Boosting Student Achievement: The Impact of Comprehensive School Reform on Student Achievement

Kevin Booker, RAND
Betheny Gross, Center on Reinventing Public Education
Dan Goldhaber, Center on Reinventing Public Education

information contact
Betheny Gross
206-616-2092
betheny@u.washington.edu

June 12, 2006

Center on Reinventing Public Education
Daniel J Evans School of Public Affairs• University of Washington, Box 353055 • Seattle WA 98195-3055
206.685.2214 • www.crpe.org

CRPE working paper # 2006-4

Abstract:

Over the last 10 years, federal, state, and local education agencies have devoted considerable resources to support the implementation of Comprehensive School Reform (CSR) models. In addition to the 1.6 billion federal dollars distributed through the Comprehensive School Reform Demonstration (CSRD) project and its successor, the CSR program, states and districts have made CSR adoption a central reform strategy for their lowest-performing schools. Despite a policy commitment that extends across multiple education sectors, researchers have failed to find consistent evidence that CSR adoption leads to improvement in student performance, with much of the work criticized for weak methodological approaches. The work presented here employs the most promising analytic techniques available for non-experimental studies to investigate the effects of implementing CSR models on student achievement. With a dataset of Texas student achievement scores, we test the hypothesis that receiving CSRD funding has a positive effect on student performance at the award schools.

Keywords:

Education policy; school reform; student achievement

Acknowledgements:

The authors gratefully acknowledge funding for this study from the U.S. Department of Education. All opinions expressed in this paper represent those of the authors and not the institutions with which they are affiliated or the Department of Education. All errors are solely the authors' responsibility. The authors would also like to thank anonymous reviewers, Jane Hannaway of the Urban Institute, and seminar participants at the 2006 American Educational Finance Association conference for their thoughtful feedback, and Carol Wallace for her editorial suggestions.

THE EVOLUTION OF COMPREHENSIVE SCHOOL REFORM

Comprehensive School Reform (CSR) and whole-school reform have been a significant factor in federal education reform policy since the 1994 reauthorization of Title I. With the 1994 reauthorization, the U.S. Department of Education departed from its previous policy of requiring schools to use Title I funds for programs and materials that specifically targeted low-income students and permitted schools with high concentrations of poverty to use these funds for school-wide improvement efforts. In 1998, the Department of Education expanded its commitment to school-wide improvement by joining a growing reform movement in which independent developers devised researched-based school designs that aligned school governance, curriculum, and practice (popularly known as Comprehensive School Reforms) via its initiation of the Comprehensive School Reform Demonstration (CSRD) project.

The CSRD project provided three-year grants for schools to implement one of several Comprehensive School Reform (CSR) designs. Today this grant program has evolved from a demonstration project into a reform policy embedded within the 2001 No Child Left Behind Act and continues to provide schools with funding to implement CSR designs (Tushnet, Flaherty and Smith, 2004). CSR has become a leading strategy endorsed by the Department of Education, and between 1998 and 2005, federal grants totaling more than $1.8 billion were distributed to more than 6,700 schools across the country to support their CSR implementation (Department of Education, 1998-2005).[1] However, it is not only at the federal level that CSR designs are being used to address the educational issues of the most-challenged students. Countless other schools nationwide have adopted these designs using a variety of other funding sources with the encouragement of their states and districts. For instance, the Abbott v. Burke decision regarding the financing of New Jersey schools requires the state's lowest-performing districts, which are also the districts serving the highest concentrations of low-income and minority students, to fund and adopt a CSR design in their lowest-performing schools.[2]

Although CSR designs aim to improve learning for all students, the Department of Education, in keeping with this policy's Title I roots, gave favored status to schools with high

---

[1] All statistics on CSR funding have been obtained from the SEDL CSR database, available online at http://www.sedl.org/csr/awards.html.

[2] The Abbot v. Burke decision that established the link between funding in low-performing districts and reform models was initially handed down in 1997, but the courts continued to rule on the case through 2003. For a summary of these decisions see http://www.state.nj.us/njded/abbotts/dec/. Although schools have the option to select from among 13 approved models, if they do not select a model they are required to implement *Success for All*. For more detailed information see http://www.state.nj.us/njded/abbotts/wsr/models.htm.

concentrations of poverty when making grants. A study of the implementation and impact of these CSR grants found that schools serving high concentrations of low-income and minority students received a higher share of grants (SEDL, 2003; Tushnet et al., 2004). While the allocation of funds is important in considering equity, it is only a secondary question to the central issue of student achievement: what we really want to know is whether the grants for CSR and the CSR designs these funds bring to schools actually improve student learning and student performance. Furthermore, we want to know if these funds and reform designs have had particular success with low-income students and other students who have traditionally struggled in American public schools.

A meta-analysis by Boreman et al (2003) shows that educational researchers have produced a great deal of literature on both the effects of specific designs and the general effect of CSR on school outcomes. That is, we have learned whether specific designs or schools implementing one of a variety of reforms can be linked to gains in school-level, aggregate scores. However, little research has explored the impact of the CSR grants policy itself. It remains to be seen if the policy approach of providing financial assistance for a broad range of reform strategies while leaving the selection of strategies to the local schools and districts will result in across-the-board improvement. And because the Department of Education does not mandate specific changes or provide oversight, the success of this policy is largely a function of the local selection and implementation processes.

Unfortunately, thus far little attention has been paid to the overall impact of the government's CSR policy on student achievement, and no one has explicitly explored the differential impact of these grants and reform designs on individual students. This gap in the research can be largely attributed to the difficulty of obtaining student-level achievement data. Examples of studies that do examine student-level outcomes include those by Bifulco, Duncombe, and Yinger (2005), Bloom, Ham, Melton, and O'Brien (2001), and Supovitz and May (2004). While each of these studies provides the field with valuable methodological and conceptual insights (we discuss the findings from these studies in greater detail below), none have a sample that allows for the examination of the variation of impact across different types of students.

The goal of the research we present in this paper is to determine the extent to which CSR grants have improved students' achievement in schools receiving awards. Specifically, we examine the policy impacts on individual student achievement rather than school-level aggregate scores in order to generate more precise estimates of the impact of CSR funding on students in these schools and to differentiate the effects of CSR across different types of students. Given that this is an evaluation of the policy of funding schools to adopt CSR designs, and schools have great flexibility in determining which CSR design they adopt (nationally, the adoption of more than 600 externally developed designs and more than 700 home-grown models have been subsidized with federal CSR funds[3]), it is important to stress that, here, we are examining the *average impact of receiving federal CSR funds*, not the impact of CSR model implementation on students. The effect of model implementation is also worthy of focus, as some research has found considerable variation in the implementation of CSR designs across schools and has also found that the level of implementation has implications for the designs' success in the school (Datnow, 2000; Desimone, 2002; Vernez, Karam, Mariano, and DeMartini, 2004). This research also is limited in that it speaks only to the effects of receiving CSR funding. Given the popularity of CSR designs nation-wide, many schools have implemented a CSR even in the absence of an award. Unfortunately, we cannot identify all schools that implement any one of the hundreds of designs available.

THE VALUE OF STUDENT-LEVEL ANALYSIS IN CSR RESEARCH

The use of student-level data for evaluating CSR offers two very powerful technical advantages and one conceptual advantage over the school-level data that most previous CSR studies have relied on. First, analysis with student-level data allows researchers to control for the mobility of students. The issue of student mobility, especially of low-income, urban, and minority students, is well documented in educational literature (Engec, 2006; Hartman, 2006; Rumberger, 2003). And mobility is a particular concern in the evaluation of CSR, because CSR designs take time to mature in a school (Tushnet et al., 2004) and require that students be exposed to a specific curriculum or set of instructional practices in order to benefit from the designs. Students that transition in and out of schools cannot be expected to benefit in the same way as students with more consistent exposure. Therefore, we have good reasons to believe that it is necessary to explicitly account for mobility – something that cannot be done with only school-level performance indicators – in order to fully understand the effect of a CSR design on students who are exposed to the program.

[3] Statistics regarding the number of CSR designs adopted with federal CSR funding are taken from http://www.sedl.org/csr/awards.html.

The second technical advantage of student-level data is that it allows for the use of controls to account for unobservable differences in both school and student populations with models specified in terms of school and student fixed effects (or campus-student spell effects). While the school-level data that is most often used for CSR evaluations allows for an accounting of unobservable school characteristics, it is only with student-level data that we can take the extra step of controlling for unobservable student-level characteristics. This extra step is important because even if we compare two schools that appear to have similar populations based on observable student characteristics, there could still be unobservable differences between schools or in student ability and new research shows that failure to control for these differences can lead to biased estimates of schooling variables (Clotfelter, Ladd, and Vigdor, forthcoming). Student-level data allows us to control for this directly using student fixed-effects models or similar models referred to as spell-effects models.

Finally, student-level data provide a conceptual advantage by giving us the opportunity to examine the differential impact of CSR designs across students of different types. The Department of Education conceived of these CSR implementation grants as a means to improve achievement in high-poverty schools. While most CSR developers indicate that all students will benefit from their designs, many tout the specific benefit their programs provide to schools' most challenging populations. For example, publicity materials for *Success for All* (a popular CSR design nationwide) state, "Our top priority is the education of disadvantaged and at-risk students in pre-K through grade eight."[4] Although CSR designs can vary rather substantially in pedagogical approach (e.g., constructiveness versus direct instruction) as well as the level of prescription with regard to curriculum and assessment, many incorporate instructional practices such as tutoring, additional time for reading instruction, and regular performance assessments that are of particular value to students who have traditionally struggled in school. Thus, we might reasonably expect these programs to impact some students more than others. Although Slavin and Madden (2001) report that *Success for All* has been shown to be especially effective with schools' lowest-performing students, overall, limited research is available on the differential effects of CSR.

Using primarily exploratory case analysis, researchers examining CSR implementation have explored how well core elements of CSR designs meet the needs of special populations (including English Language Learners (ELL), low-income students, and minority students), and

---

[4] See http://www.successsforall.net

have found instances in which the CSR designs alone are inadequate for their needs. Cooper and Jordan (2003) suggest that a tandem strategy of emphasizing increased recruitment of African American male teachers along with CSR design implementation will be necessary to meet the unique needs of African American male students – CSR design implementation alone is insufficient. In a study examining the state administration of CSR designs and its impact on ELL students, Hamman, Zuliani, and Hudak (2001) find that the CSR designs failed to provide specific accommodations for ELL students, suggesting that the programs actually provide little additional benefit to this population. Finally, Koh, Siegel, and Robertson (2003) investigate teacher perspectives of three comprehensive school reform models (*Accelerated Schools, Roots and Wings*, and *Voices of Love and Freedom*) in an effort to determine the designs' impacts on special education. In this study, special education teachers reported that, although the models promoted inclusion and social interaction for special education students, the designs' fast-paced curriculum and lack of provisions for making modifications to meet the particular needs of special education students were weaknesses of these reform approaches.

While these case analyses provide useful information regarding the extent to which CSR designs explicitly address the concerns of special student populations, we cannot make generalized statements about their impact on these special populations from this research. We hope our analysis of federal CSR funding in Texas schools will fill an important gap in the understanding of this issue.

We also hope our research extends the work of previous researchers who have responded to criticism surrounding previous CSR evaluations by using student level data. In one example of student-level CSR research, May and Supovitz (2004) examine the impact of the *America's Choice* design in Rochester. Using data spanning nearly a decade, they test whether the growth trend of students in *America's Choice* schools differed from that of similar students in non-*America's Choice* schools. While this analysis effectively controls for the mobility of students, it does not control for all school fixed effects, potentially introducing bias into the estimates of model effects. Bloom et al. (2001) offer an interrupted time series approach to studying the effects of the *Accelerated Schools* CSR design on student achievement. This approach tests whether students' performance trends differ after being exposed to a CSR design relative to before they were exposed, which effectively accounts for selection bias at the student level. Bloom et al.'s analysis, however, is limited to students in only eight schools and is therefore restricted in its generalizability and its ability to account for school selection bias. Finally, Bifulco et al. (2005) offer a valuable study of *More Effective Schools*, *Success for All*, and the *School Development Program* in New York City schools, using a school fixed-effects model to account for school-level selection bias. While this previous student-level work has contributed a great deal to our understanding of the effects of CSR and the methods that will allow us to capture these effects, it has not been able to fully account for the potential student- and school-selection bias in the same model and with a large-scale dataset.

DATA AND METHODOLOGY

Our analysis of federal funding for CSR compares a school-level analysis—the traditional level of analysis for CSR research—of the grants' effects on reading and math achievement to a variety of student-level analyses. These student-level analyses explore the impact of CSR on individual students, the differential effect across student subgroups, and the differential effects based on a student's exposure to federally funded CSR schools. We pursue these analyses with data from Texas that span the 1996-97 through the 2003-04 school years. During this time 181 Texas elementary schools received grants to implement CSR designs, with statewide awards totaling more than $60 million. As can be seen in **Table 1**, sixty-five schools received awards in March of 1999 and 117 schools received awards in July of 2001. Our data include observations on all public school students in Texas in grades 3 through 8. The data provide student math and reading scores on the statewide standardized exam (until 2001-02 this was the TAAS exam; in 2002-03 the state switched to the TAKS exam). The data also contain student, family, and program characteristics including gender, ethnicity, eligibility for a free or reduced-price lunch (which is used an indicator of economically disadvantaged status), limited English proficiency, and participation in special education. Each student has a unique identification number, which allows us to follow students as they switch schools and progress through the educational system.

(INSERT TABLE 1 ABOUT HERE)

In order to make the student test scores comparable across years and to remove any statewide trends in scores, we normalize the scores within grade and year to a standardized Z-score with a mean of zero and a standard deviation of one.[5] The estimation sample includes

---

[5] Our representation of student achievement with a Z-score implies that the annual mean of all scores within a grade is zero even though student achievement has been noted as increasing statewide between 1996 and 2004. We choose to use a z-score in which the total variance is the denominator instead of the within-district or -school variance for two reasons: (1) this z-score indicates how students' achievement is changing relative to their counterparts across

student observations in grades 3-8 and years 1996-97 through 2003-04. For computational tractability we only include in our data a random ten percent sample of the students who were never observed at a treated school. All students who were ever observed at a treated school are included in the estimation sample. The regressions are weighted to account for the sampling.

One concern that is raised in any educational policy study, especially one that occurred at such an active time in educational policy, is the coincidence of different policies targeting the same population of schools. Beginning in the mid-1990s, Texas operated a school accountability program that identified the state's lowest-performing schools—a population of schools that likely overlapped the schools receiving CSR awards. Identified schools suffered sanctions and received support for improvement. If identification for state accountability and CSR awards occurred at identical times in identical schools we could not convincingly disentangle the effects of the CSR award from the effects of state accountability identification. While there is an overlap in identified schools and CSR award schools the number of schools identified for low performance far exceeds the number of schools with CSR awards. Therefore, we do not feel this overlap undermines the analysis.

A second concern that arises specifically in the context of investigating CSR is that of selection. Schools have to apply to receive federal CSR funding. Not only is this process voluntary, it requires the entrepreneurial efforts of school and district administrators to complete the application and seek out a CSR reform design and thus a significant investment of time and effort. As a consequence there are good theoretical reasons to believe that schools receiving CSR awards may differ from non-awarded schools in important and unobservable ways. We attempt to account for this though the use of fixed effects and spell effects models that implicitly control for time-invariant observed and unobserved characteristics of schools and students.

Each of the models presented in this paper are student-level models that become successively more detailed as we add first school fixed effects and then student fixed effects. With these models, we address the selection bias that is inevitable with non-random samples in which the awarded schools are not necessarily comparable to non-award schools. Our basic model estimates the effect of a CSR award on the one-year change in student achievement given in equation one.

$$(1) \qquad A_{jst} - A_{jst-1} = CSR_{st} + X_{jst} + Y_{st} + \theta_{gt} + \nu_{jt}$$

---

the system and not just within their own school, and (2) it allows us to eliminate the chance that we confound statewide trends or the shift in scores due to the transition from the TAAS to the TAKS.

A_{jst} is the average achievement of student *j* at school *s* in year *t*; A_{jst-1} is the average achievement of student *j* at school *s* in year *t-1*; CSR_{st} is the CSR treatment status of school *s* in year *t*; X_{jst} is a vector of student characteristics including racial or ethnic background, gender, free or reduced lunch (FRL) status, special education status, and limited English proficient status (LEP); Y_{st} is a vector of school characteristics including percent of students identified as LEP, percent of students identified by ethnic or racial demographic (African American or Hispanic), percent of students enrolled in special education, and percent of students on FRL, average teacher experience, and student-teacher ratio; θ_{gt} captures the grade-by- effects[6]; and v_{st} is the random disturbance term. This model does not account for school or student fixed effects.

It should be noted that some debate remains over the most appropriate specification for educational achievement models. The most common specifications include an achievement-levels approach where achievement in year *t* is modeled as a function of individual and school characteristics, and one of two value-added approaches: either achievement is modeled as a function of lagged achievement and school and student characteristics (a value-added "lagged" approach), or the first difference where a one-year change in achievement is modeled as a function of school and student characteristics (a value-added "change" approach). We opted to employ a value-added approach, since it is both a more widely used empirical specification today and generally perceived to be superior to estimating achievement levels (Hanushek, 1986). While the two value-added approaches are conceptually similar, there are some important trade-offs between them that makes the question of which one to use complicated.

The change approach implicitly assumes that an individual's prior learning from the prior year does not decay, while the lagged achievement approach has the flexibility to estimate the rate at which prior learning decays. The parameter estimates from either achievement model, however, will likely be biased if there exists decay of knowledge, serial correlation of schooling characteristics (arising, for instance, from the sorting of students to teachers based on their prior achievement), or any factors affecting student achievement that are omitted from the model (Rivkin, 2005; Todd and Wolpin, 2003).

We have opted to use the change approach in specifying our models for several reasons.[7] First, the potential bias created from the change specification is predictable (depending on the

[6] These grade-by-year effects, in addition to controlling for potential systematic differences across years and grades, also account for grade retention of students.
[7] Bifulco et al. (2005) provides an example of a study that models the effects of CSR with value-added models that include lagged achievement A_{jst-1} as an explanatory variable.

degree of serial correlation and rate of decay of knowledge), consequently our CSR estimates of effects will be conservative. Second, the inclusion of a lagged measure of achievement as an explanatory variable will almost necessarily be biased in *unpredictable* ways in the face of any omitted inputs in the model. Finally, it is necessary to instrument for lagged achievement if student fixed effects are included in the model, and this creates additional complications in the spell models we introduce below. A typical instrument for the one-year lag is a two-year lagged achievement level, but we cannot use this due to data constraints. Therefore, using the change specification allows us to predict the bias in the model and be consistent across the three different models we introduce.[8] Having said all that, it's worth noting that our findings remain qualitatively similar whether we employ a lagged or change specification approach.

Our second set of models, specified in equation 2, extends from the baseline model of achievement change presented in equation 1 by including a school fixed effect.

$$(2) \qquad A_{jst} - A_{jst-1} = CSR_{st} + X_{jst} + Y_{st} + \theta_{gt} + \mu_s + \nu_{jst}$$

In these models Y_{st} is a vector of *time varying* school characteristics, and μ_s is the school fixed effect. Since μ_s captures all of the stable characteristics of schools, Y_{st} reflects only time varying school characteristics including the student economic, demographic and education concentrations described for model 1. These models substantially improve the analytic comparison from model 1. Instead of comparing the gains made by students in award schools to students outside award schools, these models estimate the effect of the CSR awards within schools. Specifically, these models estimate CSR award effects by comparing the score of students in the school before the award to student scores in the school after the award was received. By comparing the achievement of schools, or students within the schools, over time we eliminate the school selection bias that is due to the systematic targeting of awards to low-performing schools serving low-income populations as well as the bias generated by the school characteristics that cannot be easily measured. While the school fixed effects potentially eliminate much of the bias due to selection at the school level, different achievement trends across award and non-award schools prior to the award period may also bias the estimates. In the cohort of schools that received awards in 2001, we found that, on average, award schools showed a slightly lower trend in achievement than non-award schools. Therefore, these estimates should be considered lower-bound estimates of award effects.

[8] For a more detailed discussion of all these issues, see Rivkin (2005) and Todd and Wolpin (2003).

Our next models, however, account for both school and student fixed effects with spell model specifications. This specification is shown in equation 3.

$$(3) \qquad A_{jst} - A_{jst-1} = CSR_{st} + X_{jst} + Y_{st} + \theta_{gt} + \phi_{js} + \nu_{jt}$$

X_{jst} is a vector of time-varying student characteristics including free and reduced price lunch status, special education status, and limited English proficient status. Y_{st} is a vector of time varying school characteristics as described above and ϕ_{js} is the school-student spell fixed effect. While equation 2 compares the average change in student scores for all students in the school prior to the award to that of students in the school after the award, the models specified according to equation 3 compare the pre-award and post-award changes in achievement scores for individual students who were in the school both before and after the school received the award.[9] One important advantage of examining how an individual student's performance changes after her school receives a CSR award is that the estimation of the effect does not include mobile students who have only a short history with the school. The mobility of students in the context of evaluating CSR awards is important for two reasons. First, because highly mobile students often struggle academically for a variety of reasons, the performance of these students should not be fully attributed to the school's instructional program. Second, schools implementing CSR designs (with or without awards) often advertise the use of these models to attract students to their school. By estimating the award effects within students we eliminate the selection bias that results from student selection. While it is important that schools address the needs of mobile students and it is valuable to attract new students with a CSR design, our goal in this analysis was to determine the effect of CSR grants on students in the schools that received these awards. The best way to measure this effect is to hold the student and the school constant.

RESULTS

Our analyses start with the basic models of the effects of CSR funding on changes in student achievement, then move on to the school fixed-effects models and the more detailed

---

[9] As with the school fixed-effects models, the spell-effects models do not allow us to explicitly estimate the effects of stable student characteristics on performance. These factors, such as economic background, race, language history, and gender are well documented as factors associated with a student's academic performance. Our interest in this analysis, however, is to separate the effect of federal CSR funding on student performance from all of the other factors that potentially contribute to the student's performance. While the use of student fixed effects reduces the total information gleaned from the analysis, we feel this loss of information is justified because the technique allows us to minimize the bias introduced by unmeasured stable characteristics of students, such as attitudes and dispositions toward learning or family support.

spell-effects models. Across these analyses we specify CSR funding awards in a number of ways including (1) indicators specifying that a school received an award and when the school has moved into the post-award status, (2) a linear trend of years since the award was received, (3) a year-by-year effect of having received the award, and (4) the per-pupil award amount. The coefficient estimates across all of the models can be interpreted as the percent of a standard deviation in the student's raw score. A one standard deviation difference in raw score equates to approximately 8-10 points in the Texas Learning Index (TLI) scaled score – the scale used to report Texas Assessment of Academic Skills (TAAS) and Texas Assessment of Knowledge and Skills (TAKS) results to schools and students. The mean TLI score is approximately 78, with a maximum score generally ranging between 93 and 100. For reference to the proficiency scale, the base score for proficiency is set at 70. Summary statistics for the student characteristics and school characteristics used in these models are provided in **Table 2**.

(INSERT TABLE 2 ABOUT HERE)

Our baseline models estimate the average change in student achievement in reading and math as a function of school and student background information, and an indicator for a student's presence in a school receiving federal funding for CSR. Again, these models offer a very basic comparison between students in award schools to students in non-award schools. These models do not correct for selection bias inherent in the CSR application or selection process. In the math models in **Table 3**, we see that point estimates for the various specifications of CSR funding are negative and significant, indicating that the students and schools in award schools are showing lower gains in math achievement compared with students in non-award schools. The point estimates indicate that math gains for students in award schools were, on average, about 2 percent of a standard deviation (model 1) lower than their non-award school peers, with an average annual gain of about 0.7 percent of a standard deviation lower than non-award school peers (model 3). Results from model 2 suggest that even larger per-pupil awards yielded lower math gains in achievement. The estimates for reading are inconsistent, with most models showing CSR to have non-significant effects. The only statistically significant effects on reading scores are for the effect of CSR in years five through six in model 4.

In the context of these baseline models, it is worth remembering that schools self-select into CSR application status (and are selected to receive funds by the federal government). The CSR program is likely to be significantly more attractive to high-poverty (a requirement for receipt of funds), low-performing schools since schools in need of improvement are mostly likely to wish to make significant changes to their educational programs. In fact, many CSR programs specifically target these types of schools.

To help address the potential selection bias, we estimate school fixed-effects variants of the models discussed above. These models essentially control for all time-invariant characteristics of schools and would, therefore, control for the factors that drive the selection of schools for CSR funding.[10] Hopefully, these models eliminate school-level selection bias by comparing students in the school before receiving the award to students in the same school after receiving the award to generate estimates of the award's effect.

A comparison between the math results of our school fixed-effects models (which are presented in **Table 4**) to the math results for models with no fixed effects (presented in **Table 3**) illustrates the importance of addressing school-selection bias with fixed effects. In these models, the coefficients on the different specifications of CSR funding for math—although still negative and significant—are smaller and less significant than in the models with no school fixed effects, possibly reflecting the fact that many CSR designs focus on reading instruction with new reading curriculum, tutoring, and extended reading periods. Interestingly the opposite pattern is present in the reading results where the point estimates from model 1 not only become negative but statistically significant at a 90 percent level of confidence. Nonetheless, these results suggest that accounting for school fixed effects is important in assessing the effect of CSR awards.

(INSERT TABLE 4 ABOUT HERE)

Spell-effects models. While researchers studying CSR widely accept school fixed-effects models as competent models for the evaluation of CSR with large-scale data, these models still potentially reflect bias that may result from student selection of schools. As Bifulco (2002) and Bifulco et al. (2005) explain, while school fixed-effects models account for what seems to be the most conceptually significant source of bias in these evaluations, there is still some chance that parents are moving their children to schools based on the presence of CSR programs in the school. While this effect is likely much less pronounced than the school-selection bias, it is nevertheless still a potential source of bias.

Models accounting for this additional source of bias would include both school and student fixed effects, but we are unaware of any research that has done so; this may be explained, at least in part, by the fact that it is not possible to estimate such models without a large sample

[10] This specification, however, cannot account for the bias that would result if CSR award schools displayed a different trend than non-award schools prior to the award periods.

of CSR schools that include student-level information spanning several years of time. Fortunately, the structure of the Texas data is such that we have the opportunity to explore models that account for both school fixed effects and student fixed effects. The models we explore are known as spell-effects models and yield results much like the traditional fixed-effects models with school and student fixed effects.

We tested a series of models beginning with a model of the effect of CSR funding on the one-year achievement gains of students in these schools, and then moved to more detailed models exploring the differential impact of federal CSR funding on different types of students. It is important to remember that, with spell-effects models, the students contributing to the estimate of the effect of some event (for example, receiving an award or moving to post-award status) are those enrolled in the school both before and after the event occurs. For example, the students contributing to the estimate of the effect of federal CSR funding are those who were in award schools both before and after the school had funding.

In **Table 5**, we present the results of multiple spell model specifications that reflect a school in award and post-award status, the length of time since the school received the award, and the per-pupil award amount. The findings from these models are qualitatively different from the school fixed-effects specification results presented in **Table 4** and present a strikingly different picture from the baseline model results we present in **Table 3**. Specifically, we see that CSR funding has had an impact on student performance in reading and math, that CSR funding has had a continuing effect on students' achievement, and that to a small extent the effect of funding is a function of the per-pupil amount of the award. These differences in findings reflect the fact that the spell-effects models compare the performance of individual students in schools that received funding to *their own performance* before the school received funding, while the school fixed-effects model compares the overall performance of students in the school after the school received the award to their overall performance before it received the award, and the baseline models simply compare the performance of students in schools that have or have not received CSR funding.

(INSERT TABLE 5 ABOUT HERE)

The most concise picture of the CSR funding policy impact is a specification that captures both the effect of receiving the awards (which are granted for three years) and the effect of moving to post-award status. Our estimates for this model, identified as model 1 in **Table 5**, show that after schools received an award the students in the school gained, on average, 2 percent of a standard deviation more in math and 3.5 percent of a standard deviation more in reading than students in non-award schools. These gains amount to less than a one-point increase in the Texas Learning Index (TLI) scores for both reading and math. The effects, however, continue as the school moves into post-award status. Here, award students out-gained non-award students in math by almost 11 percent of a standard deviation while reading scores increased by 7 percent of a standard deviation.[11]

The next specifications (models 2-4) illustrate the long-term effect of CSR funding, with an estimate of the average annual increase for students after the school has received an award and an estimate of the year-by-year increases seen in student scores for the first six years after award receipt. Model 3, which tests a linear trend in post-award school performance, estimates that students in award schools improve on their non-award peers by 4 percent of a standard deviation and 3 percent of a standard deviation for math and reading, respectively. Model 4, which tests long-term effects but does not assume a linear trend, estimates the year-by-year increase seen in student test scores for the first six years after a school received an award.[12] These models show that gains by students in award schools bettered their peers from non-award schools in both math and reading in years three through six. In math, the impact of awards on student gains reached almost 29 percent of a standard deviation in year six. Students saw the greatest effects of CSR awards in reading in years four through six, where impacts reached as high as 18 percent of a standard deviation (around two points on the TLI).

The last specification in these models examines the effect of the award amount. For these we included one variable giving the log per-pupil award amount and another indicating the maximum log per-pupil award for schools in post-award status. We listed the log award for school not receiving awards as zero. Model 2 in **Table 5** indicates that the per-pupil amount of the award is positively associated with gains in student performance in math but only in the post-award period. In the long term, students in CSR award schools showed better gains that students in non-award schools and their gains are improved with the size of the school's award.

Our last series of spell-effects models explores the extent to which funding affected student subgroups differently. The ability of schools to adequately address the needs of student subgroups defined by economic advantage, minority status, and special educational needs has

[11] Some may argue that it seems unreasonable to expect gains in performance during the first funding year. Therefore, we estimated models that tested the effect of CSR funding starting one year after the receipt of an award. We did not find a substantial difference from the models described in this paper.
[12] We only tested up to six years after an award because that is the maximum length of time a school could have had an award in the years for which we have data.

received considerable attention now that the No Child Left Behind Act holds schools accountable for the performance of student subgroups in addition to overall school-level performance. The intention of this requirement is to force schools to give equal attention to the needs of even the most under-represented students in a school. Given this growing concern over the performance of student subgroups and the original intention of the CSRD policy to improve achievement in high-poverty schools, it is important to investigate the policy's impact on major student subgroups. In the analysis described below, we investigate how CSR funding impacted reading and math achievement for students in five subgroups: students receiving free or reduced-price lunch (FRL), students identified as African American, students identified as Hispanic, students receiving special education services, and students classified as Limited English Proficient (LEP).

Table 6 presents results that parallel to some extent the story illustrated by those shown in **Table 5**: being in an award school improves math and reading gains for many students; we can detect a linear-improvement trend reflecting increasing performance relative to students in non-award schools each year after schools are awarded CSR funding; and the amount of the award is positively associated with performance, though the effect is relatively weak. More importantly, however, these results show that federal CSR funding affected different subgroups of students differently. In this analysis we examined three different specifications of CSR funding: the first estimates the effect of a school receiving an award and moving into the post-award status interacted with each student subgroup; the second estimates the linear trend for the years since a school received an award interacted with the student subgroups; and the third estimates the effect of per-pupil funding during the award period and during the post-award period interacted with student subgroups. It is worth noting here that, when we compare the effects for the baseline group shown in the first row of **Table 6** to the effects for the similar models without interactions presented in **Table 5**, the main effects in the interaction models are smaller for all model specifications. This indicates that some portion of the effect seen in the simpler models from **Table 5** can be attributed to the effect of CSR funding on the subgroups investigated in the various interaction models, thus CSR funding impacted students in the subgroups differently than the baseline group of students.

(INSERT TABLE 6 ABOUT HERE)

Three findings from the student-subgroups models stand out: (1) CSR funding most consistently impacted LEP students; (2) CSR funding showed greater distinction in reading performance between baseline students and subgroup students; and (3) special education students did not realize the positive CSR award effects seen in other subgroups. It is immediately obvious that LEP students showed the most consistent gains on non-award students across the model specifications. This subgroup showed gains above the baseline award group for each of the three specifications and in both reading and math. Of all subgroups, LEP students saw the largest positive impact of the CSR award in each of the specifications. Importantly, the CSR award impacted LEP students' reading scores the most, perhaps the subject most fundamental to their future academic success. For example, LEP students in award schools improved against non-award students by nearly 9 percent of a standard deviation in math, but a more noteworthy increase was seen in reading, where they improved by nearly 34 percent of a standard deviation after the award was received, relative to non-award students. The effect of the CSR award on student gains in the post-award years was 53 percent of a standard deviation above gains in non-award schools – about 5 points on the TLI.

Our examination of the subgroups also highlights the distinction between baseline and subgroup students in their experience with schools receiving CSR funding. Baseline students showed no statistically significant changes in reading or math gains when their schools were awarded CSR funds or moved to post-award status (model 1). We found no association between per-pupil award and reading achievement for baseline students (model 2). We also found no linear improvement trend for baseline students in schools receiving federal CSR funds for math and a negative trend for reading (model 3). By contrast, each subgroup except special education students and FRL students saw gains in reading across all of the models. It is well known that CSR programs tend to emphasize reading instruction, especially in the elementary and middle school grades. These results suggest that the attention given to reading instruction provides the most benefit to ethnic- and language-minority students.

It is important to note that the only subgroup with negative and zero gains in reading performance across the specifications was special education students. The special education subgroup students saw statistically significant and negative effects in reading when their schools received CSR funding (model 1) and a negative association in per-pupil awards (model 2). We also measure a negative trend in reading (model 3), but this effect was not statistically significant. From these results it appears that special education students, who generally are not specifically addressed within most CSR designs, did not see the positive effect of CSR funding in their schools that was seen by other subgroups. This result raises questions about the value of these school-wide reforms for the school's most-challenged learners.

TESTING THE IMPORTANCE OF EXPOSURE

The analyses described above examine the average effect of being in a CSR school at different points in the *school's* history with the funding. These models do not take into account the length of time the *student* is exposed to the CSR, which can also determine the impact of the program. We explore the impact of a student's exposure with one additional specification of the models that indicate award and post-award status.

The impact of the length of a student's exposure to a CSR award school is estimated with spell-effects models that include an interaction between the number of years since the school received their award and the number of years a student had been in a CSR-funded school (up to six years). Although this specification allows for 36 interaction terms (some schools had received funding as much as 6 years prior to a student entering multiplied by the 6 years that students could be exposed to the CSR funded school), we found only a limited number of students who had more years of exposure to a CSR-funded school than years since their school was awarded a CSR grant. That is, it was relatively rare to find a student with four years of exposure to a CSR grant school who was in a school where only 2 years had elapsed since receiving the award.[13] For this reason (and brevity), we opt to report only the estimates for students who had an equal number or fewer years of exposure than the number of years since their school had received its award. That is, their school had the CSR award as long, or longer than they had been in the school.

In these analyses it is interesting to compare the CSR effects for students who are in the same CSR-funded school for different lengths of time, which equates to looking at a CSR school that had received funding three years earlier, and comparing the effect for students who had been there for two years to that for students who had been there for three years. A similar comparison could be made between students with two and three years of exposure in a school four years after the start of its funding.[14] Essentially these comparisons show whether a student's additional year

---

[13] For a student to have more years of exposure than their school had since the award would mean that the student transferred to the school from a school that had earlier received an award.

[14] We would have compared students with two and three years of exposure in a school with five years since receiving the award, but there were too few students in our sample with only two years of exposure in the fifth year award-schools to estimate an effect.

of exposure matters after accounting for the number of years since the school was awarded federal CSR funds.

Table 7 shows the coefficient estimates that compare students with different years of CSR funding exposure in schools with equal years of time since receiving the CSR award. In both math and reading, it appears that the longer a student is in a school with a CSR award, the larger effect the model has on the student's test performance. In math and reading, our results suggest that students with three years of exposure to a CSR-funded school saw larger CSR effects than did students with only two years of exposure. This is also true when we control for the length of time since the school received the award. For example, for students in schools that received the CSR award three years earlier, students with two years of exposure had an effect of 0.071 while the effect for students in their third year of exposure was 0.075. We see a similar pattern in the reading models: among students in schools that received the award three years earlier, students with two years of exposure to a CSR award saw an effect of 0.019 while students in their third year of exposure saw an effect of 0.028. This pattern also holds when we consider schools four years after receiving the award: students with more exposure see larger effects.

(INSERT TABLE 7 ABOUT HERE)

THE FINANCIAL IMPACT

At the beginning of this paper we noted that the federal government has distributed over \$1.8 billion nationwide for schools to implement CSR programs. What students have gained from this investment is of critical importance in evaluating its success. While our results varied across math and reading, for different types of students, and for schools at different stages of implementation, we can offer some general estimates of the cost-effectiveness of these programs in Texas. From the inception of federal CSR funding in 1998 through the 2002-03 school year (the last year for which we have data), Texas schools serving 3rd through 8th grade students received \$61.7 million in CSR awards. During this time, there were 631,771 'student years' in Texas CSR-funded schools. Therefore, approximately \$98 was spent on each student year.

Looking at **Table 5**, model 4, the average annual effect in math was approximately 11 percent of a standard deviation (treating the non-significant effects as zero) while the average annual effect on gains in reading was 5.2 percent of a standard deviation (treating non-significant effects as zero). These effects are equivalent to one standard deviation in math costing \$890 per student and one standard deviation in reading costing $1884 per student. Compared to cost-effectiveness calculations for other reform programs, these numbers seem quite reasonable. Goldhaber (forthcoming) calculated that one standard deviation gain from students' having a teacher that is certified by the National Board for Professional Teaching Standards costs $7,300 and Levin and McEwan (2001) estimated that a similar gain from class size reduction costs approximately $1,900. It is important to remember that these figures are only estimates and only account for costs to contract and establish a CSR program in a school. In reality, the costs of CSR extend beyond the cost of the contract and well into the future when one considers that these programs often require the staffing of a full-time instructional facilitator, smaller class sizes, tutoring programs, or an annual re-supply of instructional and assessment materials. Therefore, these estimates should be considered conservative. Of course the findings in **Table 6** suggest that CSR funding has differential impacts on different types of students, so the gross estimates of cost effectiveness presented in **Table 7** mask the fact that there seem to be distributional consequences associated with the CSR funding. This in turn suggests that the program is considerably more cost effective when it comes to educating minority and LEP students and less cost effective when it comes to impacting the achievement of white students.

CONCLUSIONS

The Office of Management and Budget released profiles of major federal programs and concluded that, "The Department has not determined a direct relationship between funding levels (of CSR) and (the program's) performance goals" (2002). To date, the available research cannot even answer the most basic policy questions: Does the program produce positive effects? Does it impact the targeted population? Do the results justify the cost? This paper attempts to address these questions. Taking advantage of a powerful dataset that contains longitudinal student-level data, we find that CSR funding has impacted students' performance and has had effects on targeted student subgroups. Our research also sheds some light on the magnitude of the policy's impact on student performance. Some of the largest impacts (as high as 2 points on the TLI) were seen by LEP students in reading, which is presumably a critical improvement for these English language learners. Moreover, our estimates suggest that CSR offers similar, if not better, cost effectiveness to other well-recognized reform programs.

Our study also shows the importance of methodology in determining the effects of the CSR policy. As we illustrate with the three levels of analysis, researchers can come to very different conclusions based on the aggregation of the data and the efforts to account for school and student effects.

It is very important to note, however, that this research reflects the average effect of the funding policy for schools that receive CSR awards, and does not account for the extent to which the schools implement the CSR designs they consequently adopt. Given that implementation varies widely across schools and has been found to have important implications on the impact of the policy (Desimone, 2002), our effects estimates reflect the average effect on students' performance across schools of different degrees of implementation. While it is possible that higher-implementing schools might actually see more substantial effects, it is also possible that the effects might be very different for CSR schools not receiving federal CSR funding. Although the positive effects for most students show the promise of this reform, our research raises concerns about the extent to which this policy addresses the needs of special education students, a subgroup that is not explicitly targeted in the original CSRD program, but which is a targeted population in NCLB. While additional research can only improve our understanding of current CSR policy, this study provides policy analysts at the Department of Education valuable insights into what has been accomplished thus far and the types of students on which CSR funding has had the greatest impact.

References

Bifulco, R. (2002). Addressing self-selection bias in quasi-experimental evaluations of whole-school reform: A comparison of methods. *Evaluation Review* 26(5), 545-72.

Bifulco, R., W. Duncombe, and J. Yinger. (2005). Does whole-school reform boost student performance? The case of New York City. Journal of Policy Analysis and Management 24, 47-72.

Bloom, H.S., S. Ham, L. Melton, and J. O'Brien. (2001). Evaluating the Accelerated Schools approach: A look at early implementation and impacts on student achievement in eight elementary schools. New York, NY: Manpower Demonstration Research Corporation, 127.

Borman, G. D., G. M. Hewes, L. T. Overman, and S. Borwn. (2003). Comprehensive school reform and achievement: A meta-analysis. Review of Educational Research, 73, 125-230.

Clotfelter, C. T., H. F. Ladd and J. L. Vigdor. (Forthcoming). Teacher-student matching and the assessment of teacher effectiveness. NBER Working Paper #11936. Journal of Human Resources.

Cooper, R., and W. J. Jordan. (2003). Cultural issues in comprehensive school reform. Urban Education, 38, 380-397.

Datnow, A. (2000). Implementing an externally developed school restructuring design: enablers, constraints, and tensions. Teaching and Change, 7(2), 147-71.

Department of Education, U.S. (1998-2005). Budget for fiscal years 1998-2005, edited by Office of Elementary and Secondary Education. Washington, DC: U.S. Department of Education.

Desimone, L. (2002). How can comprehensive school reform models be successfully implemented? Review of Educational Research, 72(3), 433-79.

Engec, N. (2006). Relationship between mobility and student performance and behavior. Journal of Educational Research, 99(3), 168.

Goldhaber, D., and E. Anthony. (forthcoming). Can teacher quality be effectively assessed? national board certification as a signal of effective teaching. Review of Economics and Statistics.

Hammann, E. T., I. Zuliani, and M. Hudak. (2001). English language learners, the comprehensive school reform demonstration project, and the role of state departments of education. Providence, RI: The Education Alliance at Brown University.

Hartman, C. (2006). Students on the move. Educational Leadership, 63, 20-24.

Koh, Myung-Sook, and Janna Siegel Robertson (2003), "School Reform Models and Special Education." *Education and Urban Society* 35, pp. 421-442.

Levin, H. M., and P.J. McEwan. (2001). Cost-Effectiveness Data. 2nd Ed. Thousand Oaks: Sage Publications.

Office of Management and Budget. (2002). Detailed information on the comprehensive school reform assessment. (available online at *ExpectMore.gov* http://www.whitehouse.gov/omb/expectmore/detail.10000184.2005.html).

Rivkin, S. (2005). Cumulative nature of learning and specification bias in education research. Mimeo, Amherst College.

Rumberger, R. W. (2003). The causes and consequences of student mobility. The Journal of Negro Education, 72(1), 6-21.

SEDL (Southwest Educational Development Lab). (2003). Follow-up study of schools implementing comprehensive school reform in the southwest. Evaluation Report. Austin,

TX: Southwest Educational Development Lab, pp 101. Available online at http://www.sedl.org/pubs/change93/csr-followup-report.pdf.
Slavin, R., and N. A. Madden. (2001). One million children: Success for All. Thousand Oaks, CA: Corwin.
Supovitz, J. A., and H. May. (2004). A study of the links between implementation and effectiveness of the America's Choice comprehensive school reform design. Journal of Education for Students Placed at Risk, 9(4), 389-419.
Todd, P.E., and K. I. Wolpin. (2003). On the specification and estimation of the production function for cognitive achievement. The Economic Journal, 113, F3-F33.
Tushnet, N. C., J. Flaherty, and A. Smith (2004). Longitudinal assessment of comprehensive school reform program implementation and outcomes: First-year report. Washington, DC: U.S. Department of Education.
Vernez, G., R. Karam, L. Mariano, and C. DeMartini. (2004). Assessing the implementation of comprehensive school reform models. Santa Monica, CA: RAND Education

Table 1. CSR awards for Texas elementary schools – 1998-99 to 2003-04

School year	Number of schools with new awards	Date Awarded
1998-99	64	3/99
1999-00	0	--
2000-01	0	--
2001-02	117	7/01
2002-03	0	--
2003-04	0	--

Table 2. Summary Statistics for Student and School Characteristics

	All students and Schools in 2002		All students and schools in award status in 2002	
Student Characteristics	**Mean *n≈240,000***	***(Std. Dev.)***	**Mean n≈*30,000***	***(Std. Dev.)***
Standardized math raw score	-0.0834389	*(1.012629)*	-.0910416	*(1.005219)*
Standardized reading raw score	-0.0870476	*(1.021978)*	-.1519462	*(1.030687)*
Change in standardized math raw score	-0.0094735	*(0.7103914)*	.0079223	*(0.7223418)*
Change in standardized reading raw score	0.0028511	*(0.7527784)*	.0193601	*(0.7729936)*
Percent of students in award schools	0.1262979	*(0.3321855)*	-	-
Percent of students in post award schools	0.151771	*(0.3587996)*	-	-
Average per pupil CSR award	21.36621	*(62.64205)*	169.1725	*(77.87684)*
Percent female students	0.5096734	*(0.4999074)*	.5125378	*(0.4998509)*
Percent black students	0.1638243	*(0.3701168)*	.1284747	*(0.3346231)*
Percent Hispanic students	0.4893154	*(0.4998869)*	.6162359	*(0.4863095)*
Percent LEP students	0.0833097	*(0.2763503)*	.1154644	*(0.3195869)*
Percent disabled students	0.5832151	*(0.4930276)*	.6823897	*(0.4655546)*
Percent special education students	0.0394927	*(0.1947643)*	.0331964	*(0.179152)*
Percent of students in charter schools	0.0058364	*(0.0761732)*	.003935	*(0.0626067)*
School Characteristics	**Mean N≈5,360**	***Std. Dev***	**Mean N=135**	***Std. Dev.***
School wide mean raw math score	-0.0278851	*(0.4705576)*	-0.2166719	*(0.4313355)*
School wide mean raw reading score	-0.0223578	*(0.4543849)*	-0.2627846	*(0.391409)*
Percent of Schools in Award Status	0.025149	*(0.1565922)*	-	-
Percent of Schools in Post Award Status	0.0203055	*(0.1410564)*	-	-
Per Pupil Award	-		215.3649	*(113.5426)*
Percent Black	0.1404623	*(0.2015179)*	0.1605852	*(0.230606)*
Percent Hispanic	0.4041165	*(0.3212807)*	0.5935556	*(0.3268117)*
Percent LEP	0.1490598	*(0.1872605)*	0.2606074	*(0.2215241)*

Table 3. Baseline models of the effects of CSR funding on changes in student achievement

	Model 1		Model 2		Model 3		Model 4	
	Math	Reading	Math	Reading	Math	Reading	Math	Reading
Intercept	0.3439**	0.2842**	0.3442	0.2841**	0.0344**	0.2842**	0.3444**	0.2839**
Robust standard error	*0.0350*	*0.0263*	*0.0350*	*0.0263*	*0.0350*	*0.0263*	*0.0351*	*0.0263*
Award	-0.0197**	0.0026						
	(0.0082)	*(0.0057)*						
Post-award [1]	-0.0340**	0.0018	-0.0072**	0.0005				
	(0.0107)	*(0.0075)*	*(0.0024)*	*(0.0017)*				
Log per-pupil award[2]			-0.0035**	0.0009				
			(0.0014)	*(0.0012)*				
Years since award					-0.0073**	0.0007		
					(0.0022)	*(0.0015)*		
1 year since award							-0.0075	0.0003
							(0.0100)	*(0.0080)*
2 years since award							-0.0363**	0.0038
							(0.0124)	*(0.0085)*
3 years since award							-0.0162	0.0043
							(0.0135)	*(0.0090)*
4 years since award							-0.0299**	0.0156
							(0.0146)	*(0.0116)*
5 years since award							-0.0487**	-0.0327**
							(0.0191)	*(0.0127)*
6 years since award							-0.0229	0.0221*
							(0.0148)	*(0.0114)*
N	1,639,550	1,639,555	1,639,550	1,639,555	1,639,550	1,639,555	1,639,550	1,639,555
Adjusted R^2	0. 0111	0.5628	0. 0111	0.0084	0. 0111	0.0084	0.0111	0.0084

[1]Post-award is a dummy variable identifying schools that received an award but are now beyond the three-year award period. For models of log per-pupil award, post-award is defined as the natural log of the highest annual per-pupil award when the school was in the three-year award period.
[2]Log per-pupil award is defined to be zero for schools not receiving an award.
**Indicates significance at a 95 percent level of confidence.

Table 4. The effect of CSR funding on changes in student achievement with school fixed-effects

	Model 1		Model 2		Model 3		Model 4	
	Math	Reading	Math	Reading	Math	Reading	Math	Reading
Intercept	0.4654**	0.1473**	0.4652	0.1471**	0.4663**	0.1479**	0.4660**	0.1469**
Robust standard error	*0.0912*	*0.0708*	*0.0912*	*0.0708*	*0.0911*	*0.0708*	*0.0912*	*0.0708*
Award	-0.0150*	-0.0048*						
	(0.0079)	*(0.0071)*						
Post-award[1]	-0.0241**	-0.0079*	-0.0055**	0.0019				
	(0.0114)	*(0.0100)*	*(0.0025)*	*(0.0021)*				
Log per-pupil award[2]			-0.0029*	0.0008				
			(0.0016)	*(0.0014)*				
Years since award					-0.0044**	-0.0013		
					(0.0022)	*(0.0019)*		
1 year since award							-0.0047	-0.0076*
							(0.0985)	*(0.0082)*
2 years since award							-0.0311**	-0.0031
							(0.0124)	*(0.0101)*
3 years since award							-0.0098	-0.0032
							(0.0127)	*(0.0104)*
4 years since award							-0.0226	0.0073
							(0.0153)	*(0.0148)*
5 years since award							-0.0381**	-0.0411**
							(0.0182)	*(0.0130)*
6 years since award							-0.0107	0.0097
							(0.0172)	*(0.0128)*
N	1,639,550	1,639,555	1,639,550	1,639,555	1,639,550	1,639,555	1,639,550	1,639,555
R^2	0.0322	0.0208	0.0322	0.0208	0.0322	0.0208	0.0322	0.0208

[1]Post-award is a dummy variable identifying schools that received an award but are now beyond the three-year award period. For models of log per-pupil award, post-award is defined as the natural log of the highest annual per-pupil award when the school was in the three-year award period.
[2]Log per-pupil award is defined to be zero for schools not receiving an award.
**Indicates significance at a 95 percent level of confidence.

Table 5. Spell-effects models of CSR funding on student Math and Reading performance

	Model 1 Award status		Model 2 Per-pupil award amount		Model 3 Linear trend effect of award		Model 4 Year-by-year effect of award	
	Math	Reading	Math	Reading	Math	Reading	Math	Reading
Award	0.0226	0.0355*						
Robust standard error	*(0.0288)*	*(0.0258)*						
Post-award[1]	0.1062**	0.0732*	0.0212*	0.0143				
	(0.0564)	*(0.0154)*	*(0.0125)*	*(0.0097)*				
Log per-pupil award[2]			0.0043	0.0072				
			(0.0059)	*(0.0052)*				
Years since award					0.0433**	0.0296**		
					(0.0207)	*(0.0138)*		
1 year since award							0.0219	0.0309
							(0.0277)	*(0.0244)*
2 years since award							0.0256	0.0574
							(0.0447)	*(0.0381)*
3 years since award							0.0991	0.0916*
							(0.0675)	*(0.0493)*
4 years since award							0.1597**	0.1287**
							(0.0818)	*(0.0629)*
5 years since award							0.2151**	0.0975
							(0.0984)	*(0.0696)*
6 years since award							0.2875**	0.1865**
							(0.1316)	*(0.0862)*
N	1,736,151	1,721,195	1,736,151	1,721,195	1,736,151	1,721,195	1,736,151	1,721,195
R^2	0.4911	0.4823	0.4911	0.4823	0.4911	0.4823	0.4911	0.4823

[1]Post-award is a dummy variable identifying schools that received an award but are now beyond the three-year award period.
[2]Log per-pupil award is defined to be zero for schools not receiving an award.
*Indicates significance at a 95 percent level of confidence.

Table 6. The interaction of CSR funding with student subgroups

	Model 1 Award Status				Model 2 Per-pupil award amount[1]				Model 3 Linear trend effect	
	Math		Reading		Math		Reading		Math	Reading
	Award	Post-award	Award	Post-award	Award	Post-Award	Award	Post-Award		
Baseline group	-0.0030	-0.0112	-0.0194	-0.0817	-0.0012	-0.0053	-0.0039	-0.0186	-0.0069	-0.0230
Robust standard error	*(0.0365)*	*(0.0815)*	*(0.0303)*	*(0.0565)*	*(0.0076)*	*(0.0168)*	(0.0061)	(0.0115)	*(0.0222)*	*(0.0159)*
Interaction with FRL	0.0127	0.0295	0.0053	0.0339**	0.0034	0.0065	0.0017	0.0663	0.0080	0.0107*
	(0.0243)	*(0.2846)*	*(0.0200)*	*(0.0322)*	*(0.0050)*	*(0.0063)*	(0.0043)	(0.0075)	*(0.0060)*	*(0.0059)*
Interaction with African American	0.0327	0.1107**	0.1105**	0.2719**	0.0063	0.0255	0.0224**	0.0614**	0.0614*	0.0747**
	(0.0585)	*(0.1014)*	*(0.0526)*	*(0.0934)*	*(0.0127)*	*(0.0220)*	(0.0114)	(0.0202)	*(0.0357)*	*(0.0270)*
Interaction with Hispanic	0.0149	0.1345*	0.0267	0.1325**	0.0030	0.0311*	0.0050	0.0297**	0.0600**	0.0607**
	(0.0457)	*(0.0754)*	*(0.0333)*	*(0.0650)*	*(0.0096)*	*(0.0166)*	(0.0069)	(0.0140)	*(0.0253)*	*(0.0167)*
Interaction with Special Education	0.0022	0.0520	-0.0971*	-0.1303	0.0014	0.0126	-0.0206*	-0.0271	0.0076	-0.0303
	(0.0532)	*(0.0806)*	*(0.0556)*	*(0.0917)*	*(0.0013)*	*(0.0183)*	(0.0120)	(0.0204)	*(0.0213)*	*(0.0210)*
Interaction with LEP	0.0872**	0.0602	0.3441**	0.5321**	0.0193**	0.0148	0.0759**	0.1193**	0.0392**	0.1490**
	(0.0344)	*(0.0684)*	*(0.0584)*	*(0.1091)*	*(0.0078)*	*(0.0153)*	*(0.0127)*	*(0.0249)*	*(0.0139)*	*(0.0268)*
N	1,653,799		1,639,697		1,774,253		1,760,774		1,653,799	1,639,697
Adjusted R^2	0.5008		0.4929		0.5008		0.4929		0.5008	0.4929

[1]Log per-pupil award is defined to be zero for schools not receiving an award.
*Indicates significance at a 90 percent level of confidence.
**Indicates significance at a 95 percent level of confidence.

Table 7. The effect of students' exposure to federally funded CSR schools

	Math		Reading	
	I Students with two years of exposure	II Students with three years of exposure	III Students with two years of exposure	IV Students with three years of exposure
School in third year after receiving the award	0.0707** *(0.0235)*	0.0755** *(0.0240)*	0.0193** *(0.0123)*	0.0212** *(0.0212)*
School in fourth year after receiving the award	0.0553** *(0.0213)*	0.0761** *(0.0231)*	0.0283** *(0.0166)*	0.0501** *(0.0520)*

**Indicates significance at a 95 percent level of confidence.