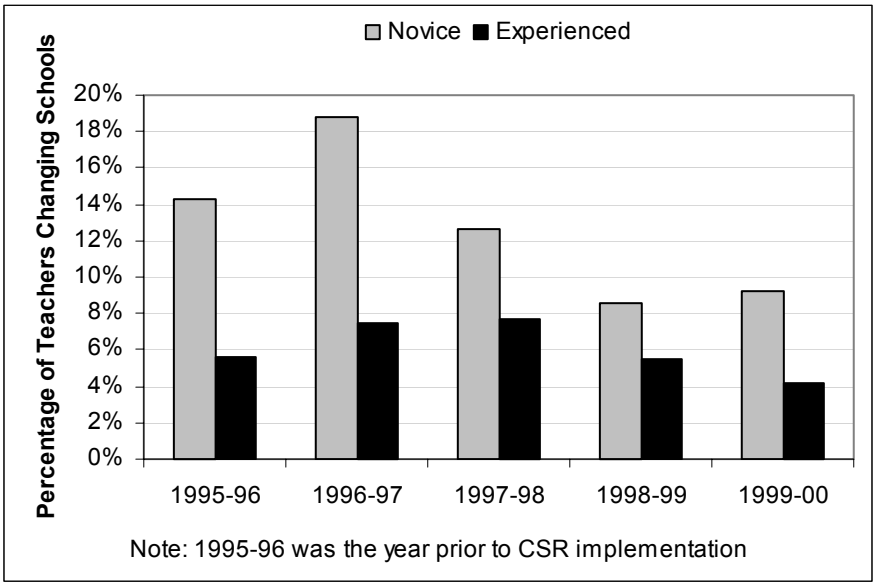


Cross-School Migration and Teacher Characteristics

Historically, novice teachers have always been more mobile than their more experienced peers. Novice teachers tend to be younger and more socially mobile. In this section we examine the frequency of school migration by the experience level of the teachers. As in prior CSR reports, we divide teachers into two groups: Novice teachers (those with 0–3 years of experience), and Experienced teachers (with 4 or more years of experience.)

As shown in Figure C13, novice teachers are more mobile than their more experienced peers, but the difference in teacher mobility rates varies with the year under study. In the year prior to CSR, novice teachers were approximately 2.5 times as likely as their experienced colleagues to change schools. While the migration rates for both groups increased for the first year of CSR, the ratio of the two rates remains constant—novices continue to migrate at a rate 2.5 times greater than experienced teachers. It is by the second year of CSR that we see a marked difference between the two groups—the migration rate for novices declines from 19 percent to 13 percent, while the rate for experienced teachers climbs slightly from 7 percent to 8 percent. There is a uniform decline in migration rates in the following year (1998–99),. During the last year under study, there is a somewhat puzzling increase in novice migration, while experienced teacher migration continues to decrease.

Figure C13–
Percentage of teachers changing schools by experience level



Conclusions — Cross-School Migration

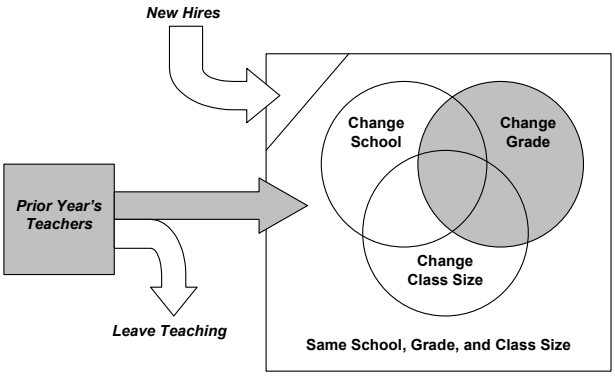
While there is evidence of an increase in cross-school migration with the onset of CSR, this migration accounts for a relatively small proportion of teachers — less than 8 percent in most cases. Furthermore, these school movers were not dramatically more likely to be moving from large to small classes than their school-staying peers.

In general, the onset of CSR coincided with an increase in trends that already were present. Teachers who changed schools had always on average moved to higher-SES schools — this trend increased with the onset of CSR. While novice teachers have always been more mobile than more experienced veterans, this discrepancy increased when CSR was started. Before CSR there had been detectable differences in teacher mobility when schools were grouped by SES quartile — these differences were exaggerated during the first two years of CSR, and returned to their pre-CSR levels shortly thereafter.

The one migration change seemingly unaffected by CSR was the relative rate of migration between urban and non-urban boundaries. These migration rates held relatively constant over the five years studied in this report.

It appears that the spike in demand for new teachers "stirred the pot" more vigorously, but there was little qualitative change from previous patterns. The migration trends that had previously existed were temporarily excited, but not fundamentally changed.

Cross-Grade Migration



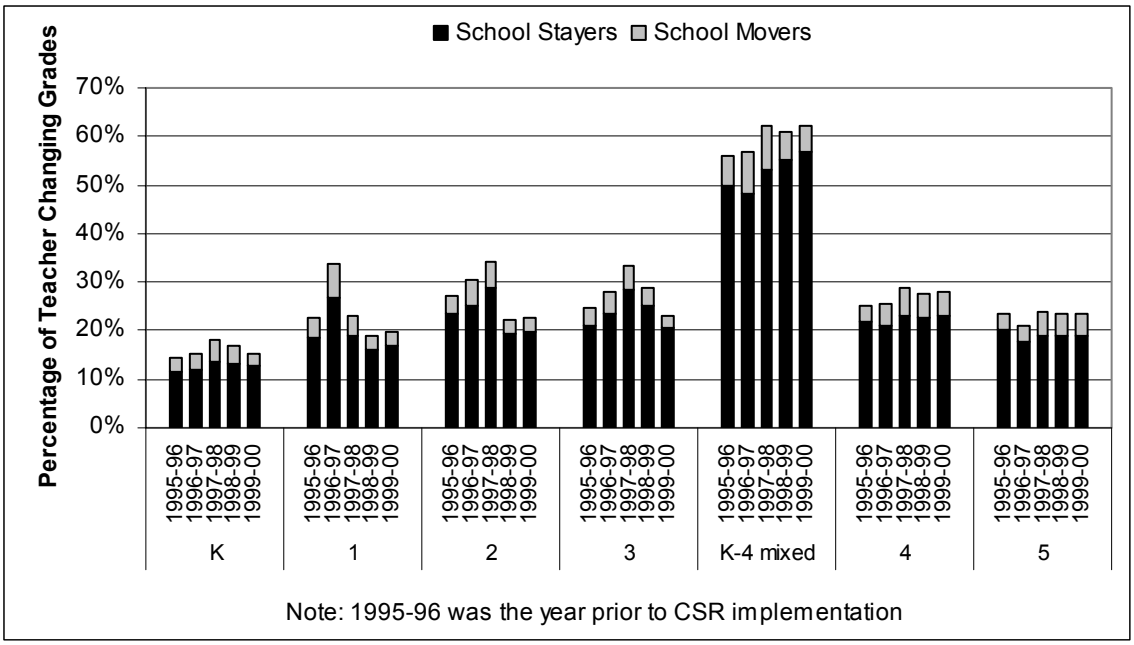
The Class Size Reduction act applies to specific grade levels. Schools applying for CSR funds were expected to fully reduce their first grade classes, after which they could use funds for second grade classes. Once first and second grade class sizes were completely reduced, schools could elect to reduce kindergarten and third grade class sizes as well.

This section examines teacher migration patterns between grade level assignments. The previous section found that with the onset of CSR, cross-school migration dynamics intensified for two years and settled back into old patterns. In this section, we apply a similar analysis to the movement between grade assignments. One would expect this migration to be more sensitive to the impact of CSR, because CSR explicitly targets particular grade levels at particular times.

Figure C14 shows the distribution of grade-moving K–5 teachers by both year and prior job assignment. Each bar in Figure C14 represents the percentage of the teaching force at a particular grade level and year that was not teaching at the same grade level the previous year. We also show the relative proportion of teachers who changed both grade and school by dividing the bar into two portions: the black section represents those teachers changing grades within the same school, while the gray section represents those changing both grade and school.

In the year prior to CSR implementation, there already was a considerable amount of grade changing in the teaching profession. The percentage of teachers coming from another grade assignment ranged from a low of 15 percent (kindergarten) to a high of 56 percent (K–4 mixed grades). This alone is worth noting—the mixed-grade classrooms tend to be highly volatile. More than half of the mixed-grade teachers in any given year were previously teaching a single-grade classroom. Not surprisingly, grades 4 and 5 show a relatively constant rate of migration from other grades.

Figure C14–
Percentage of teachers migrating from other grades, by year and current assignment



Beginning with 1996–97, we observe a significant increase in the percentage of first grade teachers migrating from other grade levels, climbing from a base rate of 23 percent in 1995–96 to 34 percent in 1996–97. However, by the second year of CSR implementation, that rate is reduced back to 23 percent, and to even slightly lower rates in following years.

Grades K, 2 and 3 show a more gradual increase in movement over the first two years of CSR implementation. This is not surprising because schools must have fully implemented CSR in first grade before reducing the size of other grade levels. Grades 2 and 3 show similar growth in grade migration rates, although second grade migration drops sharply in 1998–99, while third grade shows a more gradual decline.

The grade assignment showing the lowest level of grade migration is kindergarten. This must be interpreted with some caution, however. The PAIF employment database does not differentiate between kindergarten and pre-kindergarten teachers—the same job assignment category captures all of these teachers. Many of these K teachers, then, may in fact be in pre-K positions that were not eligible for CSR funds.

Teacher Migration from Grades 4 and 5

Of particular interest is the migration out of non-CSR impacted grade levels into K–3 teaching assignments. In particular, it has been hypothesized that the onset of CSR would spark a wave of migration from grades 4 and 5 into K–3. We examine the specific migration patterns of grade 4 and 5 teachers in Table C4, and find little support for this

hypothesis. Movement from grade 4 to grade 3 begins at a base rate of 3 percent in 1994–95, and gradually increases to a peak of 8 percent in 1997–98. Movement from grade 4 to grade levels lower than grade 3 does not change significantly. The migration rates of grade 5 teachers into grade 3 are less pronounced, peaking at 5 percent in 1997–98.

In all, it appears that the onset of CSR had a very modest impact on the migration of grade 4 and 5 teachers to CSR classrooms.

Table C4–
Specific grade level changes for grades 4 and 5 teachers

		Current Year									
		1994–95		1995–96		1996–97		1997–98		1998–99	
Current Grade	Next Year's Assignment	N	Pct.	N	Pct.	N	Pct.	N	Pct.	N	Pct.
4	K	63	1%	78	1%	85	1%	64	1%	104	1%
	1	81	1%	169	2%	108	1%	81	1%	96	1%
	2	143	2%	215	2%	292	3%	183	2%	169	2%
	3	258	3%	429	5%	602	7%	630	8%	448	5%
	K–4 mixed	422	5%	406	5%	272	3%	242	3%	180	2%
	4	5,975	68%	5,766	66%	4,647	51%	5,049	62%	5,320	64%
	5	449	5%	342	4%	319	4%	327	4%	337	4%
	Other	770	9%	708	8%	566	6%	663	8%	577	7%
	(Not Tracked)	621	7%	611	7%	2,140[8]	24%	964	12%	1,114	13%
4 Total		8,782	100%	8,724	100%	9,031	100%	8,203	100%	8,345	100%
5	K	80	1%	84	1%	108	1%	84	1%	105	1%
	1	60	1%	145	2%	75	1%	52	1%	63	1%
	2	97	1%	149	2%	187	2%	121	2%	106	1%
	3	141	2%	201	2%	330	4%	348	5%	247	3%
	K–4 mixed	103	1%	126	1%	108	1%	61	1%	58	1%
	4	313	4%	337	4%	343	4%	309	4%	299	4%
	5	5,828	70%	5,744	68%	4,556	53%	4,879	64%	5,065	65%
	Other	1,084	13%	1,102	13%	836	10%	879	12%	799	10%
	(Not Tracked)	577	7%	548	6%	2,041[8]	24%	897	12%	1,046	13%
5 Total		8,283	100%	8,436	100%	8,584	100%	7,630	100%	7,788	100%

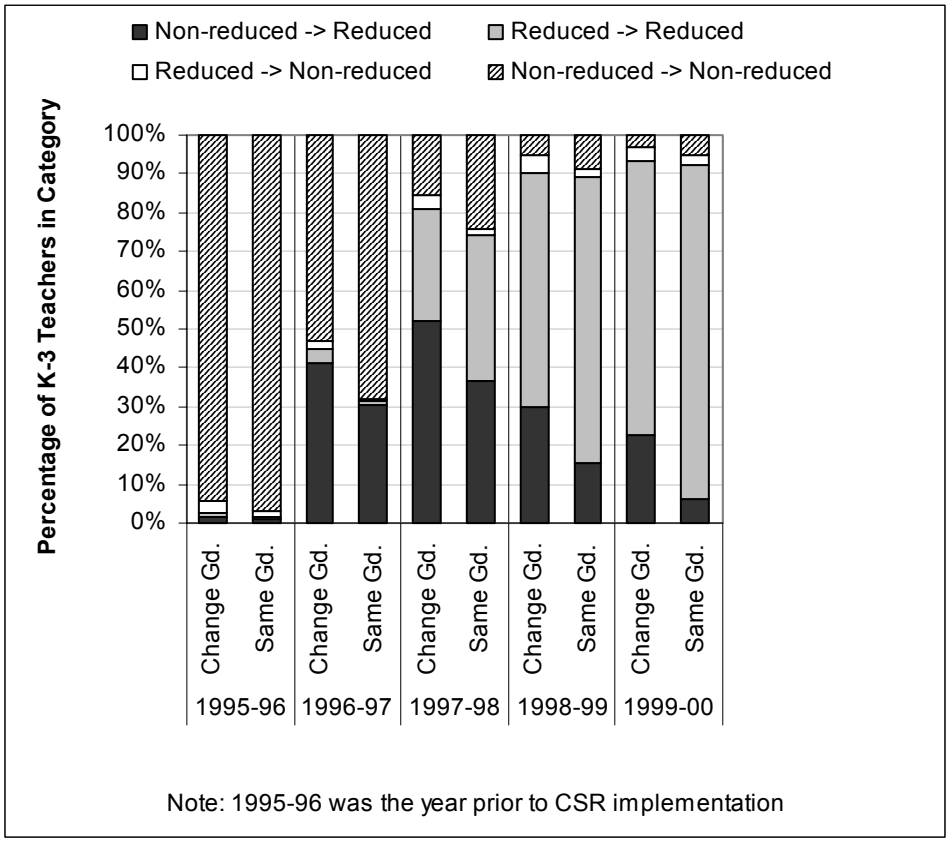
Cross-Grade Migration and Class Size

It is widely assumed that the introduction of small classes for targeted grades would result in cross-grade migration. Indeed, this is what we observe in the data. Figure C15

[8] This sudden increase in the percentage of teachers "not tracked" is due to a disproportionate loss of tracking ID numbers in the 1997-98 school year. An increase in the number of teachers not tracked this year decreases the observed percentage remaining in the same grade level, although the true percentage may not have changed significantly. See the section ""Methods and Data Sources" for details.

shows the distribution of class size changes broken down by the grade change status of teachers. We see that, predictably, teachers who change grade also move from larger to smaller classes in a higher proportion than those who do not change grades. This effect of approximately 15 percentage points remains relatively constant over the life of class size reduction. Also notable is the tendency for grade stayers to remain in large classes in slightly larger numbers. Although not shown in this figure, there is no more than a slight difference when novice teachers are compared to experienced teachers.

Figure C15–
Class size change status by grade change status

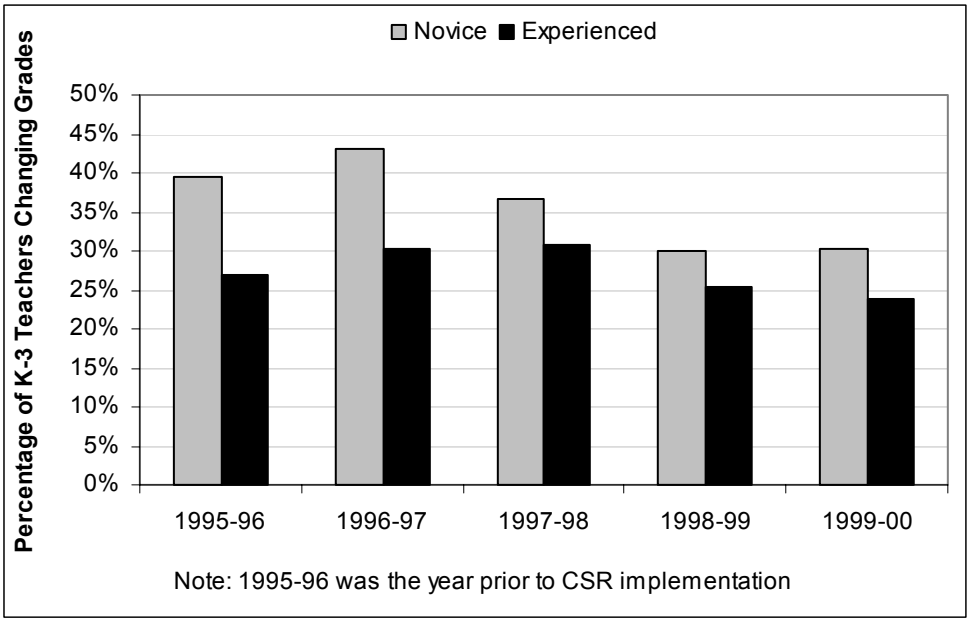


Cross-Grade Migration and Teacher Characteristics

As with cross-school migration, novice teachers are more likely to change grade assignments than their more senior colleagues. Figure C16 shows the difference in grade-changing rates broken down by the experience level of the teachers. During the first year of CSR, novices change grades roughly 1.5 times as frequently as experienced teachers, a ratio similar to that in the year prior to CSR. By the second year of CSR, the rate of novice grade change begins to approach that of experienced teachers, and both groups' migration rates decline for the remaining years under study.

As with the analysis of school migration above, it appears that the onset of CSR temporarily increases a tendency that was already present in the grade migration patterns—novices are slightly more mobile than their more experienced peers. Furthermore, referring back to data presented in Figure C16, it appears that there is little "displacement" of junior teachers by more senior teachers.

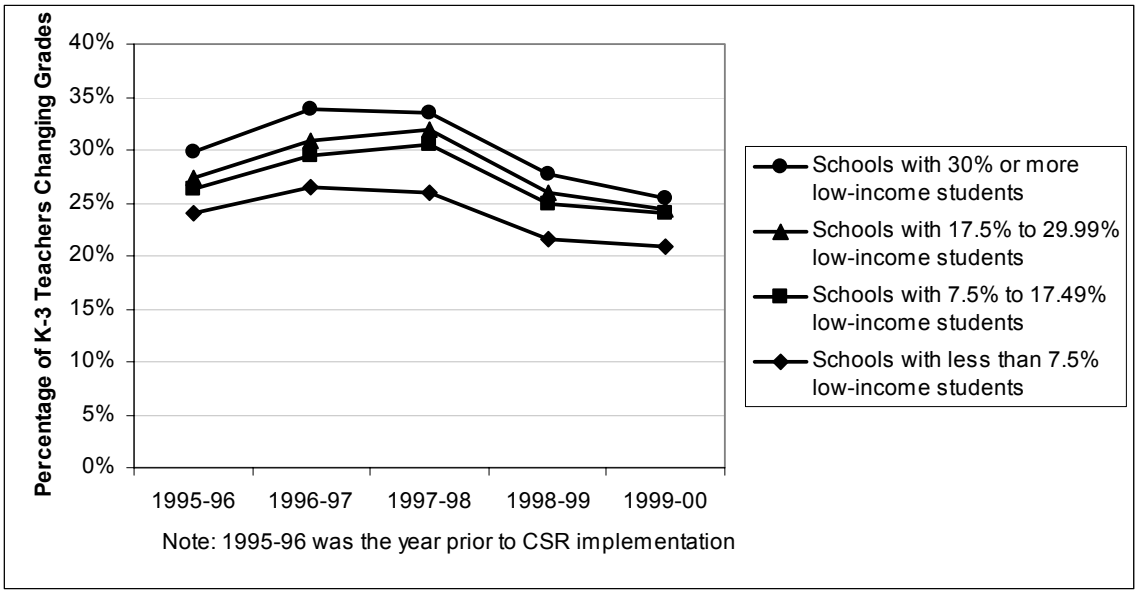
Figure C16–
Percentage of teachers changing grade assignment, by teacher experience



Cross-Grade Migration and School Characteristics

It was not surprising in a previous section to discover that schools with the highest percentage of low-income students experienced the greatest degree of cross-school migration. Low-SES schools are known to have a higher teacher turnover rate. Our data also show, however, that there are detectable differences in cross-grade migration when schools are viewed by SES quartile.

Figure C17–
Percentage of Teachers Changing Grades, by SES Quartile of School



In Figure C17 we show the cross-grade migration rates by SES quartile. The schools with the lowest percentage of low-income students show the least cross-grade movement, with cross-grade migration rates increasing with the percentage of low-income students. These differences remain fairly constant over the five years studied, although the overall migration rates rise and fall with the onset of CSR.

Of course, cross-grade migration in general is associated with younger, less experienced teachers, and in previous reports we have shown that lower SES districts tend to have a disproportionate share of novice teachers. The migration data presented here may be nothing more than a reflection of that tendency.

Conclusions — Cross-Grade Migration

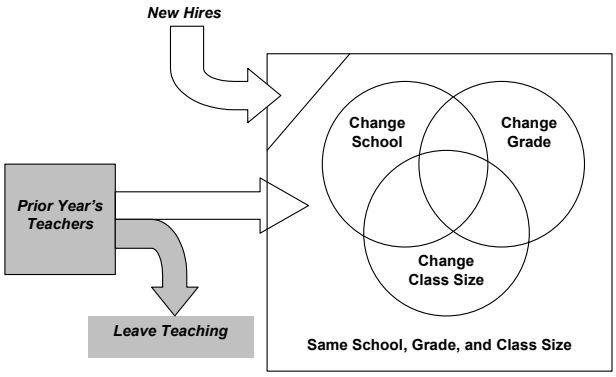
Clearly, the rate of cross-grade migration dwarfs that of cross-school movement. As with cross-school movement, novice teachers are generally more mobile than their senior peers, and lower-SES schools show a slightly higher degree of cross-grade churn than do higher-SES schools.

Grade changers are more likely to change from a large class to a small class than are those who remain in the same grade assignment. This trend is strongest during the first two years of CSR implementation, and then tapers off. By 1998–99, the overall proportion of teachers in small class assignments tops 90 percent for both grade movers and stayers.

The first two years of CSR implementation coincide with a peak in the K–3 grade migration rates, indicating that there was indeed some movement into K–3 grades from

other grades. This movement, however, is transitory, and returns to pre-CSR levels within a couple of years.

Teacher Attrition



Prior research has documented the problem of new teacher attrition. Roughly 50 percent of newly hired teachers leave the field within the first five years of service. Our surveys indicate that K–3 teachers strongly prefer to work with a reduced-size class. Does this preference translate into greater longevity of service?

Teacher attrition is most dramatic within the first five years of employment. Thus we concentrate on teachers newly entering the profession when CSR was first implemented. Taking the cohort of new K–3 hires in the 1996–97 school year, we track their progress through the following three years for which we have data, differentiating between those who teach in large and small classes.

It is important to consider the socio-economic status of the school when examining teacher attrition. First and most obvious is that lower-SES schools tend to have a higher turnover rate than their wealthier counterparts. Second, and for this study more importantly, higher-SES districts were able to implement CSR more quickly than lower-SES districts. By attempting to control for the effects of SES, we hope to show more cleanly the effect of class size on teacher attrition.

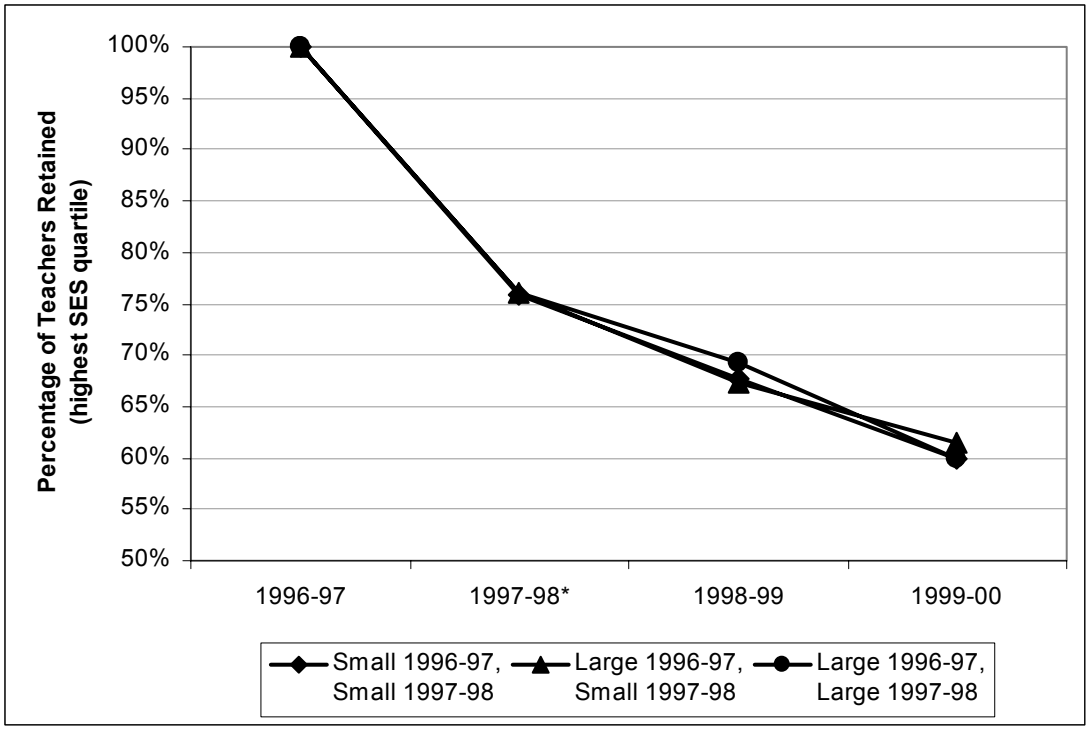
The results are surprising. We find that in the top and bottom SES quartiles, exposure to a small class makes little difference in teacher attrition, while in the middle two quartiles, there is a significant relationship with CSR exposure. This is illustrated in the following series of charts.

Figure C18 shows the teacher retention rate for newly hired K–3 teachers in the 1996–97 year in the highest SES quartile of schools. In 1997–98 there are two data points, not easily visible on the chart, separating those teachers who were hired into a small class in 1996–97 from those who were hired into a large class. In 1998–1999 the group of teachers who had been hired into a large class is split once again, this time into those who remain in a large class for their second year, and those who were able to teach in a

small class for their second year of teaching. These three trend lines are then extended to the 1999–00 school year.

As one can see, by the 1999–00 school year this is little discernable difference between those who taught in small and large classes. The overall retention rate is between 60 and 62 percent (understanding that missing data account for some of this attrition).

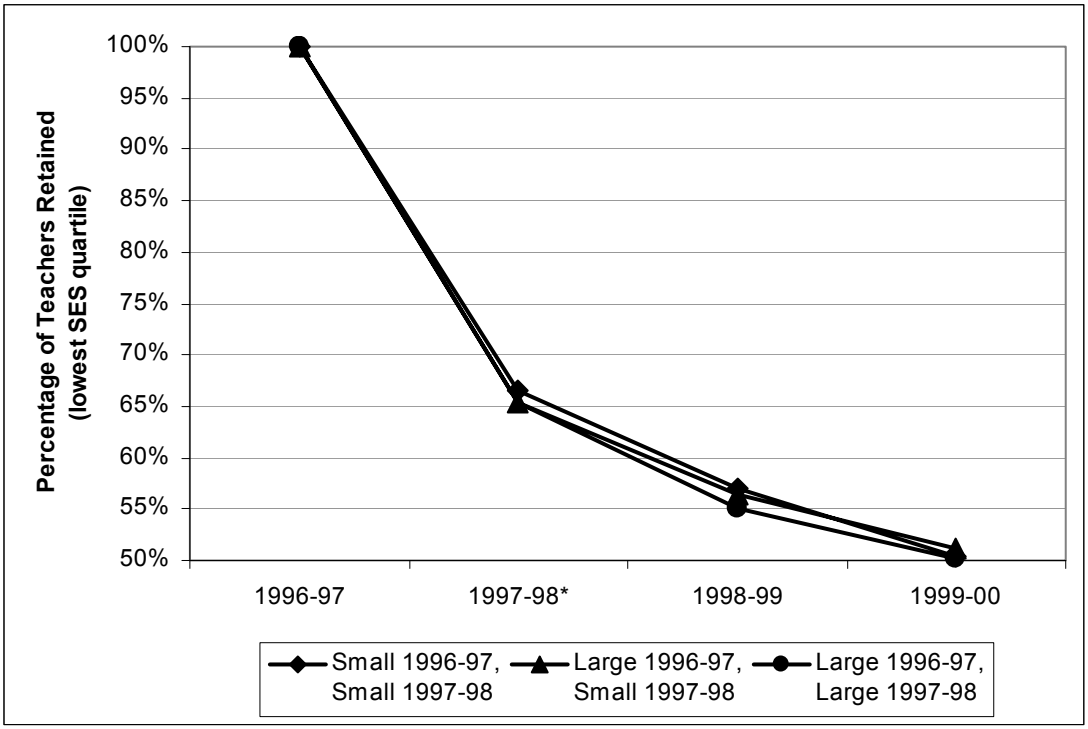
Figure C18–
1996–97 new K–3 teacher retention for highest SES quartile, by year and CSR assignment*



* The year 1997–98 data contain a disproportionate number of missing teaching identifiers. Thus it appears that many of these teachers have left the field. Many of these teachers are actually teaching, but we are unable to link them with the prior year's data.

At the other extreme, Figure C19 illustrates the retention statistics for the lowest SES quartile. Once again exposure to a reduced-size classroom plays very little role in influencing retention. By the year 1999–00 there is less than a 1 percentage point difference among all groups. Note, however, that the retention rate for the lowest SES quartile is a full 11 percentage points lower than that of the highest SES quartile (Figure C18).

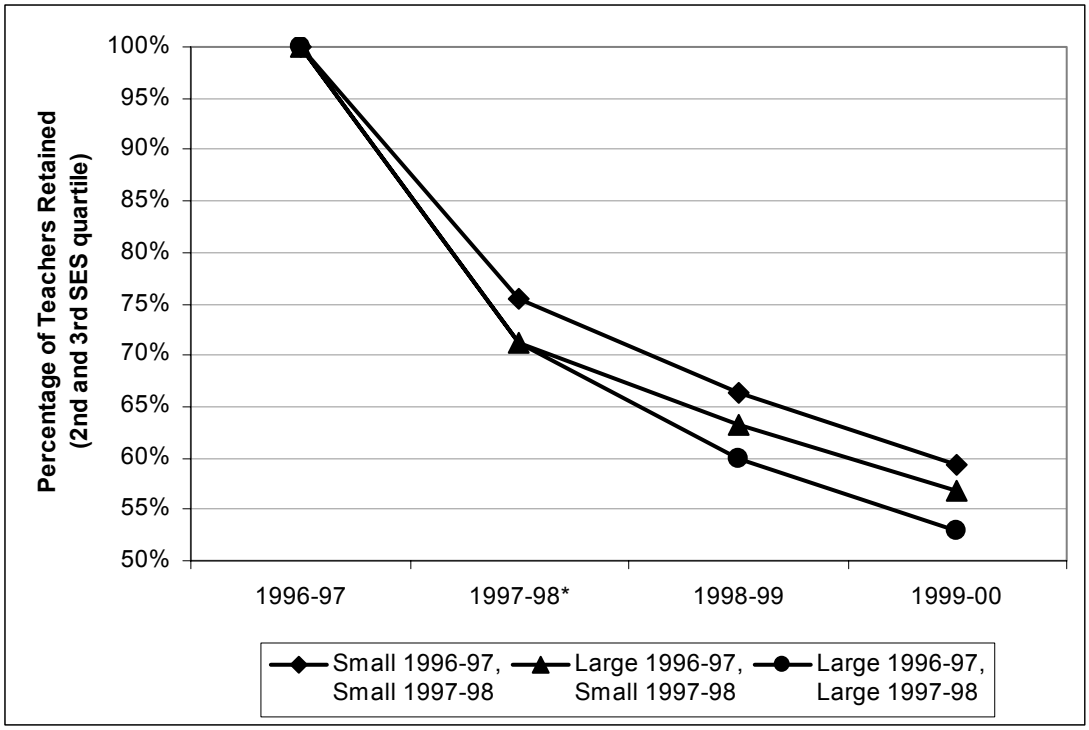
Figure C19–
1996–97 new K–3 teacher retention for lowest SES quartile, by year and CSR assignment*



* The year 1997–98 data contain a disproportionate number of missing teaching identifiers. Thus it appears that many of these teachers have left the field. Many of these teachers are actually teaching, but we are unable to link them with the prior year's data.

The effects of CSR on teacher retention are best illustrated in the middle two quartiles, combined in Figure C20. Here we see that even one year after hire (1997–98), there is a noticeable difference in the retention rate between those who were hired into a small class, and those who were not (75% vs. 71%). In 1998–99, we split the trajectory of those originally hired into a large class into two pieces. At this point those who have been exposed to a large class for their first two years of teaching have the worst retention rate (60%), while those with two years of small class teaching fare the best (66% retention.) In the final year studied, these retention rates maintain their relative ranking, with six percentage points separating those who had at least two years of small class teaching from those who did not.

Figure C20–
1996–97 new K–3 teacher retention for 2nd and 3rd SES quartile, by year and CSR assignment*



*The year 1997–98 data contain a disproportionate number of missing teaching identifiers. Thus it appears that many of these teachers have left the field. Many of these teachers are actually teaching, but we are unable to link them with the prior year's data.

We note in passing that the effect of class size on teacher retention is most detectable for novice teachers. With more experienced teachers, the effect of class size on retention is negligible.

Conclusions

There are modest but measurable changes in teacher migration patterns coinciding with the implementation of CSR. As stated in the introduction, one cannot make causal claims based on correlational time-series data. Other concurrent changes in the California policy environment may also account for some of the migration patterns observed over time. Some of these policy changes may serve to augment migration patterns due to CSR, others may counteract these effects.

The cumulative impact of migration from three sources—other schools, other grades, and new hires—is considered in Figure C21. For each year and grade level, the horizontal bar represents the total teaching force that year. Beginning with the left-hand side, the first bar segment represents the proportion of the teaching force that was in the same grade assignment and school the previous year. With the exception of the K–4 mixed grade classes, that figure is always greater than 50 percent of the teaching force.

The next bar segment represents those teachers who were previously teaching at the same school, but have changed grade assignments. When combined with the left-most bar, the total represents all teachers who have stayed at the same school, whether or not at the same grade level. Thus we see that in most cases, more than 80 percent of the teachers in a given grade level and year were at the same school the previous year.

The third and fourth bar segments combine to indicate all of those teachers who have changed schools. The third bar shows teachers who have changed both grades and schools, while the fourth indicates those who changed schools but retained the same grade assignment. Lastly, the right-most bar indicates teachers who were newly hired in a given year.

This summary chart provides a graphic answer to three of the four research questions asked in the beginning of this report:

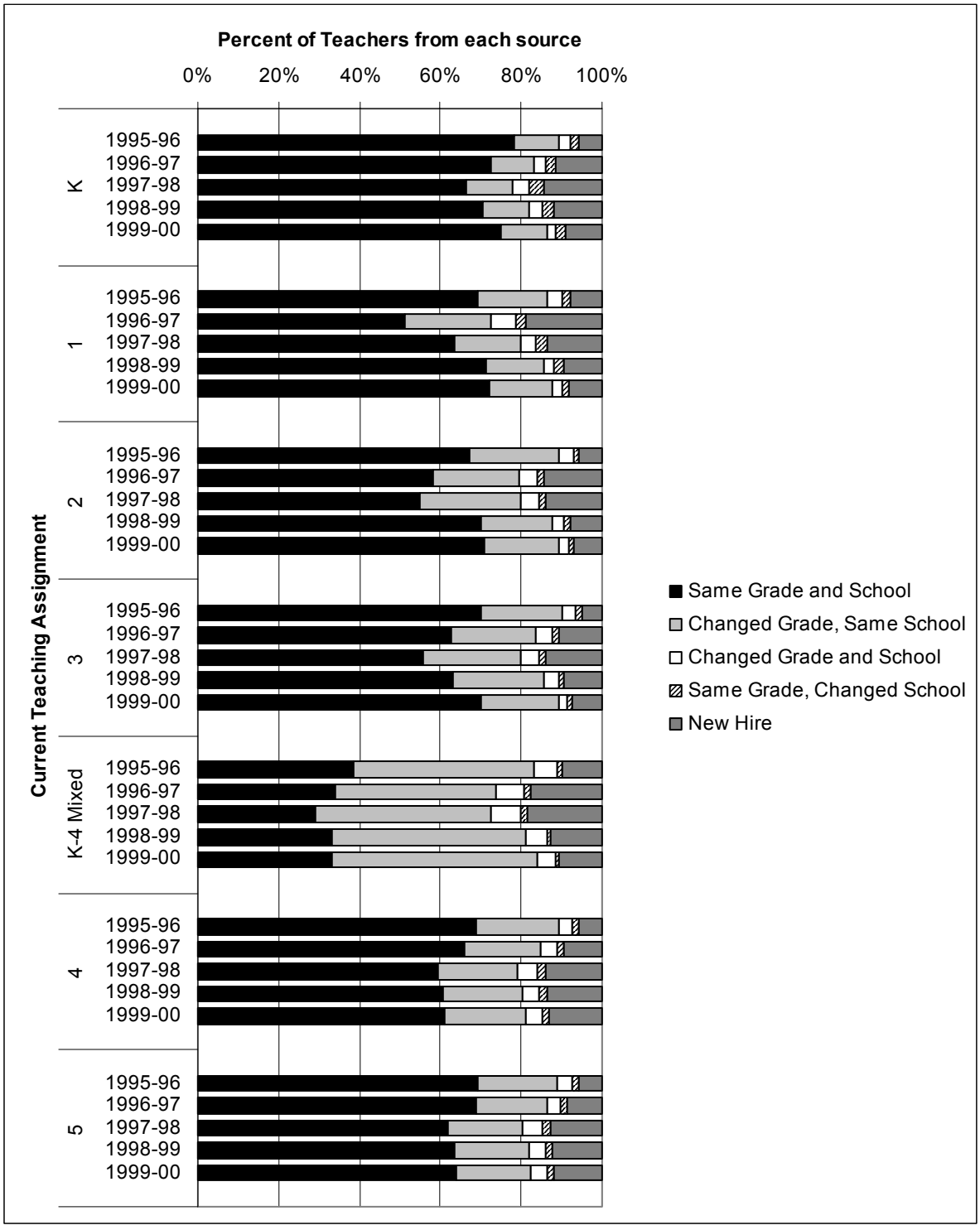
What percentage of newly created K–3 classrooms are being filled by new hires?

As shown by the right-most bars in Figure C21 (as well as the entirety of Figure C2), there is a marked increase in the percentage of new hires among the K-5 teaching force coincident with the onset of CSR. The percentage of new K-5 hires more than doubles in the first year of CSR, from an average of 5.5% to 12.3%. After the 1996-97 school year, the percentage of K-3 teachers who are new hires begins to taper off, returning to nearly the pre-CSR percentage in our last year of data, 1999-2000.

Table C5–
Combined K-5 hiring rates

1995-96	1996-97	1997-98	1998-99	1999-00
5.5%	12.3%	10.9%	7.8%	6.6%

Figure C21–
Sources of current teachers by prior teaching assignment and prior school



To what degree are teachers changing schools with the onset of CSR?

While there was a noticeable increase in K-5 cross-school migration coinciding with the onset of CSR, the overall rates were not at high as many had expected. The greatest increment in cross-school migration occurred with grade 1 teachers in 1996-97, where the migration rate increased from 6.5% to 10.3% (see Figure C3 for a detailed chart). The overall K-5 cross school migration rates are listed in Table C6, and may be graphically seen in the third and fourth bars from the left in Figure C21.

Table C6–
Combined K-5 cross-school migration rates

1995-96	1996-97	1997-98	1998-99	1999-00
5.7%	7.4%	8.1%	6.0%	4.8%

Cross school migration rates peaked in 1997-98 at 2.4 percentage points higher than the pre-CSR baseline of 5.7%. By the last year of data analyzed, the cross school migration rate had returned to a level lower than that of the pre-CSR year.

Within these relatively low migration rates we observe some changes in the school demographics for school movers, as well as a slightly greater tendency for school movers to be in a reduced-size classroom than their school-staying peers.

To what degree are teachers changing grade level assignments with the onset of CSR?

Although we observe a rise in cross-grade migration for grades 1, 2, and 3 (see Figure C14), the overall cross-grade migration rates are remarkably stable over time. In particular, there is no evidence of a massive influx of teachers from non-CSR impacted grades into the K-3 teaching force. This can also be graphically observed by noting the relatively stable lengths of the second and third bars from the left in Figure C21.

Table C7–
Combined K-5 cross-grade migration rates

1995-96	1996-97	1997-98	1998-99	1999-00
25.2%	25.5%	27.0%	23.7%	23.0%

Does teaching in a reduced-size class predict teacher attrition?

As is predicted from prior research, SES appears to be a strong determinant of teacher attrition. Within the middle two SES quartiles, however, we do note a moderate difference (6 percentage points – see Figure C20) in attrition rates between those with two years of reduced-size classes and those without. Without a longer time span of data it is difficult to draw any substantial conclusions. However, the data are suggestive of a

relationship between reduced-size classroom teaching and attrition, and additional research on future panels of data is warranted.

Final Conclusions

In the conclusion to Chapter 4 of the year 3 evaluation report, we speculated that the disparities in teacher qualification across socio-economic strata were due to a combination of differential hiring (i.e., lower-income schools were more likely to hire less qualified teachers) and migration of more qualified teachers to higher-income schools.

We are now in a position to address this speculation.

In general, the onset of CSR coincided with an increase in many trends that were already present in the data prior to CSR. For example, teachers who changed schools on average have always moved to higher-SES schools – this trend increased with the onset of CSR. While novice teachers have always been more mobile than more experienced veterans, this discrepancy increased when CSR was implemented.

Before the onset of CSR the majority of veteran teachers in K-5 were either teaching the same grade level in their current school, or were changing grade levels but remaining in their same school. With the advent of CSR the majority of K-5 teachers continued to be teachers who had been in the same school the prior year. Most of the newly created vacancies were filled through new hires, not massive teacher migration.

The key conclusion of this report, however, is that while cross-school and cross-grade migration rates did increase with the onset of CSR, these increases were not of the magnitude expected by many policy-makers. There was no wholesale abandonment of grades 4 and 5 by teachers desiring a reduced-size class, nor was there a massive migration from lower-income schools to early-adopting higher-income schools.

Thus, we are in a position to answer the key question motivating this investigation, namely, why does the variation in K-3 teacher qualifications across socio-economic strata increase dramatically with the onset of CSR? Overall, these analyses suggest that the gap in teacher qualifications among schools increased because schools serving lower-income, minority, or EL students were less able to hire qualified teachers to fill their new positions, rather than because of a rapid increase in transfers among qualified teachers already in the schools.

Supplement to Appendix C

Detailed Grade Migration Tables

These tables tally the grade level transitions for teachers with valid ID's. Along the left hand margin is the prior year's grade assignment, while across the top row is the current year's assignment. Numbers are expressed as both absolute counts of current employment at a grade level, as well as a percentage of the specific row category.

Year	Prior Year's Grade		K	1	2	3	K-4 mixed	4	5	Other	Un-known	Total
1995-1996	K	School Movers	235	76	39	24	53	16	6	57	0	506
			46.4%	15.0%	7.7%	4.7%	10.5%	3.2%	1.2%	11.3%	0.0%	100%
		School Stayers	8,504	328	160	68	366	43	54	132	727	10,382
			81.9%	3.2%	1.5%	0.7%	3.5%	0.4%	0.5%	1.3%	7.0%	100%
	1	School Movers	79	200	67	38	71	23	14	77	0	569
			13.9%	35.1%	11.8%	6.7%	12.5%	4.0%	2.5%	13.5%	0.0%	100%
		School Stayers	343	6,888	492	126	721	74	48	159	752	9,603
			3.6%	71.7%	5.1%	1.3%	7.5%	0.8%	0.5%	1.7%	7.8%	100%
	2	School Movers	32	71	115	49	63	20	17	87	0	454
			7.0%	15.6%	25.3%	10.8%	13.9%	4.4%	3.7%	19.2%	0.0%	100%
		School Stayers	133	340	6,427	352	741	106	58	152	682	8,991
			1.5%	3.8%	71.5%	3.9%	8.2%	1.2%	0.6%	1.7%	7.6%	100%
	3	School Movers	35	40	56	127	55	43	32	120	0	508
			6.9%	7.9%	11.0%	25.0%	10.8%	8.5%	6.3%	23.6%	0.0%	100%
		School Stayers	77	90	298	6,415	745	301	106	181	669	8,882
			0.9%	1.0%	3.4%	72.2%	8.4%	3.4%	1.2%	2.0%	7.5%	100%
	K-4 mixed	School Movers	52	68	47	44	97	32	17	107	0	464
			11.2%	14.7%	10.1%	9.5%	20.9%	6.9%	3.7%	23.1%	0.0%	100%
		School Stayers	395	714	848	767	2,849	406	81	372	596	7,028
			5.6%	10.2%	12.1%	10.9%	40.5%	5.8%	1.2%	5.3%	8.5%	100%
	4	School Movers	20	22	37	42	35	127	68	171	3	525
			3.8%	4.2%	7.0%	8.0%	6.7%	24.2%	13.0%	32.6%	0.6%	100%
		School Stayers	43	59	106	216	387	5,848	381	599	618	8,257
			0.5%	0.7%	1.3%	2.6%	4.7%	70.8%	4.6%	7.3%	7.5%	100%
	5	School Movers	16	24	24	36	30	47	135	226	0	538
			3.0%	4.5%	4.5%	6.7%	5.6%	8.7%	25.1%	42.0%	0.0%	100%
		School Stayers	64	36	73	105	73	266	5,693	858	577	7,745
			0.8%	0.5%	0.9%	1.4%	0.9%	3.4%	73.5%	11.1%	7.4%	100%
	Other	School Movers	75	94	77	91	105	103	138	2,929	3	3,615
			2.1%	2.6%	2.1%	2.5%	2.9%	2.8%	3.8%	81.0%	0.1%	100%
		School Stayers	126	124	137	195	296	532	900	21,988	3,266	27,564
			0.5%	0.4%	0.5%	0.7%	1.1%	1.9%	3.3%	79.8%	11.8%	100%
	New Hires	School Stayers	547	656	469	391	607	420	392	1,314	0	4,796
			11.4%	13.7%	9.8%	8.2%	12.7%	8.8%	8.2%	27.4%	0.0%	100%
	Unknown	School Movers	0	0	1	0	3	0	0	6	0	10
			0.0%	0.0%	10.0%	0.0%	30.0%	0.0%	0.0%	60.0%	0.0%	100%
		School Stayers	480	488	369	357	418	308	293	2,431	17	5,161
			9.3%	9.5%	7.1%	6.9%	8.1%	6.0%	5.7%	47.1%	0.3%	100%
1995-1996 Total			11,256	10,318	9,842	9,443	7,715	8,715	8,433	31,966	7,910	105,598
			10.7%	9.8%	9.3%	8.9%	7.3%	8.3%	8.0%	30.3%	7.5%	100%

Year	Prior Year's Grade		K	1	2	3	K-4 mixed	4	5	Other	Un-known	Total
1996-1997	K	School Movers	291	172	60	18	60	18	9	71	0	699
			41.6%	24.6%	8.6%	2.6%	8.6%	2.6%	1.3%	10.2%	0.0%	100%
		School Stayers	8,453	693	171	81	270	27	46	126	705	10,572
			80.0%	6.6%	1.6%	0.8%	2.6%	0.3%	0.4%	1.2%	6.7%	100%
	1	School Movers	69	355	78	37	70	20	11	90	0	730
			9.5%	48.6%	10.7%	5.1%	9.6%	2.7%	1.5%	12.3%	0.0%	100%
		School Stayers	293	7,328	462	120	447	56	32	164	697	9,599
			3.1%	76.3%	4.8%	1.3%	4.7%	0.6%	0.3%	1.7%	7.3%	100%
	2	School Movers	44	132	183	47	90	23	13	85	0	617
			7.1%	21.4%	29.7%	7.6%	14.6%	3.7%	2.1%	13.8%	0.0%	100%
		School Stayers	155	473	6,739	317	607	94	35	127	690	9,237
			1.7%	5.1%	73.0%	3.4%	6.6%	1.0%	0.4%	1.4%	7.5%	100%
	3	School Movers	31	81	73	144	64	45	24	115	1	578
			5.4%	14.0%	12.6%	24.9%	11.1%	7.8%	4.2%	19.9%	0.2%	100%
		School Stayers	92	192	392	6,390	631	255	85	188	647	8,872
			1.0%	2.2%	4.4%	72.0%	7.1%	2.9%	1.0%	2.1%	7.3%	100%
	K-4 mixed	School Movers	69	145	78	62	110	34	17	110	1	626
			11.0%	23.2%	12.5%	9.9%	17.6%	5.4%	2.7%	17.6%	0.2%	100%
		School Stayers	417	1,139	968	873	2,357	380	76	306	583	7,099
			5.9%	16.0%	13.6%	12.3%	33.2%	5.4%	1.1%	4.3%	8.2%	100%
	4	School Movers	20	47	51	75	41	142	46	160	1	583
			3.4%	8.1%	8.7%	12.9%	7.0%	24.4%	7.9%	27.4%	0.2%	100%
		School Stayers	58	122	164	354	365	5,624	296	548	610	8,141
			0.7%	1.5%	2.0%	4.3%	4.5%	69.1%	3.6%	6.7%	7.5%	100%
	5	School Movers	14	36	38	56	23	58	128	229	0	582
			2.4%	6.2%	6.5%	9.6%	4.0%	10.0%	22.0%	39.3%	0.0%	100%
		School Stayers	70	109	111	145	103	279	5,616	873	548	7,854
			0.9%	1.4%	1.4%	1.8%	1.3%	3.6%	71.5%	11.1%	7.0%	100%
	Other	School Movers	112	230	141	122	139	146	163	2,849	8	3,910
			2.9%	5.9%	3.6%	3.1%	3.6%	3.7%	4.2%	72.9%	0.2%	100%
		School Stayers	142	373	219	246	322	522	865	22,003	3,295	27,987
			0.5%	1.3%	0.8%	0.9%	1.2%	1.9%	3.1%	78.6%	11.8%	100%
	Unknown	School Movers	0	1	0	0	1	0	0	4	0	6
			0.0%	16.7%	0.0%	0.0%	16.7%	0.0%	0.0%	66.7%	0.0%	100%
		School Stayers	900	1,502	1,032	751	847	579	476	3,124	29	9,240
			9.7%	16.3%	11.2%	8.1%	9.2%	6.3%	5.2%	33.8%	0.3%	100%
	New Hires	School Stayers	1,192	2,477	1,530	1,016	1,066	718	640	2,258	0	10,897
			10.9%	22.7%	14.0%	9.3%	9.8%	6.6%	5.9%	20.7%	0.0%	100%
	1996-1997 Total		12,422	15,607	12,490	10,854	7,613	9,020	8,578	33,430	7,815	117,829
			10.5%	13.2%	10.6%	9.2%	6.5%	7.7%	7.3%	28.4%	6.6%	100%

Year	Prior Year's Grade		K	1	2	3	K-4 mixed	4	5	Other	Un-known	Total
1997-1998	K	School Movers	398	118	64	22	51	12	6	72	0	743
			53.6%	15.9%	8.6%	3.0%	6.9%	1.6%	0.8%	9.7%	0.0%	100%
		School Stayers	7,449	521	274	99	266	41	37	114	2,873	11,674
			63.8%	4.5%	2.3%	0.8%	2.3%	0.4%	0.3%	1.0%	24.6%	100%
	1	School Movers	144	401	137	59	104	17	13	102	0	977
			14.7%	41.0%	14.0%	6.0%	10.6%	1.7%	1.3%	10.4%	0.0%	100%
		School Stayers	375	8,752	667	179	741	65	38	151	3,649	14,617
			2.6%	59.9%	4.6%	1.2%	5.1%	0.4%	0.3%	1.0%	25.0%	100%
	2	School Movers	57	95	228	69	51	33	18	84	0	635
			9.0%	15.0%	35.9%	10.9%	8.0%	5.2%	2.8%	13.2%	0.0%	100%
		School Stayers	154	361	7,250	342	568	92	46	98	2,935	11,846
			1.3%	3.0%	61.2%	2.9%	4.8%	0.8%	0.4%	0.8%	24.8%	100%
	3	School Movers	35	48	66	178	56	46	33	112	0	574
			6.1%	8.4%	11.5%	31.0%	9.8%	8.0%	5.7%	19.5%	0.0%	100%
		School Stayers	96	132	508	6,148	436	198	79	148	2,525	10,270
			0.9%	1.3%	4.9%	59.9%	4.2%	1.9%	0.8%	1.4%	24.6%	100%
	K-4 mixed	School Movers	60	90	84	56	101	39	21	80	0	531
			11.3%	16.9%	15.8%	10.5%	19.0%	7.3%	4.0%	15.1%	0.0%	100%
		School Stayers	307	758	1,030	855	1,822	295	67	258	1,687	7,079
			4.3%	10.7%	14.6%	12.1%	25.7%	4.2%	0.9%	3.6%	23.8%	100%
	4	School Movers	24	19	47	77	33	140	66	134	0	540
			4.4%	3.5%	8.7%	14.3%	6.1%	25.9%	12.2%	24.8%	0.0%	100%
		School Stayers	61	89	245	525	239	4,507	253	432	2,140	8,491
			0.7%	1.0%	2.9%	6.2%	2.8%	53.1%	3.0%	5.1%	25.2%	100%
	5	School Movers	21	13	40	63	28	69	157	196	0	587
			3.6%	2.2%	6.8%	10.7%	4.8%	11.8%	26.7%	33.4%	0.0%	100%
		School Stayers	87	62	147	267	80	274	4,399	640	2,041	7,997
			1.1%	0.8%	1.8%	3.3%	1.0%	3.4%	55.0%	8.0%	25.5%	100%
	Other	School Movers	115	134	153	152	139	174	186	2,422	0	3,475
			3.3%	3.9%	4.4%	4.4%	4.0%	5.0%	5.4%	69.7%	0.0%	100%
		School Stayers	204	312	398	415	381	545	769	18,296	8,544	29,864
			0.7%	1.0%	1.3%	1.4%	1.3%	1.8%	2.6%	61.3%	28.6%	100%
	Unknown	School Movers	0	1	0	2	0	0	0	11	0	14
			0.0%	7.1%	0.0%	14.3%	0.0%	0.0%	0.0%	78.6%	0.0%	100%
		School Stayers	1,478	1,790	1,786	1,343	904	942	858	3,931	13	13,045
			11.3%	13.7%	13.7%	10.3%	6.9%	7.2%	6.6%	30.1%	0.1%	100%
	New Hires	School Stayers	1,197	1,341	1,365	1,143	839	793	660	1,873	0	9,211
			13.0%	14.6%	14.8%	12.4%	9.1%	8.6%	7.2%	20.3%	0.0%	100%
	1997-1998 Total		12,262	15,037	14,489	11,994	6,839	8,282	7,706	29,154	26,407	132,170
			9.3%	11.4%	11.0%	9.1%	5.2%	6.3%	5.8%	22.1%	20.0%	100%

Year	Prior Year's Grade		K	1	2	3	K-4 mixed	4	5	Other	Un-known	Total
1998-1999	K	School Movers	362	78	48	27	39	18	11	70	1	654
			55.4%	11.9%	7.3%	4.1%	6.0%	2.8%	1.7%	10.7%	0.2%	100%
		School Stayers	8,839	409	173	101	249	44	50	193	1,419	11,477
			77.0%	3.6%	1.5%	0.9%	2.2%	0.4%	0.4%	1.7%	12.4%	100%
	1	School Movers	144	323	79	64	62	29	14	108	0	823
			17.5%	39.2%	9.6%	7.8%	7.5%	3.5%	1.7%	13.1%	0.0%	100%
		School Stayers	474	9,922	634	245	781	78	59	220	1,652	14,065
			3.4%	70.5%	4.5%	1.7%	5.6%	0.6%	0.4%	1.6%	11.7%	100%
	2	School Movers	71	72	227	83	65	42	23	129	2	714
			9.9%	10.1%	31.8%	11.6%	9.1%	5.9%	3.2%	18.1%	0.3%	100%
		School Stayers	206	390	9,607	604	841	104	56	208	1,622	13,638
			1.5%	2.9%	70.4%	4.4%	6.2%	0.8%	0.4%	1.5%	11.9%	100%
	3	School Movers	32	41	56	187	46	41	32	133	2	570
			5.6%	7.2%	9.8%	32.8%	8.1%	7.2%	5.6%	23.3%	0.4%	100%
		School Stayers	115	106	330	8,424	505	243	84	193	1,305	11,305
			1.0%	0.9%	2.9%	74.5%	4.5%	2.1%	0.7%	1.7%	11.5%	100%
	K-4 mixed	School Movers	55	64	64	38	60	21	11	86	0	399
			13.8%	16.0%	16.0%	9.5%	15.0%	5.3%	2.8%	21.6%	0.0%	100%
		School Stayers	374	879	865	834	1,954	259	52	354	823	6,394
			5.8%	13.7%	13.5%	13.0%	30.6%	4.1%	0.8%	5.5%	12.9%	100%
	4	School Movers	16	22	28	60	18	154	60	152	0	510
			3.1%	4.3%	5.5%	11.8%	3.5%	30.2%	11.8%	29.8%	0.0%	100%
		School Stayers	48	59	155	570	224	4,895	267	511	964	7,693
			0.6%	0.8%	2.0%	7.4%	2.9%	63.6%	3.5%	6.6%	12.5%	100%
	5	School Movers	10	17	30	41	9	46	140	215	1	509
			2.0%	3.3%	5.9%	8.1%	1.8%	9.0%	27.5%	42.2%	0.2%	100%
		School Stayers	74	35	91	307	52	263	4,739	664	896	7,121
			1.0%	0.5%	1.3%	4.3%	0.7%	3.7%	66.5%	9.3%	12.6%	100%
	Other	School Movers	87	70	74	136	62	144	156	2,393	5	3,127
			2.8%	2.2%	2.4%	4.3%	2.0%	4.6%	5.0%	76.5%	0.2%	100%
		School Stayers	154	146	180	360	202	585	807	18,935	4,330	25,699
			0.6%	0.6%	0.7%	1.4%	0.8%	2.3%	3.1%	73.7%	16.8%	100%
	Unknown	School Stayers	1,049	1,034	915	941	573	655	609	2,999	10	8,785
			11.9%	11.8%	10.4%	10.7%	6.5%	7.5%	6.9%	34.1%	0.1%	100%
	New Hires	School Stayers	1,054	971	790	917	523	782	664	1,699	0	7,400
			14.2%	13.1%	10.7%	12.4%	7.1%	10.6%	9.0%	23.0%	0.0%	100%
	1998-1999 Total		13,164	14,638	14,346	13,939	6,265	8,403	7,834	29,262	13,032	120,883
			10.9%	12.1%	11.9%	11.5%	5.2%	7.0%	6.5%	24.2%	10.8%	100%

Year	Prior Year's Grade		K	1	2	3	K-4 mixed	4	5	Other	Un-known	Total
1999-2000	K	School Movers	271	84	39	31	36	15	15	62	0	553
			49.0%	15.2%	7.1%	5.6%	6.5%	2.7%	2.7%	11.2%	0.0%	100%
		School Stayers	9,583	430	200	155	293	78	62	143	1,535	12,479
			76.8%	3.4%	1.6%	1.2%	2.3%	0.6%	0.5%	1.1%	12.3%	100%
	1	School Movers	76	204	73	50	47	23	23	95	3	594
			12.8%	34.3%	12.3%	8.4%	7.9%	3.9%	3.9%	16.0%	0.5%	100%
		School Stayers	417	9,929	592	204	689	101	72	218	1,670	13,892
			3.0%	71.5%	4.3%	1.5%	5.0%	0.7%	0.5%	1.6%	12.0%	100%
	2	School Movers	45	92	166	53	54	38	24	80	1	553
			8.1%	16.6%	30.0%	9.6%	9.8%	6.9%	4.3%	14.5%	0.2%	100%
		School Stayers	205	434	9,623	515	846	146	82	180	1,644	13,675
			1.5%	3.2%	70.4%	3.8%	6.2%	1.1%	0.6%	1.3%	12.0%	100%
	3	School Movers	29	44	51	168	42	58	36	140	1	569
			5.1%	7.7%	9.0%	29.5%	7.4%	10.2%	6.3%	24.6%	0.2%	100%
		School Stayers	153	127	388	9,582	597	341	137	240	1,693	13,258
			1.2%	1.0%	2.9%	72.3%	4.5%	2.6%	1.0%	1.8%	12.8%	100%
	K-4 mixed	School Movers	42	59	40	35	46	23	22	67	1	335
			12.5%	17.6%	11.9%	10.4%	13.7%	6.9%	6.6%	20.0%	0.3%	100%
		School Stayers	326	803	899	771	1,856	228	46	176	799	5,904
			5.5%	13.6%	15.2%	13.1%	31.4%	3.9%	0.8%	3.0%	13.5%	100%
	4	School Movers	22	16	29	46	28	132	45	143	0	461
			4.8%	3.5%	6.3%	10.0%	6.1%	28.6%	9.8%	31.0%	0.0%	100%
		School Stayers	82	80	140	402	152	5,188	292	434	1,114	7,884
			1.0%	1.0%	1.8%	5.1%	1.9%	65.8%	3.7%	5.5%	14.1%	100%
	5	School Movers	10	6	15	28	15	56	119	189	2	440
			2.3%	1.4%	3.4%	6.4%	3.4%	12.7%	27.0%	43.0%	0.5%	100%
		School Stayers	95	57	91	219	43	243	4,946	610	1,044	7,348
			1.3%	0.8%	1.2%	3.0%	0.6%	3.3%	67.3%	8.3%	14.2%	100%
	Other	School Movers	63	65	72	76	46	134	151	2,100	9	2,716
			2.3%	2.4%	2.7%	2.8%	1.7%	4.9%	5.6%	77.3%	0.3%	100%
		School Stayers	207	198	206	332	213	579	732	19,079	4,653	26,199
			0.8%	0.8%	0.8%	1.3%	0.8%	2.2%	2.8%	72.8%	17.8%	100%
	Unknown	School Movers	0	0	0	0	0	0	1	6	0	7
			0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	14.3%	85.7%	0.0%	100%
		School Stayers	871	881	800	771	431	624	557	2,820	49	7,804
			11.2%	11.3%	10.3%	9.9%	5.5%	8.0%	7.1%	36.1%	0.6%	100%
	New Hires	School Stayers	830	781	692	703	408	803	650	1,630	0	6,497
			12.8%	12.0%	10.7%	10.8%	6.3%	12.4%	10.0%	25.1%	0.0%	100%
	1999-2000 Total		13,327	14,290	14,116	14,141	5,842	8,810	8,012	28,412	14,218	121,168
			11.0%	11.8%	11.6%	11.7%	4.8%	7.3%	6.6%	23.4%	11.7%	100%

Appendix D

The Relationship Between Teacher Characteristics and Student Achievement in Reduced-size Classes: A Study of 6 California Districts

Jamie Shkolnik, Hiro Hikawa, Marika Suttorp, J.R. Lockwood, Brian Stecher, George Bohrnstedt

Study Objectives

The Class Size Reduction year 1, year 2, and year 3 evaluation reports examined the relationship between class size reduction (CSR) and SAT-9 test scores. The year 1 and year 2 results suggested a positive relationship between being in a reduced-size class and academic achievement, although the size of the effect was somewhat smaller than the one found in the Tennessee Student/Teacher Achievement Ratio (STAR) study. However, unlike the Tennessee study, the California evaluation did not find a greater class-size effect for inner city and minority students than for other children.

Were CSR's effects in California offset by other changes that occurred during implementation? Schools serving high percentages of at-risk students have been the most adversely affected by declines in teacher qualifications brought on by CSR; a disproportionate percentage of teachers without full credentials, with limited teaching experience, and without Master's degrees have ended up in these schools since the implementation of the reform. Might we have found greater CSR effects for schools with high percentages of at-risk students if they had teachers with higher credentials? Using state-level data from the National Assessment of Educational Progress, Darling-Hammond finds that the effects of teacher characteristics are far greater on achievement than those associated with class size. In particular, she finds that the best predictor of 4th and 8th grade reading and math achievement scores is the percentages of teachers who have both a state certification and the equivalent of a major in their teaching field[1] (Darling-Hammond, 2000).

Unfortunately, the data collected by the state do not allow one to link teacher characteristic variables (such as years of experience, credential status or education level) with individual student SAT-9 scores. As a result, we were not able to examine the relationship between the qualifications and experience of California teachers and the

[1] For elementary teachers who teach multiple subjects to the same group of students, this means an elementary teaching degree.

achievement of their students using statewide data. However, each district prepares annually the Professional Assignment Information Form (PAIF) for the California Basic Educational Data System (CBEDS). CBEDS PAIF details the assignments, credentials, and experience of each teacher. Each district also receives the SAT-9 data from California's state testing program. By working directly with districts, we were able to assemble data that contain teacher and student characteristics as well as student achievement information.

We conducted this study with data obtained from six large districts because the effect of the implementation of CSR on the teacher shortage was magnified in large districts.[2] These districts hired more less-experienced, not-fully-credentialed, and less-educated teachers than smaller districts in order to meet the increased demand for teachers in newly created classrooms. The purpose of this study is to examine the importance of these teacher characteristics in promoting student achievement in reduced-size classes.

Sample

The first phase in this study involved contacting a subset of the largest districts in the state to determine (a) the availability of and ease of obtaining the data necessary for this task and (b) their willingness to participate in this study. We chose large districts because we wanted to be certain there were enough teachers with varying degrees of preparation to have sufficient statistical power to detect reasonable differences in student achievement. In particular we were interested in getting districts with teachers who had emergency credentials. Giving priority to those districts with large enrollments, we also selected candidate districts to maximize heterogeneity with regard to district demographic characteristics (e.g., region of the state, percentage English Learners, and poverty levels). The candidate districts were contacted regarding the availability of the requisite data and their willingness to participate in the study. We were ultimately restricted to using districts that were willing and able to provide the data necessary for our analyses.

We contacted eight districts, six of which provided useable data; one district could only provide data for a small percentage of the third graders, and one district refused to participate.

Data

We obtained data from each participating district on SAT-9 test scores, teacher characteristics, student demographic characteristics, and classroom demographics.

Data from each district included:

[2] These 6 districts account for approximately 6 percent of the total enrollment in California public schools.

1. *Math and reading SAT-9 scores* were obtained for all third graders in 2000–2001, as well as the 1999–2000 second grade scores (pretest scores) for those same students. Conversations with districts indicated that the 1999–2000 school year was the first for which data could be provided. Testing begins in second grade by law in California so there were no prior scores for 2000–01 second graders. We used SAT-9 scale scores for all but one of the six districts; in one district we used normal curve equivalent (NCE) scores since this district did not provide scaled scores.[3]

2. *Characteristics of third grade teachers* include assignment, credential status, education level, years of experience, and gender for the school year when the students were in third grade. Districts supply these data annually as part of the CBEDS PAIF data collection.

3. *Student characteristics* include gender, race/ethnicity, special education status, English learner status, and poverty status (eligible for free or reduced price lunch).

4. *Classroom variables* include the percentage of Black, Hispanic, Asian, and Other students in the classroom. The percentage of students in special education, the percentage eligible for free or reduced price lunch, the percentage English learners, and the percentage female are also included where available.

Characteristics of the Data, by District

The following tables present means for the two outcomes (math and reading) used in the analysis. Some students had scores for math but not reading, or reading but not math. The samples were restricted to students who had scores for both years, which therefore excluded students who switched districts between second and third grade. Tables D1 and D2 show means, standard deviations, minimums, maximums, and numbers of observations for the math and reading scores, respectively. The numbers in two tables are very similar, because the most students took both tests both years.

[3] We use scaled scores because they are comparable among grade levels. Even though we have only third graders in the sample, we use their previous year score in the analysis.

Table D1– Math Scores: Demographics for the Six Districts for Third Graders, 2000–2001

	District A (N=2,241)				District B (N=3,663)				District C (N=3,924)				District D (N=5,389)				District E (N=7,157)				District F* (N=3,449)			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Test Score (SAT-9 Scaled Score)																								
Math 2001	608.83	42.69	497	767	594.09	43.53	485	767	597.59	36.87	494	767	590.84	40.30	478	767	611.07	44.49	488	767	50.40	20.23	1	99
Math 2000	579.19	39.51	449	708	568.16	42.40	420	731	570.31	37.35	389	747	562.90	40.46	443	708	577.97	44.29	429	731	45.46	20.84	1	99
Change in Math Scores	29.64	30.52	-108	136	25.93	30.35	-160	141	27.28	27.83	-141	150	27.93	28.18	-96	176	33.10	30.25	-118	163	4.94	15.17	-59	58.9
Student Variables																								
Female (1= Female, 0= Male)	50%	0.5			49%	0.5			49%	0.5			50%	0.5			49%	0.5			49%	0.5		
Asian (1= Asian, 0= Not)	3%	0.18			17%	0.37			3%	0.17			18%	0.38			10%	0.3			2%	0.14		
Black (1= Black, 0= Not)	0.3%	0.06			47%	0.5			0%	0.05			11%	0.31			17%	0.38			17%	0.38		
Hispanic (1= Hispanic, 0= Not)	47%	0.5			26%	0.44			93%	0.25			52%	0.5			52%	0.5			61%	0.49		
Other Race (1= Other, 0= Not)	9%	0.28			3%	0.17			1%	0.09			1%	0.12			5%	0.21			2%	0.14		
Special Ed. (1= SE, 0= Not)	6%	0.24			7%	0.25			0%	0.07			9%	0.28			7%	0.26			N/A	N/A		
English Learner (1= EL, 0= Not)	23%	0.42			35%	0.48			68%	0.47			46%	0.5			55%	0.5			42%	0.49		
Free/Reduced Price Lunch (1= F/RPL, 0= Not)	52%	0.5			68%	0.47			86%	0.35			N/A	N/A			67%	0.47			82%	0.39		
Teacher Variables																								
Full Credential (1= Full, 0= Not)	94%	0.24			80%	0.4			74%	0.44			99%	0.11			72%	0.45			85%	0.35		
Years Teaching = 1	6%	0.23			5%	0.23			10%	0.31			3%	0.16			9%	0.29			9%	0.29		
Years Teaching = 2	11%	0.31			8%	0.28			11%	0.31			3%	0.18			10%	0.3			9%	0.29		
Years Teaching = 3	5%	0.23			8%	0.28			14%	0.35			1%	0.12			8%	0.27			12%	0.32		
Years Teaching = 4–9	21%	0.41			26%	0.44			29%	0.46			29%	0.45			30%	0.46			28%	0.45		
Master's Degree or Higher	30%	0.46			34%	0.47			35%	0.48			7%	0.26			26%	0.44			37%	0.48		
Female (1= Female, 0= Male)	94%	0.24			81%	0.4			83%	0.37			83%	0.38			86%	0.35			81%	0.39		
Classroom Variables																								
% Black	0.34	1.51	0	10	46.71	35.70	0	100	0.24	1.18	0	11	10.91	12.55	0	100	17.74	14.92	0	100	17.27	15.59	0	100
% Hispanic	46.90	27.70	0	100	27.41	34.87	0	100	93.03	11.99	50	100	51.73	27.48	0	100	51.19	28.63	0	100	60.64	25.55	0	100
% Asian	3.56	5.43	0	33	16.40	26.42	0	100	3.09	6.77	0	50	17.45	20.88	0	100	9.75	13.64	0	95	2.07	4.54	0	50
% Special Education	6.57	7.57	0	100	7.18	8.23	0	100	0.57	2.21	0	100	9.19	9.25	0	100	6.98	10.78	0	100	N/A	N/A	N/A	N/A
% Free/Reduced Price Lunch	52.87	28.08	0	100	67.54	25.65	0	100	85.93	19.50	0	100	N/A	N/A	N/A	N/A	64.15	25.70	0	100	81.70	17.87	0	100
% English Learner	22.68	29.53	0	100	35.71	36.07	0	100	68.64	28.59	0	100	45.21	29.84	0	100	53.42	30.76	0	100	41.65	31.42	0	100
% Female	49.46	11.53	0	75	48.63	11.31	0	100	49.09	13.22	0	100	49.47	11.59	0	100	49.56	10.39	0	100	49.23	14.05	0	100

* SAT-9 scaled scores were not available for District F. Normal Curve Equivalent (NCE) scores were used.

Table D2– Reading Scores: Demographics for the Six Districts for Third Graders, 2000–2001

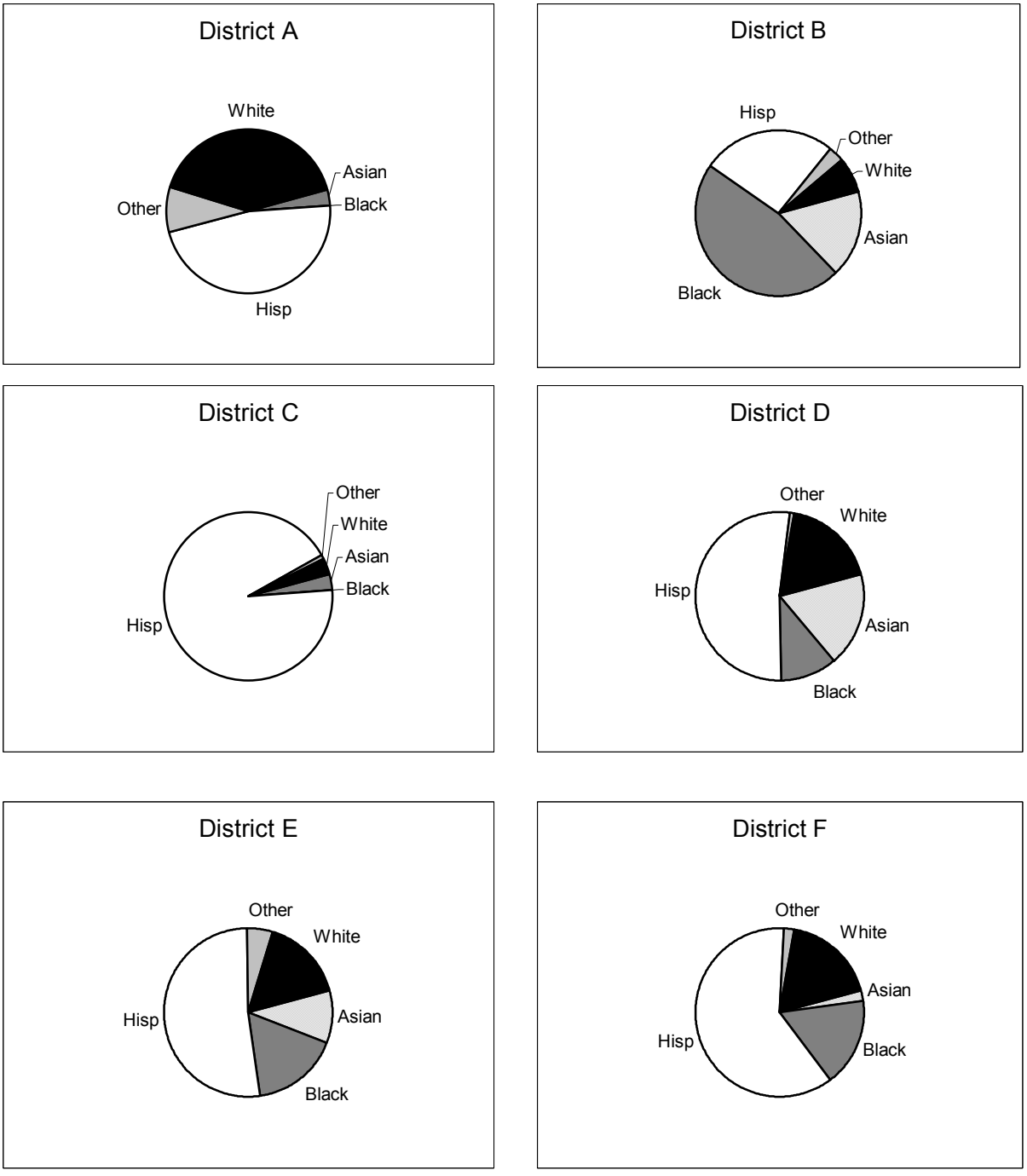
	District A (N=2,020)				District B (N=3,453)				District C (N=3,773)				District D (N=5,091)				District E (N=6,965)				District F* (N=3,357)			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Test Score (SAT-9 Scaled Score)																								
Reading 2001	616.25	42.47	511	775	600.48	43.84	468	775	588.47	34.80	498	753	591.59	40.39	495	775	607.01	42.46	488	775	40.17	17.58	1	99
Reading 2000	585.08	40.15	485	759	566.30	39.84	485	737	561.19	36.38	421	730	563.30	39.26	479	759	576.11	39.73	485	759	41.82	16.93	1	99
Change in Reading Scores	31.17	24.17	-75	139	34.19	26.17	-104	130	27.28	24.70	-141	126	28.29	23.69	-90	122	30.90	24.14	-71	146	-1.65	11.45	-47	61
Student Variables																								
Female (1= Female, 0= Male)	51%	0.5			49%	0.5			49%	0.5			50%	0.5			49%	0.5			49%	0.5		
Asian (1= Asian, 0= Not)	3%	0.18			17%	0.38			3%	0.18			18%	0.39			10%	0.3			2%	0.14		
Black (1= Black, 0= Not)	0.2%	0.05			47%	0.5			0%	0.05			11%	0.31			17%	0.38			17%	0.38		
Hispanic (1= Hispanic, 0= Not)	44%	0.5			26%	0.44			93%	0.26			51%	0.5			51%	0.5			61%	0.49		
Other Race (1= Other, 0= Not)	9%	0.28			3%	0.17			1%	0.09			1%	0.12			5%	0.21			2%	0.14		
Special Ed. (1= SE, 0= Not)	6%	0.23			7%	0.25			0%	0.07			8%	0.28			7%	0.25			N/A	N/A		
English Learner (1= EL, 0= Not)	19%	0.39			35%	0.48			67%	0.47			46%	0.5			54%	0.5			42%	0.49		
Free/Reduced Price Lunch (1= F/RPL, 0= Not)	49%	0.5			67%	0.47			85%	0.35			N/A	N/A			66%	0.47			81%	0.39		
Teacher Variables																								
Full Credential (1= Full, 0= Not)	95%	0.21			81%	0.4			74%	0.44			99%	0.11			72%	0.45			85%	0.35		
Years Teaching = 1	6%	0.23			6%	0.23			10%	0.31			3%	0.16			9%	0.29			9%	0.29		
Years Teaching = 2	10%	0.31			9%	0.28			11%	0.31			3%	0.18			11%	0.31			9%	0.28		
Years Teaching = 3	5%	0.21			8%	0.28			14%	0.35			1%	0.12			8%	0.27			12%	0.33		
Years Teaching = 4–9	22%	0.41			25%	0.44			30%	0.46			29%	0.45			30%	0.46			28%	0.45		
Master's Degree or Higher	30%	0.46			34%	0.47			35%	0.48			8%	0.27			26%	0.44			37%	0.48		
Female (1= Female, 0= Male)	94%	0.23			81%	0.39			83%	0.37			83%	0.38			86%	0.35			81%	0.39		
Classroom Variables																								
% Black	0.36	1.55	0	10	46.97	35.69	0	100	0.24	1.19	0	11	10.93	12.42	0	100	17.69	14.84	0	100	17.20	15.50	0	100
% Hispanic	44.16	26.23	0	100	26.73	34.55	0	100	92.86	12.20	0	100	51.40	27.44	0	100	50.98	28.66	0	100	60.62	25.58	0	100
% Asian	3.66	5.46	0	33	16.52	26.59	0	100	3.14	6.81	0	50	17.64	20.92	0	100	9.82	13.71	0	95	2.06	4.50	0	50
% Special Education	6.45	6.71	0	30	7.28	8.19	0	100	0.57	2.23	0	100	9.03	8.90	0	100	6.88	10.40	0	100	N/A	N/A	N/A	N/A
% Free/Reduced Price Lunch	50.37	27.53	0	100	67.21	25.87	0	100	85.68	19.74	0	100	N/A	N/A	N/A	N/A	64.00	25.78	0	100	81.51	17.95	0	100
% English Learner	19.65	26.75	0	100	35.09	35.88	0	100	68.22	28.78	0	100	45.07	29.73	0	100	53.26	30.75	0	100	41.57	31.54	0	100
% Female	49.73	11.29	21	75	48.78	11.38	0	100	49.08	12.94	0	100	49.50	11.53	0	100	49.60	10.33	0	100	49.17	13.95	0	100

* SAT-9 scaled scores were not available for District F. Normal Curve Equivalent (NCE) scores were used.

Student Characteristics

The six districts have substantially different population mixes of students. (See Figure D1.) District A is mostly White and Hispanic, District B is almost half Black, District C is almost entirely Hispanic, and Districts D and E are very similar: about half Hispanic with a mix of Black, White, and Asian students. District F is over half Hispanic with mostly Black and White students making up the remainder.

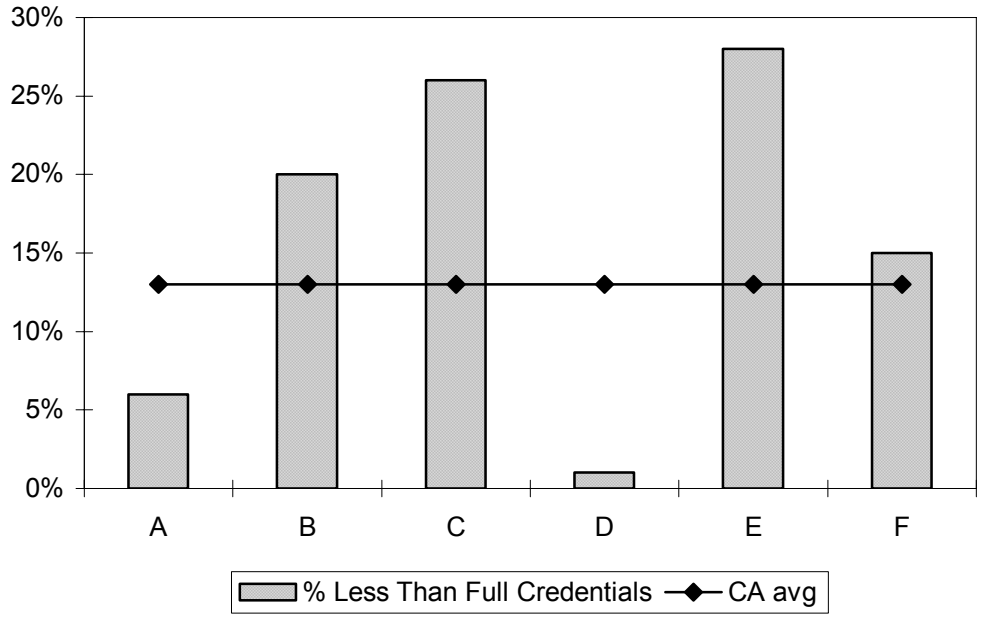
Figure D1–
Race/Ethnicity of Third Grade Students, by District



Teacher Characteristics

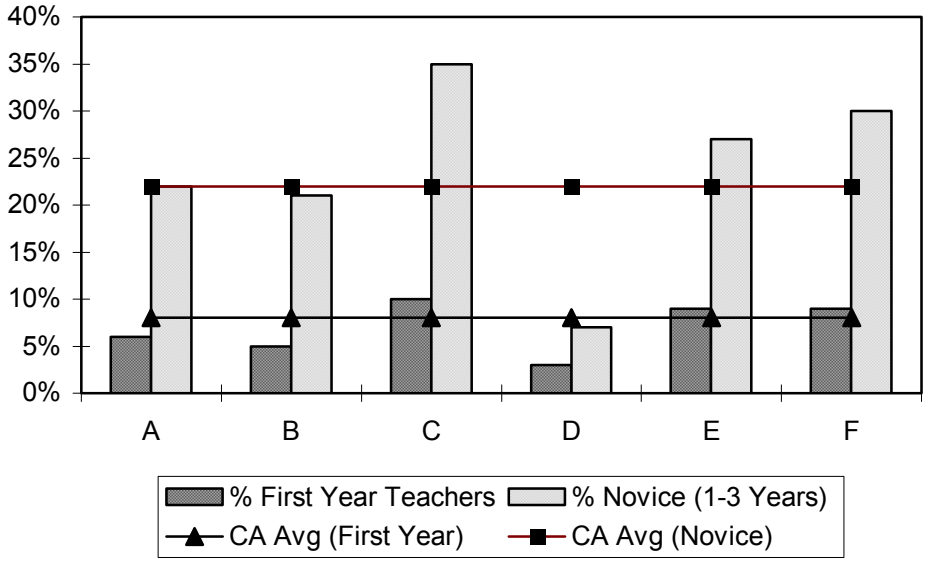
There is also a wide range of teacher characteristics across the districts in our sample. We show the percentage of third grade students who have teachers with particular characteristics (e.g. credential status, years of teaching experience, and education level). Since almost all classes have the same number of students (20 students), the percentage of students taught by a particular type of teacher is similar to the percentage of teachers of that type in the district. District D has the smallest percentage of students with teachers who do not have full credentials, only 1 percent; in contrast, District E has the highest percentage of teachers without full credentials—28 percent (see Figure D2 below). By way of comparison, the year 3 evaluation (see Stecher and Bohrnstedt, 2002) reported that 13 percent of all California K–3 teachers lacked full credentials in the 2000-2001 school year.

Figure D2–
Percentage of Third Grade Students with Less than Fully Credentialed Teachers, by District, 2000–2001



District D not only has the lowest percentage of less than fully credentialed teachers, but also has the lowest percentage of first year teachers (3 percent), and the lowest percentage of novice teachers, defined as teachers in their first three years of teaching (7 percent). (See Figure D3.) District C has the highest percentage of first year teachers in the 2000–2001 school year (10 percent) and the highest percentage of novice teachers (35 percent). In comparison, year 3 of the CSR evaluation found a state average of 8 percent first year teachers and 22 percent novice teachers in the 2000-2001 school year. Districts above the state average in percentage of first year teachers are also above average in the percentage of teachers in their first three years of teaching (novice teachers), indicating that districts that have had to hire in 2000–2001 also had a lot of hiring to do in the two prior years.

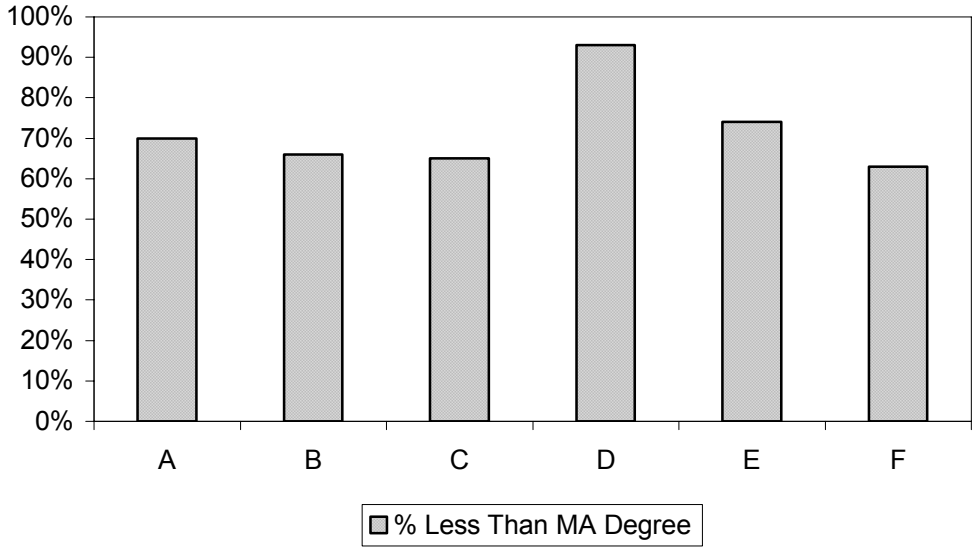
Figure D3–
Percentage of Third Grade Students with First-year and Novice Teachers, by District, 2000–2001



While District D had the lowest percentage of less than fully credentialed and inexperienced teachers, it surprisingly had the highest percentage of teachers with less than a Master's degree (93 percent). (See Figure D4.) The percentage of teachers with less than a Master's degree is very high compared to Districts B, C, and F, which had 66, 65, and 63 percent, respectively. Comparable numbers are not available from the year 3 evaluation.[4]

[4] In year 3 of this evaluation, we separated teachers into two groups: those with a bachelor's degree only and those with education beyond a bachelor's degree. So the lower of the two education levels includes teachers with credits beyond the bachelor's degree but no Master's degree. In this study, to obtain consistency among all of the districts, it was necessary to divide the teachers into two different groups: those with a Master's degree, and those without a Master's degree.

Figure D4–
Percentage of Third Grade Students with Teachers with Less than a Master's Degree, by District, 2000–2001



District A has the highest percentage of female teachers, 94 percent. The other five districts range from 81 to 86 percent female. The state average for K–3 teachers is 91 percent female. (See Table D1.)

Average Test Scores

Figure D5 shows a bar graph that compares 2001 third grade math scores with the prior year scores for the districts that provided scaled scores. Steady increase in math scores over the two years can be observed in all the five districts. The relative performance of districts is unchanged across the two years, showing the strong relationship that the 2001 test scores have with the prior year scores.[5] The comparable gain for reading scores is shown in Figure D6. It should be also noted that the five districts have average math and reading scores lower than the California state average in both years, except for District A's 2001 reading scores.

[5] District F is not shown in these graphs because they only provided normal curve equivalent scores instead of scaled scores, but the pattern of scores was similar.

Figure D5–
Math SAT-9 Scaled Scores for Second Grade (2000) and Third Grade (2001), by District

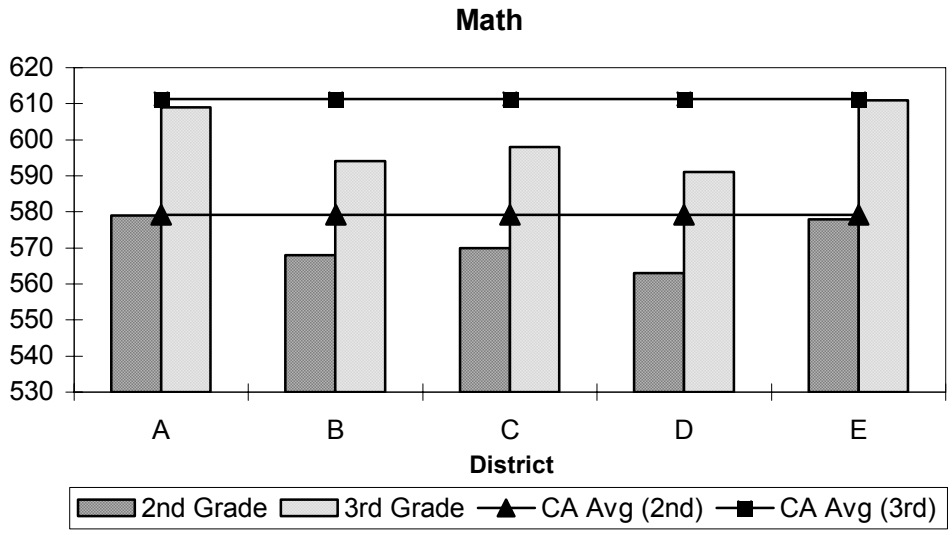
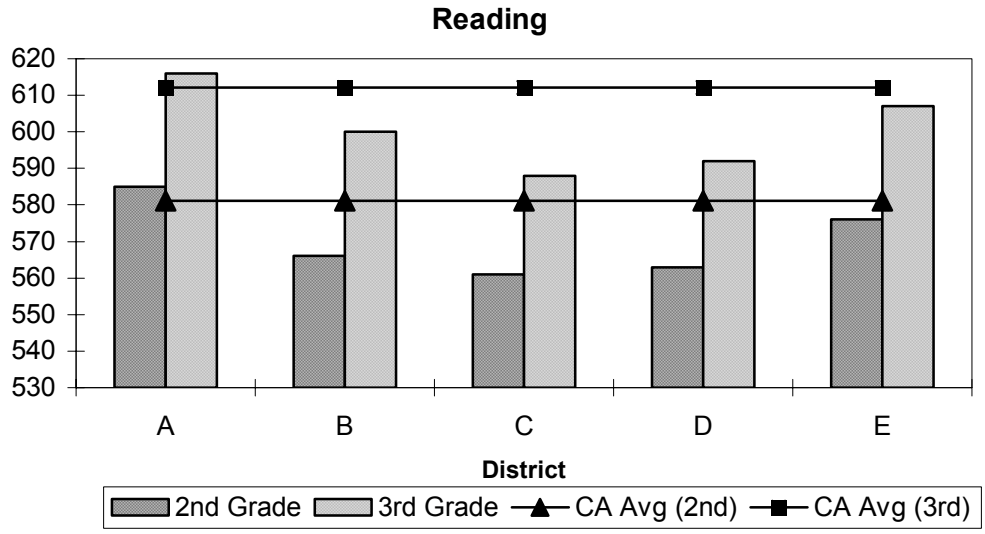


Figure D6–
Reading SAT-9 Scaled Scores for Second Grade (2000) and Third Grade (2001), by District



District E has the greatest increase in average math test scores between the two years. (See Table D1 for the average difference in math test scores.) In the graph of reading test scores it is District B that shows the highest average test score gains. (See Table D2 for the average difference in reading test scores.)

We now move to a discussion of the model that was used to assess student achievement in 2001 taking into account achievement in 2000 as well as other student, teacher, and classroom characteristics.

Model

We examined the relationship between achievement and teacher characteristics in two ways. We first modeled third grade student test scores in 2001 as a function of the 2000 second grade test score and student, teacher, and classroom characteristics:

$$(1)\quad Score_{2001} = \alpha_1 + Score_{2000} * \beta_1 + Student * \delta_1 + Teacher * \theta_1 + Classroom * \lambda_1 + \varepsilon_1$$

We also modeled gain scores (2001 SAT-9 score minus 2000 SAT-9 score) as a function of student, teacher, and classroom characteristics:

$$(2)\quad Score_{2001} - Score_{2000} = \alpha_2 + Student * \delta_2 + Teacher * \theta_2 + Classroom * \lambda_2 + \varepsilon_2$$

where *Student* stands for student characteristics, *Teacher* stands for teacher characteristics, and *Classroom* stands for classroom characteristics. All variables are described in Section 3. The error term is represented by ε.

We show the results from the first set of analyses resulting from equation (1) in the main text; interested readers can find the gain score results (equation (2)) in Tables A and B.

Given our interest in examining the effects of teacher characteristics net of not only student and classroom characteristics, but school effects as well, we used Hierarchical Linear Modeling (HLM) with random intercepts for classrooms and constant slopes. HLM (Raudenbush and Bryk, 2002) is a common analytical approach for nested data. Responses from individuals nested in the same unit (e.g. students within the same classroom) may tend to be more similar than responses from individuals who do not share a unit. Such correlation, if not properly accounted for (e.g. by OLS), can bias standard error estimates and lead to improper inferences. HLM introduces random effects that try to capture this extra correlation structure. Because we were primarily interested in classroom-level variables (e.g. teacher characteristics), we used random intercepts at the classroom level to improve our chances of making valid inferences.

The teacher characteristic variables examined for their effects on student achievement in 2001 were: Full Credential or not, Years Teaching, and Master's Degree or Higher or not. We controlled for student characteristics (gender, race/ethnicity, special education status, English learner status, and eligibility for free or reduced price lunch) and classroom variables (percentage of girls in the class, percentage of students in each race/ethnicity group, percentage in special education, percentage of English learners, and percentage of students eligible for free or reduced price lunch).

Results

None of the three teacher characteristics examined (years of experience, credential status, and education level) had a statistically significant relationship with achievement in more than two of the six districts. First year teachers in two of the six districts had students with 2001 test scores that were on average lower when compared to their colleagues with 10 or more years of experience (the omitted variable in the regressions reported in Tables D3 and D4 below). First year teachers had a negative effect on math achievement in all 6 districts, however the coefficient was only statistically significant at the 5 percent level in two of the districts (B and F). First year teachers had a negative effect on reading achievement in 5 of the 6 districts, again with only two districts having statistically significant coefficients (B and D).

Years of Teaching Experience

We note that most of the estimated regression coefficients for years of experience are negative, although many are not significant at the 5 percent level. This suggests that we could make more efficient inferences if we aggregated results across districts to create pooled confidence intervals for the coefficients. For each of reading and math, and for each of the 4 classes of years of experience (1, 2, 3, and 4 to 9, which are relative to teachers with 10 or more years of experience), we estimated pooled regression coefficients as follows. First, we assume that for district j, the regression coefficient β_j can be modeled as

$$\beta_j = \beta_0 + \gamma_j + \varepsilon_j$$

where γ_j are i.i.d. and normally distributed as $(0,\tau^2)$ and ε_j are independent of the γ_j and independent of one another with mean zero normal distributions. We assume that the variances of these distributions, denoted σ_j^2, are known and equal to the squared standard errors reported from the district-specific regression models. This model implies that the regression coefficients have a common mean across districts, and variance equal to $(\tau^2+\varepsilon_j)$. It allows for the possibility that there are district random effects that center the sampling distribution of the regression coefficient for that district, conditional on γ_j, at $\beta_0+\gamma_j$.

Using the observed regression coefficients $\hat{\beta}_j$, we estimate τ^2 using restricted maximum likelihood, a method that reduces bias in estimating variance components for random effects models when sample sizes are small. Our analyses are based on five districts (we exclude the NCE results for District F), so this technique is particularly appropriate. The estimates of τ^2 (denoted $\hat{\tau}^2$) for each subject and level of experience are given below:

	Years of Teaching Experience			
	1	2	3	4 to 9
Reading	2.31	0.00	0.00	1.03
Math	2.28	8.50	0.00	0.00

Estimates of zero imply that there appears to be no extra variability among districts in the estimated regression coefficients beyond what is reasonable given the level of sampling variance within districts. Alternatively, positive estimates suggest additional inter-district variability, which is important to account for in the pooled estimation to make appropriate inferences. Given estimates of τ^2 for each type of regression coefficient and the σ_j, we estimate the pooled regression coefficient (which is the restricted maximum likelihood estimate of β_0 in the model above) using a weighted average

$$\sum_j w_j \times b_j$$

where

$$w_j = \frac{\frac{1}{\hat{\tau}^2 + \sigma_j^2}}{\sum_j \left(\frac{1}{\hat{\tau}^2 + \sigma_j^2} \right)}$$

That is, each estimated regression coefficient is weighted by a factor proportional to its estimated precision, giving more weight to regression coefficients estimated with less variance. Because the estimator is a linear combination of independent normal random variables, its variance is readily available and is equal to

$$1 \Big/ \sum_j \frac{1}{\hat{\tau}^2 + \sigma_j^2}$$

This reduces to the usual variance of the sample mean when $\hat{\tau}^2 = 0$ and σ_j^2 are all the same. Also note that this variance is larger (i.e. more conservative) than what would be the variance for the usual pooled estimator where $\hat{\tau}^2$ identically equal to zero. By explicitly allowing for extra variability among districts, we are guarding against mistaken overconfidence in our inferences.

Figures D7 and D8 summarize our findings for math and reading, respectively. They plot the pooled estimated regression coefficients along with error bars representing two standard errors of the pooled estimate, using the formulas presented above. The estimates are in terms of scale score points on the respective 2001 SAT-9 examination. The figures suggest that there is a persistent negative relationship for first-year teachers, as the approximate 95 percent confidence intervals for the pooled first-year coefficient are entirely below zero (zero represents equivalence to teachers with 10 or more years of experience). Curiously, there is also a significant negative relationship for third year

teachers. Even though the confidence intervals for 2 years and 4 to 9 years contain zero, their point estimates are negative as well. In all, the pooled results suggest an overall negative effect of having less experienced teachers. The patterns are remarkably similar across the two subjects, suggesting that real structure is being captured.

Figure D7–
Math: Pooled Estimated Regression Coefficients for Years of Teaching Experience, with 95% Confidence Intervals

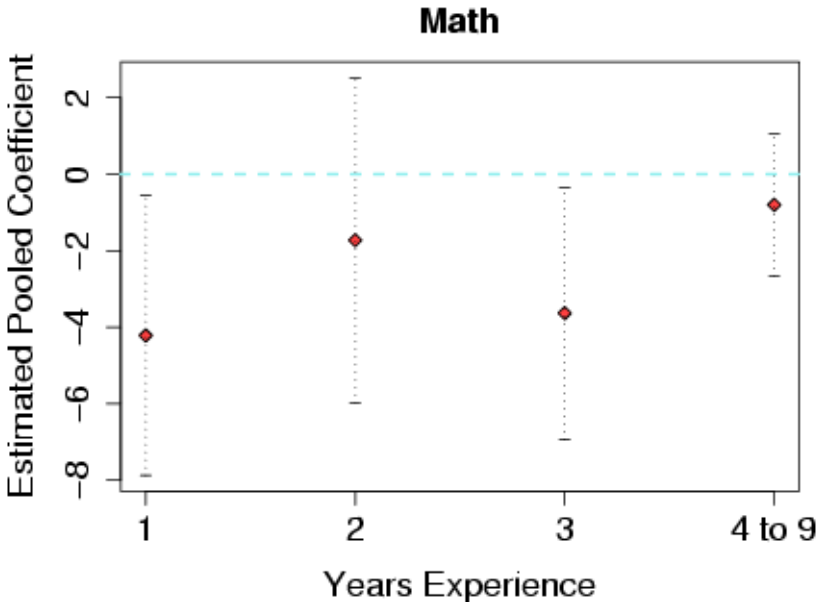
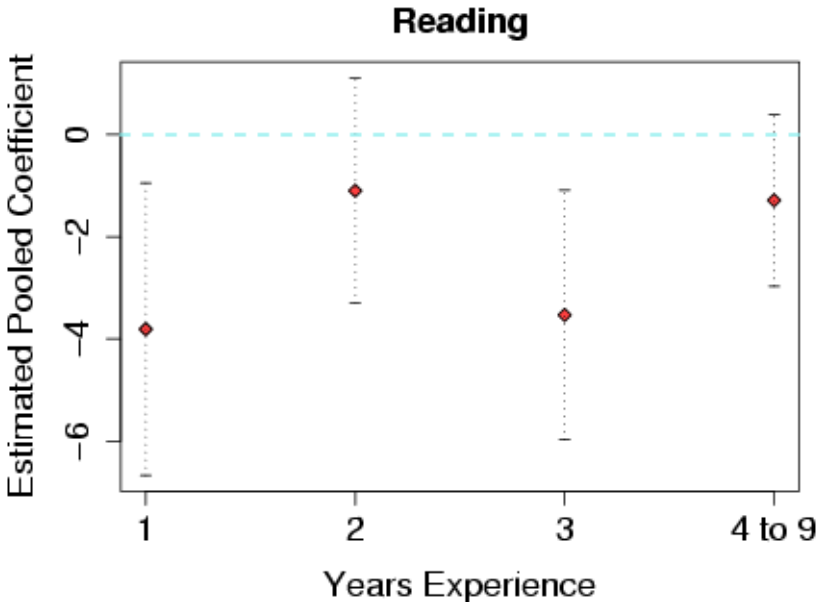


Figure D8–
Reading: Pooled Estimated Regression Coefficients for Years of Teaching Experience, with 95% Confidence Intervals



When interpreting these results, it is important to bear in mind the inability to make definitive causal inferences because of the observational nature of the data. The results imply that ceteris paribus, students with more experienced teachers tend to perform better than students with less experience teachers. However, there are many potential sources of bias and the results may be capturing the effect of unmeasured covariates. For example, the more experienced teachers may be able to select into the best schools and/or classrooms, and the covariates of the model may not fully capture the abilities of the students in these classrooms. The ostensible "third year slump" could be result of cohort variability and the particular timing of the study relative to CSR programmatic deployment in California, rather than predictive of a general pattern expected of teachers in their third year.

Teacher Education and Credential Status

The other teacher characteristic variables (teacher education and credential status) were not statistically significant when taking student and classroom characteristics into account. Only district E had a statistically significant relationship with student achievement, and then, only for math. Of the six districts, District E also has the lowest percentage of teachers with full credentials, 72 percent. Similarly, only one district had a statistically significant coefficient for teachers with masters degrees or higher, showing a negative relationship between reading scores and the highly educated teachers.

Student Characteristics

Most of the student characteristics were significantly related to achievement gains. For both math and reading, many of the gender, race/ethnicity, special education, English learner, and free/reduced price lunch variables were statistically significant. A student's prior year score was also a large and statistically significant predictor of student achievement consistently across districts. Some of the classroom demographic variables such as the percent of various types of students in the classroom were statistically significant.

Table D3– Regression Analysis– Hierarchical Linear Model (HLM): Dependent Variable SAT-9 Third Grade Math Scaled Score 2001.*

Math Scaled Score 2001	District A			District B			District C			District D			District E			District F**		
Variable	Coef	T	P	Coef	T	P	Coef	T	P	Coef	T	P	Coef	T	P	Coef	T	P
Intercept	203.10	14.34	<0.001	220.62	20.63	<0.001	294.08	20.69	<0.001	209.88	20.56	<0.001	220.67	29.73	<0.001	19.75	5.10	<0.001
Math Score 2000	0.72	48.66	<0.001	0.67	58.10	<0.001	0.59	47.65	<0.001	0.68	71.64	<0.001	0.69	83.14	<0.001	0.70	62.83	<0.001
Student Variables																		
Female (1= Female, 0= Male)	-1.16	-1.14	0.253	2.95	3.68	<0.001	0.09	0.12	0.905	0.75	1.18	0.239	0.79	1.35	0.178	1.04	2.50	0.013
Asian (1= Asian, 0= Not)	9.94	3.36	0.001	5.87	2.43	0.015	2.55	0.77	0.439	-0.49	-0.34	0.733	2.14	1.40	0.163	3.71	2.29	0.022
Black (1= Black, 0= Not)	-5.17	-0.57	0.572	-11.35	-5.29	<0.001	4.74	0.59	0.553	-4.29	-3.22	0.001	-7.43	-6.19	<0.001	-1.93	-2.58	0.010
Hispanic (1= Hispanic, 0= Not)	-4.90	-3.54	<0.001	-3.73	-1.54	0.125	-1.02	-0.40	0.693	-2.21	-2.12	0.034	-3.44	-2.76	0.006	-1.86	-2.70	0.007
Other Race (1= Other race, 0= Not)	-3.60	-1.80	0.073	-6.18	-2.05	0.041	-1.83	-0.39	0.699	-8.56	-3.09	0.002	-1.64	-0.98	0.325	1.33	0.87	0.386
Special Education (1= Special Ed, 0= Not)	-8.66	-4.02	<0.001	-7.23	-4.37	<0.001	-11.89	-1.98	0.048	-9.69	-8.22	<0.001	-8.09	-6.46	<0.001	.	.	.
English Learner (1= EL, 0= Not)	-1.60	-0.86	0.391	-5.14	-3.34	0.001	-9.93	-9.13	<0.001	1.26	1.34	0.181	1.68	1.60	0.111	1.64	2.64	0.008
Free/Reduced Price Lunch (1= F/RPL, 0= Not)	-1.90	-1.48	0.139	-0.76	-0.76	0.449	-1.77	-1.32	0.186	.	.	.	-3.29	-4.25	<0.001	-1.61	-2.69	0.007
Teacher Variables																		
Full Credential	2.05	0.35	0.725	0.62	0.21	0.836	-1.09	-0.57	0.568	10.86	1.61	0.108	4.46	2.26	0.024	1.59	1.06	0.289
Years Teaching = 1	-5.58	-1.00	0.317	-9.97	-2.19	0.029	-0.31	-0.11	0.915	-7.09	-1.58	0.115	-3.77	-1.21	0.225	-3.89	-2.14	0.032
Years Teaching = 2	-11.35	-2.26	0.024	-5.76	-1.32	0.185	0.61	0.23	0.816	0.66	0.14	0.887	1.71	0.59	0.554	-0.59	-0.31	0.756
Years Teaching = 3	-10.99	-1.80	0.072	-4.47	-1.11	0.266	-2.69	-1.09	0.274	-0.15	-0.02	0.981	-3.31	-1.04	0.298	2.10	1.23	0.218
Years Teaching = 4–9	-5.93	-1.63	0.103	0.83	0.32	0.749	0.05	0.03	0.977	-1.27	-0.70	0.482	-0.88	-0.45	0.655	1.35	1.05	0.295
Master's Degree or Higher	2.81	0.82	0.415	-3.27	-1.43	0.153	1.30	0.79	0.427	0.77	0.26	0.793	-1.25	-0.68	0.497	-0.18	-0.16	0.871
Female	-1.27	-0.23	0.820	2.02	0.79	0.429	2.26	1.19	0.233	-0.49	-0.23	0.816	0.69	0.32	0.748	2.54	1.89	0.059
Classroom Variables																		
% Female	0.15	1.31	0.192	0.12	1.52	0.129	0.05	1.26	0.208	0.04	0.63	0.531	0.13	1.96	0.050	0.03	0.95	0.343
% Asian	0.04	0.15	0.885	0.16	1.44	0.149	-0.09	-0.47	0.637	-0.12	-1.69	0.090	-0.13	-1.24	0.214	-0.01	-0.10	0.924
% Black	0.02	0.02	0.987	-0.02	-0.28	0.781	-1.03	-1.54	0.123	-0.15	-2.18	0.029	-0.24	-3.56	<0.001	-0.04	-0.78	0.434
% Hispanic	0.00	0.01	0.993	0.01	0.10	0.920	-0.17	-1.18	0.236	-0.12	-2.03	0.042	-0.27	-2.95	0.003	-0.06	-1.35	0.178
% Special Education	-0.07	-0.47	0.636	0.02	0.28	0.777	0.45	2.02	0.043	-0.13	-2.21	0.027	-0.10	-1.93	0.053	.	.	.
% English Learner	0.12	1.44	0.150	-0.06	-0.61	0.543	-0.14	-3.69	<0.001	-0.02	-0.41	0.683	0.06	0.81	0.418	0.00	-0.03	0.976
% Free/Reduced Prince Lunch	-0.30	-3.79	<0.001	-0.10	-1.79	0.073	-0.01	-0.14	0.891	.	.	.	0.06	1.64	0.101	-0.01	-0.16	0.870

* Shaded areas indicate statistical significance at the 5% level.

** SAT-9 scaled scores were not available for District F. Normal Curve Equivalent (NCE) scores were used.

Table D4– Regression Analysis- Hierarchical Linear Model (HLM): Dependent Variable SAT-9 Third Grade Reading Scaled Score 2001.*

Reading Scaled Score 2001	District A			District B			District C			District D			District E			District F**		
Variable	**Coef**	**T**	**P**	**Coef**	**T**	**P**	**Coef**	**T**	**P**	**Coef**	**T**	**P**	**Coef**	**T**	**P**	**Coef**	**T**	**P**
Intercept	141.39	12.48	<0.001	171.96	16.38	<0.001	328.67	26.95	<0.001	177.64	21.53	<0.001	193.87	30.09	<0.001	12.73	5.37	<0.001
Reading Score 2000	0.81	54.20	<0.001	0.78	60.54	<0.001	0.54	41.28	<0.001	0.76	78.97	<0.001	0.76	88.06	<0.001	0.76	63.48	<0.001
Student Variables																		
Female (1= Female, 0= Male)	-1.70	-1.70	0.090	3.30	4.20	<0.001	0.62	0.88	0.379	0.61	1.00	0.315	0.18	0.35	0.725	0.38	1.03	0.303
Asian (1= Asian, 0= Not)	-3.95	-1.38	0.168	-6.88	-2.96	0.003	-8.69	-2.95	0.003	-9.58	-6.99	<0.001	-3.92	-2.89	0.004	0.04	0.03	0.977
Black (1= Black, 0= Not)	-11.20	-1.11	0.266	-11.82	-5.68	<0.001	-1.84	-0.26	0.795	-8.74	-6.86	<0.001	-7.43	-6.98	<0.001	-1.73	-2.63	0.009
Hispanic (1= Hispanic, 0= Not)	-4.02	-3.01	0.003	-8.81	-3.74	<0.001	-7.84	-3.42	0.001	-5.40	-5.43	<0.001	-3.53	-3.18	0.002	-1.89	-3.13	0.002
Other Race (1= Other race, 0= Not)	-4.42	-2.27	0.023	-3.51	-1.20	0.229	-9.98	-2.37	0.018	-6.10	-2.33	0.020	-6.83	-4.67	<0.001	-1.35	-1.01	0.314
Special Education (1= Special Ed, 0= Not)	-6.86	-3.13	0.002	-5.97	-3.65	<0.001	-7.14	-1.34	0.180	-9.90	-8.63	<0.001	-8.49	-7.50	<0.001	.	.	.
English Learner (1= EL, 0= Not)	-3.30	-1.76	0.079	-3.34	-2.21	0.027	-13.74	-13.33	<0.001	-2.35	-2.59	0.010	-3.99	-4.25	<0.001	-1.27	-2.30	0.021
Free/Reduced Price Lunch (1= F/RPL, 0= Not)	-4.48	-3.55	<0.001	-1.39	-1.42	0.156	-1.73	-1.45	0.149	.	.	.	-4.20	-6.08	<0.001	-2.20	-4.15	<0.001
Teacher Variables																		
Full Credential	1.11	0.28	0.781	3.00	1.21	0.226	-2.58	-1.66	0.098	0.50	0.11	0.913	1.11	0.85	0.393	0.20	0.23	0.815
Years Teaching = 1	0.40	0.11	0.912	-8.02	-2.15	0.032	-1.60	-0.68	0.499	-8.50	-2.71	0.007	-3.20	-1.56	0.118	-1.55	-1.45	0.147
Years Teaching = 2	-5.05	-1.61	0.107	-1.87	-0.52	0.603	-1.55	-0.72	0.470	0.44	0.14	0.887	0.14	0.08	0.940	-1.45	-1.32	0.188
Years Teaching = 3	0.72	0.18	0.856	-6.27	-1.87	0.061	-2.06	-1.02	0.306	-3.16	-0.71	0.475	-5.33	-2.52	0.012	0.45	0.46	0.647
Years Teaching = 4–9	3.24	1.43	0.153	-0.32	-0.15	0.883	-1.79	-1.17	0.241	-1.41	-1.17	0.243	-3.20	-2.47	0.013	-0.94	-1.26	0.208
Master's Degree or Higher	-0.05	-0.02	0.983	-3.99	-2.09	0.036	0.31	0.23	0.815	1.36	0.69	0.491	-0.78	-0.64	0.521	-0.46	-0.73	0.463
Female	0.78	0.22	0.824	5.42	2.55	0.011	1.72	1.12	0.265	0.06	0.04	0.966	1.06	0.75	0.456	1.24	1.58	0.113
Classroom Variables																		
% Female	0.18	2.33	0.020	0.11	1.71	0.087	-0.02	-0.43	0.667	0.01	0.22	0.825	0.03	0.55	0.582	0.05	2.24	0.025
% Asian	-0.11	-0.62	0.535	0.02	0.26	0.794	-0.06	-0.44	0.661	-0.07	-1.53	0.127	-0.12	-1.72	0.086	-0.11	-1.61	0.107
% Black	-0.56	-0.98	0.328	-0.11	-1.72	0.085	-1.05	-1.97	0.048	-0.14	-2.88	0.004	-0.21	-4.62	<0.001	-0.04	-1.30	0.194
% Hispanic	-0.09	-1.32	0.186	0.03	0.25	0.799	-0.19	-1.85	0.065	-0.09	-2.11	0.035	-0.19	-3.08	0.002	-0.03	-1.20	0.232
% Special Education	0.05	0.37	0.710	-0.06	-0.77	0.439	0.28	1.45	0.147	-0.02	-0.44	0.660	-0.14	-3.56	<0.001	.	.	.
% English Learner	0.13	2.29	0.022	-0.14	-1.88	0.061	-0.10	-3.31	0.001	-0.02	-0.63	0.530	0.04	0.69	0.492	0.01	0.33	0.744
% Free/Reduced Prince Lunch	-0.09	-1.81	0.071	-0.09	-2.00	0.046	0.01	0.11	0.913	.	.	.	-0.03	-1.21	0.225	-0.01	-0.59	0.556

* Shaded areas indicate statistical significance at the 5% level.

** SAT-9 scaled scores were not available for District F. Normal Curve Equivalent (NCE) scores were used.

Conclusions

Student characteristics were related to achievement gains in all cases, and teacher characteristics were generally not related to achievement gains. The one exception was years of teaching experience, where pooling results across districts indicated a significantly negative effect for first (and third) year teachers compared to teachers with ten or more years of experience. In fact, the estimated effects for all experience classes relative to the most experienced teachers were negative, for both math and reading. The other teacher characteristics, teacher education and certification status, appeared to be unrelated to student achievement after controlling for student and classroom characteristics.

It is important to distinguish between teacher quality and teacher characteristics. Rivkin and Hanushek (2002) look at the relationship between student performance and differences among teachers. They find that there are quality differences among teachers but that the differences are not explained by the commonly used measures such as teacher education, experience, and certification.

The results of our study do not imply that teacher characteristics do not matter, but that the teacher variables that are easily measured and conveniently obtained do not seem to matter. There are likely to be many variables that can be observed by the school district such as teacher motivation and energy, that are very important and also difficult to measure.

Discussion

The purpose of our study was to examine how teacher characteristics are associated with student achievement scores in reduced-size classes in large school districts. The results from our analyses on six of California's large districts indicated that students with less experienced teachers do not perform as well as students with teachers who have ten or more years of experience. Other teacher characteristics did not appear to be associated with student achievement scores in reduced-size classes.

Since this research was only conducted using 6 districts, generalization of results is somewhat limited. Nonetheless, these districts represent approximately 6 percent of California third graders.

By 2000–2001, the fifth year of class size reduction in California, (but the first year we have available data). Therefore, this study may miss the possible effects of the initial inflow of new and less than fully credentialed teachers that may have occurred mainly in the first three years of CSR implementation. Furthermore, one year's time may not be sufficient to see effects due to differences in teacher characteristics.

Supplement to Appendix D

Table D5– Regression Analysis- Hierarchical Linear Model (HLM): Dependent Variable SAT-9 Third Grade Math Gain Score 2000–2001.*

Math Gain Score	District A			District B			District C			District D			District E			District F**		
Variable	Coef	T	P	Coef	T	P	Coef	T	P	Coef	T	P	Coef	T	P	Coef	T	P
Intercept	36.73	3.10	0.002	13.01	1.56	0.120	33.65	2.61	0.010	9.33	1.12	0.263	24.19	4.38	<0.001	-1.82	-0.44	0.664
Student Variables																		
Female (1= Female, 0= Male)	-0.01	-0.01	0.995	3.58	4.03	<0.001	2.21	2.50	0.013	2.37	3.37	0.001	2.08	3.24	0.001	1.43	3.13	0.002
Asian (1= Asian, 0= Not)	5.00	1.58	0.115	3.91	1.46	0.144	-1.70	-0.45	0.651	-0.03	-0.02	0.984	-0.08	-0.05	0.963	3.00	1.68	0.093
Black (1= Black, 0= Not)	2.22	0.23	0.822	-0.87	-0.37	0.712	2.45	0.27	0.788	1.64	1.12	0.263	-2.61	-1.99	0.046	0.92	1.12	0.262
Hispanic (1= Hispanic, 0= Not)	-1.57	-1.06	0.288	-1.06	-0.39	0.695	2.57	0.88	0.380	0.99	0.86	0.389	-1.50	-1.10	0.272	-0.16	-0.22	0.827
Other Race (1= Other race, 0= Not)	1.19	0.55	0.579	-1.31	-0.39	0.695	5.40	1.01	0.314	-5.14	-1.68	0.094	-0.68	-0.37	0.709	3.85	2.29	0.022
Special Education (1= Special Ed, 0= Not)	-1.89	-0.83	0.409	-0.82	-0.45	0.651	-6.51	-0.95	0.341	-2.27	-1.77	0.077	-2.22	-1.62	0.105	.	.	.
English Learner (1= EL, 0= Not)	1.12	0.56	0.575	1.84	1.09	0.275	0.99	0.84	0.403	2.15	2.05	0.040	2.19	1.90	0.058	1.56	2.28	0.023
Free/Reduced Price Lunch (1= F/RPL, 0= Not)	1.03	0.75	0.455	0.44	0.40	0.692	-2.14	-1.41	0.158	.	.	.	-0.66	-0.77	0.440	-0.60	-0.91	0.362
Teacher Variables																		
Full Credential	4.69	0.75	0.451	0.54	0.17	0.864	-2.70	-1.31	0.191	11.06	1.64	0.102	4.01	1.96	0.050	1.94	1.17	0.241
Years Teaching = 1	-4.01	-0.67	0.501	-7.87	-1.63	0.104	-3.64	-1.16	0.245	-3.48	-0.77	0.440	-3.82	-1.18	0.237	-3.90	-1.95	0.052
Years Teaching = 2	-10.39	-1.93	0.054	-5.41	-1.17	0.240	-1.16	-0.41	0.685	0.71	0.15	0.877	1.89	0.63	0.529	0.15	0.07	0.943
Years Teaching = 3	-9.10	-1.39	0.165	-2.97	-0.70	0.486	-3.39	-1.27	0.203	3.75	0.58	0.560	-2.98	-0.90	0.368	3.20	1.70	0.089
Years Teaching = 4–9	-5.44	-1.40	0.163	1.56	0.57	0.569	-1.01	-0.50	0.616	-0.61	-0.34	0.735	-0.37	-0.18	0.855	0.90	0.63	0.528
Master's Degree or Higher	1.96	0.53	0.594	-3.64	-1.50	0.134	0.87	0.49	0.625	-0.06	-0.02	0.983	-1.12	-0.58	0.559	0.31	0.26	0.795
Female	-3.07	-0.51	0.608	2.30	0.85	0.397	1.51	0.74	0.460	-0.50	-0.24	0.811	0.73	0.33	0.743	2.82	1.91	0.056
Classroom Variables																		
% Female	0.12	0.92	0.357	0.11	1.33	0.185	0.04	0.81	0.415	0.03	0.55	0.581	0.04	0.56	0.578	0.01	0.36	0.718
% Asian	-0.08	-0.27	0.786	0.14	1.18	0.238	-0.06	-0.32	0.752	0.05	0.68	0.495	-0.01	-0.13	0.894	0.00	0.00	0.998
% Black	0.14	0.13	0.893	0.06	0.69	0.492	-0.42	-0.58	0.560	0.01	0.20	0.845	-0.08	-1.22	0.221	0.00	0.01	0.991
% Hispanic	0.02	0.15	0.880	0.04	0.30	0.763	-0.02	-0.16	0.872	0.06	1.04	0.300	-0.11	-1.11	0.266	-0.07	-1.28	0.201
% Special Education	-0.05	-0.29	0.770	0.13	1.43	0.154	0.38	1.51	0.131	-0.04	-0.68	0.496	0.02	0.33	0.742	.	.	.
% English Learner	0.09	1.05	0.294	0.05	0.48	0.630	-0.06	-1.35	0.179	0.00	0.05	0.957	0.08	1.00	0.317	0.02	0.76	0.445
% Free/Reduced Prince Lunch	-0.28	-3.33	0.001	-0.06	-1.13	0.258	-0.03	-0.38	0.707	.	.	.	0.10	2.40	0.017	0.04	0.93	0.353

* Shaded areas indicate statistical significance at the 5% level.

** SAT-9 scaled scores were not available for District F. Normal Curve Equivalent (NCE) scores were used.

Table D6– Regression Analysis- Hierarchical Linear Model (HLM): Dependent Variable SAT-9 Third Grade Reading Gain Score 2000–2001.*

Reading Gain Score	District A			District B			District C			District D			District E			District F**		
Variable	Coef	T	P	Coef	T	P	Coef	T	P	Coef	T	P	Coef	T	P	Coef	T	P
Intercept	31.78	4.39	<0.001	34.87	5.17	<0.001	33.38	3.54	<0.001	29.33	5.03	<0.001	43.03	11.99	<0.001	-2.49	-1.05	0.297
Student Variables																		
Female (1= Female, 0= Male)	-2.28	-2.20	0.028	2.12	2.61	0.009	0.27	0.33	0.741	-0.11	-0.18	0.859	-0.75	-1.36	0.173	-0.24	-0.62	0.534
Asian (1= Asian, 0= Not)	-5.74	-1.93	0.054	-5.86	-2.42	0.015	-9.44	-2.77	0.006	-6.84	-4.71	<0.001	-4.42	-3.08	0.002	-0.38	-0.25	0.801
Black (1= Black, 0= Not)	-6.24	-0.60	0.550	-5.00	-2.35	0.019	-11.13	-1.36	0.174	-5.46	-4.06	<0.001	-4.18	-3.74	<0.001	-0.48	-0.70	0.486
Hispanic (1= Hispanic, 0= Not)	-1.74	-1.26	0.207	-5.85	-2.40	0.017	-4.20	-1.59	0.113	-3.19	-3.03	0.002	-1.59	-1.36	0.175	-0.87	-1.37	0.171
Other Race (1= Other race, 0= Not)	-2.51	-1.25	0.213	0.84	0.28	0.782	-8.02	-1.66	0.098	-2.85	-1.03	0.304	-6.16	-3.99	<0.001	-0.65	-0.46	0.647
Special Education (1= Special Ed, 0= Not)	-2.52	-1.12	0.263	-2.19	-1.30	0.194	-8.98	-1.46	0.144	-4.63	-3.87	<0.001	-4.31	-3.63	<0.001	.	.	.
English Learner (1= EL, 0= Not)	0.79	0.41	0.682	2.35	1.53	0.126	1.47	1.36	0.174	-0.76	-0.79	0.430	-2.19	-2.21	0.028	-1.30	-2.24	0.025
Free/Reduced Price Lunch (1= F/RPL, 0= Not)	-2.56	-1.97	0.049	-0.003	0.00	0.997	-0.40	-0.29	0.773	.	.	.	-2.20	-3.03	0.003	-1.30	-2.34	0.019
Teacher Variables																		
Full Credential	0.33	0.08	0.936	2.63	1.03	0.304	-4.50	-2.73	0.006	-0.51	-0.11	0.915	0.62	0.48	0.632	0.24	0.25	0.799
Years Teaching = 1	1.29	0.35	0.728	-5.59	-1.45	0.146	-3.14	-1.26	0.209	-6.36	-1.97	0.049	-3.43	-1.68	0.093	-1.61	-1.41	0.158
Years Teaching = 2	-4.49	-1.40	0.163	-1.72	-0.46	0.643	-2.64	-1.16	0.245	0.56	0.18	0.860	0.24	0.13	0.900	-1.12	-0.96	0.339
Years Teaching = 3	2.92	0.72	0.472	-5.16	-1.50	0.135	-2.13	-1.00	0.318	-1.75	-0.39	0.700	-5.15	-2.43	0.015	0.79	0.75	0.451
Years Teaching = 4–9	4.16	1.79	0.073	0.60	0.27	0.787	-1.95	-1.21	0.228	-0.83	-0.67	0.501	-3.33	-2.58	0.010	-1.05	-1.33	0.184
Master's Degree or Higher	-0.65	-0.30	0.766	-4.30	-2.19	0.029	-0.22	-0.15	0.879	1.30	0.65	0.518	-0.80	-0.66	0.509	-0.57	-0.85	0.393
Female	0.09	0.03	0.980	5.69	2.60	0.010	1.14	0.69	0.488	0.06	0.04	0.966	0.74	0.52	0.605	1.29	1.56	0.119
Classroom Variables																		
% Female	0.16	2.03	0.043	0.12	1.81	0.070	-0.03	-0.68	0.499	0.00	0.07	0.944	-0.03	-0.67	0.504	0.04	1.90	0.057
% Asian	-0.25	-1.46	0.146	0.02	0.19	0.853	-0.02	-0.11	0.914	0.04	0.73	0.468	-0.05	-0.79	0.427	-0.11	-1.58	0.113
% Black	-0.66	-1.12	0.262	-0.06	-0.90	0.367	-0.44	-0.78	0.433	-0.03	-0.61	0.544	-0.09	-2.06	0.040	-0.01	-0.38	0.702
% Hispanic	-0.10	-1.46	0.145	0.05	0.51	0.607	0.02	0.14	0.890	0.05	1.16	0.246	-0.08	-1.30	0.194	-0.02	-0.72	0.473
% Special Education	-0.01	-0.08	0.939	-0.03	-0.41	0.683	0.28	1.28	0.201	0.05	1.00	0.319	-0.08	-1.96	0.050	.	.	.
% English Learner	0.14	2.42	0.016	-0.08	-0.97	0.332	-0.01	-0.27	0.785	0.01	0.32	0.748	0.08	1.59	0.111	0.03	1.41	0.157
% Free/Reduced Prince Lunch	-0.06	-1.17	0.242	-0.07	-1.46	0.145	0.02	0.36	0.720	.	.	.	-0.03	-1.20	0.230	0.01	0.53	0.593

* Shaded areas indicate statistical significance at the 5% level.

** SAT-9 scaled scores were not available for District F. Normal Curve Equivalent (NCE) scores were used.

References

Darling-Hammond, L. (2000). Teacher Quality and Student Achievement: A Review of State Policy Evidence. *Education Policy Analysis, 8,* 1.URL http://epaa.asu.edu/epaa/v8n1/**.**

Raudenbush, S.W. and Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods.* Newbury, CA: Sage Publications

Rivkin, S.G. and Hanushek, E.A. (forthcoming). Prepared for Research Seminar on Instructional and Performance Consequences of High-poverty Schooling, The Charles Sumner School Museum and Archives, Washington, D.C. March 11, 2002.

Stecher, B.M. and Bohrnstedt, G.W. (Eds.). (2002). *Class size reduction in California: Findings from 1999–00 and 2000–01*. Sacramento, CA: California Department of Education.

Appendix E

State Class Size Reduction Measures

Reprinted by permission from ECS Information Clearinghouse
Updated June 1999 *(Additional updates by WestEd Staff May 2002)*

The following table targets states that have attempted to limit the student-teacher ratio to at most 20:1. However, several "marginal" class size reduction measures that do not meet that limit also are included.

State	Category (type)	Year Enacted	Description	Notes	Funding
AK	mandate	1997 amended 1998	K–4 targeted; 20 students per class.	Legislation authorizes formation of smaller classes and provides funding for those schools choosing to do so.	1999-00 funding: $1.58 million.
AL	mandate	1997 amended 1998	State board resolution sets a timetable and limits. K-3, 18 students per teacher.	Classes with aides reviewed as an exception by the state supt. of education.	Through the 1995 Foundation Program Plan; legislature approved $127 million to pay for up to 900 new teachers to help reduce class size
CA	voluntary/ incentive Cal. Chap. 6.10, §52120	1996	Legislation authorized formation of smaller classes and provided funding for those schools choosing to do so. Initial targets: 20 in K-3; grade 4 added in 97-98 Additional $200 million for 8,000 additional classrooms, either through remodeling or use of portables. The appropriation for new facilities is a one-time provision, while class-size reduction funds are expected to be included annually in the state budget.	Legislation also mandated independent evaluation by 3-28-98. Approximately 20,000 new teachers were needed to accommodate the smaller class sizes, which prompted the governor to sign a bill relaxing teacher certification requirements. Raises concerns about districts hiring unqualified teachers. Other unintended consequences: a surge of teachers moving from "less-advantaged" to more desirable districts to fill newly created staff positions; a shortage of substitute teachers; supervision and training of non-certificated teachers, creating a problem for higher education teacher training programs.	$1 billion 96-97 ($650 per student in smaller classes), $200 million for facilities $1.5 billion 97-98 ($800 per student)

State	Category (type)	Year Enacted	Description	Notes	Funding
				State-sponsored study by the Class Size Reduction Research Consortium reports that massive spending has resulted in, at best, marginal gains for third graders; teacher quality severely affected. (www.pacificresearch.org/press/opd/1999/99-07-07td.html)	
FL	voluntary No law, just funding	1996	Targets K-3 classrooms with a priority to Kindergarten and 1st grade; 20 students per teacher or 20+ (no more than 30 students) if a full-time aide is provided.	State Supreme Court approved a citizen initiative to appear on the 2002 November ballot, which would create a statewide cap for class size in public schools; initial cost estimates range between $2.5 and $12.5 billion. 1998 statewide study conducted by the Florida Department of Education's Office of Policy Research: "The relationship of school and class size with student achievement in Florida: An analysis of statewide data." www.firn.edu/doe/bin00048/home0048.htm.	1999-00 funding: $100 million.
IL	mandate (initially a pilot program) 105 Ill.Comp.Stat. 5/2-3.51	1988 1997	K-1 target of 18 students; 2-3 grades target of 20 students. Reading Improvement Block Grant Program authorized grants to improve reading instruction through several measures, one of which is to reduce class size K-3.	Teachers have reported improved student behavior, higher test scores and more efficient classrooms. However, program evaluations indicate a weak relationship between lower class size and student achievement, but significant improvement in teachers' morale and attitudes.	1999-00 funding: $5 million.
IN	pilot initially Ind. Code §21-1-29-1 mandate statewide 88-89 Ind. Code §§1-1-30-1 to 1-1-30-9	1981 1988	"Prime Time" program; programs decided at local level, while funds allocated at district level. 1995-99: 18-20 students. 1999-00: 15-18 students, depending on at-risk index.	Of 293 districts, only 3 do not participate.	Through funding formula 1995: $77 million
IA	grants IOWA CODE §§ 256E.2 to 256E.6	1999	Early intervention block grant program with goals to provide resources necessary to reduce class sizes in basic skills to 17:1 for K-3. Overall aim is improvement in reading instruction.	Flexibility in how funds used (not limited to class size reduction), but districts must develop class size management plan with goals of 17:1 for grades K-3. Must integrate plan into required comprehensive school improvement plan. Dollars received must supplement, not supplant. Requires annual public reporting on reading proficiency levels and class size. Limited evaluation of School	1999-00, $10M; 2000-01, $20M; 2001-03, $30M each year. Allocation formula targets low-income districts.

State	Category (type)	Year Enacted	Description	Notes	Funding
				Within a School program targeted at at-risk 7th-10th graders, which included reduced class size; increased performance between 1989-92 on progress toward graduation, attendance and dropout measures. (ERIC ED371045)	
LA	mandate LA.Rev. Stat. Ann. §17:174	1986	K-3 classes not to exceed 20 unless authorized in writing by the state superintendent.	Students above the maximum not to be counted for funding purposes. No provision of this measure to take effect until funds appropriated specifically by the legislature. Districts that meet requirement can use CSR funds to hire certified teachers for other grades.	1999-00 funding: $31.94 million
ME	voluntary/grants ME. Rev.Stat. Ann.tit. 20, §4252	1989	Local units may elect to target class size within one or more grades, K-3. Recommendation of 15 to 1, with a maximum of 18 to 1.		competitive grant program
MD	Ann. Code of M.D. Sec. 5-212	1999	Requires districts to submit plans and reports describing how they will use additional funds for any/all of several areas; one option is reducing 1st and 2nd grade reading program to no more than 1:20; another is reducing math instruction – 7th grade- to no more than 1:20.	Funds appropriated if, in the opinion of the state superintendent, the plan meets conditions prescribed by the legislature.	1999-00 funding: $11.6 million.
MN	mandate M.N. Stat. §126C.12	1993 1999	Learning and Development Program. State's program strives to reduce class size to 17 students in K-6. Significantly expanded by an additional $100M over 2 years; districts' must first target K-1.	As of FY2003, CSR revenue will fund additional teaching staff only in grades K-3.	State funding of $134.8 million, plus $2.9 million from another program. Additional funding of $98 million over 1999-00 and 2000-01 school years is directed at K-3 grades.
NC	voluntary N.C. Gen. Stat. §115C-301	1993 1995, 1997	Measure targeted to K-2, with a 1:23 ratio. Pilot in Burke County Schools, 1991+	Funded 1:23 for each grade, but allowing administrative units to use dollars to reduce K-2 or to hire reading teachers within K-2 or otherwise reduce the ratio within kindergarten through 2nd.	foundation
NV	mandate Nev. Rev. Stat. §388.700	1989 revised 1993, 1995	Legislature limited class size in K-3 to 15 (core subjects) School districts and licensed personnel association(s) must develop plan to reduce class sizes in grades 1-3 within limits of available financial support.	Legislature appropriated $450,000 for professional development. A questionnaire revealed that principals, teachers and parents believe smaller class sizes are associated with new teaching practices, increased teacher-student interaction, positive student attitudes toward learning and improved grades. Districts reported that fewer	Special revenue fund for class-size reduction Nev. Rev. Stat. §388.730 1999-00 funding: $82.9 million.

State	Category (type)	Year Enacted	Description	Notes	Funding
				special education referrals and less teacher absenteeism were associated with class-size reductions. More in-depth evaluations show student achievement levels remained the same when small classes were compared with larger classes (tested over a three-year period). In some districts, however, students in smaller classes (1-20) did significantly better in reading and moderately better in math than students in classes of 21 and over.	
NY	mandate	1997	Class Size Reduction Program: Goal is to reduce class size in grades K-3 to 20 students beginning in school year 1999-00.	Funds can be used for teacher salaries and benefits and for one-time startup costs for each new classroom; funds cannot be used for professional development or for new buildings.	1999-00 funding: $75 million; 2000-01: $150 million; 2001-02: 225 million.
OH	voluntary O.H. 122nd Gen. Assem. Am. Sub. H.B. 650	1997	To reduce class size in grades K-1.	As of December 1999, targeted funding for large districts (i.e. Columbus) to assist low-performing schools in reducing class-size from 25 to 15; 30% of school district qualify for state funding. 2000 report issued by the Legislative Office of Education Oversight; findings relate to implementation issues rather than achievement gains.	$131 million. Districts receive funding through the Disadvantaged Pupil Impact Aid fund.
OK	mandate 70 Okl. St. @ 18-113.1	1990 2002	Targets grades K, 1-3, 4-6. No more than 20 students may be regularly assigned to a teacher. With the exception of certain conditions (these vary by grade levels above), fiscal and accreditation penalties apply for noncompliance. Target expanded to include a limit of 140 students per day for each teacher of grades 7-12.	If limitations exceeded after the first 9 weeks of the year, no fiscal penalty applies. Physical education, music, vocational not subject to limitation. If classrooms are not available and district meets certain guidelines (has maximum millage allowable or voted indebtedness within 5 prior years), then district not penalized.	Funding addressed through foundation program.
RI	voluntary/ grants R.I. Gen. Laws §16-67-2	1987 (eff. 88-89); re-enacted 1996	Districts encouraged to reduce class size to no more than 15 in grades K-3 (The Literacy Program).		Educational Improvement block grants R.I. Gen. Laws §16-5-31 (3)
SC	mandate S.C. Code Ann. §59-20-40 mandate S.C. Code Ann. § 59-139-10	1977 1993	To qualify for funds, each district required to attain 21 to 1 average pupil-teacher ratio in basic skills of reading and mathematics (grades 1-3); districts may apply to the state board for waivers (phased in from 1979 to 1983) Early Childhood Development and Academic Assistance requires		Funding is addressed through foundation program (Kindergarten weighted 1.30; primary 1-3, 1.24) 1999-00 funding: $34.16 million. State addressing facilities

State	Category (type)	Year Enacted	Description	Notes	Funding
	voluntary/ grants	1997	districts to design long-range plans which may include reduction in K. pupil-teacher ratio (the class size component here is voluntary, but the plan is mandatory) 1-3 goal is to reduce student-teacher ratio to 15-1 over time.	Matching grant program (districts match state funds based on ability to pay). Priority given to districts with most critical needs and with higher percentages of low-income students.	need through a $750 million school facility bond bill for FY2000.
SD	voluntary/ grants S.D. Codified Laws § 13-14-8.1	1993	Youth-at-risk funds (grants) offered as incentives for reducing class sizes in K-3 to 15 or less.		grants for up to 3 years
TN	pilot Tenn. Code Ann. §49-6-3501 mandate 1985 Tenn. Pub. Acts, Ch. 463, 1	1984 1985	Demonstration centers (operated by local boards) established with class maximum enrollment 17. Two hundred teaching positions were funded by the department of education. Every public school system required to have a policy that pupil-teacher ratios not exceed ratio prescribed. Within a building, the average of any grade level cannot exceed the average, although any individual class within the unit may exceed the average (but not the maximum). K-3 avg: 20 (maximum of 25).	Purpose of the demonstration projects and centers was to study the effects of reduced pupil-teacher ratio on the achievement of students in public school. First study began in 79 elementary schools in 1985. Greatest gains in inner-city small classes. Classes with teacher aides achieved slightly higher scores than regular classes, but differences were not statistically significant. (Project STAR - Student Teacher Achievement Ratio) Longitudinal study funded in 1990 (Lasting Benefits Study) -- see p. 6 for details.	All but 5% of costs paid by the department of education. Funding provided through the foundation program (weighting).
TX	mandate Tex. Educ. Code Ann. §25.112 §25.111	1984 1995	School district may not enroll more than 22 students in K-4 classes. Stipulates ratio of not less than one teacher to each 20 students in average daily attendance (K-4). As of 2002, the target is expanded to include pre-K limits of 15:1.	Numerous exceptions apply. 1991 study conducted by Harvard University professor, Ronald Ferguson: "Using data from more than 800 district representing over 2.4 million students in Grades 1 through 7, and using student/teacher ratio as a measure of class size, Ferguson found that student achievement fell as the student/teacher ratio increased for every student above an 18 to 1 ratio." ("Paying for public education: New evidence on how and why money matters," Harvard Journal on Legislation 28 (2), 1991, pp. 465-98)	unknown

State	Category (type)	Year Enacted	Description	Notes	Funding
UT	mandate Utah Code Ann. §53A -17a-124.5	1992	Through use of appropriations, districts must reduce average class size in grades K-4, with emphasis on K-2. Must use 50% of allocation to reduce class size in K-2, with emphasis on improving reading skills. If average class size is below 18 in K-2, may petition the state board for waiver to use its allocation for reduction in other grades.	20% of district's allocation may be used for capital facilities projects that will help to reduce class size. 2000 legislative audit found that almost 50% ($148M of $344M) of CSR spending could not be identified because it was co-mingled with other funds; districts cannot account for CSR expenditures.	Funding formula (weighted pupil units) allocated $46,311,678 in 1997 to be dispersed over four years (ending with fiscal year beginning July 1, 2000); 1996: $19,544,621; 1995: $18,632,768; 1994: $15,451,271; 1993: $11,053,098; 1992: $4,389,540
VA	voluntary Va. Code Ann. §22.1-199.1	1996	Legislature established long-term goal of reducing pupil-teacher ratio and class size for K-3 in those schools with high or moderate concentrations of at-risk students.	1992 study of 31 elementary schools in Fairfax found that first graders that had been placed in smaller classes (average size 15) had a 75% passing rate in second grade, compared to 54% of those that had been in larger classes (average size 22). (http://www.picket.com/class/research.htm)	State funding based on the incremental cost of providing the lower class sizes according to the greater of the division average per-pupil cost of all divisions or the actual division per-pupil cost. Local districts must provide matching funds based on the composite index of local ability to pay. State Board of Education to budget accordingly.
WA	Vol. Public Law 106-113 I-728 ballot initiative	1986 2000	K-4 (Prior to 1999-00, targeted grades were K-3). To enhance staffing in grades K-4 by funding an additional 4.2 certificated instructional staff (CIS) per 1,000 FTE students in grades K-3 (over the minimum of 49 CIS) and an additional 7.2 CIS in grade 4 (over the minimum of 46 CIS). Provides an additional $184 million for class size reduction.	In 1998-99, participation rate was 96.4%.	$99 million.
WI	voluntary/ grants 1995 Act 27 Chapter 118.43	1995	Student Achievement Guarantee in Education (SAGE); districts eligible to enter 5-year achievement guarantee contract with Dept. of Public Instr. on behalf of one school if minimum of 30% low-income students and no preschool-grade 5 grant on behalf of that school. (Also implements curricular and programmatic reqmts.)	Targets K, 1st grade in 98-99; adds grade 2 in 99-2000; adds grade 3 in 2001-2003. Class size reduction is one of several reqmts. for grants; schools must also extend hours of operation, provide rigorous curriculum, create staff development and accountability programs and pass annual review.	Finance formula funds reduction in class size to 1: 15 in each SAGE classroom.