Class-size reduction and teacher quality: Evidence from California

Jennifer Imazeki
San Diego State University
Department of Economics
5500 Campanile Drive
San Diego, CA 92182
jimazeki@mail.sdsu.edu
(619) 594-5012

Abstract: This chapter examines the effect of class-size reduction policies on the distribution of teacher quality across schools. In California, teacher quality, as measured by education, experience and credentials, fell dramatically in the wake of a state-wide class-size reduction policy. In particular, the pre-existing gap in teacher qualifications across students of different socioeconomic backgrounds grew wider and has persisted over time. After giving an overview of the changes in California's teacher labor market, this chapter explores the possibility that the way in which the policy was financed may have contributed to lower teacher quality in the highest-need districts.

Introduction

Class-size reduction (CSR) is an enormously popular policy. With strong support from parents and teachers, policymakers in over twenty states have instituted class-size reduction programs in the last decade, often coordinated with federal funding, and generally targeted at schools serving low-income students (U.S. Department of Education, 2000). At the same time, researchers are still in disagreement about whether smaller classes significantly improve student learning and even if they do, whether the gains are large enough to justify the cost. This conflict between political popularity and uncertain results has policymakers in the difficult position of deciding whether or not to adopt (or continue funding) CSR programs in the face of tighter budgets and an unstable economy.

The current trend to reduce class sizes gained much of its momentum with Tennessee's class-size experiment during the late 1980's. Project STAR (Student/Teacher Achievement Ratio) provided the most believable evidence to date that smaller classes significantly improve student learning (Finn and Achilles, 1999). SAGE (Student Achievement Guarantee in Education), a smaller program in Wisconsin, also produced similar results (Molnar et al, 1999). But these programs were relatively narrow in focus. In 1996, California became one of the first states to implement class-size reduction statewide. Unfortunately, the adoption of CSR on such a large scale led to a substantial deterioration in teacher qualifications that likely mitigated the positive effects of smaller classes. Of particular concern was the realization that CSR may have contributed to educational inequities between students of different socioeconomic status since schools serving the most disadvantaged students were often the last to implement the program and had to hire the least-qualified teachers (Bohrnstedt and Stecher, 2000). Throughout the state,

many districts have not seen the expected gains and a few are already abandoning the program (Hayasaki, 2002).

The lack of concrete results, the drop in overall teacher quality, and the increasing disparities in teacher qualifications have been discouraging for CSR supporters. However, California's experience may contain valuable information about how other states can design and implement more successful CSR policies. For example, many of California's teacher labor market problems have been attributed to the speed and scale of implementation, at a time when there was already a teacher shortage (see Bohnstedt and Stecher, 2002; Jepsen and Rivkin, 2002). In this chapter, I will also discuss the possibility that the way in which the policy was *financed* may have played a role, particularly in the distribution of teachers across schools. With a better understanding of what has happened in California, policymakers in other states may be able to avoid these pitfalls and design more effective CSR policies.

To further that understanding, this paper examines the impact of California's class-size reduction policy on the teacher labor market, with a particular focus on the distribution of teachers across schools serving different student populations. Section II reviews what we know about the effectiveness of teachers and smaller classes. Sections III and IV give a synopsis of California's experience with CSR and the changes in the teacher labor market. Using data from the California Department of Education, I reiterate the findings of other researchers (e.g., Bohnstedt and Stecher, 2000, 2002a; Jepsen and Rivkin, 2002) that teacher quality, as measured by experience and education, did fall significantly after implementation of CSR and there are large inequities in how these measures are distributed across students of different socioeconomic background. However, the data also show that a) these measures are slowing returning to pre-CSR levels, and b) there was substantial inequality in these measures even prior to CSR. While

the gap in teacher quality between schools serving high- and low-income students did grow slightly wider after CSR, particularly the proportion of teachers lacking a full credential, the effect of CSR on the distribution of teachers should be kept in perspective. On the other hand, there is no doubt that low-income students were considerably worse off, in terms of teacher qualifications, after the program began and the gaps have persisted over time. In Section V, I discuss how the way in which CSR was funded may have contributed to the problems with the teacher labor market. I show that policies that do not account for differences in salary costs and hiring needs will exacerbate inequities. Section VI presents preliminary results of a simple simulation that suggests more disadvantaged districts in California do, indeed, receive less CSR revenue per new teacher needed, making it more difficult to attract high-quality teachers.

The Effectiveness of Teachers and Class-Size Reduction

Deterioration in the preparation and qualifications of teachers, particularly in schools and districts with large concentrations of traditionally disadvantaged students, could have important implications for the success of CSR. If there is a negative impact on student performance because of lower teacher quality this may offset the positive gains of smaller classes and greatly reduce the effectiveness of the original policy. In California, the media and policymakers have focused on the decline in teacher experience, education and credentials that followed CSR. Thus, throughout this paper, these are the measures of teacher quality that will be discussed. However, it is important to point out that while several studies have found that teachers matter, in the sense that there are large and statistically significant differences in the performance of students with different teachers (e.g., Hanushek, Rivkin and Kain, 1998; Ehrenberg and Brewer, 1994), the production function literature has been somewhat inconsistent in measuring the effect

of specific, observable variables (see Hanushek, 1986). It is possible that a decline in characteristics such as education and experience will have relatively little impact on actual student learning.

On the other hand, the conflicting research may be because the relationship between these variables and student performance is more complex than many studies reflect. For example, average teacher experience in a school may not appear to affect student learning because the effect of teacher experience may be highly non-linear. Murnane (1975) finds that the first few years of experience matter, at least for children in inner-city schools, and Murnane and Phillips (1981) find evidence that having more than 9 or 15 years of experience may matter, but the marginal effects of experience between these thresholds appears to be negligible. There may also be differential effects of teachers on different types of students – at least one paper (Summers and Wolfe, 1977) finds that low-performing students do better with *less* experienced teachers.

It is also important to point out that the impact of lower teacher education or experience on student performance need not be large in order to offset the positive impact of smaller classes. Tennessee's Project STAR, which randomly placed primary grade students in classes of either 13-17 or 22-25 students, is considered the most conclusive evidence on the effectiveness of CSR. Several studies of the STAR experiment have found that smaller classes do increase student performance, particularly for low-income and minority students (Finn and Achilles, 1999; Krueger, 1999), with student performance higher in the smaller classes by 0.15-0.4 standard deviations. At the same time, CSR is also one of the most expensive of education reforms and among the least cost-effective (Levin, 1988; Grissmer, 2002). Furthermore, these studies do not

consider any impact on teacher labor markets;[1] any decline in teacher qualifications resulting from CSR would further reduce the cost-effectiveness of the policy.

Early evaluations of California's experience are consistent with prior research in that gains from reduced class sizes appear to be small (Bohrnstedt and Stecher, 2002b). It can be very difficult to isolate the effects of CSR, given that California instituted numerous reforms during the same time period (for example, a highly-publicized accountability system and limits on bilingual education). However, early studies of the program suggest that the drop in teacher qualifications is indeed offsetting the benefits of smaller classes (Jepson and Rivkin, 2002). In the next sections, we take a closer look at what actually happened with the teacher labor market in California and whether the structure of the CSR policy might have contributed to the problems.

California's K-12 school system and CSR

California has long been in the spotlight for issues of school finance, beginning with the landmark Serrano v. Priest equalization case in 1971 and continuing with the passage of Proposition 13, which effectively eliminated all local control of the property tax. In response to Serrano, California's system of school finance was reformed to equalize per-pupil spending across districts and it has been highly successful in doing so. However, this equality was largely achieved by reducing spending in high-spending districts more than increasing spending in low-spending districts, so that during the 1980's and 90's, average per-pupil spending in California fell roughly fifteen percent relative to the rest of the country (Sonstelie, Brunner and Ardon, 2000).

[1] Such studies rarely consider the effect of smaller classes interacting with other inputs but instead calculate the effect of reducing class sizes, all else equal.

Districts responded to these declining revenues by increasing class sizes (Sonstelie, Brunner and Ardon, 2000). By 1993-94, California's average class size for grades K-6 was one of the highest in the country, at an average of 27.7 (U.S. Department of Education, 1996). At the same time, student achievement was falling: in 1994, California ranked second to last among all states for 4th-grade reading scores on the National Assessment of Educational Progress, and in the bottom fifth in the nation for 8th-grade math (U.S. Department of Education, 1997b).

The dismal state of elementary and secondary education brought increasing pressure on California policymakers for reforms. In 1996, with research emerging on the success of Tennessee's Project STAR class-size reduction experiment, and with increasing state revenues from an economic upswing, California became one of the first states to adopt a state-wide class-size reduction program. Under the policy, schools receive a per-pupil grant for each K-3 student in a class of 20 students or less, taught by a certificated teacher. In the first year of the policy, the grant was $650, or $325 for students in small classes for half a day. In 1997-98, this was increased to $800/$400, with a cost of living adjustment in each of the following years.[2] Funding for capital and facilities is allocated separately. Schools are required to reduce classes in grade 1 first, then grade 2, and then may choose to reduce classes in either grade 3 or kindergarten.[3] To qualify for funding, schools must keep classes at or below an average of 20.4.

California's policy applies to any school serving K-3 students. Although technically, the program is voluntary, ninety-four percent of eligible districts (839 of 895 districts) implemented the program at some level in the first year, though not in all schools or all grades. By the fourth year, all but nine eligible districts were participating (California Department of Education, 2002).

[2] The grant amounts were $832 and $416 in 1998-99; $844 and $422 in 1999-00; $855 and $428 in 2000-01 (California Department of Education, 2002).

The size and speed of implementation led to several problems, some anticipated and some not. The two primary issues were the large quantities of physical and human capital needed on relatively short notice. Physically, districts needed space for the new classes. Some were able to buy or lease portable classrooms, if they had the land on which to put them, but many schools resorted to turning libraries, teacher lounges and even cafeterias into classrooms (Bohrnstedt and Stecher, 2000). Districts also needed people to staff these new classes – in the first two years of the program, 23,000 new teachers[4] were added to California's K-3 teaching force (see Table 1). Some of these new hires were experienced teachers (e.g., teachers who had left the profession and were enticed to return, or who were hired from other states); however, a large number had never taught before. The shortage of certified teachers willing to teach led many districts to seek emergency waivers for new hires, thus allowing inexperienced teachers into the classroom while working toward their full teaching credential.

<Insert Table 1 about here>

Table 1 shows summary statistics for key variables in four years. 1995-96 was the year just preceding the adoption of CSR; 1996-97 and 1997-98 were the first two years of implementation and the years with the greatest hiring activity; 2000-01, the fifth year of the program, is included to give some idea of whether the changes brought about by CSR are long-run outcomes or primarily transitional problems. All of the data are from databases collected annually by the California Department of Education. Characteristics of teachers are from the Professional Assignment Information Form (PAIF), which contains information on the characteristics and assignments of individual teachers. For much of the analysis that follows, the

[3] Classes serving multiple grades are allowed under the program but second- or third-grade students in combined classes are only counted after all students in the previous grade (i.e., first or second, respectively) are in reduced classes.

[4] Throughout this paper, one 'teacher' refers to one full-time equivalent (FTE).

PAIF data have been aggregated to create school-level variables. The teacher characteristics of interest include years of experience, level of education and whether or not a teacher has a full teaching credential.[5] Although CSR had some impact on teacher qualifications at all levels, the analysis here is limited to K-3 teachers only.[6] Data on school and district characteristics are from the School Information Form and Aid to Families with Dependent Children. Due to problems with the PAIF data, all schools in the Los Angeles Unified School District are excluded from the analysis. The largest district in the state, Los Angeles educates over 235,000 K-3 students and employs roughly twelve percent of the K-3 teaching force. Because of its size, diverse student population, and the urban environment of many of its schools, Los Angeles faced the toughest challenges in implementing CSR. All of the trends discussed below are similar whether or not Los Angeles is part of the analysis, but the gaps between high- and low-poverty schools tend to be larger when Los Angeles is included.

The averages for the teacher characteristics in Table 1 are over all K-3 teachers (excepting those in Los Angeles Unified). Both 1996-97 and 1997-98 are included throughout the analysis because although the policy began in 1996-97, many schools were not able to implement the policy beyond first grade until the second year, or later. Thus, the increase in the total teaching force was just as large in 1997-98 as in the first year and the decline in teacher qualifications just as sharp.

Table1 shows that teacher qualifications such as experience, education and credentials all fell dramatically in the wake of CSR. The proportion of teachers with less than three years of

[5] In response to CSR, there were many teachers who began teaching K-3 that had previously taught in other subject areas. These teachers may have already had a credential for their previous subject area while *also* needing a waiver or emergency credential to teach K-3. Therefore, rather than focusing on teacher with waivers, some of whom may be quite experienced and highly qualified, I use the percent of teachers who lack a full credential of any kind.

[6] The movement of teachers within and between districts led to openings in all grades, not just those grades affected directly by CSR. The patterns of teacher qualifications for all teachers in all grades are similar to the patterns for K-3 teachers alone, though the changes are smaller.

experience doubled and those lacking a full credential tripled. These declines have already been well-documented by other evaluations of CSR (e.g., Borhnstedt and Stecher, 2000, 2002a; Jepsen and Rivkin, 2002[7]). However, the data from 2000-01 also suggest that some of these characteristics are slowly returning to their pre-CSR levels. The simple passage of time will accomplish part of this – by 2000-01, teachers hired in the first years of the program would have four or five years of experience, and would have acquired their credential, which would also require additional units beyond a bachelors degree. To the extent that experience and training increase teacher quality, the teaching force appears to be slowly improving.

On the other hand, the proportion of teachers lacking a full credential in 2000-01 was still double the level it was in 1995-96, suggesting that schools were continuing to struggle to find experienced, qualified teachers. This may be due partly to the high turnover rate for new teachers. Teachers are most likely to leave a position in their first few years of teaching (U.S. Department of Education, 1997a) so by 2000-01, districts may already have needed to replace teachers hired in the early years of the CSR program. It could also reflect the fact that although the greatest hiring activity occurred in the first two years of the program, many schools did not *fully* implement CSR, in all eligible grades, until the third, fourth or fifth years (California Department of Education, 2002), and these teachers may still be working toward their credentials. In addition, California has only recently adopted policies to address the teacher shortage (such as incentives for recruitment and retention in low-performing schools). Thus, it may still take a few more years before the teacher labor market settles into its long-run equilibrium.

[7] There are minor differences between the numbers in the analysis presented here and those in the CSR Consortium reports and in Jepsen and Rivkin, primarily due to differences in how socioeconomic groups are defined and how the averages are weighted (e.g., Jepson and Rivkin weight averages by student enrollment).

It is also important to point out that experience, education and certification, per se, may not fully reflect teacher "quality". With the enormous increase in the teacher labor force during the first few years of CSR, there were surely many teachers employed who might not have been hired otherwise. If teachers hired during the first years of implementation are of lower quality overall (e.g., in unobservable characteristics like motivation and creativity), they may continue to be of lower quality even after acquiring additional experience, education or credentials.

Distribution of teacher characteristics

Clearly, the qualifications of K-3 teachers, in terms of education and experience, fell for the state overall in the years following CSR. But much attention has also been directed at the *distribution* of the problem. Evaluations of the project have repeatedly pointed out that it was primarily large urban schools and those with more disadvantaged student populations (e.g., low-income, limited English proficient, disabled) that had the hardest time filling new teaching positions and these schools saw the largest declines in teacher qualifications (Bohrnstedt and Stecher, 2000; Jepsen and Rivkin, 2002[8]). I therefore turn to the distribution of teacher qualifications across different types of schools. Schools were divided into five groups based on the socioeconomic status (SES) of the students where socioeconomic status is measured by the percent of students in the federal free and reduced price lunch program.[9] Table 2 shows the poverty ranges and number of schools, teachers and students in each quintile group.

<Insert Table 2 about here>

[8] It should be noted that the general trends discussed in this section are also covered in these studies; the reader is referred to them for more detailed analysis of the data.

[9] Other ways to assess the socioeconomic status of students might be to use the percent of students that are non-white or limited English proficient. These variables are highly correlated with the percent low-income and all results are similar when using these other measures. For the rest of the paper, SES and income are used interchangably.

Media coverage of the inequities in teacher qualifications has often focused on the size of disparities after CSR but it is important to point out that teachers were distributed far from equally even before CSR. Figure 1 shows the distribution of three K-3 teacher characteristics in 1995-96. Schools with the highest concentrations of poor students had six percentage points more inexperienced teachers than schools with the lowest concentrations of poor students (14.8 percent vs. 8.8 percent), and double the proportions of teachers with only a bachelor's degree, and of teachers lacking a full credential. As will be discussed in more detail later, schools with the most low-income students also had slightly higher pupil-teacher ratios[10] and paid slightly higher beginning salaries.

<Insert Figure 1 about here>

Figures 2 through 4 show how the distribution of teacher qualifications changed following CSR, both immediately (1996-97 and 1997-98) and in the longer-run (2000-01). There is no question that schools with the largest concentrations of low-income students suffered a large decline in teacher qualifications. Schools in the lowest SES quintile saw a 12.9 percentage point increase in the proportion of inexperienced teachers, and a 6.7 percentage point increase in the proportion of teachers with only a bachelor's degree. Most notably, by 1997-98, the percentage of teachers lacking a full credential almost tripled in the lowest-income schools, to just over twenty percent.

<Insert Figures 2, 3, 4 about here>

From a certain perspective, one could argue that the effect of CSR on the *distribution* of teachers is perhaps not quite as bad as it has been made out to be. As mentioned, there were already large disparities in teacher qualifications between rich and poor schools prior to the start

[10] Pupil-teacher ratios are only an approximation of class size. Because teachers may spend part of their time engaged in non-classroom activities, such as lesson planning, pupil-teacher ratios tend to be lower than actual

of CSR. After CSR, schools in the highest SES quintile also saw large declines in teacher qualifications so that the gaps in the proportion of teacher's with a bachelor's degree or less, and the percentage of inexperienced teachers, stayed roughly the same. That is, while the *levels* for all schools deteriorated, a substantial portion of the *differences* already existed.

On the other hand, the gap in the proportion of teachers lacking a full credential doubled in size: by 1997-98, the lowest-income quintile had almost four times as many non-credentialed teachers as schools in the highest quintile. In addition, because the size of school and student income are correlated, schools in the lower SES quintiles tend to be bigger; thus, even if the teacher quality measures change by similar percentages, that may represent a greater *number* of teachers (serving a greater number of students) in the lower-income schools.[11]

Finally, it is disturbing to note that the gaps are persisting, and even increasing, over time. Although it is encouraging that teacher qualifications are slowly improving overall, they are improving slightly faster for higher SES schools. This could be explained, in part, by patterns of implementation. Schools with the most low-income students were the last to fully implement in all eligible grades (Bohrnstedt and Stecher, 2002a), thus they were still hiring for new positions in later years. As these teachers gain experience and earn their credentials, the gaps in these measures may shrink further. It is likely that some of these problems could be mitigated by phasing in implementation, beginning with low-income schools.

CSR and teacher labor markets

---

number of students per class.

[11] Because the unit of analysis here is the *school*, each quintile represents a different number of teachers and students. In particular, poverty, size and urban environment all tend to be correlated so that the schools in the higher-poverty quintiles are also likely to be larger and part of the larger urban districts. To explore whether these large urban districts are driving the patterns here, the analysis was repeated excluding four of the largest districts: San Diego Unified, San Francisco Unified, Long Beach Unified and Fresno Unified (Los Angeles Unified has already been excluded). The patterns were roughly similar. Detailed data tables are available from the author.

California's experience with CSR has clearly been difficult. Many schools are still struggling with staffing problems, or recovering from the difficulties experienced during implementation. With a weaker economy straining state and local budgets, many districts are struggling to keep their small classes intact. At least a few districts have already abandoned the program in some grades (Seaton, 2002; Hayasaki, 2002) while others have only retained their programs through huge fundraising efforts by parents and teachers (Borgatta, 2002). For other states that are considering adopting large-scale class-size reduction policies, an important question is whether California policymakers could have done things differently to avoid some of the problems. Other researchers have stressed that the speed and scale of implementation was a large factor (Bohrnstedt and Stecher, 2002b; Jepson and Rivkin, 2002). In this section, I consider whether the way in which the program was financed may have contributed to the problems with the teacher labor market. I will discuss why certain districts are more likely to experience teacher shortages and lower teacher quality in general, and why California's CSR policy was likely to exacerbate these inequities.

Even without the greater demand created by CSR, districts with many disadvantaged students are likely to have a harder time hiring high-quality teachers because they face the dual challenge of higher demand and lower supply.[12] On the demand side, these districts tend to be larger so there are simply more teaching positions to be filled. There also tend to be higher attrition rates out of such districts as teachers leave for positions in more attractive schools as well as other professions (see Imazeki, 2001; Theobald and Gritz, 1996).

On the supply side, it is often the case that fewer teachers want to work in these districts, for a variety of reasons: they are more difficult teaching jobs because the students have higher

[12] See Imazeki, 2002, for a fuller discussion of the supply and demand problems of districts serving large populations of high-need students.

needs and there is less support from parents and the community, they are often in locations associated with higher crime and higher costs of living, and the schools themselves are often physically less appealing places to work (U.S. Department of Education, 1997b). As a result, teachers of a given quality require higher wages to teach in these districts (Rosen, 1987). If wages are not high enough to fully compensate teachers for the harsher working conditions, teacher quality will be lower.

Because of these difficulties, schools and districts with the most disadvantaged students have the hardest time attracting and retaining high-quality teachers. A class-size reduction policy can then exacerbate these problems. To see how, first consider a district's choice about the quantity and quality of teachers hired. Let us assume that the district chooses the number and quality of teachers in order to maximize school quality, subject to the district budget and the wages that must be paid to teachers. Both better teachers and smaller classes increase school quality. Note that choosing the quantity of teachers is equivalent to choosing class size, assuming student enrollment is exogenous and relatively fixed. In California, the district budget is also exogenous, determined by the state, and is roughly equal per pupil across districts.

The wage that district *i* must pay teacher *j* can be broken down into two parts, a base wage and a premium for ability:

$$(1) \qquad W_{ij} = w_i + p_i A_j$$

The base wage, w_i, can be considered the price of additional teachers of some base quality level while the premium for ability, p_i, can be considered the price of additional 'units' of teacher quality, A_j. For example, if teacher quality is measured by experience, then A_j is the number of

years of experience that teacher *j* possesses, w_i is the salary that district i must pay for a new, inexperienced teacher (with $A_j = 0$) and p_i is the premium that must be paid (per year of additional experience) for a more experienced teacher.[13]

For the sake of simplicity, assume that the only input is teachers so that the entire budget is used for teacher salaries. The wage bill for a district that hires N_i teachers will be equal to the district budget and can be represented as in equation 2:

$$(2) \qquad W_i = w_i N_i + p_i \overline{A} = B_i$$

where $\overline{A}$ is total teacher ability in the district (i.e., the sum of A_j over all teachers) and B_i is the district budget. Thus, given a set budget, a district must trade off quantity and quality: buying additional teachers (higher N) to have smaller classes must mean a reduction in the total quality of teachers bought (lower $\overline{A}$). Specifically, hiring one more teacher of basic quality will cost the district w_i and they must therefore give up w_i/p_i units of total ability. In equilibrium, each district chooses the combination of class size and teacher quality that maximizes school quality, without exceeding the district budget.

If every district faces the same budget and the same prices (i.e., if B_i, w_i and p_i are the same for all districts), then every district will choose the same combination of class size and teacher quality.[14] However, as discussed above, some districts may need to pay higher wages in order to attract a teacher of even the most basic quality, and may need to pay higher premiums to

[13] It should be noted that because of union contracts and the structure of teacher salary schedules, in reality, the only measures of teacher quality actually captured by A_j are usually experience and education.

[14] I am also assuming that all districts have the same production function for school quality. If there are differences in the marginal productivity of inputs (i.e., smaller classes and teacher quality have different effects on school quality in different districts), this will also lead to different choices of teacher quality and quantity. However, the basic implications of the model are the same.

attract teachers of higher quality. Thus, in the model here, the base wage and the ability premium are both treated as district-specific. Two districts may have similar budgets but face different prices and they will end up with different combinations of teacher quality and quantity. It is straightforward to show that a district that must pay a higher base wage and/or a higher ability premium can end up with both fewer teachers (larger classes) and lower-quality teachers.

A policy requiring a particular class size (assuming it is lower than a district would choose on its own), without any increase in the district budget, must lead to a reduction in teacher quality. If a district needs to hire, say, five additional teachers, they must spend $5w_i$ simply to get teachers of basic quality, and must therefore reduce teacher ability by $5w_i/p_i$. However, this trade-off only occurs because the district budget remains the same. If the budget were to increase by at least $5w_i$, the district could buy the additional teachers without necessarily reducing total teacher ability. However, note that although *total* ability may remain the same, *average* teacher ability can still fall because the number of teachers (N_i) has increased. For example, if all the new teachers are inexperienced (A_j=0 for each of the new teachers), then the sum of total experience of all teachers ($\overline{A}$) may remain the same but average experience ($\overline{A}$/N) will fall.

Consider two districts, A and B, that have the same student enrollment but district A has characteristics that require paying teachers higher wages. That is, $w_A>w_B$.[15] Prior to the implementation of CSR, district A has fewer teachers than district B ($N_A<N_B$), and lower ability teachers ($\overline{A}_A<\overline{A}_B$), though both districts have classes larger than twenty students. Under CSR, both districts hire enough teachers to reduce class sizes to twenty students. Because district A started out with larger classes (but the same number of students), it must hire more teachers

[15] We could also assume $p_A>p_B$ but it is not necessary for the results that follow.

while also paying a higher wage. Thus, simply to maintain the same ability level, district A will require a larger increase in their budget. However, both districts receive the same amount of money from the state (since enrollment is the same and the policy gives each district a set amount per pupil). Even if this amount is sufficient to maintain teacher quality in district A, the same amount can buy both the additional required teachers *and* higher quality in district B.[16] Thus, the gap in teacher quality is likely to *increase* under this policy.

Furthermore, this simple model assumes that wages are fixed. When CSR is implemented on a large scale, as in California, it is possible that the simultaneous increase in demand for teachers in all districts will lead to changes in w_i and p_i. In particular, the increased opportunities for teachers in more-desirable districts may cause a fall in the supply of teachers of any given quality to less-desirable districts, and a consequent rise in the salaries required to attract a teacher of any given quality into these districts. This further increases the cost of the CSR policy for these districts. Thus, even in districts where salaries are increasing, teacher quality may still fall and the gap in teacher quality could increase even more.

Did financing matter?

The key issue is that the way in which California's CSR policy is financed does not take into account differing salary costs and hiring needs across schools and districts. By giving the same grant per pupil to all, and beginning implementation for everyone at the same time, such a policy could exacerbate inequities in the distribution of teacher quality. There is an increasing awareness among researchers and policymakers that state policies to distribute *general* school aid should consider the differential costs of educating students with different needs (see

[16] Or conversely, if the amount is sufficient to maintain teacher ability in district B, it will be insufficient for maintaining teacher ability in district A.

Duncombe, Ruggiero and Yinger, 1996; Reschovksy and Imazeki, 2001). But this may be equally important for financing *specific* school reforms as well.

To show how California's policy distributes CSR revenue, I simulate full implementation and compare the amount of CSR revenue that a district was eligible for in the first year of the program to the cost of hiring additional teachers.[17] In the first year of CSR, 1996-97, districts were eligible to receive $650 per K-3 student in a class of twenty students of less. Thus, the maximum amount of revenue that a district was eligible to receive was equal to $650 multiplied by K-3 enrollment. The number of additional teachers needed depends on pre-CSR class sizes; the larger the average class, the more teachers needed to reduce classes down to twenty. To calculate how many additional teachers a district would need to hire, I assume that the number of teachers that a district would need under CSR is equal to total K-3 enrollment in 1996-97 divided by twenty.[18,19] The difference between this number and the number of teachers employed in the previous year (1995-96) is used as a rough estimate of the additional teachers needed to implement CSR fully. It is then straightforward to calculate how much CSR revenue a district would receive per additional teacher needed.

Calculated in this way, seventy-eight districts (nine percent) would not need to hire any new teachers. This does not necessarily mean that these districts already had K-3 classes of twenty or less. However, most did have class sizes much closer to twenty than the average

[17] Although the distributional analysis in the previous section was at the school level, the simulation here is at the district level. Although CSR revenue was technically distributed to schools, CSR revenue is only reported at the district level, and teacher salaries and other budgetary decisions are all set at the district level.

[18] Note that this gives the number of teachers required to have a *pupil-teacher ratio* of 20 to 1. The number of teachers required for *average class sizes* of twenty could be slightly higher (see footnote 10). Thus, this simulation gives only a lower bound on the number of teachers needed.

[19] An alternative would be to calculate the additional teachers needed at each school and sum up the total for the district (this is possible with the data). However, this may overstate the number of new teachers needed because in response to CSR, students and teachers were often re-allocated across schools within districts to keep classes as close as possible to twenty and minimize the number of new teaching positions. The qualitative pattern of results is similar using either method.

California district and did not implement CSR in the first year. They are generally much smaller than other districts (the largest has seventeen K-3 teachers), and are distributed across SES quintiles fairly evenly.

Table 3 shows the results for the remaining districts. On average, salaries for new teachers vary only slightly across SES quintiles, though they are still highest in the lowest-income districts.[20] However, the amount of CSR revenue per new teacher needed is over $10,000 less for the districts with the most low-income students than for the districts with the fewest. This may help explain the increasing gap in teacher experience and credentials discussed earlier; higher-SES schools received relatively more CSR revenue per new teacher needed and thus could afford to hire more experienced teachers.[21] This is consistent with the fact that in schools serving the highest-income students, forty percent of new hires were experienced teachers while in schools serving the lowest-income students, only twenty-four percent of new hires had previously taught elsewhere.

<Insert Table 3 about here>

This simulation is highly simplified and is intended only to give a very general picture of the distribution of CSR revenues. There is no adjustment for changes in enrollment and in that sense, may overestimate the number of new teachers needed specifically for CSR in districts with growing enrollments (i.e., some teachers would have been hired even in the absence of CSR). On the other hand, it almost certainly underestimates the number of new teachers *actually* hired because many schools reduced classes closer to nineteen, instead of twenty, in order to allow for fluctuations in enrollment (California Legislative Analyst's Office, 1997). The state's

[20] It should be noted that when socioeconomic status is measured with the percent minority students or limited English proficient, there are larger differences in salaries across quintiles. Districts with larger populations of minority and LEP students pay higher salaries.

policy is that classes may be larger than twenty on any given day but cannot exceed an average of 20.4 students; to ensure funding, many schools kept classes well under the maximum.

It should also be noted that although the numbers in Table 3 imply that even the lowest-SES districts would have received adequate CSR revenue to cover the minimum salary of new teachers, this does not necessarily mean that it was adequate to cover the full costs of implementation. In addition to basic salaries, CSR revenue also needed to cover non-salary costs such as benefits and training for new teachers, and in later years, increased expenditures for recruiting and hiring bonuses. Together with the tendency of districts to hire more teachers than anticipated, this meant that, in many cases, the cost of CSR implementation was *not* fully covered by revenue from the state. In surveys conducted by the CSR Consortium, the majority of superintendents reported that CSR revenue did not cover the full costs of the program in the first year (Bohrnstedy and Stecher, 2002a). This was mitigated somewhat in later years when the per-pupil grant was increased; however, it still did not fully cover costs in many districts, particularly with salary costs increasing over time as new teachers gain experience and education, thus moving up the salary ladder.

Conclusion

California's experience with a state-wide program of class-size reduction has not been trouble-free. Difficulties recruiting the large number of teachers needed to implement the program led to declining teacher quality across the state, and particularly in schools with larger populations of high-need students. If this deterioration in teacher qualifications offsets the gains from smaller classes, policymakers must surely question whether a comprehensive CSR program

[21] It also possible that the additional revenue allowed higher-SES schools to offer bonuses, or spend more on recruiting, however data on specific expenditures for such items are not available.

is worth the expense. Indeed, several California districts have begun to ask this very question. If California's experience with CSR is indicative of wide-scale CSR programs in general, this may not bode well for other states that are considering such policies. On the other hand, a better understanding of California's experience may provide valuable lessons about how to structure and implement CSR more successfully.

This chapter has provided an overview of the changes in the distribution of teachers under CSR. Teacher quality, as measured by experience, education and certification, fell across the state immediately after the implementation of the program, and the drop was worse in disadvantaged districts. There is some indication that the situation is improving over time, as hiring slows and as recently-hired teachers gain experience and earn their credentials. However, the gaps between high-SES and low-SES schools are still persisting. And it should not be forgotten that there was a good deal of inequity even before the program began; thus, even returning to pre-CSR levels still means disturbing disparities in the education being provided to students of different income levels.

The way in which California's policy is financed may have contributed to the decline in teacher qualifications. A policy that gives a per-pupil grant for each student in a small class, without accounting for salary levels or starting class sizes, may "buy more" for some districts than others. A simple simulation shows that, relative to the costs of the program, schools serving lower-income students receive less CSR revenue under this financing scheme, because they start out with larger class sizes and must pay slightly higher salaries.

The popularity of class-size reduction seems unlikely to wane anytime soon. Many states already target CSR programs at low-income students, and parents and teachers continue to advocate for wider adoption of these policies. As these programs are expanded to include more

schools, the way they are financed should recognize the differential costs of educating students with different needs. Hopefully, California's experience will help policymakers to maximize the effectiveness of CSR policies elsewhere.

References

Bohrnstedt, G.W., and Stecher, B.M. (2000). *Class size reduction in California: The 1998-99 evaluation findings*. Sacramento: California Department of Education.

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

Bohrnstedy, G.W. and Stecher, B.M. (2002b). *What we have learned about class size reduction in California*. Sacramento: California Department of Education.

Borgatta, T. (2002, May 8). Donations to maintain K-3 class-size limits in Irvine. *Los Angeles Times*, p. M1.

California Department of Education (2002). Class size reduction. Available: http://www.cde.ca.gov/classsize/.

California Legislative Analyst's Office (1997, February 12). Class size reduction. Available: http://www.lao.ca.gov/class_size_297.html.

Duncombe, W. D., Ruggiero, J., and Yinger, J. M. (1996). Alternative approaches to measuring the cost of education. In H.F. Ladd (ed.), *Holding schools accountable: Performance-based reform in education* (pp. 327-356). Washington, D. C.: The Brookings Institute.

Ehrenberg, R.G., and Brewer, D. J. (1994). Do school and teacher characteristics matter? Evidence from High School and Beyond. *Economics of Education Review 13*, 1-17.

Finn, J. D., and Achilles, C. M. (1999). Tennessee's class size study: Findings, implications, misconceptions. *Educational Evaluation and Policy Analysis, 21*, 97-109.

Grissmer, D. (2002). Cost-effectiveness and cost-benefit analysis: The effect of targeting interventions. In H. M. Levin and P. J. McEwan (eds.), *Cost-effectiveness and educational policy* (pp. 97-110). New York: Eye on Education.

Hanushek, E.A. (1986) The economics of schooling: Production and efficiency in public schools. *Journal of Economic Literature, 24*, 1141-1177.

Hanushek, E. A., Rivkin, S. G., and Kain, J. F. (1998) Teachers, schools and academic achievement. NBER Working Paper Number W6691. Cambridge, MA.

Hayasaki, E. (2002, May 18). Class-size reduction initiatives faltering. *Los Angeles Times*, p. A1.

Imazeki, J. (2002). Teacher attrition and mobility in urban districts: Evidence from Wisconsin. In Roelke, C. and Rice, J. K. (eds.), *Fiscal Issues in Urban Schools*. Greenwich, CT: Information Age Publishing, Inc.

Imazeki, J. (2001, March). *Moving On or Moving Out? Determinants of Job and Career Changes for Teachers*. Paper presented at the meeting of the American Education Finance Association, Cincinnati, OH.

Jepsen, C., and Rivkin, S. (2002). *Class size reduction, teacher quality, and academic achievement in California public elementary schools*. San Franicisco: Public Policy Institute of California.

Levin, H. M. (1988). Cost-effectiveness and educational policy. *Educational Evaluation and Policy Analysis 10*, 51-69.

Krueger, A. B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics, 114*, 497-532.

Molnar, A., Smith, P., Zahorik, J., Palmer, A., Halbach, A., and Ehrle, K. (1999). Evaluating the SAGE program: A pilot program in targeted pupil-teacher reduction in Wisconsin. *Educational Evaluation and Policy Analysis 21*, 165-177.

Murnane, R.J. (1975). *The impact of school resources on the learning of inner city children*. Cambridge: Ballinger.

Murnane, R. J., and Phillips, B. R. (1981). What do effective teachers of inner-city children have in common? *Social Science Research, 10*, 83-100.

Reschovsky, A., and Imazeki, J. (2001). Achieving educational adequacy through school finance reform. *Journal of Education Finance 26*, 373-396.

Rosen, S. (1986). The theory of equalizing differences. In O. Ashenfelter and R. Layard (Eds.), *Handbook of Labor Economics, Volume 1* (pp. 641-692). Amsterdam: Elsevier Science Publishers BV.

Schevitz, T. (2002, June 23). Small classes produce little gains so far. *The San Francisco Chronicle*, p. A1.

Seaton, D. (2002, April 18). Class sizes to grow: the Riverside board decides to change the pupil-teacher ratio for third grades to save cash. *The Press-Enterprise*, p. B01.

Sonstelie, J. , Brunner, E.J., and Ardon, K. (2000). *For better or worse? School finance reform in California*. San Francisco: Public Policy Institute of California.

Summers, A. A., and Wolfe, B. L. (1977). Do schools make a difference? *American Economic Review, 67*, 639-652.

Theobald, N. D., and Gritz, R. M. (1996). The effects of school district spending priorities on the exit paths of beginning teachers leaving the district. *Economics of Education Review 15*, 11-22.

U.S. Department of Education, National Center for Education Statistics (1996). *SASS by state, 1993-94 Schools and Staffing Survey: Selected results.* Washington, D.C.: author.

U.S. Department of Education, National Center for Education Statistics. (1997a). *America's teachers: Profile of a profession, 1993-94*. Washington, D.C.: author.

U.S. Department of Education, National Center for Education Statistics (1997b). *State indicators in education 1997*. Washington, D.C.: author.

U.S. Department of Education (2000). The class-size reduction program: Boosting student achievement in schools across the nation. Available: http://www.ed.gov/offices/OESE/ClassSize/class.pdf

Table 1. Teacher Characteristics in California Public Elementary Schools

	1995-96	1996-97	1997-98	2000-01
K-3 schools	4461	4503	4535	4692
K-3 teachers (total FTEs)	53,637	64,277	76,569	82,071
K-3 enrollment	1,571,774	1,610,435	1,619,035	1,635,556
School Characteristics:				
Average K-3 FTEs	14.81	17.55	20.65	21.93
Average K-3 enrollment	352.10	357.16	356.69	348.29
Percent of students in free and reduced price lunch	48.75%	50.05%	50.28%	49.28%
Teacher Characteristics:				
Average teacher experience	13.97	12.55	11.91	12.33
Bachelor's degree or less	13.0%	15.8%	18.2%	18.0%
Master's degree or more	25.0%	23.4%	22.8%	25.1%
Lacking a full credential	4.6%	7.7%	12.1%	10.6%
New Hire (first year in district)	8.4%	19.2%	17.2%	9.7%
Inexperienced teachers (less than three years experience)	11.3%	19.3%	23.0%	12.8%
Experienced teachers (ten or more years)	57.6%	51.3%	48.4%	49.3%

Table 2. Socioeconomic Status Quintiles

SES Quintile	1995-96	1996-97	1997-98	2000-01
	K-3 Enrollment			
1	288,177	296,382	297,438	303,132
2	285,834	293,504	293,583	299,671
3	302,716	300,784	302,416	298,253
4	325,167	332,415	331,440	345,858
5	369,313	385,368	393,152	387,657
	K-3 FTEs			
1	10283.8	12522.4	14659.5	15834.7
2	10065.9	12311.5	14307.4	15612.5
3	10575.0	12282.8	14529.6	15374.2
4	11266.4	13303.3	15786.5	17753.8
5	12803.1	15244.1	18423.4	19791.7
	K-3 Schools			
1	893	901	907	939
2	892	901	907	939
3	892	900	907	938
4	892	901	907	938
5	892	900	907	938
	Percentage of students in free and reduced price lunch program			
1	0 - 19.3	0 - 19.3	0 - 19.2	0 - 17.2
2	19.3 - 38.9	19.3 - 40.6	19.2 - 40.8	17.2 - 39.7
3	38.9 - 57.9	40.6 - 60.3	40.8 - 61.0	39.7 - 60.4
4	57.9 - 78.4	60.4 - 80.2	61.0 - 80.3	60.5 - 79.8
5	78.4 - 100	80.2 - 100	80.4 - 100	79.8 - 100

Figure 1.
Teacher Characteristics, by Student Poverty
1995-96

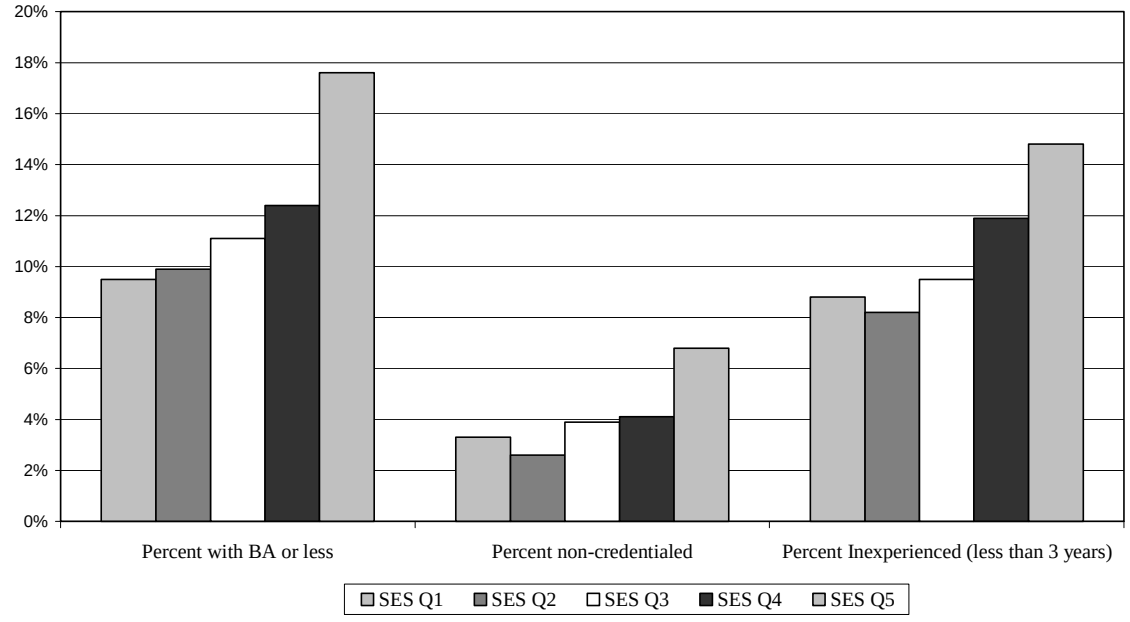

Figure 2.
Percent of Teachers with Bachelor's Degree or Less

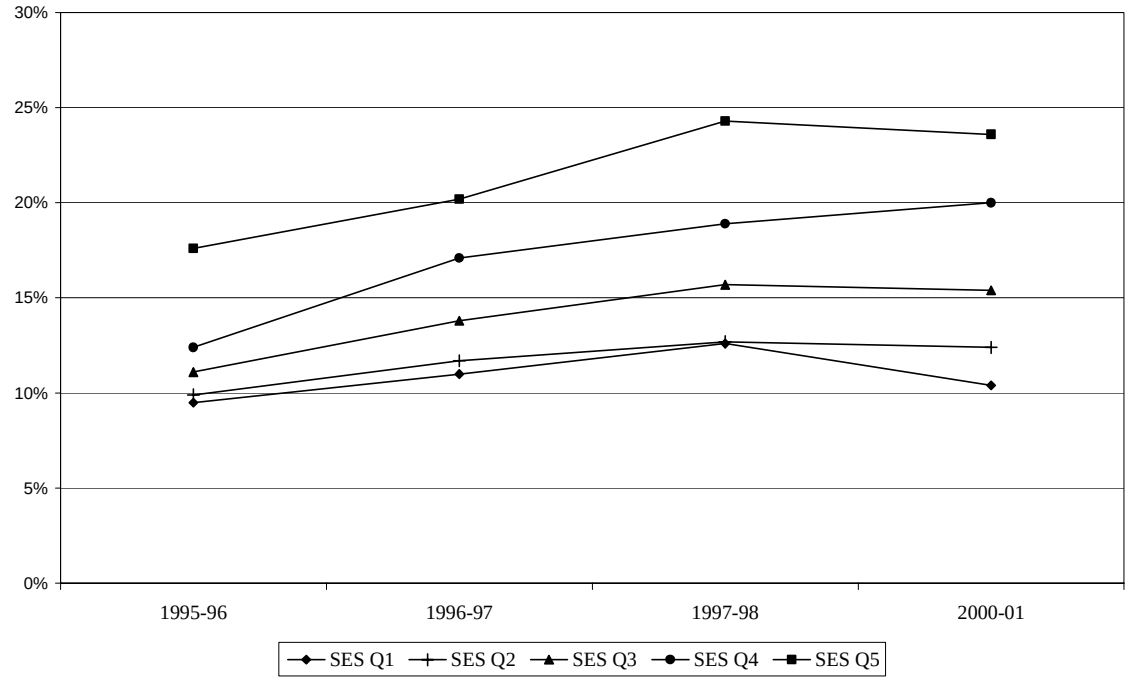

Figure 3.
Percent of Teachers with Less than Three Years Experience

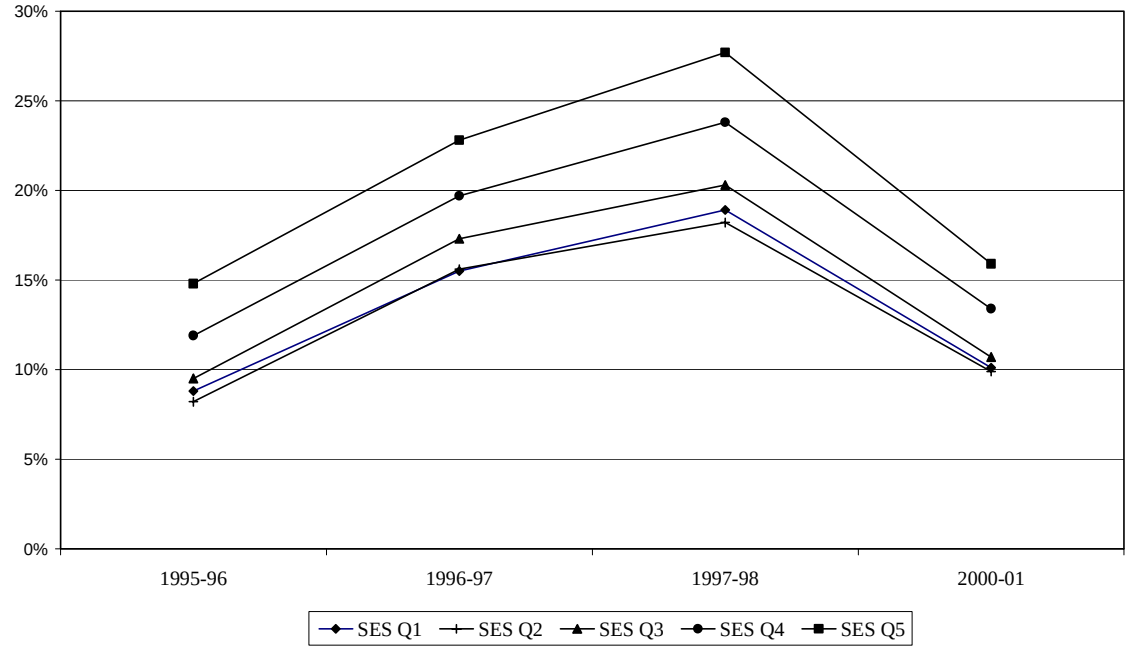

Figure 4.
Percent of Teachers Lacking Full Credential

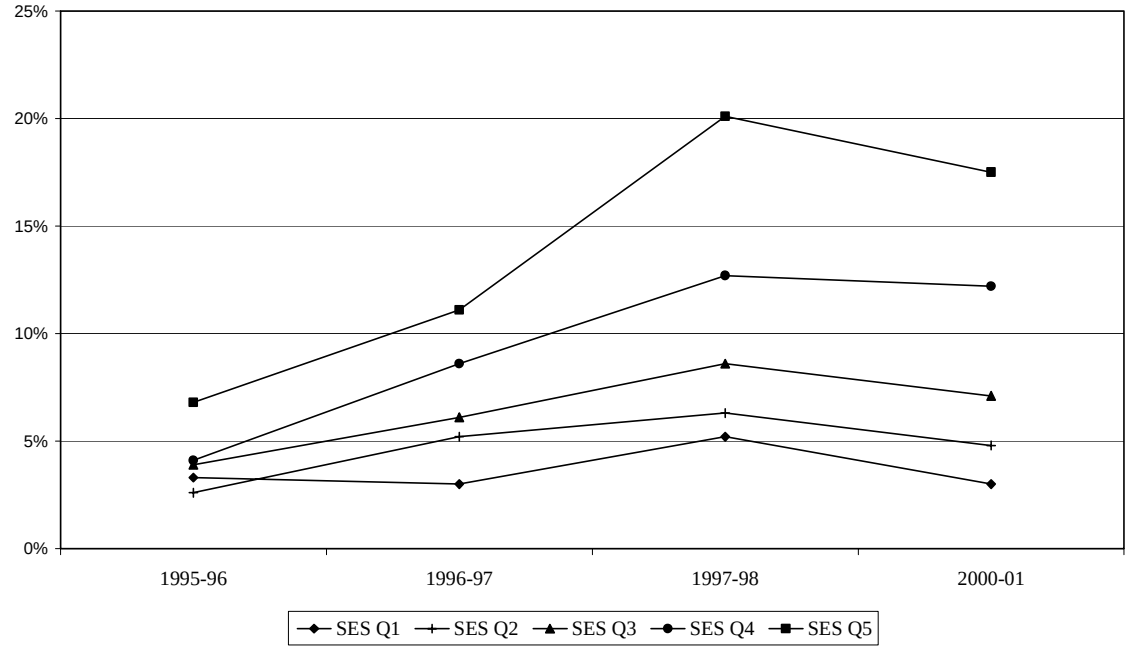

Table 3. Simulation of Full CSR Implementation

SES Quintile	1995-96 FTE	FTE for full implementation of CSR	Additional FTE needed	CSR revenue per new FTE	Minimum Salary
1	49.81	73.53	23.73	$ 55,815.42	$ 26,989.56
2	63.75	95.78	32.04	$ 45,681.53	$ 26,760.56
3	72.26	109.71	37.45	$ 50,777.37	$ 26,614.96
4	96.34	145.13	48.79	$ 45,426.58	$ 26,773.74
5	63.67	97.56	33.90	$ 43,546.48	$ 27,032.68