# Seeing Success: The Impact of Implementing Model Practices on Student Outcomes

**Betheny Gross**
**Lydia Rainey**
**Dan Goldhaber**

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

**April 21, 2006**

*CRPE working paper # 2006-2*

**Abstract**

This paper evaluates the effect of implementing one of three prominent CSR models on student achievement and discipline outcomes, using a matched sample of Florida model and non-model schools. Longitudinal fixed-effects models to analyze school-level reading and math scores showed inconsistent effects of CSR participation and the use of practices endorsed by models on student achievement and discipline outcomes. Although we find that longer participation with the CSR designs increased math achievement, we do not find that increasing the use of the model practices increases student achievement or has consistent effects on student discipline. Overall, this analysis offers only weak support to the assumption that CSR is a generalizable policy approach for improving low-performing schools.

**Introduction**

In 1998, the U.S. Department of Education responded to growing concern about the nation's long history of piecemeal and seemingly ineffective school reform efforts by creating the Comprehensive School Reform Demonstration (CSRD) Project– also referred to as the Obey-Porter Act.[1] For the decades leading up to the CSRD Project, which borrows its name from the national movement for Comprehensive School Reform (CSR), school reform efforts seemed to be full of many ideas such as shared planning time, block scheduling, new textbooks, curriculum programs, testing programs, and targeted assistance programs, but with very little effort to coordinate these reform ideas into a comprehensive plan for schools. Not surprisingly, schools found themselves with strategies that failed to address root issues and, at times, even conflicted with each other. Educational reformers, who were frustrated with the lack of improvement from all of these "good ideas," started to pull compatible reform ideas together into more coherent programs that would address deficiencies across the whole school. The rationale was that schools (in particular, low-performing schools) needed a more comprehensive strategy that would provide guidance on the best way to organize staff and students, coordinate curriculum, and deliver instruction.

Since the beginning of the CSRD Project in 1998, the Department of Education has distributed approximately \$1.8 billion for the implementation of CSR models (Department of Education, 1998-2005), and, with the U.S. Congress, reaffirmed its support of the CSR movement by including funding for CSR programs in the 2001 No Child Left Behind Act.[2] The CSRD Project provides schools with funding to adopt research-based reform programs that offer a comprehensive set of strategies for addressing several elements of the school, including the organization of teachers and students, internal governance, curriculum, and instructional practices. These reform programs, though they vary in content and were developed by many different designers, are collectively referred to as Comprehensive School Reform designs.

---

[1] Leading up to the passage of CSRD, these designs were largely promoted by the New American Schools organization, an independent reform organization focused on research-based practice
[2] Elementary and Secondary Education Act of 2001, Title I Part F

The various CSR designs – *Success for All, Core Knowledge, Direct Instruction, Accelerated Schools, America's Choice,* and the *Comer Model* are some of the most popular – promote divergent strategies for improving schools' climates and increasing student learning. As part of its investment in CSR, Congress mandated that the Department of Education fund research studies of the impact of CSR designs on student outcomes. In this paper, we present findings from one of the six large longitudinal studies that were commissioned to satisfy this mandate.[3] We estimate the impact of CSRD funding and participation in three prominent school reform designs – *Core Knowledge* (CK), *Direct Instruction* (DI), and *Success for All* (SFA) – on student outcomes, including students' test score achievement, absences, suspensions, and in-school violence, in a sample of schools from Florida. In our work, we make explicit efforts to link the use of practices endorsed by these CSR designs to academic and behavioral outcomes of students and employ analytic techniques that account for many of the technical challenges raised in the evaluation of CSR. In the end, we find that CSR implementation and practices have only a weak association with student outcomes at the school level.

**The Challenges of Evaluating CSR Models and Their Impacts**

Confidence in the promise of CSR designs extends beyond the federal level, as many states and districts encourage the adoption of one or more models. For example, the 1997 Abbott v. Burke (school funding) decision in New Jersey not only requires schools in several of the state's lowest-performing districts to adopt a CSR design, but further mandates state and district funding for these programs.[4]  A number of large districts across the country, including Galveston, Texas, and Houston, Texas, have included *Success for All* in district-wide initiatives,[5] and Polk County, Florida, and Baltimore, Maryland, included *Core Knowledge* in their district reform programs.[6] To date, over 5,000 public schools in the United States have obtained federal CSRD funds to

---

[3] For an overview of the six research projects refer to http://www.ed.gov/offices/OERI/csrresearch.html.
[4] 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.
[5] Houston introduced *Success for All* into its Title I schools, which fed into the high school initiative, *Project Grad.*
[6] The Baltimore Curriculum Project involved joint support district-wide for *Core Knowledge* and *Direct Instruction.*

implement a CSR model—and countless additional schools have used other funding sources for their own model implementation.[7]

Despite the large community of CSR design providers and this enormous commitment of resources and support for CSR, consistent evidence of the impact of CSR on student outcomes remains elusive largely due to the conceptual and analytic challenges of studying these reforms. Researchers have made considerable efforts to evaluate CSR, but this field of inquiry has historically posed some very difficult conceptual and analytic challenges that have made much of the existing empirical work vulnerable to criticism. Some of the most difficult are (1) the conceptual challenge of identifying and evaluating a movement and policy that is comprised of so many different reform approaches, (2) the logistical challenge of giving the reform a sufficient length of time to reveal its effects, (3) the technical and conceptual difficulty of evaluating a reform that relies so heavily on the quality of implementation for its success and (4) the technical challenge of eliminating bias when schools self-select their participation. These issues, along with the typical educational research challenges of appropriately defining student outcomes, obtaining reliable data on student outcomes, and context variability, have left the reform community with a very muddy picture of CSR effects. Nonetheless, this previous work has provided the current cohort of CSR researchers with valuable insights into the critical conceptual issues that must be addressed in our research, as well as the analytic improvements that are needed to clarify our understanding of the impacts of this reform.

Defining the intervention

A fundamental challenge facing this research is posed by variation in reform approach across designs. The CSR movement is represented by a wide assortment of designs that are each unique in the combination of strategies and practices they endorse, the training they provide, and the follow-up strategies they use to monitor ongoing progress. Schools are encouraged to select one (or more) design that best suits their school's needs. But this variation poses challenges to researchers' efforts to evaluate the CSR movement and the chance that all designs will be subject to high-quality evaluation.

---

[7] The number of schools receiving federal CSRD funds is reported by SEDL in their CSR Awards Database, which is available online at http://www.sedl.org/csr/awards.html.

With CSRD-funded schools using nearly 700 externally developed CSR models, a design-by-design approach to evaluating CSR can only focus on a limited number of programs, and thus, the reform community is left to rely on the developers' own evaluations as evidence of success. Although evaluations of individual models are important to model developers and those interested in adopting the model, this approach implies that many models have gone, and will continue to go, without rigorous evaluation.

Efforts to examine a group of reform designs collectively—such as Berends, Bodilly, and Kirby's (2002) study of the New American Schools models—potentially broaden the scope of the available literature and may help to assess the movement as a whole. However, multi-design studies often must still select only a handful of programs to facilitate data collection, and cannot differentiate the weak from the strong programs when the designs are evaluated simultaneously. It is clear that this conundrum will not be resolved in a single study and will therefore not be resolved in this work. Researchers must decide which approach they will take – multi-design versus individual design – and be clear about the limitations of the approach they have chosen.

Timing is Crucial

Among the technical challenges posed by this research, the issue of timing is one of the most basic issues in evaluating reform. One of the most consistent comments in CSR, or any reform for that matter, is that it takes time for effects to be fully realized. Model developers typically argue that the implementation process alone can take three to five years (Tushnet, Flaherty, & Smith, 2004). Timing has been a particular issue in evaluations of federal funding for CSR which, until recently, have been constrained to time periods that are too short to make firm conclusions about the policy's impacts. For example, an evaluation of federal funding for CSR—which failed to find positive effects of the policy—looked for impacts only three years after the first distribution of funds (Tushnet et al., 2004). The authors warn that finding effects in this short period of time would be unreasonable.

Timing, however, is no longer an issue for researchers. By 2005, the CSRD policy had been in effect for six years, while many of the CSR movement's leading programs had been in schools

for at least ten years, and in some cases much longer.[8] Our study design takes advantage of both the extended time for which the policy has been in place and the longevity of reform models by focusing on mature models in Florida, a state with a long history of outcomes data on its schools.[9]

Implementation matters

Researchers and CSR observers have repeatedly raised the importance of understanding implementation in evaluations of CSR research. The fact that these designs have been implemented across the full spectrum of school contexts introduces wide variation in the quality of implementation across schools. Boosted by the funding from the federal CSRD policy, and as a result of the active promotion of CSR through organizations like the New American Schools and the Clearinghouse on Comprehensive School Reform (now known as the Center for Comprehensive School Reform and Improvement), these designs have been introduced across schools varying in geographic regions, prior performance levels, and student demographics. Consequently, it is now common for schools to receive implementation support from design consultants who work remotely from their organization's home base and who have only restricted access to the design's core developers. Furthermore, the wide array of schools to which these CSR designs are applied have unique circumstances that must be addressed, which often leads to alterations in the design. Not surprisingly, schools do not implement models with the same measure of fidelity (Vernez, Karam, Mariano, & DeMartini, 2004). The quality of training, commitment to the model by teachers and administrators, resources available, and negotiated modifications to the model play an enormous role in determining the models' effects (Bacevich, Le Floch, Stapleton, & Burris, 2005; Cooper, 1998; Datnow & Stringfield, 2000; Desimone, 2002; Kurki, Aladjem, & Carter, 2005; Murphy & Datnow, 2003; Rowan, Harrison, & Hayes, 2004; Supovitz & May, 2004)

---

[8] We use the age of the most prominent models according to the *Catalog of School Reform* (published by the Northwest Regional Education Laboratory), which provides snapshot descriptions of 21 prominent CSR models.
[9] All three models in this study have a lengthy history in school reform. *Core Knowledge* has been available as a curriculum model since 1986. *Direct Instruction*, as an instructional model, has been in existence since 1968 with later revisions in order to make the model a CSR. *Success for All* has been in schools since 1993 (Northwest Regional Education Laboratory, Catalog of School Reform. Available online at http://www.nwrel.org/scpd/catalog/index.shtml)

While the focus on implementation is not new to reform research, CSR, with its emphasis on research-based practices, adds yet another dimension to this focus on what schools are actually doing in reform. While we argued earlier that the variation in the specifics of designs has been an important part of the movement but a challenge to evaluation, it can also be argued that the actual variation between models is limited, and this too has hampered our ability to effectively evaluate CSR designs. The reality is that most CSR designs are actually a variation on the theme of well-known instructional and organizational strategies. For example, a structured tutoring program for students falling behind is a core feature of the *Success for All* design; across-subject and across-grade curriculum coordination is key to the *Core Knowledge* design; and standardized instructional practice is central to the *Direct Instruction* model. Each of these strategies is widely discussed, present to some extent in many other CSR designs, and widely used in schools that are not formally participating in any CSR design. As such, parsing out the specific effects of CSR designs presents a significant challenge when studying their effects on student outcomes unless we know the instructional and organizational practices being used in both CSR implementing schools and their non-implementing counterparts.

Measuring school practice, unfortunately, is technically challenging and expensive. Some of the more sophisticated approaches to measuring practice have involved teacher logs (Rowan et al., 2004). Surveys of instructional and organizational practice in the context of CSR research involve identifying key components of the reform, capturing these components with valid questions, and, finally, administering the survey in a manner that yields reliable responses (Vernez et al., 2004).

Examples of different approaches to capturing practice and implementation can be seen in previous research. The recent Study of Instructional Improvement (SII) conducted by Rowan, Harrison, and Hayes (2004) represents the most intensive effort to capture implementation and incorporate this information into an evaluation of school reform designs. Through logs kept by more than 500 teachers, the authors showed that literacy curriculum content varied widely and as a function of the CSR design used in the teachers' school. Bifulco, Duncombe, and Yinger (2005) measure CSR implementation using model provider reports of their schools' implementation in an evaluation of whole school reform in New York City. RAND's study of

New American Schools models used "audits" by research teams to estimate each school's level of implementation (Bodilly, 1998). AIR's study of CSR employed more traditional surveys to capture implementation and a describe implementation as highly variable (Kurki et al., 2005; Zhang, Shkolnik, & Fashola, 2005). Given the great importance of practice and the inherent difficulty in measuring it, our study devoted considerable resources to developing measures of practice in both CSR schools and control schools.

Controlling for bias

Perhaps the most significant challenge in evaluating the impacts of CSR is dealing with the sampling bias. CSR as a reform strategy and as a component of federal education policy intended to address low-performing schools serving low-income students is an attractive program to certain types of schools, specifically schools that have struggled for years to improve student performance and those serving high concentrations of low-income and often high concentrations of minority students. Not surprisingly, these types of schools are more likely to seek funding for, and adopt, a CSR design. In addition to the bias generated by this "self-selection," districts and states are also targeting low-performing schools and requiring schools to adopt CSR designs or providing schools with resources to pursue designs.

Sampling bias is, therefore, a significant concern for the evaluation of CSR designs. The most rigorous strategy for eliminating this bias is through the use of randomized trials. One such study of *Success for All* is underway, with first-year results showing only modest effects (Borman et al., 2005) but second-year results showing that SFA schools are improving at a faster rate than control schools (Borman et al., in press). However, randomized trials require significant funding and rely on participating schools' agreement to let their reform strategies be determined by the experimental format – something many principals, school boards, and parents are unwilling to do.

Because randomization is a limited option, researchers must be able to correct for bias in non-randomized data. An examination of the implementation of three different CSR designs in New York City by Bifulco, Duncombe, and Yinger (2005) judiciously attempts to address this issue of bias. In this study the authors compare multiple analytic approaches, including a matched

sample to improve comparisons, strategies to control for all fixed characteristics of schools, an instrumental variables analysis that accounts for systematic bias in a school's search for CSR models, and strategies that account for bias from student mobility. The authors provide a comparative discussion of these methodological strategies. Although they find that none of the three whole-school reforms show consistent results across student- and school-level analyses, and that only one of the three examined reforms show positive and statistically significant effects on student achievement in the student-level models (Bifulco et al., 2005), this work provides guidance to the research community on the direction that should be taken in future research to control for bias.

**Data and Analytic Approach**

We attempt to address the analytical challenges described above in a variety of ways. First, , this study was designed from the outset to cover a longitudinal range over which we could reasonably expect to see effects if they occurred. Our evaluation of the CSRD policy looks across the five years following the first allocation of funds for CSR implementation,  while our examination of specific models extends to five years and includes schools that have been implementing models from between one and twelve years. We account for instructional and organizational practice variation by surveying teachers and principals on the implementation of CSR model practices in implementing schools as well as non-implementing schools. Measuring practice across all schools in the study allows us to investigate the impact of practices favored by the designs, which gives us an indication of whether the reform developers have espoused effective practices.

We make efforts to reduce bias in our sample and estimates. Like Bifulco, Duncombe, and Yinger (2005), we selected schools for the study using matching techniques that improve the balance of schools in the sample. We further reduce bias with modeling techniques that control for all the observed and unobserved characteristics of schools that do not change over time but which may influence a school's decision to adopt a model (Allison, 2005).

Our analysis is laid out in two stages– an evaluation of the federal CSRD policy and its impacts on schools receiving funding, and a joint evaluation of the impact of three important CSR

designs on schools that have purchased one of these CSR designs. In this analysis we pursue three questions:

1.  What is the effect of receiving federal CSRD funds on school-wide academic outcomes (4th grade reading and 5th grade math) and student behavioral outcomes (absenteeism, in-school suspension, and out-of-school suspension) in Florida elementary schools receiving CSRD funds?

2.  What is the average impact of participation in one of three prominent CSR designs – *Direct Instruction*, *Core Knowledge*, or *Success for All* – on academic outcomes and student behavioral in a sample of Florida elementary schools?

3.  What is the impact of increasing the use of practices endorsed by the CSR providers on student academic and behavioral outcomes?

Florida provides a valuable site for study because many of its schools have adopted CSR models and because the state has consistent outcomes data spanning several years.[10] The data we employ includes school-wide average scores on the 4th grade reading and 5th grade math sections of the Florida Comprehensive Achievement Test (FCAT), the percent of students recording more that 20 absences, in-school suspensions, or out-of-school suspensions, and the number of violent acts recorded in the school obtained from the Florida State Department of Education; school-wide averages for student background, teacher characteristics, and organizational characteristics provided by the Florida State Department of Education; and data on CSRD Project funding obtained from the Southwest Educational Development Laboratory (SEDL).

In addition, we conducted a survey of teachers and principals from which we obtained information on schools' instructional and organizational practice. In consultation with the three CSR design developers, we designed a survey that would capture the extent to which teachers in the school are engaged in the practices endorsed by the CSR designers. In this survey, we take the unique step of crafting many of the questions regarding the core components endorsed by the three CSR designers in general language that would be relevant for our control schools as well.

---

[10] The full study, which included schools in both Florida and Texas, examined four different comprehensive reform designs. However, we could not gather an adequate sample of Florida schools using the *Accelerated Schools* model, leaving only three different models to be examined in this analysis.

In so doing, our survey captures the extent to which our control schools are engaged in the same organizational and instructional practices that are expected in CSR design schools.

Although we approach each of the three primary research questions with slightly different analytic models, each of the models (which are explained in more detail below) resembles a school fixed-effects model. The key feature of a school fixed-effects approach is that all stable (time invariant) school characteristics, including unmeasurable characteristics, are accounted for in the model.[11]  These models allow researchers to focus on within-school change over time, thus enabling them to see the impact of receiving funding from CSRD, adopting a model, moving into different stages of adoption, or deepening implementation while controlling for all else. The specific approaches and data that we used for each section of analysis are detailed below.

Analyzing the effect of receiving federal CSRD funds

Early efforts by the Department of Education to analyze the effects of receiving CSRD funds on student achievement found no measurable effects, but concluded that further research using a longer time horizon – more than three years – would be necessary to make a more definitive conclusion about the impact of the CSRD policy on student achievement (Tushnet et al., 2004). The analysis presented in our paper extends the time horizon to six years and includes data from 1998-1999 (the first school year in which CSR funds were given) to 2003-2004.

We generated our sample by matching schools that received funding in the first year CSRD funds were distributed ('cohort-one schools') to schools that never received funding. The match, known as a propensity score match, was based on the school's probability of receiving CSRD funds and controlled for school background characteristics and generated a sample of 79 treatment and control schools (158 total schools).[12] For each school, we obtained Florida state

---

[11] In the context of longitudinal models, within-school effects are those effects on outcomes that derive from variation in schools over time (e.g., the effect of increasing poverty rates on school-wide outcomes). Across-school effects are effects that derive from variation across different schools (e.g., the effect of being a high-poverty school.) Importantly, these models estimate the within-school effects while controlling for *all* of the stable characteristics of schools – including unmeasurable characteristics. Controlling for these characteristics is important because long-term, stable features of schools such as organizational capacity, the nature of the student population, or the nature of the school's district, often drive design CSR adoption decisions and can be difficult to measure.

[12] We employed a "greedy match" algorithm using SAS code written by Bergstralh and Kosanke (1995).  The "greedy match" with propensity scores ranks all treatment and control schools according to their propensity score. The matching algorithm then matches the treatment school to the control school with the closest propensity scores.

assessment data, school background data, and CSR funding information from 1999 to 2004. The mean values and test statistics for selected background characteristics, given in **Table 1**, show that the matched sample of unfunded schools mirrors the characteristics of funded schools with no significant differences in means detected for any of the school characteristics.

**(Table 1 about here)**

Our goal for this analysis was to determine whether schools receiving CSRD funding posted greater gains after receiving the awards than did their unfunded counterparts. It is important to note here that we do not attempt to capture the extent to which schools actually *implemented* the designs for which they used these CSRD funds. This analysis provides only an average effect of the CRSD policy of funding schools to use CSR designs. We focus our attention on an analytic model that captures gains in achievement over the five-year span between the school's first year of funding and the end of the 2003-2004 school year, and allows us to see the year-by-year differences in achievement. The model is defined as follows:[13]

$$y_{st} = \mu_t + \gamma_1 X_{st} + \gamma_2 D_{st} F_s + \alpha_s + e_{st} \quad \textbf{(1)}$$

$Y_{st}$ = the outcome for school $s$ at time $t$

$\mu_t$ = time-varying intercept (represented with dummy variables for each year above 1999, the baseline year)

$X_s$ = vector of time-varying school-level background indicators

$(D_{st})\ (F_s)$ = a vector containing the interaction terms between time and CSR funding

$\alpha_s$ = school fixed effects

$\varepsilon_{st}$ = random disturbance

---

The maximum difference in propensity scores was set at 0.01. Once a match is made neither the treatment nor the control school can be matched to another school. The algorithm matched 81 of the 94 cohort-one Florida elementary to Florida schools that never received funding between 1999 and 2004, yielding a total of 162 schools for the sample. After the elimination of two funded schools that closed between 1999 and 2004 and their matched schools, the final sample was reduced to 158.

[13] We estimate this model with the SAS generalized linear model process, which models school fixed effects and provides standard error estimates that account for the clustering of observations within schools (Allison, 2005). Standard errors that account for clustering can be retrieved from OLS estimation by calculating Huber-White robust standard errors. The estimates from the SAS GLM procedure (PROC GLM) yield identical coefficient estimates to OLS models, and standard errors very similar to the robust standard errors that can be computed in traditional OLS. The advantage of PROC GLM is that the adjustments needed for fixed-effects models require very little code, and the ability to account for clustering is built into the procedure.

The time-varying intercepts ($\mu_t$) give the average yearly difference between each year and the baseline year (1999) for student outcomes in the unfunded schools. The coefficients on the interaction between funding and the dummy variables for time give the disparity between funded and unfunded schools in the yearly change from the baseline year and, thus, give the yearly effect of receiving funding.

<u>Analyzing the effects of implementing one of three leading CSR designs</u>

In our second set of analyses we focus on three specific CSR designs – *Core Knowledge, Direct Instruction,* and *Success for All* – and are interested in the basic effect of bringing a CSR design and design team in to the school as well as the effect of the level of implementation of the various design components of the CSR. The three CSR designs that we focus on are among the most widely and longest used CSR models. Specifically, of the more than 1200 external and homegrown models employed by schools with CSRD funds, 16 models have been employed by 47% of all schools receiving funding.[14] These three models are among these 16 most popular models; *Success for All* is the leading externally developed reform model implemented by schools receiving funding nationwide. In Florida, these three models have been used in 14% of all schools that have received federal CSRD funding.[15]

We first look at the trends from 1999 to 2004, the complete span over which we have data on Florida schools. We then turn our attention to trends from 2002 to 2004, the years in which we surveyed teachers and principals. As was the case with the policy evaluation described above, careful attention was given to selecting a sample of schools for this investigation. However, to minimize the impact of multiple studies on schools, the Department of Education—which funded this study along with five other major studies of CSR—asked us to collaborate with the other research teams in our sample selection strategy, which ultimately limited our flexibility in

---

[14] All figures regarding the models used by funded schools are according to the Southwestern Educational Development Laboratory CSR Awards Database, which can be accessed on-line at http://www.sedl.org/csr/awards.html.

[15] According to the SEDL CSR Award Database, the ten most frequently implemented reforms by schools receiving CSRD funding include *Success for All* (477), *Lightspan* (312), *America's Choice* (275), *Accelerated Schools* (273), *Co-nect* (191), *High Schools That Work* (178), *Coalition of Essential Schools* (176), *Effective Schools* (173) , *Renaissance Learning* (169), *Direct Instruction* (161). *Core Knowledge* ranks 15th, with 98 CSRD-funded schools implementing its design.

selecting schools. The schools in our sample were drawn at random from lists of participating schools in Florida provided by the three CSR design developers. Using data from 2002 and a multivariate matching approach, we matched each sample school that participated in a model to a corresponding school not using the model, based on student demographics, enrollment, and urbanicity. The final sample included 185 Florida schools, with representation from each of the three models as well as from the control schools.[16] We pooled the sample of schools across the models in order to capture enough power to detect small effects.

In addition to the outcomes and background data used in the policy evaluation described above, we used data from surveys of principals and teachers collected by our research team each spring in 2002, 2003, and 2004. We developed these surveys in collaboration with CSR design developers and designed them to capture the extent to which schools used the providers' preferred practices. A unique aspect of our survey is that, whenever possible, we wrote questions regarding design practices in general language that could be interpreted by non-design schools, thus allowing us to capture the use of design practices in all of the sample schools, regardless of whether a model is used or the use of a specific model. From the survey responses we generated composite scores that reflect the extent to which each school in the sample used the practices endorsed by each of the three design providers. A complete list of the items used to generate practice scores for each of the designs is provided in **Table A1** of the appendix.

The model specification used for the evaluation of implementation is different from that used for analysis of federal funding for CSRD described above. Because all CSRD schools in the sample received funding in the same year, an examination of the change in outcomes from the baseline year is informative for the policy evaluation. However, when the CSR start dates vary, as they do in our sample of CSR implementers, the baseline years (1999 and 2002) do not have a conceptual relevance. Instead, we employ multi-level trend models that account for clustered observations, control for a time trend, and separate within-school effects from across-school effects by first centering the time-varying independent variables around the school mean, and then entering these means into the model (Allison, 2005; Horney, Osgood, & Marshall, 1995).

---

[16] The sample of schools included 29 *Core Knowledge Schools*, 57 *Direct Instruction Schools*, 35 *Success for All* schools, and 64 control schools.

Equation 2 illustrates this centering technique and analytic model for evaluating CSR implementation:

$$y_{st} = \gamma_0 + \gamma_1 t + \gamma_2 \left(X_{st} - \bar{X}_s\right) + \gamma_3 \left(P_{st} - \bar{P}_s\right) + \gamma_4 \bar{X}_s + \gamma_5 \bar{P}_s + \alpha_s + e_{st} \quad \textbf{(2)}$$

$Y_{st}$ = the outcome for school $s$ at time $t$

$X_{st}$ = vector of time school-level background indicators

$P_{st}$ = vector indicating a school's participation in the model or a series of vectors indicating the number of years of participation

$\varepsilon_{st}$ = random disturbance

$\rho_0$ = random error around the intercept

$\rho_1$ = random error around the growth term (slope)

In the analysis presented in this paper, participation in CSR designs is represented with a simple time-varying vector, in which a school's participation in a reform model is indicated with a one, and with multiple variables indicating different stages of participation. To account for variation in practice we replaced the participation indicators with a measure of a school's practice as it relates to the three models investigated in this study. While keeping the same structure as the models of participation, we only include the three survey years, 2002-2004, in the models of practice, whereas in the participation models we utilize all years of data.

While the analysis which represents implementation with the years of design participation is similar to previous work, we are able to do this analysis over a much longer period of time than previous studies. In addition, by measuring the extent to which all schools use practices endorsed by CSR designs, we are able to assess the specific contribution of the CSR designer's coordination and oversight of their program.

**Results**

A common theme running through these results is that, while the use of a CSR design shows some notable positive effects on student outcomes in some specifications, these effects are neither robust across model specifications, nor are the positive findings very large. The core

ideas that back CSR reforms may be research-based, and early pilot implementations may have shown great promise, but ultimately in this sample, these strategies fail to provide evidence of school-level effects under the *typical* conditions in which these programs—which have now been scaled-up significantly—are implemented. The results that follow describe a very similar story to other large studies of CSR, and raise questions about the generalizability of CSR for the improvement of low-performing schools.

Evaluating the Federal CSRD policy

Our models examining schools that received CSRD funding provide no evidence that this intervention improved school-level student performance in reading or math, although there is modest evidence that funded schools improved in measures of student discipline. The model estimates associated with the time dummy variables, given in **Table 2**, reflect the change made by *non-funded* (control) schools, while the estimates of the interaction between time and funding indicates difference in change between unfunded and funded schools. The first two columns in **Table 2** show that unfunded schools, on average, increased their math and reading scores throughout the five years in this study with schools in 2004 scoring, on average, 37 and 21 points higher in 2004 than 1999 in reading and math, respectively. Because we found no significant interaction effects, it appears that all schools improved at approximately the same rate.

**(Table 2 about here)**

Changes in school, student, and teacher characteristics seemed to play a role in changes in student achievement, both in these models of achievement and in the models of CSR implementation that will be presented in later sections. For the most part, the impact of these background variables corresponds with commonly held expectations that indicators of more challenging student populations and school contexts are associated with lower gains—suggesting that CSRs have not overcome the traditional issues that have historically challenged schools.[17] Estimates of these effects are given in **Table A2** of the appendix.

---

[17] Increases in the percent of minority students, or in the percent of students receiving free or reduced-price lunch, correspond with lower achievement gains in both reading and math, while increases in the average experience of teachers is associated with greater gains in student performance. Interestingly, increases in a school's expenditure on regular education appear to be associated with lower achievement, indicating that lower-performing schools and

The discipline measures listed in **Table 2** including rates of violent acts, percent of students with high absences, and the share of students receiving in-school or out-of-school suspension show inconsistent trends. These models indicate that, on average, schools reduced their violent acts by 0.069 between 1999 and 2004, and the percent of students with high absences declined by 1.2 percent between 1999 and 2004. However, we find no statistically significant difference between funded and unfunded schools in these outcomes. The results show a different story for in-school and out-of-school suspensions (ISS and OSS, respectively). Here, unfunded schools show no statistically significant change in their rates of ISS or OSS over the six years of the study. Funded schools, however, *reduced* their rate of ISS, at least early on (in 2000 and 2002), but showed other periods in which they *increased* their rates of OSS (2003 and 2004). In 2004, the change between the 1999 and the 2004 OSS rates for funded schools was on average almost 1.4 percentage points higher than their unfunded counterparts. Considering that the 1999 rate of OSS for funded schools was 3.9, such a change amounts to an almost 36 percent increase for funded schools. It should be noted that the application of ISS and OSS are at the discretion of school faculty and administration and can be differentially applied across schools. Therefore, changes in these rates can be difficult to interpret. We include these discipline indicators, however, because they have important implications for the amount of time students are in their classrooms. ISS students, while in the building, often are not with their classes receiving instruction; OSS students are not even in the building. These variables reflect students' level of engagement in their instruction as well as the overall climate for students.

Background characteristics do not, by and large, appear to show consistent effects across the four indicators of discipline in these models or the models of CSR implementation presented in subsequent sections. Results for background characteristics are given in **Table A3** of the appendix.

Overall, these results do not present an overwhelmingly positive or negative picture of the impact of CSRD funding from the first cohort of schools in Florida. Schools receiving funding

---

districts are likely are being addressed with additional funding. Changes in the percent of disabled students and percent of students with limited English proficiency have different effects across the math and reading models.

appear to have increased their student achievement in reading and math and reduced their per capita acts of violence and high absenteeism at about the same rate as similar, unfunded schools. While it appears that they reduced their rate of ISS at certain points in time, they appear to have increased their rate of OSS overall.

*Participation models*

The analysis of our participation models reveals that participation in three prominent CSR models – DI, CK, or SFA – corresponds with modest gains in achievement scores and modest improvements in discipline indicators at the school level. The results presented in **Table 3** show the impact of schools' participation in CSR irrespective of the length of implementation (see Columns I and III), and the impact of different lengths of participation (see Columns II and IV) on student achievement, which we have defined in three stages. Stage one schools are those in their 1$^{st}$ and 2$^{nd}$ year of implementation and at the point where they are being introduced to and trained in the model practices. Stage two schools include those in their 3$^{rd}$, 4$^{th}$, and 5$^{th}$ years of implementation. These are schools with experience in the model practices but often still receiving regular support from their providers. This is the stage at which most reform experts expect to see improvements. The third stage includes schools that have been implementing the models for more than five years.

In this discussion, we focus on the estimates of within-school effects of CSR implementation because these estimates reflect the change in school-level student outcomes that occurs after a school adopts a model or moves to a different stage of implementation. The multi-level models used in this analysis produce estimates of across-school effects, which in many cases offer an indication of how the overall status of schools is associated with student outcomes. These across-school effects must be viewed with caution, as they could still be correlated with the error in the model and therefore reflect bias. We discuss the across-school effects only when they offer an interesting insight into the difference between CSR schools and non-CSR schools.

By looking at the coefficients on time we see that, on average, schools across the sample increased their reading and math scores over the five years of the study at a rate of just over five points in reading (Column I) and almost 4 points in math (Column II). Adopting one of the three

study models is not associated with increases in student reading performance. The models of math achievement, however, reveal that schools contracting with a CSR provider had an additional boost of almost four points in their 5th grade math achievement as they moved into the second stage of implementation. No additional jump in performance is seen for schools entering the third stage of implementation. Just as with the policy evaluation models, we find that student and school background factors account for most of the explained variation in student scores. These estimates can be seen in the appendix **Table A4**.

Consistent with the achievement results, the effects of CSR participation on discipline indicators (given in **Table 3**, Columns V-VII) varies across indicators and stages of implementation and do not yield a uniformly positive picture. Notably, schools contracting with a CSR provider reduced the percent of students with high absentee records by 0.79 percent upon adopting the CSR models, although additional improvements in absenteeism do not continue into subsequent stages (see Column VI). Schools also saw a modest decline in the rate of violent acts as they entered the second stage of implementation. Somewhat discouraging results appear in the models of ISS and OSS in Columns X and XII. On average, schools entering the third stage of implementation saw an increase in their rate of ISS of 0.96 percent and an increase in OSS of 0.96 percent.

**(Table 3 about here)**

*Implementation models*

The final set of models examines the practice of schools and explores whether engagement in instruction, organizational, or governance practices endorsed by the CSR design (which we refer to as 'model-preferred practice') had improved student achievement or discipline rates. As explained earlier, we designed this analysis to account for the fact that schools engaged in a formal contract with CSR design providers may be engaged in many of the same practices that schools with no formalized CSR design affiliation employ to improve student performance. Through our surveys, we measured the extent to which *all* schools in our sample utilized the strategies endorsed by the design providers examined in this study, and tested whether greater usage of model-preferred practice impacted outcomes across the entire sample. Our primary interest is to learn if increasing design-preferred practices increases student outcomes and if

actually contracting with a CSR design provider gives an additional boost to the impact of the practices.

**Tables 4 and 5** give selected coefficient estimates for models of student achievement and discipline indicators.[18] In addition to the effect of increasing practice, we also explore the interaction between contracting with a CSR provider and the use of design-preferred practices. Because our survey measured the use of the model practices for all schools, these analyses are limited to practices that can be described in a general language and do not include information about the use of specific curriculum materials or practices that are unique to the model.

Overall, we find no evidence that increasing the use of any of the preferred practices for CK, DI, or SFA in model and non-model schools is associated with increases in student achievement. Nor do we find evidence that increasing practice while contracting with a model provider is associated with increases in student achievement (See **Table 4**).

The models of reading and math achievement show that increasing the level of CK, DI, and SFA-preferred practice fails to show statistical significance, and it appears that contracting with a CSR design provider provides no additional boost to the impact of practice on student achievement. Moreover, the interaction terms not only fail to show statistical significance for all models but are, in fact, negative (though not significant) for SFA. These analyses offer no evidence that increasing the use of model-oriented practices over the three survey years increased student performance or that participation with the design provider enhanced the effect of engaging in design-preferred practices.

It is interesting to note that the across-school effects in our results show that schools with higher average usage of SFA practices were generally higher-performing in reading and math by almost 17 and 17 points, respectively. Schools with higher average CK practice were higher-performing in reading by 14 points. Again, we caution that while these models give estimates of the difference in performance across schools with different average design-oriented usage, they cannot suggest a causal relationship between SFA or CK practice and student achievement

---

[18] Full model estimates are provided in the appendix to this paper. **Tables A5** and **A6**.

outcomes. We cannot tell whether the use of these practices accounts for the higher performance, or if there is something about higher-performing schools that compels them to make use of these practices.

There is a chance that the estimates from these models do not reveal the full effect of contracting with the CSR design developer. If model participation enhanced the impact of model-oriented practice, we would expect the coefficient on the interaction term to be significant.

**(Table 4 about here)**

Given that the across-school components of the model reveal a positive and sizable relationship between CK and SFA practice and student performance, it is possible that our three-year time frame is not long enough for schools to present *enough* change in their practice to generate changes in student performance. Alternatively, the effects of changing practices may not be realized for a year or more after the change, and our time frame does not permit us to test for lagged effects. Both possibilities should be considered when designing future studies.

**(Table 5 about here)**

The evidence presented in **Table 5** on the use of these practices and our discipline indicators is mixed. On the positive side, increasing the use of CK practice by one unit is associated with a 1.7 percentage point decline in ISS. We also see some added benefits to actually participating with a CSR design. For example, contracting with CK providers offered an added benefit in reducing the rate of violent acts by 0.05 (Column I). Contracting with DI providers gave an added benefit in decreasing high absenteeism among students by more than 2 percentage points (Column V). On the negative side of the story, we see that increasing DI practice is associated with increases in student absenteeism – although, as mentioned above, participating in DI appears to counteract the increases in absenteeism seen in non-participating schools. In addition, increasing DI practice appears to be associated with increases in ISS.

The implementation analysis presented in this paper makes some important steps forward: first, by estimating implementation from the perspective of teachers and principals, and second, by accounting for the use of these practices across both control and model schools. It is clear from the mixed results, however, that implementation analysis of CSR will need further development and investigation. While researchers, no doubt, continue to improve their instruments for measuring implementation, future work should also look to extend the time horizon in order to capture lagged effects and permit more time for practices to change.

**Conclusions**

Comprehensive School Reform was introduced as an approach that promised research-based practice and a coherent set of strategies to support schools and coordinate their reform efforts. The logic of this reform approach has considerable appeal and led to the development of hundreds of reform models that have now been implemented in thousands of schools. After years of research on CSR, it is clear that no single study can address the variation in models, and contexts in which they are applied, to answer the basic question: "Does it work?"

The approach that must be taken by the reform and research community to answer this million (perhaps billion) dollar question is to build a substantial body of high-quality research that examines the question from a variety of perspectives, including design-specific studies, multiple-design studies, studies that focus on the CSRD-funded schools, and studies that examine the wider community of schools that are adopting CSR designs. These studies should also collectively look across all of the relevant contexts that potentially influence the effectiveness of these policies. Most importantly, this work must employ methodological strategies that address the fundamental issues in this field of inquiry and must provide the field with defendable accounts of CSR. Although the body of research will, no doubt, have some inconsistencies, when we step back to look at such a collection of high-quality research the patterns across studies will be revealed.

While much of the earliest published research was criticized for inadequate methodological approaches, this work, along with more recent research, offers theoretical and analytic insights that advance the field of CSR research. This study contributes to the larger body of research as a

multiple-design evaluation of CSRD funding and of the implementation of CSR designs under typical circumstances for schools. In addition, we use analytic techniques that attempt to address several of the complications posed by this type of research. While we acknowledge that improvements can be made in our evaluation of design practices, this measuring of the practices of schools as they relate to specific CSR designs is an important step forward. In our investigation of funding and implementation, we find evidence that these CSR models have had some effects on student math performance and some positive effects on student discipline, though these results were not consistent across our investigation. This study, while certainly not intended to be the last word in CSR research, provides only weak support for the popular belief that CSR, as a collective reform movement, is a generalizable approach for improving low-performing schools.

**Appendix**

## Table A1:  Variables Used to Measure Design Practices

| | Variable Label | CK | DI | SFA |
|---|---|---|---|---|
| Principal | Parents always participate on school steering committee | $X^{123}$ | | $X^{123}$ |
| Principal | Principal met with staff member ?? times for CSRM this year | | $X^{23}$ | $X^{23}$ |
| Principal | Staff member spends time coordinating school wide improvement programs as required by model | | $X^{23}$ | $X^{23}$ |
| Principal | Principal disagrees "state and/or district policies and regulations impede school's efforts to improve student performance" | | $X^{123}$ | |
| Principal | District gives the school all the support it needs to implement school-wide programs | $X^{123}$ | $X^{23}$ | $X^{23}$ |
| Principal | School assesses students on Reading multiple times per marking period | $X^{123}$ | $X^{123}$ | $X^{123}$ |
| Principal | Times principal met with external consultant | | $X^{23}$ | $X^{23}$ |
| Principal | School has external consultant who assist in implementing school wide improvement programs | | $X^{23}$ | |
| Principal | At least 75% of other reading teachers are certified | | | $X^{123}$ |
| Principal | Percent of parents attending special events | $X^{123}$ | | $X^{123}$ |
| Principal | Percent of parents attending school committees or working groups | $X^{123}$ | | $X^{123}$ |
| Principal | Percent of parents volunteer | $X^{123}$ | | $X^{123}$ |
| Principal | Percent of parents attending education workshops | $X^{123}$ | | $X^{123}$ |
| Principal | Students are assigned to reading classes based on current academic performance | | $X^{123}$ | $X^{123}$ |
| Principal | A parental involvement working group meets weekly | $X^{123}$ | | $X^{123}$ |
| Teacher | Teachers agree  'reading curriculum is well aligned with state standardized test' | $X^{123}$ | | |
| Teacher | Student tests are used to assign students to reading class at least every 6-8 weeks | | $X^{123}$ | $X^{123}$ |
| Teacher | Student tests are used to assign students to reading groups at least every 6-8 weeks | | $X^{123}$ | |
| Teacher | Teachers agree 'there is consistency in curriculum/instruction among teachers in the same grade' | | $X^{123}$ | |
| Teacher | Students 'work collaboratively in groups or pairs during reading instruction every school day ' | | | $X^{123}$ |
| Teacher | Teachers consult the year long plan/pacing guide on a daily basis | | $X^{23}$ | |
| Teacher | Teachers agree 'curriculum/instruction materials are well coordinated across levels' | $X^{123}$ | $X^{123}$ | |
| Teacher | Teachers agree 'teachers in the school emphasize immediate correction of student academic errors' | | $X^{123}$ | |
| Teacher | Teachers interact formally with external consultant for implementation of improvement programs | $X^{123}$ | $X^{23}$ | $X^{23}$ |
| Teacher | Teachers receive weekly formal feedback on teaching from district staff | $X^{123}$ | $X^{123}$ | |
| Teacher | Teachers receive weekly formal feedback on their teaching from contractor | $X^{123}$ | $X^{123}$ | $X^{123}$ |
| Teacher | Teachers receive weekly formal feedback on their teaching from parents | $X^{123}$ | | |
| Teacher | Teachers receive weekly formal feedback on their teaching from principals | $X^{123}$ | $X^{123}$ | |
| Teacher | Teachers receive weekly formal feedback on their teaching from school staff | $X^{123}$ | $X^{123}$ | $X^{123}$ |
| Teacher | Classes have 20 or fewer students | | | $X^{123}$ |
| Teacher | Students with lowest reading skills are placed in smaller reading groups | | $X^{23}$ | $X^{1}$ |
| Teacher | Times teacher met with external consultants this year | $X^{123}$ | $X^{23}$ | $X^{23}$ |
| Teacher | Times teacher met with facilitator this year | $X^{123}$ | $X^{23}$ | $X^{23}$ |
| Teacher | Teachers formally meet weekly to develop or review student assessments | | | $X^{123}$ |
| Teacher | Teachers formally meet weekly to discuss instruction | $X^{123}$ | | $X^{123}$ |
| Teacher | Teachers formally meet weekly to assess  'school needs, set school goals, implement plans to meet goals, develop/review assessments, discuss instructional strategies, develop or revise curricula' | $X^{123}$ | | |
| Teacher | Teachers contribute to the development of the year long plan or pacing guide | $X^{123}$ | | |
| Teacher | Teachers use year long plan to minimize curriculum overlap | $X^{123}$ | | |
| Teacher | At least 96% of students return homework signed by parents | | $X^{123}$ | $X^{123}$ |
| Teacher | Reading groups have no more than 4 students for SFA and 9 students for DI | | $X^{123}$ | $X^{123}$ |
| Teacher | Teachers assign 20 min reading homework every school day | | | $X^{123}$ |
| Teacher | Teachers teach reading to student in small groups majority of the time | | $X^{123}$ | |

| | Variable Label | CK | DI | SFA |
|---|---|---|---|---|
| Teacher | Students in reading class or groups are at about the same reading skill level | | $X^{123}$ | $X^{123}$ |
| Teacher | Teachers review student scores with school coach after all assessments | | $X^{123}$ | $X^{123}$ |
| Teacher | Teachers review student scores with external coach after all assessments | | $X^{123}$ | |
| Teacher | Teachers review student scores with principal after all assessments | | $X^{123}$ | $X^{123}$ |
| Teacher | Teachers 'usually' or 'always' follow closely an explicit word-for-word text or script for presenting reading lessons | | $X^{123}$ | |
| Teacher | Teachers require parents to sign reading homework | | | $X^{123}$ |
| Teacher | Time teachers teach reading per day | | $X^{123}$ | $X^{123}$ |
| Teacher | Tutored students receive tutoring every school day | | $X^{123}$ | $X^{123}$ |
| Teacher | Percent (90th percentile) of students receiving tutoring | | $X^{123}$ | $X^{123}$ |
| Teacher | Students receive supplemental tutoring in reading | | $X^{123}$ | $X^{123}$ |
| Teacher | Tutored students receive one-on-one tutoring | | | $X^{123}$ |
| Teacher | Teachers use year long plan or pacing guide and usually keep-up with it or move faster | $X^{123}$ | $X^{123}$ | $X^{123}$ |

**Table A2: The Effect of Background Characteristics on Student Achievement for Models of CSRD funding**

|  | READING GRADE 4 | MATH GRADE 5 |
|---|---|---|
| Percent minority | -.232** | -.166** |
|  | (.0260) | (.0308) |
| Percent free or reduced lunch (FRL) | -.420** | -.491** |
|  | (.0404) | (.0480) |
| Percent of students disabled | .280** | -.400** |
|  | (.0850) | (.100) |
| Expenditure on regular education/ per $100 | -.293** | -.356** |
|  | (.0598) | (.0707) |
| Enrollment | -.00278 | -.00602** |
|  | (.00247) | (.00290) |
| Percent of teachers with advanced degrees | .166** | .0668 |
|  | (.0449) | (.0534) |
| Average years of experience of teachers | .584** | .401** |
|  | (.141) | (.166) |

**Table A3:  Effect Of Background Characteristics on Behavioral Outcomes
in Models Of CSRD Funding**

| | Violent Acts | Absences | In-School Suspensions | Out-of-School Suspensions |
|---|---|---|---|---|
| Percent minority | .000481** | -.0227** | -.00695 | .0374** |
| | (.000110) | (.00615) | (.00752) | (.00778) |
| Percent FRL | .000179 | .0897** | .0208* | .0361** |
| | (.000172) | (.00957) | (.0117) | (.0121) |
| Percent LEP | .0000577 | -.0390** | -.0215 | -.0459** |
| | (.000225) | (.0125) | (.0154) | (.0159) |
| Percent disabled | .000843** | .0967** | -.0250 | -.00102 |
| | (.000361) | (.0201) | (.0246) | (.0255) |
| Expenditure on regular education/ $100 | .00125** | .00742 | .0257 | .0188 |
| | (.000254) | (.0142) | (.0173) | (.0179) |
| Enrollment | -.00000650 | .00140** | .000564 | -.00204** |
| | (.000104) | (.000582) | (.000713) | (.000737) |
| Percent of teachers with advanced degrees | -.000944** | -.0299** | -.0160 | -.0405** |
| | (.000191) | (.0106) | (.0130) | (.0135) |
| Teachers' average years of experience | .000405 | -.00726 | .0518 | -.0231 |
| | (.000596) | (.0332) | (.0407) | (.0420) |

**Table A4: The effects of background characteristics in on Reading and Math
in models of CSR participation**

| | I<br>4th Grade<br>Reading | II<br>4th Grade<br>Reading | III<br>5th Grade<br>Math | IV<br>5th Grade<br>Math |
|---|---|---|---|---|
| **WITHIN-SCHOOL EFFECTS** | | | | |
| Percent minority | -0.537** | -0.550** | -0.520** | -0.529** |
| | (0.142) | (0.142) | (0.148) | (0.148) |
| Percent FRL | -0.0926** | -0.0875** | 0.0262 | 0.0273 |
| | (0.0354) | (0.0354) | (0.0390) | (0.0391) |
| Percent LEP | 0.126* | 0.1263* | -0.0987 | -0.0969 |
| | (0.076) | (0.0760) | (0.0839) | (0.0839) |
| Percent disable | -0.279** | -0.284** | -0.123 | -0.119 |
| | (0.122) | (0.122) | (0.134) | (0.134) |
| Expenditure on regular education | 0.000928 | 0.00855 | -0.00132* | -0.00127 |
| | (0.000698) | (0.000701) | (0.000775) | (0.000780) |
| Enrollment | 0.00283 | 0.00240 | -0.00282 | -0.00263 |
| | (0.00396) | (0.00397) | (0.00428) | (0.00431) |
| % of teachers with advanced degrees | 0.00297 | 0.0117 | -0.0548 | 0.0582 |
| | (0.0527) | (0.0528) | (0.0579) | (0.0581) |
| Average teacher years of experience | 0.392** | 0.427** | 0.104 | 0.138 |
| | (0.196) | (0.197) | (0.218) | (0.2187) |
| % disable | -0.279** | -0.284** | -0.123 | -0.119 |
| | (0.122) | (0.122) | (0.134) | (0.134) |
| **ACROSS-SCHOOL EFFECTS** | | | | |
| Mean percent minority | -0.139** | -0.144** | -0.172** | -.179** |
| | (0.0459) | (0.0453) | (0.0562) | (0.0556) |
| Mean percent disabled | -0.685** | -0.669** | -0.935** | -0.933** |
| | (0.197) | (0.194) | (0.241) | (0.238) |
| Mean expenditure on regular education | -0.00207 | -0.00262 | -0.00211 | -0.00250 |
| | (0.00176) | (0.00175) | (0.00215) | (0.00215) |
| Mean percent FRL | -0.606 | -0.573** | -0.629** | -0.594** |
| | (0.0800) | (0.0799) | (0.0980) | (0.0981) |
| Mean percent LEP | -0.134 | -0.147 | 0.219 | 0.184 |
| | (0.1374) | (0.137) | (0.168) | (0.168) |
| Mean enrollment | -0.00780 | -0.00937 | -0.00754 | -0.00879 |
| | (0.00605) | (0.00600) | (0.00741) | (0.00737) |
| Mean percent of teachers with advanced degrees | 0.241** | 0.218** | -0.244** | 0.221** |
| | (0.0852) | (0.0847) | (0.105) | (0.104) |
| Mean years of experience for teachers | 0.319 | 0.367 | 0.185 | 0.177 |
| | (0.339) | (0.339) | (0.415) | (0.416) |

**Table A5:  The effects of background characteristics on behavioral outcomes in models of CSR participation**

| | I | | II | | III | | IV | |
|---|---|---|---|---|---|---|---|---|
| | **Absence** | | **Violence** | | **ISS** | | **OSS** | |
| **WITHIN-SCHOOL EFFECTS** | | | | | | | | |
| Percent minority | -.054** | -.0525* | -.00038 | -.00033 | -.0469 | -.0458 | .0259 | .0270 |
| | (.0300) | (.0300) | (.000649) | (.000650) | (.0461) | (.0464) | (.0355) | (.0355) |
| Percent FRL | .0629** | .0620** | .000405** | .000404** | .0376** | .0351** | .0797** | .0783** |
| | (.00761) | (.00763) | (.000159) | (.000159) | (.0119) | (.0119) | (.00915) | (.00918) |
| Percent LEP | -.00445 | -.00417 | -.00040 | -.00041 | .0427 | -.0426 | -.00504 | -.00538 |
| | (.0166) | (.0166) | (.000349) | (.000348) | (.0259) | (.0259) | (.0200) | (.0200) |
| Percent disable | .0856** | .0876** | -.00108** | -.00113** | -.0374 | -.0365 | -.0299 | -.0297 |
| | (.0264) | (.0264) | (.000553) | (.000552) | (.0411) | (.0411) | (.0317) | (.0317) |
| Expenditure on regular education | -.010 | -.007 | .000419 | -.000368 | -.000907 | .0051 | -.030 | -.027 |
| | (.0152) | (.0153) | (.000317) | (.000318) | (.0237) | (.0238) | (.0183) | (.0184) |
| Enrollment | .000811 | .000962 | -.00001 | -.00001 | .00024 | .000066 | -.00262** | -.00252** |
| | (.000859) | (.000862) | (.000018) | (.000018) | (.00133) | (.00134) | (.00103) | (.00103) |
| Percent of teacher with Advanced Degrees | .000869 | -.00089 | -.00028 | -.00027 | .0171 | -.0219 | -.0148 | -.0172 |
| | (.0114) | (.0114) | (.000238) | (.000238) | (.0177) | (.0177) | (.0137) | (.0137) |
| Average teacher years of experience | -.0698 | -.0703 | -.00029 | -.00053 | .0419 | -.0390 | -.0835 | -.0869 |
| | (.0428) | (.0429) | (.000895) | (.000896) | (.0667) | (.0668) | (.0516) | (.0518) |
| **ACROSS-SCHOOL EFFECTS** | | | | | | | | |
| Mean percent minority | -.0321** | -.0321** | .000366** | .000376* | -.00991 | -.00895 | .0212** | .0220** |
| | (.000906) | (.00897) | (.000128) | (.000128) | (.0148) | (.0148) | (.0104) | (.0103) |
| Mean percent disabled | .0311 | .0269 | -.00057 | -.00052 | .0503 | .0549 | .0365 | .0403 |
| | (.0392) | (.0388) | (.000553) | (.000551) | (.0642) | (.0641) | (.0450) | (.0447) |
| Mean expenditure on regular education | -.042 | -.030 | .000447 | .000373 | -.028 | -.029 | .0231 | .0226 |
| | (.0338) | (.0338) | (.000479) | (.000481) | (.0554) | (.0558) | (.0388) | (.0389) |
| Mean percent FRL | .112** | .108** | .000068 | .000070 | .0512** | .0490* | .0403** | .0382** |
| | (.0156) | (.0156) | (.000221) | (.000222) | (.0256) | (.0257) | (.0179) | (.0180) |
| Mean percent LEP | -.0324 | -.0338 | -.00045 | -.00039 | -.0561 | -.0477 | -.0543* | -.0475 |
| | (.0271) | (.0269) | (.000382) | (.000382) | (.0443) | (.0444) | (.0311) | (.0310) |
| Mean enrollment | -.00041 | -.00007 | -.00003* | -.00003** | -.00081 | -.00072 | -.00377** | -.00372** |
| | (.00123) | (.00122) | (.000017) | (.000017) | (.00201) | (.00202) | (.00141) | (.00141) |
| Mean percent of teachers with advanced degrees | -.0273 | -.0231 | -.00052** | -.00054* | -.0906** | -.0909** | -.0490** | -.0486** |
| | (.0171) | (.0170) | (.000242) | (.000242) | (.0280) | (.0281) | (.0196) | (.0196) |
| Mean years of experience for teachers | -.0782 | -.0964 | .00103 | .00125 | .0513 | .0756 | -.113 | -.0951 |
| | (.0683) | (.0683) | (.000969) | (.000975) | (.112) | (.113) | (.0785) | (.0787) |

### Table A6:  Effects of background variables on math achievement in models of CSR implementation

| | READING | | | MATH | | |
|---|---|---|---|---|---|---|
| | **CK** | **DI** | **SFA** | **CK** | **DI** | **SFA** |
| **WITHIN SCHOOL EFFECTS** | | | | | | |
| % Disabled | -.337 | -.349 | -.338 | -.275 | -.249 | -.266 |
| | (.370) | (.371) | (.371) | (.450) | (.451) | (.445) |
| $ Expenditures on Regular Education | .00309** | .00343** | .00348** | .00126 | .00136 | .00125 |
| | (.00134) | (.00133) | (.00133) | (.00161) | (.00159) | (.00157) |
| % Free/Reduced Price Lunch | -.0177 | -.0112 | -.0135 | .0102 | -.0123 | -.00604 |
| | (.0771) | (.0771) | (.0775) | (.101) | (.101) | (.0998) |
| % Minority | -.105 | -.0662 | -.0398 | -.576* | -.588* | -.545* |
| | (.274) | (.274) | (.275) | (.322) | (.322) | (.320) |
| % Limited English Proficiency | .168 | .171 | .134 | .0109 | -.00932 | -.0556 |
| | (.152) | (.152) | (.153) | (.197) | (.197) | (.196) |
| % Teachers with Advanced Degrees | .137 | .126 | .147 | .237* | .238* | .253* |
| | (.109) | (.110) | (.110) | (.132) | (.132) | (.130) |
| Teachers' Average Years of Experience | .0245 | -.0558 | -.0429 | .0527 | -.0454 | -.212 |
| | (.481) | (.500) | (.488) | (.593) | (.612) | (.593) |
| Total Enrollment | -.00168 | .000290 | .000477 | -.0117 | -.0102 | -.0120 |
| | (.0144) | (.0145) | (.0144) | (.0170) | (.0170) | (.0168) |
| **ACROSS SCHOOL EFFECTS** | | | | | | |
| % Free/Reduced Price Lunch | -.552** | -.593** | -.549** | -.561** | -.611** | .551** |
| | (.0664) | (.0629) | (.0620) | (.0787) | (.0789) | (.0771) |
| % Minority | -.0717 | -.0471 | -.0702* | -.143** | -.127** | -.153** |
| | (.0398) | (.0404) | (.0390) | (.0493) | (.0505) | (.0483) |
| % Limited English Proficiency | -.329** | -.317** | -.256** | -.0149 | .00312 | .0904 |
| | (.104) | (.107) | (.104) | (.130) | (.134) | (.129) |
| % Teachers with Advanced Degrees | .196** | .184** | .172** | .0955 | .0514 | .0338 |
| | (.0754) | (.0749) | (.0724) | (.0932) | (.0934) | (.0895) |
| Teachers' Average Years of Experience | .544* | .613** | .570** | .217 | .347 | .289 |
| | (.297) | (.298) | (.289) | (.367) | (.372) | (.357) |
| Total Enrollment | .00642 | .00617 | .00428 | .0100 | .00994 | .00742 |
| | (.00523) | (.00529) | (.00516) | (.00648) | (.00661) | (.00641) |
| **COV ESTIMATES** | | | | | | |
| Intercept | 90.900** | 93.332** | 85.730** | 134.40** | 141.57** | 128.13** |
| | (13.198) | (13.609) | (12.837) | (19.615) | (20.539) | (18.958) |
| dTime | 18.637** | 18.560** | 18.244** | 95.214** | 95.520** | 93.759** |
| | (7.346) | (7.395) | (7.487) | (8.934) | (9.000) | (8.818) |
| Residual | 48.516** | 48.830** | 49.560** | - | - | - |
| | (6.764) | (6.787) | (6.916) | | | |

**Table A7:  Effects of background variables on discipline outcomes
in models of CSR implementation**

| | Interaction models – Background measures | | | | | |
| | Violence per capita | | | Percent of students with more than 20 absences | | |
| | CK | DI | SFA | CK | DI | SFA |
|---|---|---|---|---|---|---|
| | **WITHIN-SCHOOL EFFECTS** | | | | | |
| % Disabled | .00153 | .00139 | .00141 | .111 | .124* | .123* |
| | (.00112) | (.00111) | (.00112) | (.0680 | (.0669) | (.0658) |
| $ Expenditures on Regular Education | -.001** | -.00088** | -.000941** | -.035 | -.038 | -.039* |
| | (.000402) | (.000395) | (.000397) | (.0246) | (.0239) | (.0235) |
| % Free/Reduced Price Lunch | .000237 | .000247 | .000266 | .0205 | .0203 | .0192 |
| | (.000248) | (.000248) | (.000246) | (.0153) | (.0150) | (.0148) |
| % Minority | -.00014 | -.00013 | -.00001 | -.0801 | -.0831* | -.0973** |
| | (.000812) | (.000807) | (.000816) | (.0496) | (.0487) | (.0482) |
| % Limited English Proficiency | -.00089* | -.00096** | -.00100** | .0140 | .0145 | .0269 |
| | (.000485) | (.000486) | (.000484) | (.0299) | (.0295) | (.0291) |
| % Teachers with Advanced Degrees | -.00067** | -.00065** | -.00065** | -.0103 | -.00578 | -.00964 |
| | (.000330) | (.000329) | (.000329) | (.0202) | (.0199) | (.0195) |
| Teachers' Average Years of Experience | .000526 | .000569 | .000135 | -.108 | -.0475 | -.0662 |
| | (.00148) | (.00152) | (.00149) | (.0912) | (.0921) | (.0891) |
| Total Enrollment | .000026 | .000040 | .000032 | .00607** | .00560* | .00585** |
| | (.000043) | (.000042) | (.000043) | (.00262) | (.00257) | (.00254) |
| | **ACROSS-SCHOOL EFFECTS** | | | | | |
| % Disabled | -.00032 | -.00029 | -.00033 | .0689* | .0736* | .0699* |
| | (.000486) | (.000471) | (.000470) | (.0401) | (.0391) | (.0391) |
| $ Expenditures on Regular Education | .000700** | -.000734** | -.000712** | -.012 | -.013 | -.016 |
| | (.000355) | (.000346) | (.000350) | (.0290) | (.0285) | (.0290) |
| % Free/Reduced Price Lunch | .000067 | .000035 | .000056 | .0986** | .103* | .100** |
| | (.000185) | (.000183) | (.000187) | (.0153) | (.0153) | (.0155) |
| % Minority | .000344** | .000363** | .000348** | -.0354* | -.0361* | -.0360** |
| | (.000126) | (.000126) | (.000126) | (.0103) | (.0104) | (.0104) |
| % Limited English Proficiency | -.00030 | -.00028 | -.00028 | -.0428* | -.0479* | -.0484* |
| | (.000303) | (.000303) | (.000307) | (.0252) | (.0254) | (.0257) |
| % Teachers with Advanced Degrees | -.00036 | -.00037* | -.00038* | -.0124 | -.0148 | -.0121 |
| | (.000225) | (.000224) | (.000225) | (.0184) | (.0184) | (.0184) |
| Teachers' Average Years of Experience | .000502 | .000374 | .000469 | -.0545 | -.0564 | -.0520 |
| | (.000855) | (.000837) | (.000841) | (.0697) | (.0687) | (.0689) |
| Total Enrollment | -.00001 | -.000968 | -.00001 | -.00030 | -.00030 | -.00020 |
| | (.000017) | (.000017) | (.000017) | (.00139) | (.00138) | (.00139) |

## Table A7 (cont.): Effects of background variables on discipline outcomes in models of CSR implementation

| | Models without interaction – Background measures | | | | | |
|---|---|---|---|---|---|---|
| | Violence per capita | | | Percent of students with more than 20 absences | | |
| | CK | DI | SFA | CK | DI | SFA |
| WITHIN-SCHOOL EFFECTS | | | | | | |
| % Disabled | .00142 | .00139 | .00140 | .115* | .120* | .118* |
| | (.00112) | (.00111) | (.00112) | (.0679) | (.0676) | (.0674) |
| $ Expenditures on Regular Education | -.000976** | -.000914** | -.000913** | -.038 | -.041* | -.042* |
| | (.000397) | (.000396) | (.000397) | (.0242) | (.0241) | (.0241) |
| % Free/Reduced Price Lunch | .000249 | .000252 | .000261 | .0202 | .0207 | .0197 |
| | (.000249) | (.000247) | (.000247) | (.0153) | (.0152) | (.0152) |
| % Minority | -.00004 | -.00012 | -.00013 | -.0829* | -.0816* | -.0809 |
| | (.000811) | (.000809) | (.000810) | (.0494) | (.0492) | (.0491) |
| % Limited English Proficiency | -.00089* | -.00095* | -.00093* | .0143 | .0166 | .0165 |
| | (.000488) | (.000484) | (.000485) | (.0299) | (.0298) | (.0297) |
| % Teachers with Advanced Degrees | -.00067** | -.00067** | -.00067** | -.0104 | -.00969 | -.00898 |
| | (.000331) | (.000329) | (.000329) | (.0202) | (.0201) | (.0200) |
| Teachers' Average Years of Experience | .000637 | .000233 | .000306 | -.111 | -.0947 | -.0929 |
| | (.00148) | (.000149) | (.00149) | (.0910) | (.0915) | (.0911) |
| Total Enrollment | .000033 | .000038 | .000036 | .00587** | .00549** | .00554** |
| | (.000042) | (.000042) | (.000042) | (.00261) | (.00260) | (.00260) |
| ACROSS-SCHOOL EFFECTS | | | | | | |
| % Disabled | -.00032 | -.00029 | -.00033 | .0689* | .0734 | .0707* |
| | (.000486) | (.000471) | (.000470) | (.0401) | (.0392) | (.0390) |
| $ Expenditures on Regular Education | .000704** | .000734** | .000704** | -.012 | -.013 | -.014 |
| | (.000354) | (.000346) | (.000350) | (.0290) | (.0286) | (.0288) |
| % Free/Reduced Price Lunch | .000064 | .000034 | .000057 | .0987** | .103** | .0998** |
| | (.000185) | (.000183) | (.000187) | (.0153) | (.0153) | (.0155) |
| % Minority | .000346** | .000364** | .000351** | -.0355** | -.0361** | -.0366** |
| | (.000126) | (.000126) | (.000126) | (.0103) | (.0104) | (.0104) |
| % Limited English Proficiency | -.00031 | -.00028 | -.00029 | -.0426* | -.0475* | -.0478** |
| | (.000303) | (.000303) | (.000307) | (.0252) | (.0254) | (.0257) |
| % Teachers with Advanced Degrees | -.00036 | -.00037* | -.00038* | -.0123 | -.0149 | -.0124 |
| | (.000225) | (.000224) | (.000225) | (.0184) | (.0184) | (.0184) |
| Teachers' Average Years of Experience | .000485 | .000396 | .000475 | -.0541 | -.0531 | -.0534 |
| | (.000855) | (.000838) | (.000840) | (.0697) | (.0688) | (.0689) |
| Total Enrollment | -.00001 | -.000951 | -.00001 | .0689* | .0734 | .0707* |
| | (.000017) | (.000017) | (000017) | (.0401) | (.0392) | (.0390) |

**Table A8:  Across-school effects of school background characteristics
on Reading and Math**

| Across-school Effects | I<br>4th Grade<br>Reading | II<br>4th Grade<br>Reading | III<br>5th Grade<br>Math | IV<br>5th Grade<br>Math |
|---|---|---|---|---|
| Mean percent minority | -0.139** | -0.144** | -0.172** | -.179** |
|  | (0.0459) | (0.0453) | (0.0562) | (0.0556) |
| Mean percent disabled | -0.685** | -0.669** | -0.935** | -0.933** |
|  | (0.197) | (0.194) | (0.241) | (0.238) |
| Mean expenditure on regular education | -0.00207 | -0.00262 | -0.00211 | -0.00250 |
|  | (0.00176) | (0.00175) | (0.00215) | (0.00215) |
| Mean percent FRL | -0.606 | -0.573** | -0.629** | -0.594** |
|  | (0.0800) | (0.0799) | (0.0980) | (0.0981) |
| Mean percent LEP | -0.134 | -0.147 | 0.219 | 0.184 |
|  | (0.1374) | (0.137) | (0.168) | (0.168) |
| Mean enrollment | -0.00780 | -0.00937 | -0.00754 | -0.00879 |
|  | (0.00605) | (0.00600) | (0.00741) | (0.00737) |
| Mean percent of teachers with advanced degrees | 0.241** | 0.218** | -0.244** | 0.221** |
|  | (0.0852) | (0.0847) | (0.105) | (0.104) |
| Mean years of experience for teachers | 0.319 | 0.367 | 0.185 | 0.177 |
|  | (0.339) | (0.339) | (0.415) | (0.416) |

## Citations

Allison, P. D. (2005). Fixed Effects Regression Methods for Longitudinal Data Using SAS. Cary, North Carolina: SAS Institute Inc.

Bacevich, A., Le Floch, K. C., Stapleton, J., & Burris, B. (2005). High Implementation in Low-achieving Schools: Lessons from an Innovative Pilot Program (AERA Paper). Washington, D.C.: American Institutes for Research.

Bergstralh, E., & Kosanke, J. (1995). Computerized Matching of Controls (No. Technical Report 56): Mayo Foundation

Bifulco, R. (2002). Addressing Self-Selection Bias in Quasi-Experimental Evaluations of Whole-School Reform: A Comparison of Methods. Evaluation Review. v26 n5 p545-72 Oct 2002.

Bifulco, R., Duncombe, W., & Yinger, J. (2005). Does Whole-School Reform Boost Student Performance? The Case of New York City. Journal of Policy Analysis and Management, 24(I), 47-72.

Bodilly, S. J. (1998). Lessons from New American Schools' Scale-Up Phase: Prospects for Bringing Designs to Multiple Schools. California.

Borman, G., Slavin, R. E., Cheung, A., Chamberlain, A., Madden, N., & Chambers, B. (2005). Success for All: First Year Results from the National Randomized Field Trial. Educational Evaluation and Policy Analysis, 27(1), 1-22.

Borman, G., Slavin, R. E., Cheung, A., Chamberlain, A., Madden, N. A., & Chambers, B. (in press). The National Randomized Field Trial of Success For All: Second-Year Outcomes. American Educational Research Journal.

Cooper, R. (1998). Socio-Cultural and Within-School Factors That Affect the Quality of Implementation of School-Wide Programs. Report No. 28. Maryland.

Datnow, A., & Stringfield, S. (2000). Working Together for Reliable School Reform. Journal of Education for Students Placed at Risk (JESPAR). v5 n1-2 p183-204 2000.

Department of Education, U. S. (1998-2005). Budget for Fiscal Years 1998-2005. In O. o. E. a. S. Education (Ed.): U.S. Department of Education.

Desimone, L. (2002). How Can Comprehensive School Reform Models Be Successfully Implemented? Review of Educational Research. v72 n3 p433-79 Fall 2002.

Horney, J. D., Osgood, W., & Marshall, I. H. (1995). Criminal Careers in the Short-Term: Intra-Individual Variability in Crime and its Relation to Local Life Circumstances. American Sociological Review, 60(5), 655-673.

Kurki, A., Aladjem, D., & Carter, K. R. (2005). Implementation: Measuring and Explaining the Fidelity of CSR Implementation (AERA Paper). Washington, D.C.: American Institutes for Research.

Murphy, J., & Datnow, A. (Eds.). (2003). Leadership Lessons From Comprehensive School Reform. Thousand Oaks Press, CA: Corwin Press.

Rowan, B., Harrison, D. M., & Hayes, A. (2004). Using Instructional Logs to Study Mathematics Curriculum and Teaching in the Early Grades. Elementary School Journal. v105 n1 p103 Sep 2004, 25.

Supovitz, J. A., & May, H. (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. v9 n4 p389-419 Oct 2004, 31.

Tushnet, N. C., Flaherty, J. J., & Smith, A. (2004). Longitudinal Assessment of Comprehensive School Reform Program Implementation and Outcomes: First-Year Report. US Department of Education, 86.

Vernez, G., Karam, R., Mariano, L., & DeMartini, C. (2004). Assessing the Implementation of Comprehensive School Reform Models. Santa Monica, CA: RAND Education.

Zhang, Y., Shkolnik, J., & Fashola, O. S. (2005). Evaluating the Implementation of Comprehensive School Reform and Its Impact on Growth in Student Achievement (AERA Paper). Washington, D.C.: American Institutes for Research.

**Tables and Figures**

**Table 1: Covariate Balance for Matched Sample and Full Sample[19]**

| Variable | | Matched Sample (N=158) | Full Sample (N=1533) |
|---|---|---|---|
| | Funded Mean | 4320 | 4308 |
| Expenditure on regular education | Unfunded Mean | 4226 | 4120 |
| | T-Statistic | -0.92 | -3.24* |
| | Funded Mean | 30.539 | 27.843 |
| Percent of Teachers with Advanced Degrees | Unfunded Mean | 29.987 | 30.251 |
| | T-Statistic | -0.31 | -0.75 |
| | Funded Mean | 11.114 | 10.544 |
| Teachers' Average Years Experience | Unfunded Mean | 11.367 | 12.108 |
| | T-Statistic | 0.52 | 3.65* |
| Percent of Students Receiving Free/Reduced Lunch | Funded Mean | 73.704 | 71.789 |
| | Unfunded Mean | 72.687 | 50.235 |
| | T-Statistic | -0.36 | -12.17* |
| Percent of Students with Limited English Proficiency | Funded Mean | 7.7747 | 5.0517 |
| | Unfunded Mean | 5.7886 | 6.522 |
| | T-Statistic | -1.33 | -0.18 |
| | Funded Mean | 707.42 | 637.23 |
| Student Enrollment | Unfunded Mean | 662.2 | 702.12 |
| | T-Statistic | -1.08 | -0.73 |
| | Funded Mean | 64.223 | 67.841 |
| Percent of Minority Students | Unfunded Mean | 62.224 | 42.288 |
| | T-Statistic | -0.41 | -8.15* |

* indicates a significant difference in the means between the funded and unfunded schools

---

[19] The balance in the matched sample is substantially improved from that in the full sample, in which statistically significant differences are apparent in several variables.

**Table 2. The effect of CSRD/CSR funding on student outcomes
for the 1999 cohort of CSRD schools**

| | Reading Grade 4 | Math Grade 5 | Violent Acts | Absences | In-School Suspensions | Out-of-School Suspensions |
|---|---|---|---|---|---|---|
| Dummy variable 2004 (Y2004) | 37.287** | 21.168** | -.0690** | -1.122** | -.213 | .618 |
| | (2.151) | (2.541) | (.00914) | (.510) | (.624) | (.645) |
| Dummy variable 2003 (Y2003) | 23.715** | 23.433** | -.0341** | -1.105** | -.0260 | -.245 |
| | (2.122) | (2.506) | (.00901) | (.503) | (.616) | (.636) |
| Dummy variable 2002 (Y2002) | 16.723** | 18.581** | -.0268** | -1.109** | .425 | .418 |
| | (2.117) | (2.501) | (.00899) | (.502) | (.614) | (.635) |
| Dummy variable 2001 (Y2001) | 12.788** | 13.206** | -.0232** | -1.035** | .0298 | .0877 |
| | (2.108) | (2.490) | (.00896) | (.500) | (.612) | (.632) |
| Dummy variable 2000 (Y2000) | 7.153** | 12.944** | -.0208** | -1.329** | -.0478 | -.575 |
| | (2.097) | (2.476) | (.00891) | (.497) | (.608) | (.629) |
| CSR funded*Y2004 | 2.510 | -1.990 | .00777 | -.332 | .331 | 1.382** |
| | (2.122) | (2.515) | (.00901) | (.503) | (.616) | (.636) |
| CSR funded*Y2003 | -1.025 | -2.422 | .00409 | -.816 | -.586 | 1.094* |
| | (2.130) | (2.524) | (.00905) | (.505) | (.618) | (.639) |
| CSR funded*Y2002 | -.637 | .469 | -.0108 | -.538 | -1.099* | .123 |
| | (2.129) | (2.522) | (.00904) | (.504) | (.618) | (.638) |
| CSR funded*Y2001 | 1.247 | 2.365 | -.00188 | -.140 | -.670 | .0164 |
| | (2.128) | (2.521) | (.00904) | (.504) | (.617) | (.638) |
| CSR funded*Y2000 | -1.735 | 1.651 | .00574 | -.110 | -1.208* | -.0609 |
| | (2.136) | (2.520) | (.00904) | (.504) | (.617) | (.638) |
| CSR funded*Y1999 | -1.043 | -.00789 | .000119 | .211 | -.483 | .128 |
| | (2.109) | (2.491) | (.00896) | (.500) | (.612) | (.632) |
| N | 943 | 939 | 944 | 944 | 944 | 944 |

** implies p<.05, * implies p<.10