When Educators Are the Learners:
Private Contracting by Public Schools*

Silke Forbes, UC San Diego

Nora Gordon, Georgetown University and NBER

This version: October 24, 2011

Abstract

We investigate decision-making and the potential for social learning among school administrators in the market for school reform consulting services. Specifically, we estimate whether public schools are more likely to choose specific Comprehensive School Reform service providers if their peer schools have performed unusually well with those providers in the past. We consider peer groups defined by common governance, geography, or district characteristics. We find strong evidence that schools tend to contract with providers used by other schools in their own districts in the past, regardless of past performance. In addition, our point estimates are consistent with school administrators using information from peers to choose the plans they perceive to have performed best in the past. Despite choosing a market with an unusually comprehensive data source on contracts between public schools and private firms, our statistical power is sufficiently weak that we cannot reject the absence of social learning.

I. Introduction

In recent years, federal elementary and secondary education policy has become increasingly activist, utilizing conditional grants as levers to prompt states, districts and schools to adopt specific policies that many historically resisted. In the most obvious example, the No Child Left Behind Act of 2002 brought accountability regimes to all states which had not yet independently adopted them. More recently, the federal Race to the Top Assessment Program has awarded grants to two large consortia of states to develop assessments aligned with the Common Core of State Standards, despite long-held opposition to anything remotely approaching a "national" test. Burch (2006) argues persuasively that accountability policies stimulated demand for goods and services sold by private vendors to public educational agencies, documenting major increases in revenues of for-profit educational services providers in the years immediately following the passage of No Child Left Behind.[1] At the same time that federal policymakers have been more willing to dictate the specifics of how state and local educational agencies should operate, however, they have chosen to regulate the private markets stimulated by these policies relatively little, relying on competition to lead to high-quality, fairly priced inputs into public education.

One of the key assumptions underlying this logic is that educational administrators have access to comprehensive--and comprehensible--information about the relative quality of the products they consume, so that low quality or overpriced vendors eventually are forced to exit the market. Whether such information exists and is used in practice, however, is an open question. School administrators could use raw data provided by agencies, analyses generated by professional researchers via the What Works Clearinghouse or other "report card" types of sources, or they could gather data—including objective or subjective

---

[1] Using annual reports, she finds annual revenue increases of 77 percent for firms specializing in test development and preparation, 46 percent for data management and analysis, 300 percent for remedial services, and 150 percent for content area specific programming over the period from 2001 to 2004.

measures—from their peers, based on their experiences with different practices or vendors. Researchers find it difficult to identify best practices because of the lack of identifying variation in adoption decisions; the limited scale of the What Works Clearinghouse (WWC) provides one direct indicator of the scarcity of high-quality research likely to be relevant to practitioners.[2] Even in cases where such research exists, the extent to which the relevant decision-makers use it to inform policy choices is unclear.[3] Alternatively, administrators could gather data in a more ad-hoc but perhaps more informative way, asking their peers in other schools to describe their experiences with various service providers. Research on social networks of educational administrators is limited, but the most comprehensive study to date, Daly and Finnigan (2009), suggests that administrators across schools have limited interactions with one another.

Large-scale empirical evidence on how school administrators choose products from private vendors is scarce because data on such choices are rarely available. In the current study, we seek to close this gap by examining a setting in which such data do exist: the Comprehensive School Reform (CSR) program in Texas public schools from 2000 to 2005. Now defunct, CSR was a federal program which awarded one-time grants to public schools to purchase consulting services from private firms or to develop their own local reform plans. These products were bundles of relatively complex services, such as teacher training and curriculum development, whose quality would have been difficult to assess prior to purchase.

---

[2] The What Works Clearinghouse, http://ies.ed.gov/ncee/wwc/, is sponsored by the Department of Education. A September 9, 2011 search of all its indexed interventions related to academic achievement, the broadest possible category, returned 71 studies, of which 51 were characterized as based on "small extent of evidence" due to limited numbers of studies and/or sample sizes. Of the remaining 20 studies categorized as based on "medium to large extent of evidence," only five were categorized as having positive effects (the remainder had potentially positive effects, no discernible effects, or mixed effects).

[3] In addition to the likelihood that results generated are not used in practice, from the school administrator's point of view, many relevant results are not generated. Of the hundreds of potential vendors for Comprehensive School Reform in Texas, for example, fewer than fifty were sufficiently evaluated to be included in research summaries, and the bulk of those studies included were characterized as insufficient to draw inference.

As a result, the first wave of schools receiving grants had to choose among products with close to unknown quality. Each school only received a single grant so there were no repeat purchases. Schools in later waves of the grant disbursal, however, could potentially learn from the experiences of other schools in previous waves. Schools chose specific CSR plans, and each plan was provided by its sole firm (the contractor). [4] Our empirical analysis relies on a comprehensive database of the CSR providers chosen by schools which purchased the service with federal grants. To generate measures of perceived plan quality, we combine these data with information on school-level pass rates from the Texas Assessment of Academic Skills (TAAS) test.

The problem of choice among experience goods—products with quality that is difficult to observe *ex ante*—is not unique to the education context. Research in disparate settings points to a role for social learning in helping agents to make these decisions, ranging from farmers in developing countries basing decisions on production technologies on the success (or failure) peers have had with them (Munshi, 2004; Conley and Udry, 2010) to individuals in the U.S. using information from their peers in choosing among employer-provided health insurance plans (Sorenson, 2006). This paper is the first to our knowledge to empirically examine how agents acting on behalf of principals base decisions on information likely gathered via social networks. The vocabulary in our case is confusing: the agents are public school administrators, most likely school principals, acting on behalf of taxpayers as they choose private contractors for school consulting services. We ask how information disseminates from perceived quality of past experiences (proxied by changes in test scores)—not just the choices themselves—of other likely "peer" schools, over a range of potential peer group boundaries defined by common governance or geography.

---

[4] Our language throughout refers to choice of a product, firm, or vendor, but the same logic would apply to the choice of a particular technology to be applied internally. Much of the social learning literature focuses on technology diffusion rather than choice of firm to supply an experience good.

We expect that administrators may gather and use information quite differently from households and firms, but the direction of the difference is ambiguous. Given the professional nature of their appointments, administrators may have skills which reduce the cost of gathering data from public sources; at the same time, agents have lesser incentives to invest in data for optimization than do principals. Furthermore, educational administrators may be more isolated than agents in the other settings of social learning in the empirical literature to date (Daly and Finnigan, 2009).

Empirical studies of social learning grapple with Manski's (1993) "reflection problem": agents may make similar choices as their peers not because they are learning from them, but because they have common characteristics or preferences unobservable to the researcher. Two institutional features of our empirical setting allow us to minimize such incorrect inference. First, all public schools in Texas participated in statewide standardized testing with publicly available school-level results throughout the time period in our sample, and second, the CSR grants were awarded in multiple waves. We therefore are able to identify the impact of a plan's perceived past *performance* on its subsequent adoption by peer schools, *controlling for its adoption by the initial school* and any correlated unobservables. Our identifying assumption is that common unobserved characteristics and preferences affect the choice of a plan, but not its correlation with test score changes once adopted (controlling for school characteristics). This strategy allows us to interpret our estimates as Manski's *endogenous effects*—the extent to which an individual's behavior varies with the *behavior* of the group—as opposed to the group's fixed characteristics (Manski's *exogenous* or *contextual effects*), or the characteristics of the individuals comprising the group (*correlated effects*).

Several important caveats apply. First, we do not observe actual patterns of communications across schools or districts. Our setting suggests several plausible

characteristics along which relevant peer groups may form: belonging to the same school district, being geographically proximate to one another (beyond school district boundaries), or belonging to the same regional professional development group (Education Service Center, or ESC). As discussed previously, we rely predominately on our use of the peer's perceived past *experience* as an independent variable to circumvent the reflection problem. A second caveat is that there is still potential for reflection if some unobserved shared characteristic between two schools gives them a strong match with a particular plan--it would make one school fare well with it and the other, regardless of its knowledge of the first school's experience, choose and thrive with that plan as well. We control for school-level demographics to minimize this possibility. Our third and most major caveat is that, despite choosing the largest state with appropriate data available, the large number of CSR vendors relative to the total number of grants limits our statistical power.

We find strong evidence that schools are much more likely to choose plans that were used within the same school district in the past, regardless of past performance. This effect could be due to a number of reasons, besides learning about the existence and the quality of the plan. For example, there might be fixed costs for either the school district or the provider (or both) of establishing a relationship with each other. We do not find that the prior presence of a plan in either a geographically neighboring district or within the same regional administrative unit has a statistically significant effect on a school's choice of plan. We also find no statistically significant effect of the *perceived quality* of the prior experience with a plan in any of the peer groups we consider. However, the magnitudes of our point estimates would be interpreted as revealing an economically meaningful role for social learning were they more precisely predicted.

Given the increased reliance on private contractors in public education for a number of services, we conclude with suggestions for future research to help isolate the relevant

mechanisms and inform related policies. The remainder of the paper is organized as follows.

Section II provides institutional background on the Comprehensive School Reform Program.

Section III describes the empirical strategy and Section IV the data sources and descriptive

statistics. Section V describes our results and a final section concludes.

II. Background on Educational Contracting and Comprehensive School Reform

School and district administrators face many contracting decisions. Burch (2006)

attributes much of the recent growth in such contracting to the policy environment created

by state-level accountability policies. Texas had already adopted such an accountability

regime by the period we study. Using 1997 data, Rowan (2002) finds over 80 percent of

school districts in the U.S. contracted with some outside for professional source for

"professional and technical services in a given year." Such purchases may come from general

revenue, or in some cases, from grants targeted to that particular expenditure category. In

most such cases, information on product quality is limited.

Our question of interest—do decision-makers learn and use relevant information

about the quality of private contractors from their peers?—requires data on purchases of

private contracting services by public schools, as well as information of the quality of the

services that were received. In this paper, we focus on a specific market for which both of

these types of data are available, the market for Comprehensive School Reform (CSR). CSR is

a bundle of school improvement products emphasizing professional development.[5] In

practice, a CSR plan could contain any or all of the following elements: curriculum, school

organization or management strategies, assessments, teaching aids such as software or

workbooks, and teacher training to go with the entire system. A joint public-private effort

early in the 1990s focused on developing effective, scalable CSR plans; a federal effort to

---

[5] Formally, the Department of Education (2007) states that "[t]he Comprehensive School Reform
program is designed to foster coherent schoolwide improvements that cover virtually all aspects of a
school's operations, rather than piecemeal, fragmented approaches to reform."

expand demand for CSR began with the Comprehensive School Reform Demonstration grants

in 1997, and became the Comprehensive School Reform grants program in 1998.[6] Under

these programs, public schools could apply for competitive federal grants to purchase CSR

plans from private CSR providers or for school districts to develop programs locally. Schools,

districts, and states purchased CSR from a variety of funding sources, including but not

limited to the federal grants program. There is no publicly-available comprehensive

accounting of which schools use CSR or of how many schools use particular CSR plans

(providers do not wish to share their client lists and public information requests are difficult

for contracts older than one or two years).[7] There is, however, a complete database of all

schools purchasing CSR plans with federal funds from 1997 to 2005, so we restrict our

attention to these plans.

Federal funds were awarded to State Education Agencies, which distributed the funds

through competitive grants to school districts that applied on behalf of specific schools within

their districts. Applications had to specify the CSR provider that the school intended to use,

and grants were awarded for use with this provider only. Research from case studies

(Datnow, 2000), and our own interviews with administrators suggest that the choice of CSR

plan – regardless of funding source – in most districts was ultimately made at the school

rather than district level, with varying levels of guidance or pressure from the district.[8] Data

on unsuccessful applications are not available but by all accounts, it appears that grants were

not systematically awarded in any way based on the provider specified in the application.[9]

---

[6] New American Schools (NAS), a joint public-private venture beginning in 1991, supported the development of the supply side of the CSR market through support and evaluation of nine specific models of school reform (see Berends *et al.* (2001) for more on NAS).

[7] Datnow (2000) reports over 6,500 schools in 45 states implementing CSR plans (with any funding source) at the time of her study.

[8] Because we study schools purchasing CSR with federal grants, districts should have less leverage over school choices. Our data reveal significant variation in plan choice within many districts (see Appendix Table 1).

[9] The state education agency staff member in charge of the program reports that no particular models of school reform were given preference in the selection process (based on scientific evidence,

Schools could receive CSR grants only once, and grants were given for one year with the possibility of being extended to up to three years with the same provider (in practice, this was almost always the case).

As Table 1 shows, the schools ultimately awarded CSR grants were disproportionately disadvantaged and low achieving, but still were a highly varied group. The program was last funded in FY2005; in most years, about $200-300 million was appropriated for the entire program nationwide, with minimum grant amounts of $50,000 per school.[10] More recently, the "transformation" model option through School Improvement Grants of Title I of the No Child Left Behind Act has created markets for similar bundles of goods (with the additional and much more drastic requirement of replacing the principal). Multiple firms formerly touting their services as CSR are now officially in the market for school improvement.

The award criteria for CSR grants were set by the state. We therefore restrict our study to a single state in order to be able to compare the choice of CSR providers across schools. Recall from our discussion at the beginning of this section that we also require information on the effectiveness of each CSR plan. We choose Texas, the largest state which conducted statewide standardized testing with publicly available school-level results during the entire time period, as the setting for our study. Grants in Texas were awarded in three distinct waves from 2000 through 2005.

---

perceived efficacy within the state, or other factors). While federal requirements post-No Child Left Behind mandated that funds be used for models with positive evidence from "scientifically-based research" the dearth of such research on these plans makes the requirement irrelevant in practice. For example, in the 2004 request for applications for the CSR – Texas High Schools Initiative Program, the application specifies that two out of 100 points are to be awarded based on the extent to which "The design of the proposed project reflects up-to-date knowledge from scientifically-based research and effective practice." In practice it is difficult to know how even these two points could be awarded given the lack of relevant scientifically-based research on nearly all plans. Overall 15 of 100 points that year were awarded based on the larger category of "quality of project services," further supporting anecdotal evidence that applications were chosen primarily based on the *school's* level of need and capacity to implement CSR plans from any given provider, rather than due to the particular provider selected.

[10] The median grant was $70,000 per school for each one of the three years. Funds were appropriated for data collection through the CSR Clearinghouse in FYs 2007 and 2008, but not for new program implementation in those years.

Overall the characteristics of the CSR market are similar to Rowan's (2002) description of the broader school improvement industry: there are many firms offering highly heterogeneous products, low concentration, little regulation and almost no objective information on firm quality. The CSR awards database provides information on award amounts, revealing significant heaping around round numbers (most strongly at $50,000, the mandated minimum) and a much wider range in per pupil grant amounts than in total grant amounts. Because we lack data on what specific goods and services are included within each contract between a school and a provider—and what objective level of "quality" is provided—we cannot meaningfully interpret the data on award amount.[11] Nationwide, 755 different plans provided federally financed CSR services to schools over the period from 1998-2005. 443 of these plans (almost 60 percent) only served a single school. Nationally, the largest CSR plan was Success For All with a market share of 6.9 percent, the second-largest plan was Lightspan with a market share of 4.5 percent, and the third largest was Accelerated Schools with a market share of 3.8 percent. For the state of Texas, where we limit our attention in this study, there were a total of 124 different plans, and the distribution of market shares was similarly skewed as at the national level. 74 of the Texas plans (again, almost 60 percent of all plans) provided CSR to only a single Texas school, 18 plans (14.5 percent of all plans) provided CSR to two Texas schools and 6 plans (4.8 percent of all plans) provided CSR to three Texas schools.

In Table 2, we present the market shares in Texas for the 16 largest plans. Most of these plans (save the largest two, AVID and Accelerated Schools) have relatively small market shares in Texas, but nearly all operate at a significant scale nationally. The larger plans are used in similar numbers across the three grant waves, whereas some of the smaller plans are

---

[11] Mean award amounts per school and per pupil by CSR provider do not align with the potential quality proxies we discuss in Table 2.

skewed towards earlier or later waves.[12] Some of the plans also specialize in certain grade

levels. We account for this in our empirical analysis by only including plans in a school's

choice set that are used at least three times nationally at the school's grade level (in any

wave).

In Figures 1 and 2, we choose Co-nect and Renaissance Learning, respectively, as

examples to map the frequency with which federal CSR grants were used to fund adoption of

these plans across all states (these two plans were picked at random among the smaller plans

in our sample). The figures show that while the plans were not very large in our sample, they

occurred with sufficient frequency and geographic diversity to suggest that they were likely

not to be supply-constrained. While we are not able to contact representatives of all plans,

personal communications with plan staff support this interpretation for AVID (Ellis, 2011).

In practice, federal dollars have funded quite heterogeneous CSR programs, ranging

from highly specific interventions to those in which it is difficult to precisely define the CSR

treatment.[13] Such great variety in CSR program design naturally leads to the question of the

heterogeneity in efficacy of these programs. This is of primary interest to us, because we

investigate the extent to which demand for CSR depends on perceived plan quality. There is

little existing large-scale quantitative evidence on this question, and less still using

experimental or quasi-experimental variation (see Borman *et al.* (2003) for a review and

meta-analysis of CSR research; see Mason (2005) for a detailed study of five programs in Los

---

[12] Most plans are used across all waves at the national level.

[13] At one end of the CSR spectrum are programs like the Coalition for Essential Schools: on its website, it describes "school reform as an inescapably local phenomenon" and acknowledges that "no two Essential schools are alike." It instead emphasizes the shared commitment to a set of principles that are both relatively uncontroversial and vague, such as personalized instruction and an atmosphere of trust and high expectations. In contrast, Success For All (SFA) is an example of a tightly scripted program, with teachers adhering to a prescribed curriculum and methods, rhyming mnemonics included. In order to create this uniformity, SFA must provide many more curricular, assessment, and training materials than would a plan emphasizing locally specific needs.

Angeles Unified School district, with particular attention to implementation).[14]

Correspondingly, and relevant to this study, there has been little technical guidance offered to schools facing the choice of CSR provider.[15] The setting we study in this paper is one in which schools actively chose CSR providers; we therefore cannot identify the true effect of the provider on achievement as distinct from the administrator(s) who chose it. Throughout the paper, our emphasis is on CSR plan quality *as perceived* by school administrators, as opposed to a true causal measure of such quality.

The last two columns of Table 2 present two measures of potentially perceived plan quality. One of these measures is a rating produced by the American Institutes of Research (1999) meant to be useful to administrators in precisely this situation. The AIR study considered the largest plans nationally; as Table 2 shows, many of the largest plans in Texas were not included. The AIR study reviewed the evidence plan by plan; in most cases it concluded not that plans were effective or not, but rather that they lacked strong research bases. The other measure is our own estimate of the correlation between the plan--as endogenously adopted--and achievement gains (essentially a flawed value-added measure).[16] Based on both potential proxies for quality, the "best-performing" plans clearly do not have the largest market share. However, this is not necessarily indicative of an information failure in the market since it is likely that neither of these measures captures the plans' true quality.

---

[14] Mason finds that, of the five programs in his study, none led to uniform improvements in achievement. He and others (Vernez et al., 2006) also found that very few schools fully implemented their CSR plans; this paper, like the bulk of the literature on CSR, studies CSR as implemented in practice rather than the ideal plans as initially developed by New American Schools and related research efforts.

[15] Borman et al. (p. 130) discuss several "practitioner-oriented reviews, or 'catalogs.'" From a school's perspective, we anticipate that they are limited in two key ways: first, they review only the largest CSR plans, which still provides no information on the many small providers serving the majority of the market, and second, the ratings typically reflect the quality of the research more than the quality of the plan (i.e., a low rating meant that the plan had not been evaluated in a causally interpretable way). For a typical such review, see American Institutes of Research, 1999.

[16] We discuss in detail how we arrive at this estimate in section III below.

In summary, the institutional background suggests that administrators would choose CSR plans similarly to how they make other educational contracting decisions. There was relatively little "scientific" research evidence upon which to base these decisions. Existing research in other institutional settings suggests that even when such evidence is available, the transaction costs of obtaining it—what Kling et al. (2011) term "comparison frictions"— serve as a significant deterrent, even when such costs are low. Social networks could thus provide a rare source of information to administrators faced with contracting decisions. The limited research available on such networks, however, suggests relatively little communications across schools, other than those facilitated by district staff (Daly and Finnigan, 2009); no such research directly investigates the extent of networks across districts, but given their limited nature within districts, we expect that the additional costs of building and maintaining networks across district lines would result in little to no communication across districts unless facilitated by state or regional administrators.

III. Empirical Approach

Our empirical approach is built on the assumption that school administrators have the objective to improve their students' achievement. This measure is consistent with an objective function formed by altruism as well as with one influenced by incentives imposed by accountability regimes such as the one relevant in our sample. We assume that a given CSR plan has an average (however noisy) effect on achievement, as well as a school-specific match component to its efficacy. We expect schools to form expectations of CSR plan quality from communications with their peers and, following the literature on social learning (e.g. Ellison and Fudenberg, 1993), anticipate that schools may rely on "rules of thumb" to limit the quantity of information to be processed because the expected benefit of the information may exceed the cost of obtaining it. Data on school choice of CSR plan and on school achievement by year were both publicly available over this time period; however, the data on

choice of plan were not publicized, and we expect the transaction costs of obtaining relevant

data from the two sources and merging it to be sufficiently high as to rule out this possibility

for nearly all administrators. While we cannot rule out specific instances of such

sophisticated research, no administrators questioned reported such behavior.

Our empirical model, the conditional logit, allows plan choice to be determined by

both school and plan attributes. The model controls for school-level fixed effects. We also

allow key school characteristics—the fraction of its students who were white, and its average

pass rate across math and reading, both measured at baseline—to affect the likelihood of

choosing each CSR plan differentially. Our independent variables of interest are measures of

school-level potential exposure to information about individual plans' observed correlations

with achievement changes in other "peer" schools.[17] Depending on our specification, we

allow exposure to information to come from various sets of peer schools, starting with other

schools in the same district, then schools in geographically neighboring districts and schools

in common administrative regions.

We estimate school plan choice using the conditional logit specification in equation

[1] below. We denote schools with subscript i and plans with j.

[1] $$Prob(Y_i = j) = \frac{e^{\beta' x_{ij}}}{\sum_{\forall j} e^{\beta' x_{ij}}}$$

Because we have a large number of plans that are chosen by very few schools, we do not

include all of them as separate choices in the conditional logit model that we estimate.

Instead, we define a category that we call "other" for plans that are infrequently chosen. The

choice set of plans for a given school thus consists of a number of individual large plans and

"other." In constructing our definition of a large plan, our goal is to identify plans which (1)

are not supply-constrained and therefore likely constitute a viable option for a school with a

---

[17] We treat the prior experience with a plan as a plan attribute that varies across schools in the
conditional logit model.

new grant,[18] and (2) have had the opportunity to generate results which new grantees may observe and use to inform their plan selection. For our main specifications, we define large plans as plans that have been used at least six times in Texas, in at least two different counties and in at least two of the three waves. Table 2 shows that 41 percent of all schools in our sample chose plans in the "other" category. Our empirical results are robust to an alternative definition of large plans that require the plan to have been used at least ten times.

Approximating plan quality as perceived by administrators poses a significant challenge. The ideal measure of *true* plan quality would capture its causal effect on school achievement, again assuming that administrators care about pass rates on the state test during a strong accountability regime; no such measure is available because very few plans have been subject to high quality evaluations and all data available on plan choice and achievement also reflects the endogenous choice of plan by the school. There have been attempts to provide administrators with summaries of available research on specific plans, but these syntheses offer little guidance other than reiterating the lack of evidence (see Borman et al., 2003 for details). Research attempting to link plan attributes to outcomes over a sample of large plans nationally concludes that a plan is not the sum of its parts (again, see Borman, 2003 for a review). We do not attempt such analysis due to this finding, combined with the vague descriptions of CSR plans and subsequent difficulty in even determining what those hedonic values should be. Given these limitations, we pursue two separate strategies to estimate *perceived* plan quality. Both strategies assume that administrators base inference on the correlation between (endogenous) plan choice and student outcomes. It is possible that some administrators are aware, intuitively if not in a formal statistical sense, that this method would not identify plan efficacy if the most effective administrators are likely to have selected the most effective plans. Lang (2010) offers dramatic anecdotal evidence suggesting

---

[18] All of these "large" providers operate in multiple states, and each of them operates across multiple non-contiguous Texas counties.

that such intuitive understanding is in short supply in this sector; we have not seen a comprehensive investigation of this issue in the literature.

In our first approach, we calculate the raw change in achievement (pass rates) before and after the CSR intervention in the peer school(s). We view this as a plausible input into a rule of thumb that administrators could use.[19] For example, when we consider other schools in the same district to be the relevant peer group, we calculate the mean pass rate change over all schools in the district which have used the plan in the past. The plan-specific attribute used in the conditional logit estimation is then the interaction of this mean pass rate change with an indicator variable for whether the plan has been used in peer schools in the past. This measure would be appropriate if the school administrator either learned directly from her peers about achievement levels before and after the plan was implemented, or if she looked up pass rates online for the relevant peer school(s). While looking up the scores online poses what Kling et al. (2011) term "comparison frictions," the still more difficult piece of information to access, which we assume comes only from social interactions, is the mapping of school to CSR plan (also available online but not via the state's website, as discussed in the following section).

In the second approach, we compute an estimate of perceived CSR plan quality based on more observations (all schools who have used the plan statewide rather than only among peer schools) and adjusted for school characteristics. Specifically, we estimate equation [2] below:

[2] $$PASS\ RATE_{ijt} = \alpha_i + X_{it}\beta + \gamma_t + \delta_j + \varepsilon_{ijt}$$

The pass rate at school i using CSR plan (equivalent to CSR model or firm) j in period t is predicted by time-varying school characteristics X_{it} as well as school-, plan- and year-specific fixed effects. The coefficients for these plan fixed effects, δ_j, are reported in Appendix Table 2

---

[19] We have explored other potential measures, such as the levels of pass rates, and standardized changes in pass rates, with qualitatively similar results.

for several specifications varying in the control variables included at the school level. We describe the data used to generate these perceived value-added measures in greater detail in the following section. Notably, most plans (including the two largest plans) have statistically significant and positive coefficients; however, as Table 2 shows, the ranking of plans by these coefficients does not align closely with the market share of the plans. This lack of correlation in the aggregate is consistent with the average administrator lacking the technical capacity to gather data from the two sources and estimate such regressions, but also with the average administrator doubting the relevance of endogenously-identified estimates. Again, in constructing the plan-specific attributes for the conditional logit estimation, we interact these plan fixed effects with prior use of the plan by a peer, to approximate a setting in which administrators only learn of these estimated effects if the plan has been used by a peer in the past. This approach would identify social learning if (1) the plan fixed effects from our statewide regression correspond to plan quality as estimated from administrators who have used the plan in their schools—a question the literature to date does not allow us to answer—and (2) administrators share their perceptions of plan quality with their peers, as defined by our various measures.

IV. Data Sources and Descriptive Statistics

We focus our analysis on a single state, Texas, for several reasons. Most importantly, because plans align themselves to some extent with state educational standards, there are significant cross-state differences in plan popularity. Furthermore, because Texas began testing students and collecting achievement data relatively early, before the start of the federal CSR program, we can establish a school's initial achievement level before the intervention. Achievement data and accountability ratings for each school in Texas for each year were public information, as was the allocation of CSR providers to schools; more than 60 percent of the Texas schools ever awarded federal grants could observe at least one year of

test score data from other schools implementing CSR with earlier grants.[20] Finally, Texas

schools were under high-stakes accountability regime at the time of the federal CSR grants.

Administrators should be most willing to bear the costs of establishing or using social

networks, or of collecting data via other mechanisms, under such a regime.

Our identification strategy requires data on school choice of CSR provider, initial

achievement levels and changes in achievement concurrent with CSR intervention,

demographics, and geographic and administrative characteristics affecting the likely

formation of peer groups. We obtain data on federal CSR grants from the Comprehensive

School Reform Awards Database. This database was collected and maintained by the

Southwest Educational Development Laboratory (SEDL), a largely federally funded not-for-

profit research firm. SEDL collected these data from state CSR administrators and

maintained an online database.[21] The federal Department of Education later took over

responsibility for this database, and then (after the program ended and no schools were

receiving new grants) removed it from the Internet.[22] The CSR Awards Database is organized

at the school level of observation; the grants were three years in duration, and no school

received more than one grant during the seven year span of the program, so this is essentially

a school-grant level of observation. For each school, we use data on the "models used"

(equivalent to the CSR firm with whom the school contracted, and in most cases, only one

model) and the first year of the grant. Each school is identified with its unique twelve-digit

National Center for Education Statistics (NCES) identification number, allowing us to link

---

[20] The first year of implementation (for first-wave schools) was the 1999-2000 school year, and accountability ratings for that year were released on August 17, 2000. The deadline for applications for the second wave of grants was December 15, 2000, and the corresponding first year of implementation was 2001-2002. Accountability ratings from the first year of implementation for second-wave schools were released on August 1, 2002, and the applications for the final wave of grants were due March 15, 2004 and October 7, 2004 (the later deadline was for high schools).
[21] For a small minority of the schools in this database, the name of the CSR provider was missing. We obtained this information through a Public Information Request from the Texas Education Agency.
[22] See http://csrprogram.ed.gov/ for details on accessing the offline data. All data in the paper are public and available from the authors upon request.

these data with the NCES Common Core of Data. From the Common Core, we use data on the racial and ethnic composition of each school.

Our final data sources are from state agencies. The Texas Education Agency's Academic Excellence Indicator System contains annual information on each school in Texas, including the percentage of participating students at each school who passed the statewide achievement test in that year ("pass rate") [23] and the percentage of students at the school who are exempt from the test.[24] Pass rates are reported separately for math and reading. We use the average of those two pass rates averaged over all schools meeting a particular peer group definition as our "naïve" measure of school performance; we also use this averaged pass rate as the dependent variable in computing a plan's perceived value-added.

To more fully explore the mechanism for dissemination of information, we have gathered data on the composition of Education Service Centers (ESCs) within Texas. These centers are essentially a regional layer between the state and local education agencies, and primarily provide professional development and technical assistance. There are 20 centers statewide, with each center serving multiple (contiguous) counties. They therefore constitute networks of administrators that are both plausible (teachers and administrators may attend professional development programs physically provided at the ESC) and observable.

In constructing our regression sample, we start with all schools in Texas which ever received a federal CSR grant. We drop any schools that are special education, vocational, or alternative. We also drop schools which are missing in one of our three data sets. Next, we drop all schools in four districts in El Paso because El Paso had a local CSR initiative which schools in these districts were very likely to choose. This yields a total of 433 Texas schools

---

[23] Pass rates describe only one part of the distribution, and administrators may well care about other points in the distribution more (for example, if the parents of higher achieving students are more politically vocal). The pass rate is the most consistently reported measure in the data, however.
[24] Cullen and Reback (2006) provide evidence suggesting that schools use exemptions strategically to improve their accountability ratings.

that receive CSR grants. Some of these schools use their CSR grants for more than one provider, and in these cases we include one observation for each school-grant combination. The total number of school-grant combinations in our sample is 497. Of these, 136 are in the first wave (1999-2000 academic year), 201 in the second wave (2001-2002 academic year) and 160 in the last wave (2004-2005 academic year). We include schools in the first wave in the sample because their choices can help identify the effect of school demographics on plan choice.

Our model includes explanatory variables that are specific to the school making the choice and explanatory variables that are specific to the plan that is chosen. The former group consists of the percentage of white students at the schools and the initial pass rate of the school. These and several other demographic characteristics of the schools in our sample are described in Table 1, Panel A. The average school was 29 percent white; its pass rate on the statewide test in spring 1998 was 78.5 percent.

The values of our plan-specific explanatory variables vary by school, conditional on plan, and relate to the different concepts of peer or neighboring school. We include these variables to reflect both the transmission of information (if "peer" administrators share access to qualitative or quantitative data) and the issue of "fit" (even if administrators have complete information, they may weight it more heavily if generated in a context similar to their own) but do not have any way of disentangling these two effects. In this vein, we have indicator variables for whether the plan was used in the past by another school in the same district (Column 1 of Table 4 shows this was the case for 4 percent of plans chosen after the first wave), by a school in a geographically neighboring district, defined as one with its centroid within a 50-mile radius of the centroid of the school's own district (true for 30

percent of plans chosen after the first wave), or by a school in a common ESC (true for 23 percent of plans chosen after the first wave).[25]

Table 3 compares measures of perceived quality available to schools with new grants at the time they chose their providers, depending upon their peer groups. Among plans previously used by schools within a choosing school's district or ESC, those plans *not* chosen were associated with slightly higher achievement gains on average, without adjusting for any covariates. For plans used by schools in geographically neighboring district, the difference in means takes the opposite sign. However, the standard deviations on these variables are quite large and the differences are not statistically significant.

V. Results

We estimate a school's choice of plan with the conditional logit model described in equation [1]; the coefficients of interest are the plan-specific attributes which depend on the past use of the specific plan among schools in the relevant peer group. All specifications also include controls for the school's initial pass rate and its percentage of white students, interacted with dummies for the plans in the choice set, but we do not report these coefficients due to space constraints. Table 4 presents the coefficient estimates and standard errors from our regressions. Panel A shows results using the change in raw pass rates as the measure of perceived quality; Panel B presents the same specifications using the estimated value-added from equation [2] as a measure of perceived quality instead.[26] We find nearly identical results in both panels.

Plans that were used within the same school district in the past are significantly more likely to be chosen by other schools in the same district. Prior experience within the own district could affect future plan choice for a number of reasons, besides learning about the

---

[25] Our results are robust to defining "geographically neighboring districts" instead as districts within a 100-mile radius.

[26] The change in raw pass rates and the value added measures enter the regression model demeaned.

existence and the quality of the plan. CSR providers may face fixed costs of establishing a presence within a given district and therefore concentrate their marketing efforts to schools within a subset of districts. District administrators may face fixed costs in dealing with each individual CSR supplier and pressure school administrators to choose correspondingly. Large districts may be able to exert market power in negotiating with providers if several schools within the district choose the same plan; in this case, the CSR provider may offer enhanced services to schools in the district such that schools would be more likely to choose that provider even without pressure from the district. We are not able to test these explanations, but any of them would favor multiple schools within the same district contracting with the same plan.

To get a sense for the magnitude of the effect of past plan use within the district, we compute the predicted choice probabilities for the largest two plans, AVID and Accelerated Schools, for the case where the district had no prior experience with the plan and compare it to the case where the district had prior experience with the plan and the quality of the experience was equal to the mean, and find quite large changes in predicted probabilities. We find for specification (B1), for example, that the choice probability for AVID would increase from 10.7 percent to 25.2 percent and the choice probability for Accelerated Schools would increase from 6.5 percent to 16.4 percent. While we expect that some past experiences are positive, we also expect others to be negative, so we do not interpret the correlation between past use of a plan within a district and increased likelihood of future adoption as necessarily indicative of evidence-based decision making.[27]

When we look at the effect of the *perceived quality* of prior experience with a plan in the same district, we find that the estimated coefficients on either of the perceived quality

---

[27] This is a direct result of each school receiving only one opportunity to choose a vendor with the grant, as opposed to a consumer-producer match observed in equilibrium, such as the employee-health plan match in Sorenson (2006),

measures we use are statistically insignificant across all specifications. Ai and Norton (2003) point out that the sign and statistical significance of the marginal effect of such interaction terms vary with the values of covariates and cannot be inferred from the model coefficient. We therefore simulate the change in choice probabilities that our model would predict for a one percentage point increase in the perceived quality of a plan (i.e., change in pass rate or observationally-determined value added), conditional on the school having any prior experience within the relevant peer group. Table 5 shows the results of these simulations, which generate large confidence intervals. None of the changes in choice probabilities that we predict are statistically different from zero, so we cannot rule out that there is no effect of improving the perceived pass rate or value added on the likelihood that a school will choose a given plan. We find that a typical increase in choice probability would be 0.43 percentage points in the case where we use the change in pass rates as the measure of perceived quality and 0.68 percentage points when we use our estimated value added measure. These effects would be interpreted as economically meaningful were they more precisely predicted, given the range of market shares across plans, of changes in pass rates (see Table 3) and of value added (see Table 2).

In specifications (A2) and (B2), we investigate the effect of the past presence of a plan and the perceived information about the plan in a geographically neighboring district, while still controlling for past experience in the school's own district. We find a positive point estimate on the dummy which captures prior presence of the plan in a neighboring district and on the perceived quality of the plan, but the coefficients are very imprecisely estimated and we cannot rule out that there is no effect at all. This is confirmed by our simulations in Table 5, which predict that choice probabilities would increase when the perceived effectiveness of a plan increases, but none of the increases in choice probabilities are statistically different from zero.

We find very similar results in specifications (A3) and (B3), where we consider the own district and the Education Service Centers as relevant peer groups, and in specifications (A4) and (B4), where we control for information that might come either from the school's own district, its geographic neighbors, or the Education Service Centers. In all of these specifications, the only statistically significant effect comes from the dummy for past presence of the plan in the school's own district.

In results available upon request, we have performed a number of robustness checks to confirm that our results are not sensitive to the way we treat small plans in the "other" category. We have varied the cutoff for the minimum number of times that a plan must have been chosen within Texas to be considered "large" from 6 to 10 (using 6 in our preferred specifications presented). We also have varied the measures of perceived plan quality for "other" plans. In the regressions we report, the plan quality for any plan in the "other" category is equal to the peer's past experience with plans in the "other" category in the past. Alternatively, we have assigned the means of past experiences across *all* plans to any plan in the "other" category, and we find that our results remain qualitatively the same. We have also investigated whether schools might learn from other schools that share similar demographic characteristics (such as having a high percentage of students with limited English proficiency, having a low initial pass rate, and being in a particularly large school district) and have not found any evidence of social learning from these alternative potential peer groups.

VI. Discussion

The question of how school administrators use data to inform decisions is more salient now than ever due to an unprecedented increase in federal funds potentially available to education contractors, most notably due to the use of American Recovery and Reinvestment Act funds to boost support for two key existing programs exponentially, adding

23

$4.35 billion to Race to the Top funds and $3 billion to Title I School Improvement Grants (SIGs) in FY2009. Rudy Crew, the former New York City schools chancellor and now a contractor paid from SIG funds, characterized the supply response to SIGs as "like the aftermath of the Civil War, with all the carpetbaggers and charlatans," (Dillon 2010). Unsurprisingly, some firms previously receiving CSR grants are now also marketing themselves as school turnaround and transformation (specific options available to schools with SIGs) specialists.[28]

More broadly, our question fits into a topical literature about the extent to which policymakers can help agents—households, firms, or employees at public agencies—optimize their decisions through dissemination of information versus more direct limitations on choice, including regulation of choice sets and setting defaults. Recent studies have shown that consumer demand does respond to highly publicized and easily digestible information— basically, front page news (see Freedman, Kearney and Lederman, forthcoming, on toy recalls and Simonsohn, 2011, on *Consumer Reports* warnings on child safety seats). The health literature has identified consumer responses to perceived provider quality in a variety of settings, even when perceived quality may be falsely informed by considering patient outcomes in the absence of appropriate adjustments for patient characteristics (for one recent example, see Howard, 2005, on patient choice of kidney transplant facility). Recent experimental work examining the use of information that consumers must actively gather themselves, either from online sources or calling public agencies, concludes that such costs pose significant deterrents (Kling et al., 2011; Hastings and Weinstein, 2008).

How government can most help agents make sound decisions, whether about technologies to be implemented from within or goods or services to be purchased from private vendors, depends on how administrators access and use information. Obtaining

---

[28] See, for example, Success for All <http://www.successforall.net/Turnaround/turnaround.html>.

access to data on such specific purchasing or policy choices at the school level is quite challenging, and unfortunately we lack sufficient statistical power here to identify clear patterns. We thus conclude by briefly discussing research strategies to help inform policy.

In order to understand how information flows, one must first identify the information itself, in this case by better understanding the educational production function. Countless studies and millions of grant dollars per year are devoted to this broad area, yet most narrowly defined interventions schools might consider still lack a research consensus based on work with large, representative samples with exogenous variation in treatment status. It would be particularly useful to consider research models which test multiple treatments against one another, as well as a control group—and research funding models oriented towards promoting this approach.

Once researchers have identified actual best practices, one could follow the approaches taken by Hastings and Weinstein (2008) or Kling et al. (2011) and randomly assign some school (administrators) to a treatment group in which they receive highly accessible, low cost information related to an upcoming policy choice, while a control group receives guidance on pursuing a higher cost route to information. If the group with more processed information responds to it by making more evidence-based choices, this could suggest a cost-effective policy. If school administrators appear unresponsive to such data, policymakers might consider more active policies such as setting defaults or regulating choice sets.

References

American Institutes for Research. 1999. *An Educators' Guide to Schoolwide Reform.*
        Washington, DC: Educational Research Service.

Berends, Mark, Sheila N. Kirby, Scott Naftel, and Christopher McKelvey. 2001.
        Implementation and Performance in New American Schools: Three Years into Scale-up.
        Washington, DC: RAND.

Borman, Geoffrey D., Gina M. Hewes, Laura T. Overman, and Shelly Brown. 2003.
        "Comprehensive School Reform and Achievement: A Meta-analysis," *Review of
        Educational Research* 73(2), 125-230.

Burch, Patricia E. 2006. "The New Educational Privatization: Educational Contracting and
        High Stakes Accountability," *Teachers College Record* 108 (12), 2582–2610.

Conley, Timothy G. and Christopher R. Udry. 2010. "Learning about a New Technology:
        Pineapple in Ghana," *American Economic Review* 100(1), 35-69*.*

Cullen, Julie Berry and Randall Reback. 2006. "Tinkering toward Accolades: School Gaming
        under a Performance Accountability System," in T. Gronberg and D. Jansen, eds.
        *Improving School Accountability: Check-Ups or Choice, Advances in Applied
        Microeconomics* 14 (Amsterdam: Elsevier Science), 1-34.

Datnow, Amanda. 2000. "Power and Politics in the Adoption of School Reform Models,"
        Educational Evaluation and Policy Analysis 22(4), 357-374.

Department of Education. 2007. "About CSR."
        http://www.ed.gov/programs/compreform/2pager.html, accessed November 29,
        2007.

Dillon, Sam. 2010. "Inexperienced Companies Chase School Funds," *New York Times*. August
        10, 2010, A11.

Dillon, Sam. 2010. "U.S. Awards Millions for School Innovations," *New York Times*. August 6,
        2010.

Ellis, Rosemary. 2011. Personal communication (e-mail) with the authors.

Ellison, Glenn and Drew Fudenberg. 1993. "Rules of Thumb for Social Learning," *Journal of
        Political Economy* 101(4), 612-643.

Freedman, Seth, Melissa Kearney, and Mara Lederman. **"**Product Recalls, Imperfect
        Information, and Spillover Effects: Lessons from the Consumer Response to the 2007
        Toy Recalls," *Review of Economics and Statistics*, forthcoming.

Hastings, Justine S. and Jeffrey M. Weinstein. 2008. "Information, School Choice, and
        Academic Achievement: Evidence from Two Experiments," *Quarterly Journal of
        Economics* 123(4), 1373-1414.

Howard, David H. 2005. "Quality and Consumer Choice in Healthcare: Evidence from Kidney Transplantation," *B. E. Journal of Economic Analysis and Policy* 5(1), Article 24.

Kling, Jeffrey R., Sendhil Mullainathan, Eldar Shafir, Lee Vermeulen, and Marian V. Wrobel. "Comparison Friction: Experimental Evidence from Medicare Drug Plans," *Quarterly Journal of Economics*, forthcoming.

Lang, Kevin. 2010. "Measurement Matters: Perspectives on Education Policy from an Economist and School Board Member," *Journal of Economic Perspectives* 24, 167-182.

Manski, Charles F. 2000. "Economics of Social Interactions," *Journal of Economic Perspectives* 14(3), 115-136.

Mason, Bryce. 2005. "Achievement Effects of Five Comprehensive School Reform Designs Implemented in Los Angeles Unified School District," *Pardee RAND Graduate School dissertation*.

Munshi, Kaivan. 2004. "Social Learning in a Heterogeneous Population," *Journal of Development Economics* 73, 185-213.

Rowan, Brian. 2002. "The Ecology of School Improvement: Notes on the School Improvement Industry in the United States," *Journal of Educational Change* 3: 283-314.

Simonsohn, Uri. 2011. "Lessons from an Oops at Consumer Reports: Consumer Follow Experts; Ignore Invalid Information," *Journal of Marketing Research* 48(1): 1-12.

Sorenson, Alan T. 2006. "Social Learning and Health Plan Choice," *RAND Journal of Economics* 37 (4), 929-945.

Vernez, Georges, Rita Karam, Louis T. Mariano, and Christine DeMartini. 2006. "Evaluating Comprehensive School Reform models at Scale: Focus on Implementation," RAND Corporation Monograph.

Figure 1

Number of schools with federal CSR grants using Co-Nect, 1997-2005
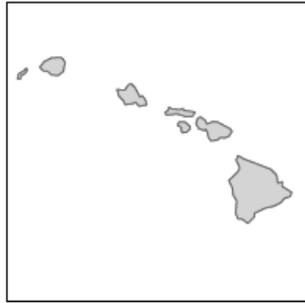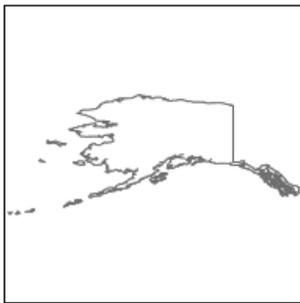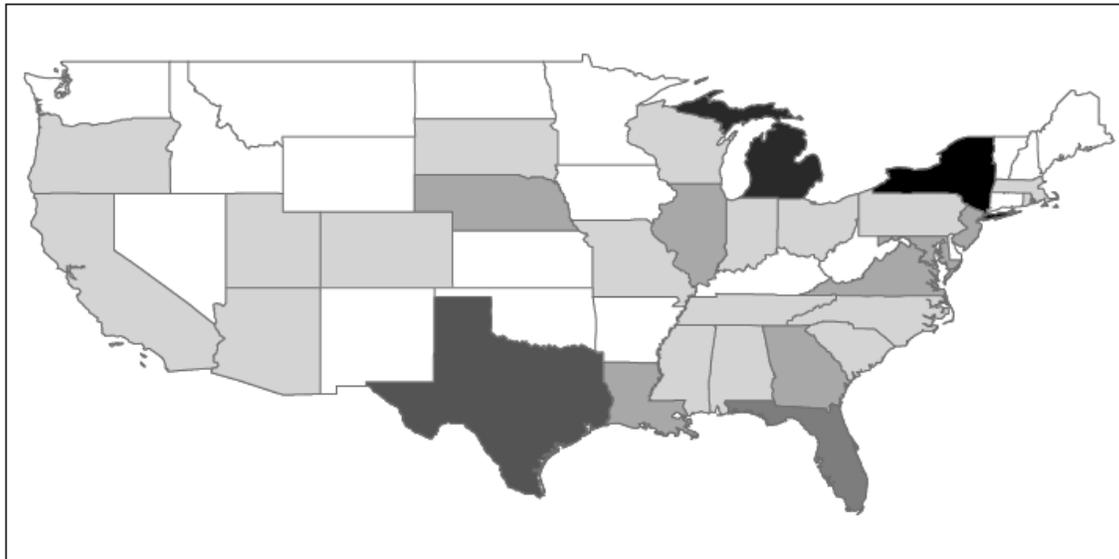
Figure 2

Number of schools with federal CSR grants using Renaissance Learning, 1997-2005

Number of Schools Using Renaissance

- 0
- 1 - 2
- 3 - 4
- 5 - 6
- 7 - 8
- 9 - 10
- 11+
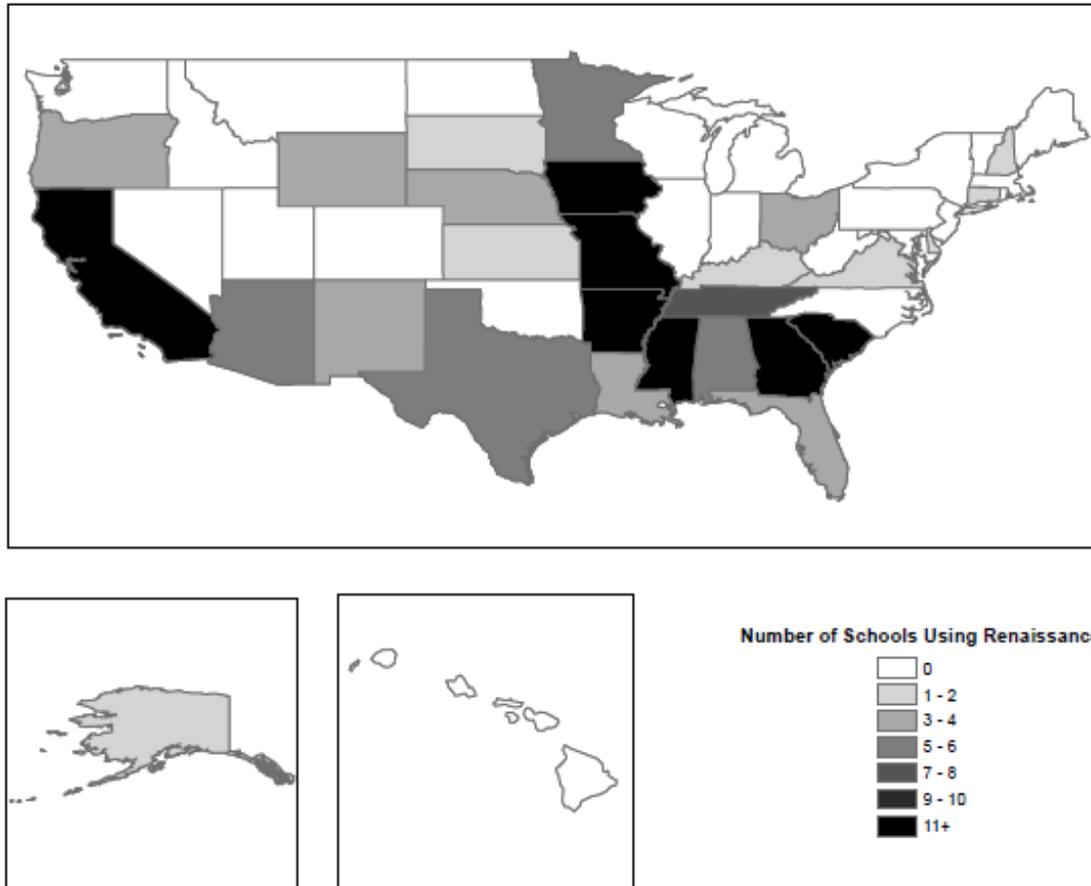
Table 1: Comparing CSR and Non-CSR Elementary Schools in Texas

Panel A: Means and Standard Deviations of Selected Variables (Spring 1998)

Variable	CSR schools	Non-CSR schools
Percent White	29.4	51.4
	(28.4)	(31.4)
Percent Hispanic	48.8	33.3
	(34.0)	(30.5)
Percent Black	20.0	13.2
	(24.4)	(19.3)
Percent Asian	1.6	1.8
	(3.0)	(3.7)
Percent Limited English Proficiency	18.5	11.3
	(20.0)	(16.8)
Percent Special Education	12.5	12.9
	(4.7)	(4.9)
School enrollment (in thousands)	0.867	0.599
	(0.633)	(0.439)
Number of schools in the district	51.2	35.5
	(72.5)	(60.7)
Observations	372	5565

Notes: Table shows variable means. Standard deviations in parentheses.
Sources: CSR Awards Database, NCES Common Core

Panel B: Distribution of Pass Rates (Spring 1998)

Percentile of pass rate	CSR-schools	Non-CSR schools
5	61.2	68.3
25	71.8	81.8
50	79.6	88.5
75	86.3	93.3
95	94.0	97.6
Mean	78.5	86.4

Source: CSR Awards Database, Texas Education Agency Academic Excellence Indicator System.

Table 2: CSR Plans in Texas and Nationally

Panel A: Plan Choice in Texas by Wave

CSR Plan	Texas Grants by Wave			Total Grants in Texas	Market Share in Texas	Total Grants Nationally	AIR rating of elementary CSR plans[1]	Estimated Value Added[2]	
	2000	2002	2005						
Accelerated Schools	30	29	26	85	17.1%	272	moderate	3.170**	(0.654)
AVID	25	18	24	67	13.5%	69		5.075**	(0.702)
Co-nect	4	10	3	17	3.4%	177	limited	9.260**	(1.409)
Success for All	5	9	2	16	3.2%	443	moderately strong	4.789**	(1.362)
High Schools That Work	0	5	9	14	2.8%	159		6.715*	(2.697)
Lightspan	6	8	0	14	2.8%	285		4.406**	(1.326)
HOSTS	0	11	2	13	2.6%	60		3.462+	(1.854)
Coalition of Essential Schools	4	7	0	11	2.2%	162	zero	2.573	(1.569)
Direct Instruction	9	1	0	10	2.0%	151	moderately strong	8.274**	(1.326)
Modern Red Schoolhouse	2	1	5	8	1.6%	101	limited	6.743**	(2.608)
Creating Independent Student-owned Strategies	6	1	0	7	1.4%	17		0.252	(1.449)
Literacy Collaborative -- Ohio State University	4	3	0	7	1.4%	101	limited	1.031	(1.673)
Breaking Ranks	0	3	3	6	1.2%	12		5.324	(4.363)
Effective Schools	3	3	0	6	1.2%	151		6.680**	(1.768)
Renaissance Learning	2	1	3	6	1.2%	160		14.811**	(2.363)
Success-in-the-Making	0	5	1	6	1.2%	29		-0.855	(2.740)
other	36	86	82	204	41.0%	--		4.446**	(0.480)
Total	136	201	160	497	100.0%	6476		Mean: 4.560	

Sources: CSR Awards Database, AIR.

[1] Zero rating means no evidence "was of sufficient quality to be counted as reliable evidence."

[2] Value-added regressions are shown in Appendix Table 2. The estimates presented here come from column 6 of that table. Standard errors are shown in parentheses.

** p<0.01, * p<0.05, + p<0.1

Table 2: CSR Plans in Texas and Nationally

Panel B: National Grants by Grade Level

CSR Plan	National Grants by Grade Level		
	Elementary	Middle	High
Accelerated Schools	158	49	17
AVID	1	33	34
Co-nect	85	53	30
Success for All	392	30	3
High Schools That Work	1	2	148
Lightspan	224	23	11
HOSTS	49	8	3
Coalition of Essential Schools	83	28	48
Direct Instruction	116	22	6
Modern Red Schoolhouse	63	20	10
Creating Independent Student-owned Strategies	3	8	4
Literacy Collaborative -- Ohio State University	99	1	0
Breaking Ranks	0	1	10
Effective Schools	86	27	27
Renaissance Learning	100	26	20
Success-in-the-Making	23	4	1

Sources: CSR Awards Database.

Table 3: Mean Pass Rate Changes, Conditional on Having Prior Experience

(Level of observation is school-plan combination; 21 plans)

Source of prior experience	Mean pass rate change for plans that were chosen	Mean pass rate change for plans that were not chosen
Own District	5.68	6.69
	(3.21)	(4.09)
	(N=79)	(N=249)
Geographic Neighbor (50 mile radius)	4.39	4.28
	(3.21)	(3.73)
	(N=227)	(N=2306)
Education Service Center (ESC)	3.80	4.02
	(3.99)	(4.33)
	(N=222)	(N=1709)

Table 4: Results from the Conditional Logit Model

Panel A: Prior Experiences Measured by Changes in Raw Pass Rate

	Variable Means [1]	(A1)	(A2)	(A3)	(A4)
Past presence of plan in own district	0.039	1.026**	1.020**	1.013**	1.022**
	(0.193)	(0.380)	(0.383)	(0.388)	(0.386)
Past presence X Change in own district pass rates	0.250	0.020	0.022	0.019	0.020
	(0.811)	(0.075)	(0.076)	(0.073)	(0.075)
Past presence of plan in geog. neighb. district	0.300		0.195		0.184
	(0.458)		(0.278)		(0.306)
Past presence X Change in geog. neighbor pass rates	1.284		0.003		0.017
	(2.027)		(0.029)		(0.040)
Past presence of plan in ESC	0.229			0.080	0.026
	(0.420)			(0.272)	(0.305)
Past presence X Change in ESC pass rates	0.912			-0.013	-0.020
	(2.061)			(0.021)	(0.029)
Observations		7,488	7,488	7,488	7,488

Panel B: Prior Experiences Measured by Plan's Value Added

	Variable Means [1]	(B1)	(B2)	(B3)	(B4)
Past presence of plan in own district	0.039	1.030**	1.035**	0.996*	1.022**
	(0.193)	(0.391)	(0.392)	(0.403)	(0.395)
Past presence in own district X Plan's value added	0.176	0.031	0.020	0.014	0.040
	(0.989)	(0.145)	(0.146)	(0.145)	(0.141)
Past presence of plan in geog. neighb. district	0.300		0.235		0.080
	(0.458)		(0.268)		(0.289)
Past presence in geog. neighb. X Plan's value added	1.549		0.095		-0.033
	(2.952)		(0.059)		(0.079)
Past presence of plan in ESC	0.229			0.105	0.190
	(0.420)			(0.268)	(0.291)
Past presence in ESC X Plan's value added	1.039			0.023	0.109
	(2.224)			(0.064)	(0.068)
Observations		7,488	7,488	7,488	7,488

** p<0.01, * p<0.05, + p<0.1

Notes: Results from conditional logit regressions. Standard errors are clustered at district level. Each specification includes plan dummies and the school's initial pass rate and its percentage of white students interacted with plan dummies. All "change in pass rates" variables are calculated only over schools in the relevant peer definition which used the particular CSR plan in the past. Changes in pass rates and value added enter the regression model demeaned.

Table 5: The Effect of Information on Choice Probabilities

(Percentage point change in choice probability as a result of increasing change in pass rate/value added by one percentage point)

	Specification		
	(A1/B1)	(A2/B2)	(A3/B3)
Prior Experiences Measured by Changes in Raw Pass Rate			
Mean	0.430	0.225	0.087
5th percentile	0.265	0.001	-0.151
Median	0.434	0.025	0.075
95th percentile	0.494	0.596	0.455
Prior Experiences Measured by Plan's Value Added			
Mean	0.680	0.995	0.401
5th percentile	0.425	0.047	0.020
Median	0.689	0.533	0.319
95th percentile	0.783	2.751	0.942

Notes: Authors' simulations, based on results from Table 4. All predicted changes in choice probability are not statistically different from zero

Appendix Table 1
CSR Plan Adoption within Large Districts over Time

District and spring of school year plans first implemented	Number of CSR schools, by year and level	School level	Plan adopted
Austin ISD			
1998	3	high	AVID
	4	middle	AVID
2004	1	elementary	Accelerated Schools
	1	elementary	Cognitively Guided Instruction
	1	middle	AVID
	3	high	AVID
Cypress-Fairbanks			
1998	3	elementary	Literacy Collaborative--Ohio State U.
	2	middle	AVID
	1	high	AVID
	1	high	locally-developed plan
2001	1	elementary	locally-developed plan
	1	middle	AVID
	1	high	AVID
Dallas ISD			
2001	2	elementary	Success for All
2004	3	elementary	3-Tier Intervention Model
	3	middle	AVID
El Paso ISD			
1998	3	elementary	Success for All
	2	elementary	El Paso Collaborative for Academic Excellence
	1	middle	El Paso Collaborative for Academic Excellence
	1	middle	Cooperative Integrated Reading and Comprehension
	1	high	Community for Learning
2001	1	elementary	El Paso Collaborative for Academic Excellence
2004	2	high	missing
Fort Worth ISD			
1998	8	middle	Direct Instruction
	1	middle	AVID
	2	high	AVID
	1	high	Direct Instruction
2001	1	elementary	Direct Instruction
	2	middle	AVID
2004	3	middle	AVID
	3	high	AVID

Sources: CSR Awards Database; authors' Public Information Request (available upon request).

Appendix Table 1 (continued)
CSR Plan Adoption within Large Districts over Time

District and spring of school year plans first implemented	Number of CSR schools, by year and level	School level	Plan adopted
Houston ISD			
1998	2	elementary	Accelerated Schools
	2	elementary	Core Knowledge
	1	elementary	Different Ways of Knowing
	1	middle	Accelerated Schools
	1	middle	Coalition of Essential Schools
	1	middle	Co-nect
	1	middle	Learner-Centered Framework
	1	high	Co-nect
2001	1	elementary	Consistent Mangement and Cooperative Discipline
	1	elementary	Core Knowledge
	1	elementary	Houston Annenberg Challenge
	2	middle	Co-nect
	1	middle	First Things First
	2	high	Co-nect
	1	high	Houston Annenberg Challenge
	1	high	Coalition of Essential Schools
2004	4	elementary	Co-nect
	2	middle	Accelerated Schools
	1	middle	Success for All
	1	middle	School Development Program
	1	high	Project CLEAR
	1	high	Houston Annenberg Challenge
	1	high	Accelerated Schools
	2	high	High Schools That Work
	1	high	Co-nect
San Antonio ISD			
1998	6	elementary	Roots and Wings
	2	elementary	Success for All
	1	elementary	Co-nect
	1	elementary	locally developed
	1	middle	Modern Red Schoolhouse
	2	middle	Capacity Building Model
	1	middle	locally developed
2001	1	elementary	HOSTS
	1	elementary	Quantum Learning
2004	1	elementary	Project Reach
	2	high	Capturing Kids Hearts
	1	high	AVID
	1	high	missing

Sources: CSR Awards Database; authors' Public Information Request (available upon request).

Appendix Table 2: Value-added Regressions

Independent variable: change in raw pass rates, averaged over math and reading for all grades reported, over CSR treatment period

Dependent variables	(1)	(2)	(3)	(4)	(5)	(6)
Any CSR	14.667**	13.834**	4.560**			
	(0.460)	(0.450)	(0.284)			
Accelerated Schools				15.296**	14.221**	5.075**
				(1.156)	(1.130)	(0.702)
AVID				13.263**	14.002**	3.170**
				(1.076)	(1.052)	(0.654)
Breaking Ranks				14.516*	14.189*	5.324
				(7.140)	(6.980)	(4.363)
Co-nect				19.900**	19.999**	9.260**
				(2.324)	(2.272)	(1.409)
Coalition of Essential Schools				12.158**	10.209**	2.573
				(2.590)	(2.532)	(1.569)
Creating Independent Student-owned Strategies				8.821**	7.544**	0.252
				(2.390)	(2.336)	(1.449)
Direct Instruction				19.427**	17.073**	8.274**
				(2.187)	(2.139)	(1.326)
Effective Schools				17.819**	15.559**	6.680**
				(2.918)	(2.853)	(1.768)
HOSTS				12.986**	11.781**	3.462+
				(3.058)	(2.990)	(1.854)
High Schools That Work				17.548**	16.380**	6.715*
				(4.452)	(4.351)	(2.697)
Lightspan				15.124**	14.841**	4.406**
				(2.186)	(2.137)	(1.326)
Literacy Collaborative -- Ohio State University				8.908**	7.825**	1.031
				(2.760)	(2.698)	(1.673)
Modern Red Schoolhouse				17.700**	16.157**	6.743**
				(4.306)	(4.209)	(2.608)
Renaissance Learning				24.799**	23.985**	14.811**
				(3.901)	(3.813)	(2.363)
Success for All				15.317**	14.521**	4.789**
				(2.248)	(2.197)	(1.362)
Success-in-the-Making				9.562*	9.929*	-0.855
				(4.523)	(4.421)	(2.740)
other				14.439**	13.123**	4.446**
				(0.787)	(0.770)	(0.480)

(continued next page)

Appendix Table 2 continued: Value-added Regressions

Dependent variables	(1)	(2)	(3)	(4)	(5)	(6)
Exempt rate		43.289**	6.001**		43.318**	5.923**
		(0.919)	(0.626)		(0.919)	(0.626)
Percent white			0.262**			0.261**
			(0.007)			(0.007)
Enrollment			-6.529**			-6.489**
			(0.556)			(0.558)
Enrollment squared			0.784**			0.781**
			(0.166)			(0.167)
1993-94 SY			-24.773**			-24.770**
			(0.121)			(0.121)
1994-95 SY			-20.196**			-20.193**
			(0.118)			(0.118)
1995-96 SY			-15.487**			-15.483**
			(0.115)			(0.115)
1996-97 SY			-10.687**			-10.684**
			(0.112)			(0.112)
1997-98 SY			-6.988**			-6.986**
			(0.110)			(0.110)
1998-99 SY			-6.208**			-6.208**
			(0.108)			(0.108)
1999-2000 SY			-4.917**			-4.917**
			(0.105)			(0.105)
2000-2001 SY			-2.575**			-2.578**
			(0.103)			(0.103)
Observations	54,512	54,512	54,512	54,512	54,512	54,512
R-squared	0.538	0.559	0.831	0.538	0.559	0.831

Standard errors in parentheses

** $p<0.01$, * $p<0.05$, + $p<0.1$