CHAPTER 1: POLICY CHOICES AND CLASS SIZE REDUCTION

INTRODUCTION AND SUMMARY
CSR receives wide support among practitioners and parents. This support exists despite its large cost and mixed findings of effectiveness in the research. Considerable research shows limited or no change in student performance after class size was reduced or in smaller classes. The Tennessee STAR experiment, on the other hand, showed immediate and long-lasting improvements in student performance after CSR. The one area of relative agreement among CSR researchers is that there is a lack of a theoretical understanding of how class size affects student performance. In this context of uncertainty about CSR implementation, policy decisions matter. Policy decisions can affect the costs of CSR, both pecuniary and non-pecuniary, and the ability of schools and districts to implement it. This chapter reviews what CSR is, the existing research on CSR, some of the recent California experiences in implementing CSR in kindergarten through third grades (K–3), and how policy choices matter.

CLASS SIZE REDUCTION: WHAT IS IT?
“There is more time to teach. The class is more cohesive and dynamic. There is room for students to spread out and complete assignments without being distracted by noise or proximity. I can work one-on-one to ensure student comprehension and success. There is harmony and learning taking place in my classroom.”

Veteran teacher in El Rio, CA, after CSR implementation in her classroom (Bartell & Weiss, 1997)

Generally CSR policy is formulated at the state or federal level to be implemented at the local level by school districts. The general form is a state or federal policy that sets a class size goal for a set of grades and/or subjects. The policy constitutes class size reduction if the existing class sizes are larger than the policy goal. Within a state there can be a large range of class sizes, with some students already sitting in small (or reduced) classes.

The ECS describes a multitude of state CSR policies, with most states providing some sort of additional funds to districts for implementation. CSR policies can vary on at least three important dimensions. The first dimension is the grades or subjects that are targeted for reduction. For example, California’s recent CSR initiative targeted first grade, then second, followed by kindergarten and third grades. But this recent focus in California on early grades came several years after a program providing incentives for reducing class
sizes in grade nine. Programs in Nevada, Iowa, and Illinois all target specific subjects including reading or “core subjects” rather than specific grades (ECS, 1999).

A second way that CSR policies vary is the mechanism that is used to get schools and districts to reduce class sizes. Many states (e.g., California, Iowa and Illinois) and the federal government provide grants to school districts (or states in the federal case) to implement CSR. These grants can vary in restrictions on the use of funds. For example, Iowa uses block grants that can be used to reduce class size, but can also be used on other measures that improve reading proficiency. In California, funds are paid on a per-student basis and a district loses funds for all students who are in classrooms with more than 20 students. Simply mandating smaller classes is an alternative mechanism used in several states including Utah, Texas and South Carolina.

A third dimension on which CSR policies vary is simply the target size of the classroom and its measurement. State CSR target class size ratios include 23, 21, 20, 18, 17 and 15 to one.

An important policy and research issue is how class size is measured. The method of measurement is an important indicator of flexibility at the district level. California has a very restrictive measurement scheme that requires that class sizes remain below a ceiling level throughout the entire school year. The California Legislative Analyst’s Office (LAO) found the state’s restrictive measurement scheme may have raised costs up to 21% because class sizes were reduced to 19:1 to ensure that class sizes did not break the 20:1 cap (Hill, 1997). Brewer et al. (1999) found moving class size measurement from school level to district level could change the yearly cost of a national CSR policy by about 15% or $900 million. These examples show how seemingly simple measurement decisions can have large effects on the final cost of CSR, reinforcing the need for CSR policy choices to be informed by their cost ramifications.

**Measuring Class Size: A Continuing Challenge**

“Class size is an administrative measure typically defined as the number of students for whom a teacher is primarily responsible during a school year”

Lewit & Baker 1997

The issues around measurement of class size extend beyond policy formation into research. Class size is a seemingly simple concept that is actually fairly difficult to consistently measure. The idealized picture of one teacher in a classroom instructing the same group of children for a day does not match the reality in many schools. For researchers and policymakers it is the students’ “experienced size that is of primary interest” (Goldstein & Blatchford, 1998). The “experienced class size” is directly related to the quantity of instruction that a student receives and the resources brought to bear on a student’s education. Difficulties in measuring this experienced size stem from three issues: the dynamic nature of classrooms, the variety of classroom models, and the lack of
precise measurements of what occurs in schools and classrooms. These factors add noise to a measure of experienced class size.

Classrooms are dynamic places, with students and adults moving in and out of the rooms and constant regrouping of students into smaller or larger groups that can be seen as “classrooms within classrooms.” The end result is that the class size students’ experience varies almost constantly. Goldstein and Blatchford suggest a continuous measure of class size as the ideal basis for research. Variation of class sizes experienced by students, compared to class sizes reported and measured, may account for some of the problems relating student performance to class size. An example of an attempt at a continuous measure can be seen in Bourke’s (1986) observational study of class size and teacher behavior. Here class size was measured every five minutes.

Another problem with measuring “experienced” class size is that classroom resource use strategies vary across and within districts and schools. This results in changes in a student’s primary instructor throughout the school day. School districts and schools vary in their use of instructional resources such as aides, pullout classrooms, resource teachers, and specialist teachers. The resource use strategies are a product of many factors including district-level policy choices, collective bargaining agreements, or school-level decisions based on school-level conditions (Lewit & Baker, 1997). Hallinan and Sorensen (1985) argue that student grouping mediates the affect of class size on student achievement with ability group size as a stronger indicator of changes in achievement than class size. For example, consider the whole school reform “Success for All” that employs all instructional personnel in a school, including specialist teachers such as librarians or art teachers, as reading instructors (Slavin et al., 1996). In the end, students experience reading class sizes that are up to a third smaller than their classes during the rest of the day. In this reform, measurements of class size based on the number of regular classroom teachers in a school will not capture the class size students experience throughout the day.

A third source of noise in class size measurement is simply the metric used to measure class size. Measurements used in research include pupil-teacher ratios, pupil-professional ratios, and class size based on the number of students assigned to a given teacher. Average class size is the most widespread of all the measures (Lewit & Baker, 1997). While there is some relationship between these measures, they are not perfectly correlated and may mask real differences in the class size student’s experience.

There is disagreement not only on what to measure when discussing class size, but also on the key question of what level to take the measurement at. Should class size be measured at the district level, at the school level, at the grade level within a school, or at the classroom level? The higher the level of measurement, the easier the measurement is to record, but it also masks wide variation in students’ experienced class size. Measurement at the district level is regularly reported and may be the most accessible since most districts have counts of students and teachers for every grade. Measurements at the school level are a commonly used metric in reporting school-level indicators, but
often averages across the grades in a school, masking differences across grades. For example, imagine an elementary school with equal numbers of sixth and first graders but with different staffing patterns for each grade, such as larger classes for sixth graders (30:1) and smaller classes for first graders (20:1). A school-level average will be between these two patterns (25:1) and would not accurately represent the experiences of either group of students. Grade-level class sizes within a school would counter these problems but are rarely provided. This may be due to the methodological difficulty of dealing with classrooms with students from different grades.

Class size is a simple concept that is difficult to measure. The dynamic nature of classrooms, variation in school-level resource strategies, and continuing questions about the metric to use reduce the ability of researchers and practitioners to describe (and evaluate) the class sizes that students experience. The cost analysis in this dissertation will use measurements of class size at the school and grade level. This measure is based on the reported class size by grade, capturing differences across grades, but still masking differences within a grade. While the differences in within grade practices are important to an individual student’s experience, they do not effect the cost of CSR.

INTRODUCTION TO CLASS SIZE RESEARCH
At least in part because of the concept’s simplicity and the availability of district- and school-level measurements, the literature on class size is extensive. There are many literature reviews and summaries of the effects of class size starting with Glass and Smith (1979) through Grissmer (1999). Despite the volume of research devoted to class size, there is little consensus on the effects of class size on student achievement. Most non-experimental data show mixed positive and negative effects associated with small classes (Hanushek, 1999). The largest CSR experiment, the Tennessee Student/Teacher Achievement Ratio (STAR) experiment, shows clear positive and long-lasting affects on student outcomes (Word, 1990, Nye, Hedges & Konstantopoulus, 1999). There is also a dearth on information on the mechanisms through which CSR affects achievement (Finn et al., 1989). Many authors have also raised concerns about the costs of CSR and the need for more cost information (Odden, 1990). The largest effort to reduce classes is occurring in California. Early evaluation of this program confirms many of the concerns and hopes of CSR’s supporters and detractors: It is expensive, but it does make limited improvements in student achievement despite very small measured differences in teacher behavior (Bohrnstedt & Stecher, 1999).

Mixed Evidence From Non-Experimental Studies
“System-wide class-size reduction would have little effect on student performance and even if it did, would cost too much money.”
Odden, 1990
As part of an overall critique on educational resource policies, several reviewers of class size studies found no clear evidence that smaller classes (or increased educational spending for that matter) are associated with improved student performance (Hanushek, 1999, 1998, 1996, 1989, 1986, Hanushek & Rivkin, 1997, Odden, 1990). Starting in the late 1980s, Hanushek and others have argued that there is little evidence supporting increased school expenditures and associated decreases in the pupil-teacher ratio (instructional staff are the largest single operational education expense\(^2\)) to improve student achievement. The general argument begins with the observation that over the last 40 years real expenditures on education have steadily increased, with pupil-teacher ratios declining from 25.6 in 1960 to 17.3 in 1990. At the same time, student achievement data over the past 25 years show no improvements. For example, high school reading, math, and science National Assessment of Education Progress (NAEP) scores have remained essentially flat since the early 1970s.

There are two general criticisms of the use of overall increases in real expenditure as a measure of the intensity of classroom resource use. First there is disagreement over the appropriate correction factor for inflation. As pointed out by William Baumol, industries where it is hard to improve productivity will suffer from higher than average inflation (Baumol, 1993). Thus estimates of increases of real expenditures that use the consumer price index (CPI) may overstate increases in education expenditures. Rothstein and Miles (1995) recommend using a “net services index” instead of the CPI-U. Using this inflation adjuster, they estimate real education spending increased 61% between 1967 and 1991 compared to 99% using the CPI-U. Hanushek and Rivkin (1997) agree that the “Baumol effect” is present in education. They find that after adjusting expenditures for inflation, increases between 1890 and 1990 in the cost of teachers account for about 43% of the overall increase in instructional expenditures. However, they also argue that over the most recent period (1970–1990) the most significant factor in increased expenditures was increased intensity in the use of teachers, with the pupil-staff ratio declining from 20.5 to 15.4.

Another criticism is simply a questioning of the relationship between expenditures and classroom resource use. Rothstein and Miles (1995) in a detailed analysis of expenditures in nine districts suggest that while overall education expenditures have increased, the share of expenditures used in regular education has declined by 21 percentage points. This does not mean that expenditures in regular education have declined, instead regular education did not grow as fast as special education. The proportion of new spending that went to regular education was 28%, while the proportion of new spending that went to special education was 38%. Hanushek and Rivkin (1997) argue that despite the increases in special education expenditures the overall program is too small to account for the overall increases in resource usage. They argue that growth in staffing for special education

\(^2\) Expenditures on instructional salaries and benefits constituted 57% of public elementary and secondary expenditures in 1996–97 (NCES 1999), Table 162.
accounts for about 40% of the overall decline in the pupil-teacher ratio between 1980 and 1990.

It is difficult to disentangle all of the sources of education expenditure increases. A significant amount of this increase, but much less than half, is due to the relative increase in the price of teachers due to the “Baumol effect.” Another significant driver of the overall increase is increases in the resources devoted to special education. It appears that regular education expenditures have increased, but the magnitude of this increase is not clear. As will be discussed later, the increases in teacher salary expenditures were not enough to maintain teachers’ salary levels relative to other occupations that require college degrees.

The increases in regular education expenditures appear to have led to smaller class sizes, without resulting in improved student test performance. These findings are bolstered by a review of over 277 econometric studies that consider pupil-teacher ratios as a factor in student achievement. Well over half are statistically insignificant, with roughly half of those that are significant showing a negative association between class size and student achievement. These are the findings one would expect if class size did not play a positive role in student achievement (Hanushek & Rivkin, 1997).

A key problem with econometric studies is the endogeneity of class size with achievement. That is, did class size cause changes in achievement, or are small classes simply more often found in schools or classrooms that have high-achieving students? For example, achievement is clearly associated with parental earnings; do parents with high earnings and high-achieving students also demand small classes? Another possibility is high-achieving students may be tracked into smaller classrooms.

Several researchers (Hoxby, 1998, Angrist & Levy, 1997, Akerhielm, 1995) have attempted to avoid this bias by using instrumental variables for class size. An instrumental variable is effective when it is correlated with the independent variable in question (class size), but is not related to the dependent variable (student achievement). These three studies yielded mixed results. Hoxby found no effect when comparing classes of 15 and 30. Angrist and Levy, as well as Akerhielm, found slightly better student performance and potentially cumulative improvements with smaller class size.

Researchers propose different reasons as to why they did not find associations between smaller class sizes and improved test scores. Hanushek (1999) argues that one issue is teacher quality. Hoxby argues that the findings of improved student achievement in the STAR study are a result of the expectation that smaller classes would result in improved performance. As an example of a “Hawthorne” effect, Hoxby argues that teachers in the STAR small classes knew they were being evaluated and this motivated them to improve student performance.
Odden (1990) argued that given CSR’s high costs and the average effect size of .1\(^3\), other less expensive and more promising reforms should be pursued. He supported targeted class size reduction to allow small group tutoring for struggling students and to target specific subjects such as reading and language arts.

**Experimental Results: More Enthusiasm**

The more enthusiastic supporters of CSR often use the findings from Project STAR in Tennessee (Nye, Hedges, Konstantopoulos, 1999, Finn & Achilles, 1990 & 1999, Pate-Bain et al., 1997, Mosteller 1995, Word et al., 1990, Finn et al., 1989). Project STAR was a state-run four-year experiment during which kindergarten through third grade students and teachers in 79 elementary schools were randomly assigned to classes that were either reduced (13–17 students per teacher), regular size (22–25 students per teacher), or regular size with a full time aide. New students and teachers were also randomly assigned to classes.

In summary, Project STAR found students in smaller classes consistently performed better than students in larger classes by an effect size ranging from .15 in kindergarten to .25 in grades 1 through 3. Minority students (mostly blacks) showed greater gains in performance than whites. No difference was found between regular classes with an aide and those without (Word et al., 1990).

The Lasting Benefits Study tracked STAR students beyond third grade and found continued positive differentials between students who had been in small classes and those who had not. Fourth grade students who had been in smaller classes were generally found to have more initiative, make more effort, and show less non-participatory behavior (Finn et al., 1989). Differences in median student test scores between students who had been in reduced classes compared to those who attended regular sized classes ranged from .18 in fifth grade through .14 in seventh grade (Finn 1998). Later analyses of student test scores in grades 4, 6 and 8 found sustained higher test scores for students that were in small classes in kindergarten through third grade in all subjects. By grade 8, the difference in test scores was at least 70% of the size of differences seen in third grade (Nye et al 1999). Differences in course enrollment and indicators of disciplinary problems were also seen in a smaller sample of STAR participants in high school. Looking at three school districts, students who had been in smaller classes took more foreign language courses, and looking at two rural school districts, students who had been in small classes took more advanced academic classes. Students who had been in small classes within a larger urban district on average spent fewer days in suspension and absent in grades 8 through 10 (Pate-Bain et al., 1997).

---

\(^3\) An effect size is the change in outcomes (in this case student achievement scores) in terms of standard deviations of the base or control group scores. So an effect size of .1 indicates that achievement scores of the treatment group were .1 standard deviations above the scores for the control group.
These seemingly consistent positive outcomes from reduced classes do raise some areas of concern. The majority of the improvements to students in smaller classes occurred in kindergarten. It is reasonable to expect that each “dose” of sitting in a small classroom might lead to increases in the differential in student performance. But this does not appear to be the case. Instead, most of the gains appeared in kindergarten and simply continued through the third grade. These findings do not lend support to wholesale reduction across all grades. Instead they are more supportive of reductions focused in kindergarten and first grade (Word et al., 1990).

Finn and Achilles (1999) recently argued that the test publisher’s scaling procedures caused the stability in the effect sizes. They used grade equivalents (GE) as an alternative scaling procedure and found students who were in small classes showed continuous improvements over students in larger classes in most subjects from kindergarten through seventh grade. Using this alternative scale, the authors argue that the positive advantages to sitting in small classes in grades K–3 were cumulative through at least seventh grade.

**A Problem for All: A Lack of Theory**

“There is a sense in which class size is a psychologically empty concept.”

Finn et. al. (1989)

A key problem with class size research is a lack of understanding of exactly what occurs in smaller classes that leads to student improvements. A consistent theory on how smaller class size affects student performance would facilitate both non-experimental and experimental studies. Finn (1998) offered a portion of the “participation-identification model” in a discussion of the benefits of smaller class sizes for students at risk.

**FIGURE 1.1 FINN’S PARTICIPATION-IDENTIFICATION MODEL RELATING CLASS SIZE AND ACADEMIC PERFORMANCE**

![Diagram of Finn's Participation-Identification Model](source: Finn, 1998)
This framework indicates that both class setting and teacher behavior affect student engagement, with changes in student engagement affecting student achievement, and achievement reinforcing student engagement. In this framework, reducing class sizes is not sufficient to maximize student performance. Optimal performance also requires changes in teacher behavior.

Unfortunately, just as improvements in student achievement are not consistently found in smaller classes, changes in teacher behavior are also not consistently found. Blanchard and Mortimore (1994) reviewed the findings to date on classroom effects and consolidated these findings into eight categories. To organize Blatchford and Mortimore’s categories, they are placed within Finn’s framework. Blatchford and Mortimore did not consistently find changes within any of their categories. Instead they concluded “research has shown that class-size reductions have little effect on teacher’s behavior...” (Blatchford & Mortimore, 1994).

Instructor behavior
- Increased individual attention for students by teachers
- Increased “quality” teaching
- Changed curriculum coverage

Student Engagement
- Increased pupil attention
- Decreased teacher time on behavior management

Class Setting
- Better student and teacher morale
- Better pupil to pupil relations

Other
- More time for teachers

The STAR results show student gains are concentrated in the first year(s) in school, with differences in performance lasting into high school. A theoretical description of these effects should describe how smaller classes affect student performance in the smaller class and also explain what changes in students allow the effects to persist into later grades. This suggests there is some sort of student socialization is facilitated by smaller classes in early grades that is not directly addressed in Finn’s framework.

An example of the challenge in finding differences in classroom behavior between small and large classrooms is the well-designed study of classroom processes in Toronto reported by Sharpson et al. (1980). Using numerous instruments and measures of teacher/student behavior and attitudes, Sharpson found “the observation of classroom process variables revealed very few effects of class size” with “virtually no changes in methods of instruction used by teachers in the different class sizes.” Sharpson argued, as many do, that teacher training is key to taking advantage of the opportunities that small classrooms present. In fact among the reviewers and researchers on class size, this seems to be one area that most agree on.
However, the lack of a theoretical base linking CSR with improvements in student behavior limits the ability of practitioners to develop guidelines on the skills that should be stressed in CSR teacher training. In Project STAR, a small cohort of second grade teachers from both reduced and non-reduced classrooms received the same training with no effect on student test scores (Word et al., 1990). Some of the school districts in Ventura County, California, conducted workshops for teachers in association with the statewide CSR efforts. A non-random survey of teachers by the Ventura County Superintendents Office (Bartell & Weis, 1997) found that variations in how training affected teachers were correlated with teacher experience. Veteran teachers (those with more than one year of experience) were more likely to change teaching practices to incorporate individualized instruction, flexible grouping, and more frequent use of assessment. First-year teachers were more likely to change classroom and time management practices as a result of training. This relatively weak evidence from Ventura County illustrates the paucity of evidence on the relationship between class size, teacher training and student performance. Clearly more effort is needed in understanding the skills teacher should learn in CSR focused training.

The lack of agreement on the efficacy of CSR and lack of understanding of teacher training needed to implement CSR point out the need for increased understanding and theorizing on the interaction of policy, teaching, and learning.

Grissmer (1999) argues that the lack of a theoretical link between class setting, teachers, parents, and student behavior hampers the ability of non-experimental studies in finding results. This leads to the mixed results seen in non-experimental results. Grissmer’s call for more sophisticated models can also be seen as simply a request for better data in doing non-experimental education studies. These more sophisticated models will require a richer understanding of students’ educational environment. They require data beyond the inputs that students and schools bring to education, to include a better understanding of the education process (Grissmer suggests focus on time-on-task), teacher quality, as well as parental involvement and behavior. These data requirements can only be met through intensive and expensive efforts, greater than what is currently devoted to most educational studies. It may be that experimental studies could become relatively less costly because they use randomization to account for difficult-to-measure constructs.

**CSR IN CALIFORNIA: CONFIRMING EVERYONE’S VIEW**

In the summer of 1996 California lawmakers found themselves in a curious position. After years of budget cutting during the recession of the early 1990s, they suddenly faced a budget surplus. The state constitution required that a majority of the surplus be allocated to education. The combination of constitutional spending requirements, recently released STAR findings, and a continued interest in CSR by teachers’ unions, came together in California’s massive early elementary CSR program. During the first year, $900 million was distributed to school districts at a rate of $650 per child in a classroom
with 20 or fewer other students for a full day or $325 per child in a class of 20 for a half day. The priority for reduction was first grade, then second, then either third or kindergarten. Two hundred million dollars of the original $900 million was distributed in the form of one-time $25,000 grants for new classrooms. During the second year of CSR implementation, school districts were given $800/$400 (full day/half day) per child in a reduced classroom with new classrooms grants going up to $40,000. By the end of the second year of implementation, almost all first and second grade students were sitting in reduced classes and about two-thirds of kindergartners and third graders were sitting in reduced classrooms. (Bohrnstedt & Stecher, 1999). Within three years, about 99% of K–3 students were sitting in reduced classrooms (Stecher & Bohrnstedt, 2000).

The implementation of CSR in California showed that CSR is relatively easy to quickly implement, but very expensive. Within three years, 29,000 additional teachers had been brought in, resulting in a total of 91,000 K–3 classrooms providing 1.7 million students with (small) measurable benefits. This can be compared to an evaluation of the Elementary and Secondary Education Act (ESEA) by Berman & McLaughlin (1978) who argue that it takes in the order of seven years of implementation for the educational effects of a new policy to become apparent. But again, CSR is very expensive. California’s CSR initiative is the most expensive state education reform in the United States at about $1.5 billion a year. This does not include additional resources devoted to CSR at the district and school level (Borhnstedt & Stecher, 1999).

California’s CSR experiences fit well into the mixed CSR literature about costs, student achievement, and teacher behavior. The achievement gains attributable to CSR are very small, with an effect size of about .1. This is at the bottom end of the STAR findings. These results fit into the range found in experimental results, but were small enough to fit the expectations of skeptics. Few changes were found in teacher classroom behavior. Teachers did not measurably modify what they taught or how they taught. Instead teachers in reduced classrooms spent a little less time disciplining students and instructing the whole class. They also spent a little more time with problem readers.

An important outcome of California CSR policies are problems associated with equitable resource allocation during implementation. A key cost and equity issue raised with the implementation of CSR in California was the availability and distribution of qualified teachers in the new classrooms. These issues are addressed extensively in Chapter 5.

One of the causes of equity problems as CSR in California was implemented was that CSR reimbursements were not tied to actual costs. CSR was a financial boon for some districts and a financial strain on others. As discussed earlier, districts received a flat dollar amount per student in a reduced classroom, regardless of the amount of effort it required to create the reduced classroom. So for districts that already had small classes, CSR was relatively cheap and in some cases free. In other cases, when class sizes were large, when rooms for extra classes were scarce, and when schools or districts were small, CSR was very expensive. Some districts were reimbursed $650 (in 1996–1997) per student when the actual cost was zero. At the same time, some districts were reimbursed
the same amount when the costs were closer to $1,000 per student (Hill, 1997). A survey of superintendents found 56% reported costs above the state reimbursement of $650 per student. At the subsequent rate of $800 per student, 43% still believed that costs of the program were beyond the reimbursement amounts (Bohrnstedt & Stecher, 1999).

A better understanding of the implementation costs of CSR could have been used to alleviate these disparities, resulting in increased efficiency in the expenditure of state resources. Reimbursing the actual cost of class size reduction would have avoided two problems. First, there was an unintended consequence of CSR dollars supporting other programs in districts where CSR costs were less than the reimbursement. Second, funding for other programs were reduced in districts where implementation costs exceed reimbursement rates.

The lack of a link between reimbursed cost and actual cost raised equity issues in terms of the rate of implementation. The implementation rate was directly tied to difficulties some districts had due to conditions at the schools, such as lack of additional space, large classes that required extra efforts to reduce, and rapid enrollment growth. Of those principals who were unable to implement CSR, 81.0% cited the lack of space for new classrooms as the cause (Bohrnstedt & Stecher 1999), despite the facilities grants that were available. A separate analysis of data on the rate of CSR implementation found that districts with the largest class sizes before CSR (greater than 30) were more likely to have slower implementation. The same study found lower implementation rates in districts with higher growth rates (Travers, 1998).

An understanding of the number of classrooms needed given the school size and existing class size in schools would assist in reducing CSR implementation constraints due to lack of space. This information would give policymakers and administrators an understanding of the range of additional space needed. Optimally the information on the range of space needed could be combined with information about the available classroom in schools to determine how to best distribute funds for new classrooms. Unfortunately, information on existing space usage is not regularly gathered so there is no good database of information. This information could be gathered as part of CSR implementation, but would need to be verifiable.

**Existing Information on CSR Costs**

There is limited available information on the costs associated with class size reduction. None of the available estimates completely capture the interactions between costs, school-level conditions (existing class size, school size), and policy choices (class size goal, grade level, and flexibility in implementation). Four estimates are available in the literature on a “cost per student” basis: Levin, Glass and Meister (1987), California Legislative Analysts Office (LAO) (Hill, 1997), Travers (1998), and Brewer et al. (1999).
Each estimates the marginal cost of CSR, or the expenditure per student for the additional classrooms needed to reduce class size to a given policy goal.

Levin’s cost estimates were made using 1980 prices that have been inflated to 1997 levels using the CPU-14, Travers’ are in 1997 dollars, and the California LAO’s are in 1996 dollars. Brewer’s estimates were made using 1998 dollars.

Levin and Travers do not take into account variation in school enrollment, though Levin’s estimates do consider existing class size. Levin estimated the cost of reducing from 35 to 30 to be $263, from 30 to 25 to be $367, and from 25 to 20 to be $548. Using a small sample of districts, Travers estimated the cost of California’s reduction to 20 from an average class size of about 29. His estimates range from about $750 to $850 per student. Both Levin and Travers produce similar cost estimates for reduction (around $800), but for different bundles of classroom resources. Levin’s estimates include facilities costs while Travers’ do not. Travers also predicts the long-term cost of CSR to be between $850 and $1000. The cost increase is due to the teacher salary scale rising automatically with experience.

Brewer et al. (1999) estimated the cost of a national CSR policy. They estimated the average cost of reduction from current class sizes to 20 to be $189, $448 for reduction to 18, and $981 for reduction to 15. This work highlights the changes in costs due to variation in policy choices about class size measurement. Brewer found that measuring class size at the district level instead of the school level decreases costs by about 16%. These estimates are the lowest of the three because calculating class size at the district level provides implementation flexibility that reduces costs, and because of the assumption that new facilities are not needed. The long-term estimates of $562 per year for a class size of 18 again include increases in expenditures over nine years due to increases in teacher salaries.

The California LAO’s estimates used district-level teacher costs, average class size, and enrollment in making their estimates. The California LAO estimated the average cost of reduction to 20 for large districts (over 20,000 students) to be $690 per student, $810 for districts with 5,000 to 20,000 students, $800 for districts with 1,000 to 5,000 students and $710 for districts with less than 1,000 students. These estimates, like Brewer’s, point out the interaction between district size and costs. Mixed in with the district size issue are the relationships between district-level teacher costs and the cost of CSR.

All of the estimates were made using different sets of assumptions about the cost of facilities and the level within the education system at which class size is measured. Levin’s estimates assume new facilities are needed for CSR, while the California LAO, Travers, and Brewer do not. Levin and Travers assume that school districts have relatively little implementation flexibility, in that class size is measured at the school

---

4 Cost of all items in 1980 $82.4, and 160.5 in 1997, 1998 Statistical Abstract of US.
level, while Brewer assumes school districts have more flexibility in implementing CSR in that class size is measured at the district level.

CONCLUSIONS FROM RESEARCH
The research is divided on the benefits of CSR. The experience in California and the Tennessee STAR experiment show that CSR is an educational reform that can be implemented rapidly, in a widespread fashion, and result in (small) improvements in student performance. This is in contrast with non-experimental literature or the literature drawn from smaller experiments that show mixed improvements in student achievement. Despite lack of agreement on achievement outcomes from CSR, there is agreement in the literature regarding the lack of a theoretical base that can explain why and how smaller classes are (or are not) important to student learning. At the same time, while it is well understood that CSR is expensive, the factors that affect cost at the school level and the relative magnitude of these effects are not documented. The implementation of CSR in California raised important equity concerns related to teacher quality and simply the ability, or even just the space, to implement the program. So while CSR is expensive to implement for all schools, it may be even more expensive for some schools based on their lack of space or disadvantages they face in competing for good teachers.

WHY POLICY CHOICES MATTER
As stated earlier, CSR is generally a policy made at the state or federal level to be implemented at the local level. The parameters of the policy formulated at the higher levels of the system (i.e., state or federal level) have very important implications as to which students are effected by CSR, how they are affected, when they are affected, the amount of the final cost for the policy, and who pays.

Policymakers must decide which grade(s) to target with a CSR policy. The current research suggests the largest effects are found in the earliest grades, K and 1 (Word et al., 1990). In deciding the grade(s) to target for CSR, policymakers begin to establish one of the key cost dimensions--how many students sit in reduced classes. Reduction in more grades increases the number of students sitting in reduced classes and raises the total cost. The second policy dimension is the class size goal, i.e., how large will reduced classroom be. The amount of the reduction is a key element in the cost of CSR.

The third issue is how class size is to be measured. Strictly enforced class size ceilings at the classroom level increase costs relative to schoolwide or district-wide averages that can allow flexibility in implementation to meet local conditions. Measuring (and enforcing) a smaller class size throughout a year is more difficult (and costly) than establishing smaller class sizes for a particular point in time in schools that have high levels of student mobility. Finally policymakers have to define the pace at which they want CSR to be
implemented. Rapid implementation may drain the market of resources needed for uniform implementation across all schools.

There are non-pecuniary costs associated with implementing a policy that are not directly measured in dollars, but are movements away from expected levels of facility quality and teacher qualifications. In California, CSR implementation has clearly been associated with declines in the quality of classroom space as classes are held in facilities not intended for instruction including storage space, gymnasiums, and libraries. Many new classrooms in California were portable or temporary classrooms. It is not clear that these represent a decrease in quality. Temporary classrooms can have newer technology, better access to bathrooms and reduced distractions because they are removed from the main buildings (personal interviews in Florida). Others view temporary classrooms as being poorly built, loud, and isolating.

Changes in teacher qualifications were clearly seen during the implementation of CSR in California. The rapid addition of over 29,000 teachers in three years led to declines in the overall experience, education, and credentialing level of the K–3 workforce. More importantly, these declines were concentrated in schools that served larger proportions of poor, minority and English language learners (ELL). The end results of the rapid CSR implementation in California were declines in the quality of classroom space and teacher qualifications. It also resulted in reduced equity in the distribution of good classroom space and qualified teachers. The teacher qualification issues will be fully addressed in Chapter 5.

There are options available to policymakers to decrease both the expenditures and the non-pecuniary costs associated with CSR. There are an array of policy decisions about who is affected, the target level for the policy, and how that target is measured. Those policy decisions affect the both the long- and short-term out-of-pockets costs as well as the non-pecuniary quality and equity costs of the policy implementation. Clarifying and enumerating some of the relationships between CSR costs and policy decisions will be the tasks of Chapters 3, 4 and 5 of this dissertation. This information should aid policymakers as they make the trade-offs inherent in the various policy options.