

School Choice and School Accountability: Evidence from a Private Voucher Program in Milwaukee, Wisconsin

Deven E. Carlson, Ph.D.
University of Oklahoma

Joshua M. Cowen, Ph.D.+
University of Kentucky

David J. Fleming, Ph.D.
Furman University

INVALSI Improving Education Conference
Rome, Italy

October 3-5 2012

Abstract

This paper considers the introduction of a performance reporting reform to private schools receiving public vouchers in Milwaukee, Wisconsin. Drawing on unique panel data collected both before and after the reform, we show that private sector performance increased dramatically when outcomes for each voucher-accepting private school were released to the general citizenry. We frame these results in the context of third-party provision of public services, and argue that our evidence suggests that market-based competition alone may not drive non-governmental providers to perform at optimal levels. Instead, such vendors may require performance-monitoring schemes similar to those faced by their governmental counterparts.

+Corresponding Author
Martin School of Public Policy and Administration
University of Kentucky
433 Patterson Office Tower
Lexington, KY 40506-0027
Phone: (859) 257-4387
Email: joshuacowen@uky.edu
<http://www.martin.uky.edu/faculty/Cowen/>

I.) Introduction

A generation of literature has documented and described the growing role of third-party entities in providing public goods and services (Kettl 2002; Salamon 1989; Stoker 1998). Numerous studies have analyzed the broad effects of this movement—whether toward privatization, contracting, non-profit management, and the like—and have returned evidence of varying degrees of success. Although this scholarship has left the normative debate over the desirability of these trends unsettled, it has clearly established third-parties as a significant presence on the contemporary governmental landscape.

In recognition of this presence, the literature has progressed from describing the growth in third-party governance to addressing a second generation of issues that have emerged as a result. One such issue concerns monitoring the quality of privately delivered goods and services. Scholars have considered both theoretically and empirically how government agencies can ensure that goods and services provided by private entities are delivered in both an efficient and high-quality manner (e.g., Osborne and Gaebler 1992; Dicke 2002; Brown and Potoski 2003, 2006; Brown, Potoski and Van Slyke 2006; Van Slyke 2007; Yang, Hsieh and Li 2009; Yang and VanLandingham 2011). One primary line of argument holds that market forces—especially competition—provide the accountability necessary to ensure high-quality good and service delivery. A second view contends that the market may not be sufficient, particularly if it lacks robust competition (Donahue 1989; Prager and Swati 1996) or if the relevant good or service is complex in nature (Brown, Potoski, and Van Slyke 2010). In these cases, additional oversight or accountability regimes may be required to ensure that alternatively delivered goods and services meet applicable quality standards.

Paralleling the growth in third-party governance has been a focus on performance management in the public sector. Although performance management systems are variable and diverse, performance measures—coupled with an incentive structure based on those measures—are central components of nearly all such systems (Heinrich and Marschke 2010). Given this centrality, the issue of performance measurement has been studied extensively. Empirical studies have focused on assessing the effects of employing these measures—and associated incentive structures—across several policy domains, including welfare services, job training programs, health care delivery, and public education (e.g., Heinrich and Choi 2007; Hill 2006; Dias and Maynard-Moody 2007; Ewalt and Jennings 2004; Heinrich 2007; Bevan and Hood 2006; Meier and O’Toole 2002, 2003). Performance measures are often shown to produce increases in outcomes, although the techniques used to achieve such gains may differ from the intended approach (Jacob 2005; Jacob and Levitt 2003).

Recently, performance measurement has begun to emerge as a strategy employed by the public sector to monitor—and hopefully increase—the quality of public goods and services delivered by private entities. These strategies may be particularly useful in cases where governments approve a set of providers, while other agencies or consumers choose between them. For example, several states have implemented rating systems for private child care facilities receiving public funding. Similarly, the federal government operates the What Works Clearinghouse, which provides information about the effectiveness of a wealth of educational policies and interventions, many of which are provided by private entities. Although implementation of performance measurement systems for private vendors is becoming increasingly common, little is known about the effects of such policies.

Within the context of education, this paper analyzes how important outcomes change in response to the implementation of performance measurement and an associated incentive system—in this case public reporting—affecting private vendors. In this we follow Carolyn Heinrich’s (2010) study of third-party vendors of supplemental tutoring programs used by the public school district serving Milwaukee, Wisconsin, to consider public support of entirely private educational delivery in the same city. Specifically, using unique data from the Milwaukee Parental Choice Program (MPCP)—the nation’s oldest and largest urban school voucher program—we assess how student achievement outcomes respond to implementation of a performance measurement policy. That policy required private schools accepting public vouchers to administer reading and mathematics assessments to all students in grades 3-8 and 10 and provide the results to the state’s supervising agency for additional analyses and subsequent release for public consumption.

Our data contain student-level outcomes before and after implementation of the policy, allowing us to gauge private school student outcomes under two conditions: 1) when market forces provide the only accountability mechanism, and 2) when market-based accountability is augmented with performance measurement and public reporting requirements. We utilize two quantitative evaluation techniques—interrupted time-series and difference-in-difference estimation—to estimate the effects of the new system. The specifics of our results are nuanced, but the basic pattern is clear: after a period of generally stability in outcomes when market forces provide the only accountability mechanism, student performance for voucher providers exhibited dramatic improvements in response to implementation of the performance measurement reform. These results have a number of important implications that we discuss in the final section of the paper.

II.) Theoretical Motivation: Third Party Governance and Performance Measurement in Education

IIA.) Third-Party Performance

A major principle underlying third-party governance is the assumption that non-governmental agencies are more effective providers of some public services either because they possess a particular specialty or because market pressures compel a sink-or-swim approach to service delivery (Gansler 2006). It is far from clear whether the empirical evidence in the extant literature supports the latter proposition in particular (Heinrich 2009; Milward and Provan 2000; Johnson and Romzek 1999; Romzek and Johnston 2005; van Slyke 2003) and over the past decade agencies' ability to monitor third-party performance has been stressed among the central questions in that literature (e.g. Dicke 2002; Brown and Potoski 2003, 2006; Brown, Potoski and Van Slyke 2006; van Slyke 2007; Yang, Hsieh and Li 2009; Yang and VanLandingham 2011). Whether third parties are more efficient may depend on the tasks and outcomes pertinent to each agency: well-defined services may be particularly well-suited to outsourcing, while in contexts with vague or highly technical services contracting may realize limited gains (Brown and Potoski 2006). Even in a context with readily defined outcomes, a comparative advantage to contracting over traditional public service delivery may have several critical limitations. As Heinrich (2010) summarizes, in addition to "complex and highly specialized" contexts, such limitations include decentralized public/private partnerships, constraints on time and financial resources, and contexts in which "there are too few suppliers (or providers) to assure a competitive market for services" (Heinrich 2010, i63)

Even in a truly competitive market of service providers, what remains uncertain in the literature is the extent to which a market-oriented accountability mechanism maximizes contract performance, or whether direct monitoring on the part of the government itself can improve

service delivery. To pose the question in more prosaic terms, is a well-functioning market *enough* to ensure a high level of service delivery? Public agencies are becoming increasingly adept at measuring their own performance. Empirical evidence on the impact of performance measurement is available from a variety of contexts, include the provision of welfare services (Heinrich and Choi 2007; Hill 2006; Dias and Maynard-Moody 2007; Ewalt and Jennings 2004), federal job training programs (e.g. Heinrich 2007), health care (Bevan and Hood 2006) and education systems (Meier and O’Toole 2002, 2003), and studies of the Government Performance and Results Act (e.g. Moynihan and Lavertu, forthcoming; Radin 2000) and the Program Assessment Rating Tool (e.g. Heinrich, forthcoming; Stalebrink 2009). Given the apparent uncertainty in the literature over whether third-party delivery actually exceeds that available within the traditional public sector, an important question concerns the potential for third-party vendors to improve their own delivery of public services in response positively to the sort of accountability-based reforms that have characterized many publicly managed agencies.

IIB.) School Vouchers as a Case Study for Third-Party Governance

Although their intellectual origins date back to economics (Friedman 1962), school voucher programs are a paradigmatic example of one aspect of third-party governance—the use of private vendors—as scholars of public administration may observe. Under these programs, private schools receive public funds to provide educational services to program participants. Such an arrangement—devolution of the responsibility for schooling to the private sector, coupled with the use of public funds to provide that schooling—is an apt illustration of prominent definitions of third-party government (e.g., Salamon 1989).¹

¹ A number of previous studies have considered issues related to privatization or third-party governance in the context of education (e.g., Heinrich 2010; Meier and O’Toole 2002, 2003), but school choice programs have yet to inform the literature to a meaningful degree.

A similar illustration has been recently utilized for third-party providers of supplemental (tutoring) educational services in Milwaukee (Heinrich 2010). We are however unaware of such an application to private school vouchers. In the case of the Milwaukee Parental Choice Program (MPCP), participating private schools receive a set amount of money from the Wisconsin Department of Public Instruction (DPI) for each voucher student enrolled in their school. During the 2011-12 school year, over 22,000 students attended approximately 100 participating private schools, and each school received approximately \$6,500 in public funds for each voucher student enrolled in their school.² With its enrollment of over 22,000 students, the MPCP educated about one-quarter of the number of students enrolled in the surrounding Milwaukee Public School system (MPS).

The MPCP has not always been of its current size and scope. Like many privatization efforts in other contexts, the MPCP started small and has expanded incrementally over time. In the initial year of the pilot program, 1990-91, there were only about 340 students enrolled in just seven participating private schools. Over time, as enrollment caps were raised and eligibility requirements were eased, school and student participation in the program grew steadily. Figure 1 presents, annually, the number of participating schools and students over the two decades the MPCP has been in existence. The figure demonstrates that the MPCP has grown from a small pilot program to become firmly established on the educational landscape in Milwaukee. Through its small start and steady expansion over time, the MPCP mirrors broader trends in the growth of third-party governance documented in previous work (e.g., Kettl 1993, 2002; Light 1999; Salamon 1989).

² During the 2011-12 school year, students had to reside in the City of Milwaukee and have a family income under 225 percent of the poverty line to be eligible for voucher receipt, although the eligibility thresholds were altered for the 2012-13 school year forward.

[Insert Figure 1 about here]

There is one feature of the MPCP program—and privatized education more generally—that differs from many other third-party governance contexts. In several contexts, such as waste management, toll collection, and social service delivery, the government contracts with a single private entity to provide a specific good or service to the public. In such cases the theoretical accountability mechanism is quite clear; low-quality good or service provision will simply result in the government contracting with a different private provider. In contrast, in the MPCP—and most other examples of privatized education (e.g., Heinrich 2010)—the state does not select one single private vendor to provide educational services, but rather approves a set of private educational institutions to offer those services. Voucher holders then choose among the selection of approved vendors.³ In cases such as this, the theoretical accountability mechanism is less clear. Does the assurance of educational quality come from parents and students only choosing to attend high-quality private schools (i.e., market-based accountability)? Does it come from government approval and monitoring of the private schools participating in the program? Ultimately, is market-based accountability enough to ensure the delivery of high-quality education from private schools participating in the MPCP, or are additional measures required?

IIC. Performance Measurement and the Market for Educational Services

Passage of the federal No Child Left Behind Act of 2001 (NCLB) effectively commenced the current era of accountability in public education. Among other provisions, the law requires states to annually administer reading and mathematics assessments to all students in grades 3-8 and one grade in high school. Results from these assessments determine whether schools and

³ Such an arrangement is not wholly unique to education. Indeed, there are other contexts, such as publicly-funded child care, in which the clients choose among a set of approved vendors.

districts will be subjected to sanctions specified in the law.⁴ In effect, NCLB imposed a performance measurement system on the nation's public schools. Researchers have devoted significant attention to studying the effects of NCLB and other accountability programs; although the debate over the extent to which these initiatives hold long-term promise for increasing educational quality continues, the results suggest that these programs have short-term positive academic impacts (Dee and Jacob 2011; Jacob 2005; Hanushek and Raymond 2005; Carnoy and Loeb 2002).

Despite NCLB's imposition of accountability on public schools across the country, private school vendors have historically been exempted from such systems, even when they accept public funds such as vouchers. These exemptions are undoubtedly the result of several factors, but perhaps the most promulgated justification is that market forces—choice and competition—provide a level of accountability exceeding that which could be provided by any governmentally designed system. As the pro-voucher Friedman Foundation summarizes in its statement on school choice: “What gives the concept of accountability real teeth is a parent's ability to choose a school freely... a parent can take a child out of a school that isn't doing the job and find another school that will.”⁵

The notion that market-based accountability can offer sufficient assurance that private vendors will provide public goods and services in a high-quality and efficient manner is not a novel one in the third-party governance literature (Savas 2000; Savas 1987; Donahue 1989; Salamon 1989). Implicitly grounded in the economic theory of perfectly competitive markets, this line of argument holds that the presence of choice and competition in the delivery of public

⁴ See Manna (2011) for a comprehensive description of the sanctions specified in NCLB that schools and districts face as a consequence of their test performance.

⁵“Are Participating Private Schools Held Accountable?” The Friedman Foundation for Educational Choice, <http://www.edchoice.org/School-Choice/School-Choice-FAQs.aspx>, retrieved 3/1/2012

goods and services will result in increased quality and efficiency, relative to an arrangement in which goods and services are delivered by a public monopoly. Indeed, the empirical literature identifies several areas in which the introduction of choice and competition into good and service delivery appears to have resulted in improved outcomes. A sampling of these areas includes waste removal, towing, data processing, street repair, and snow removal, among others (Van Slyke 2003). Applied to the MPCP, this line of argument holds that the ability of parents to choose among a set of public and private schools will spur competition—both among the private schools and between the private schools and MPS—that will result in the provision of higher-quality education, which will in turn produce superior outcomes for students attending both public and private schools. See Hoxby (2003) for a formal exposition of this argument in the context of school choice.

In cases where privatization has failed to meaningfully increase competition—a situation that may arise for a number of reasons—studies have shown that privatization efforts often fail to improve relevant outcomes (Donahue 1989; Prager and Swati 1996; Brown, Potoski, and Van Slyke 2010).⁶ Perhaps most importantly for our purposes, the recent Heinrich study (2010) found evidence that that “the government should not rely on parents and students to check the market through their choice of providers” of supplemental educational services (p. 176).

Heinrich explains this finding in part by noting the lack of parental information on third-party providers. The assumption of perfect information is, along with the assumption of perfectly competitive markets, the theoretical undergirding of market-based accountability. In the case of Milwaukee’s voucher program, it is similarly unclear whether families possess the requisite level

⁶ Studies have shown that privatization may fail to meaningfully increase competition when there are few private providers of a given good or service (Donahue 1989; Prager and Swati 1996). A special case of this scenario is the provision of goods or services that are complex in nature; in such cases few vendors possess the expertise necessary for their provision (Brown, Potoski, and Van Slyke 2010).

of information regarding educational quality to satisfy the assumption of perfect information, particularly given that voucher programs have historically been exempted from testing and accountability policies. According to a 2007 survey, almost forty percent of public school parents in the City of Milwaukee said that they had never even heard of the MPCP (Witte, et al 2008). Indeed, the best studies on the topic indicate that the level of accurate information on school characteristics—including educational quality—possessed by low-income parents participating in school choice programs is generally quite low (e.g., Schneider et al. 1998). Bridge (1978) calls the lack of parental knowledge about school options “the Achilles’ heel” of school choice plans.

It is therefore theoretically ambiguous whether the MPCP—and similar educational privatization programs—will produce quality and efficiency gains under a market-based accountability system. Despite this ambiguity, market-based accountability served as the sole accountability mechanism for the first 20 years of the MPCP. Several analyses of the program during this period found the achievement of MPCP students to exhibit little difference from that of their MPS counterparts (Witte 2000; Greene, Peterson and Du 1999; Rouse 1998; Lamarche 2008).⁷ The relatively stagnant achievement of students in the MPCP contributed to a discussion about whether market-based accountability was enough to ensure the delivery of high-quality education from private schools participating in the program, or whether additional accountability measures should be implemented.

⁷ Market-based accountability has served as the sole accountability mechanism for most other publicly-funded private school voucher programs, including those in Cleveland, Florida, and Washington DC, as well as a set of publicly supported voucher-like programs beginning in Colorado, Louisiana and Indiana. Evaluations of these programs generally found relatively little difference in the achievement of voucher participants and their publicly educated peers (Greene, Peterson and Du 1999; Howell et al. 2006; Wolf et al. 2011; Belfield 2005; Metcalf 2003; Lara, Mizala and Repetto 2011).

III.) Performance Measurement in Milwaukee

As described in the section above, previous research (e.g., Schneider et al. 1998; Bridge, 1978)—coupled with the program’s historical exemption from testing and accountability policies—provides a basis for expecting that parents seeking vouchers in Milwaukee may have lacked the requisite information to accurately judge the quality of schooling options in the MPCP, thus violating a primary assumption of the theoretical undergirding of market-based accountability. Consequently, providing parents with such information would theoretically allow families to make better schooling decisions and—through the forces of market-based accountability—realize improvements in important educational outcomes.

Statute requiring provision of such information was enacted with passage of 2009 Wisconsin Act 28. As summarized in Table 1, the requirements of 2009 Wisconsin Act 28 were phased in over the course of two school years . In the first school year—2009-10—schools were required to pay additional administrative fees to the Wisconsin Department of Public Instruction (DPI) and new schools had to obtain a more rigorous certification to participate in the MPCP. The requirements that took effect in the 2010-2011 academic year were far broader in scope. Specifically, beginning in fall of 2010, MPCP schools were required to test all voucher students in grades 3-8 and 10 in reading and math with the state’s Wisconsin Knowledge and Concepts Examination (WKCE) exam, the same test used by public schools for NCLB compliance. Additionally, MPCP schools were required to test voucher students in grades 4, 8, and 10 in English/language arts, writing, science, and social studies;⁸ all test results must be submitted to the DPI for review, analysis, and public reporting. In addition, schools were required to develop and report individual academic standards for each subject, and all teachers hired after 2010 must

⁸ It is important to clarify that private schools were only required to test students who attended their school through the voucher program. Students who attended the schools outside of the program were not required to be tested.

be certified. Finally, yearly instructional hour minimums—1,050 for elementary and 1,137 hours for junior high school—were also instituted for all MPCP schools.

[Insert Table 1 about here]

Of the various dimensions of this policy, the most important component was the performance measurement system—the mandated testing and public reporting of scores. Although no explicit sanctions are tied to school results, schools are required to participate in the new accountability program or lose their voucher funds. Since the vast majority of participating schools serve primarily voucher students (Kisida, Jensen and Wolf 2011), this implies that—for most schools—failure to participate will result in closure. Consequently, this policy arguably possesses a penalty threat for voucher schools that is equivalent to—and perhaps even greater than—that faced by public schools under No Child Left Behind.

As noted above, 2010-11 represented the first academic year in which test scores were released by the state DPI by individual school. In addition to allowing parents to observe the academic outcomes of individual private schools participating in the MPCP, the private school test scores were aggressively analyzed by outside groups in the state (e.g., Dickman and Schmidt 2012) and received significant media attention. For example, an article containing a table listing the percentage of students in each voucher school that were proficient or advanced in reading, language arts, math, science, and social studies appeared on the front page, above-the-fold in the *Milwaukee Journal Sentinel* (Richards and Hetzner 2011). Furthermore, the article allowed readers to compare the performance each individual voucher school to the Milwaukee Public Schools and state averages. A few days later, an editorial in the *Milwaukee Journal Sentinel* (2011) argued that the public dissemination of voucher school test scores was important because “parents need an annual, consistent and objective set of metrics to help them judge schools” (p.

12A). Implementation of the performance measurement system in the MPCP clearly raised the academic stakes faced by private schools serving voucher students, but did it improve these providers' performance?

IV.) Data and Sample

The primary obstacle to estimating the effects of a performance measurement policy arises from the fact that relevant outcome measures are often not available prior to implementation of the policy; their absence is precisely what necessitated implementation of the policy. In the case of the MPCP, however, a longitudinal study of the program that began in 2006 and extended through 2011 resulted in the availability of relevant outcome measures both before and after implementation of the policy described above.

Four years before the performance measurement policy became law, the MPCP was re-authorized via 2005 Wisconsin Act 125, which also contained a provision directing an independent group of researchers, of which the present authors were a part, to construct a “representative panel” of voucher students whose educational outcomes would be tracked over a period of time (Witte, et al. 2008). This representative panel of voucher students consisted of approximately 2,700 students enrolled in the MPCP during the 2006-07 school year. In addition, refresh samples of 3rd grade students in the voucher program in both 2007-08 and 2008-09. Critically, for our purposes, the law required voucher schools to allow the researchers to administer the reading and math portions of the state exam to all students selected for the representative panel.⁹

⁹ The original sample consisted of students in grades 3-8, stratified by grade, and the population of 9th graders. For the refresh samples, in each year approximately 475 MPCP 3rd graders were added to the original sample.

The law also directed the researchers to monitor a “comparable” group of traditional public students over the same time period, which they did by matching the representative panel of voucher students to students in Milwaukee’s public school system (MPS) who resided in the same census tract, had the same observable demographics, and had the same WKCE scores at baseline.¹⁰ The researchers also generated a second potential comparison group by simply drawing a random sample of MPS students.¹¹ The combined result of this data collection is a dataset that includes students not only in public schools, but also those in voucher schools in the years before and after the state’s implementation of its performance measurement policy.

From these data, we construct two distinct samples for the analyses to follow. The first of these samples consists of all students whose enrollment in MPCP was confirmed in 2008-09, 2009-10, and 2010-11. Thus, we observe sample members’ achievement outcomes for two years prior to implementation of the accountability policy and one year post-implementation. This sample allows us to analyze how—relative to their previous achievement trajectories—MPCP students’ achievement changed in response to implementation of the accountability policy in 2010. Our second analytic sample consists of all students who were either 1) confirmed in MPCP from 2008-09 to 2010-11—the same group of students comprising our first analytic sample—or 2) confirmed to be in MPS during those three school years. With this sample we are able to estimate the change in MPCP students’ achievement from 2009 to 2010—a time period over which accountability requirements changed—relative to the change in MPS students’ achievement during the same time period, in which public school students’ exposure to accountability remained constant.

¹⁰ The same procedure was used for the 2007 and 2008 refresh samples.

¹¹ The first report of the official study based on these data was presented in 2008 (Witte, et al. 2008), with subsequent yearly reports occurring through 2012.

V.) Analysis and Results

VA.) Descriptive Analysis

Prior to estimating the effect of implementing the MPCP performance measurement policy in a multivariate framework, we first provide some descriptive evidence on the issue. Using WKCE scores that are standardized by the MPS mean and standard deviation for the proper year, grade, and subject, Figure 2 plots estimated mean scores and their 95 percent confidence intervals from 2008-09 to 2010-11. The mean scores are plotted separately for reading and math. The figure demonstrates that, in math, average student achievement in MPCP was about 0.25 standard deviations below the MPS average in both 2008 and 2009, prior to implementation of the accountability policy. However, MPCP achievement exhibited a significant jump in 2010-11, which is the first year the performance measurement policy was in effect. A similar pattern is observed for reading achievement. Although initial reading achievement levels were somewhat higher than math, there is little change in average achievement from 2008 to 2009, but a substantial increase from 2009 to 2010-11.

[Insert Figures 2 and 3 about here]

Using the second analytic sample, Figure 3 presents the mean difference in achievement between MPCP and MPS students from 2008 to 2010. The figure demonstrates that, in math, the achievement of MPCP students was substantially below that of their MPS counterparts in 2008. There was very little change in this difference between 2008 and 2009, but a significant narrowing from 2009 to 2010. In reading, the average achievement of the MPCP and MPS students in our data was quite similar in both 2008 and 2009. In 2010-11, however, MPCP students significantly outperformed their MPS peers. Together, Figures 2 and 3 provide preliminary and descriptive evidence that implementation of the MPCP performance

measurement policy resulted in improved educational outcomes. To analyze this pattern further—and attempt to isolate the causal effect of implementing the MPCP performance measurement policy on student achievement outcomes—we employ two multivariate strategies.

VB.) Interrupted Time-Series Analysis

Our first strategy analyzes whether—relative to their prior achievement levels—students in MPCP schools exhibited a substantial increase in achievement in the year the performance measurement policy went into effect. We estimate two similar—yet distinct—interrupted time-series models to gain insight into this issue. The first can be written as:

$$\text{Eq. 1)} \quad A_{it} = 2010_i\pi_1 + 2009_i\pi_2 + \varphi_i + \varepsilon_{it}$$

where achievement A for student i at time t is a function of a student fixed effect, φ , and dummies for the 2009-2011 and 2010-11 school years; the 2008-09 school year serves as the reference category.¹² The model is estimated separately for reading and math achievement. The student fixed effect serves a number of important functions in this model. Specifically, it results in the parameters of the model being estimated using within-student variation. In doing so, it eliminates the threat of bias from any time-invariant factors, observed or unobserved. We interpret the π_1 coefficient as the effect of the performance measurement policy on the student achievement of voucher students in MPCP schools. The implicit assumption underlying this interpretation is that the only structural difference between 2010—when the policy was first implemented—and earlier years was the implementation of the new performance measurement policy in the MPCP schools. Although we cannot test this assumption directly, there is no evidence of any other substantial policy change taking effect that year. In addition, the reform we

¹² We do not include grade fixed effects in this model because grade and year are perfectly correlated for students not retained in grade; these students constitute the vast majority of our sample.

have described above was indeed so structural in nature—such a break with past private school operations—that we consider it the straightforward interpretation of the educational climate unique to 2010.

The model presented in equation (1) eliminates the threat of bias stemming from all time-invariant characteristics. However, it is possible that unobserved time-varying factors could bias the coefficient estimates. For example, if some factor resulted in the artificial depression of student achievement in 2009-10, then the coefficient on the 2010 dummy in equation (1) would be biased upward. To address this possibility, we estimate a second model containing lagged measures of achievement on the right-hand side. The model—commonly referred to as a value-added model—can be written as:

$$\text{Eq. 2)} \quad A_{it} = 2010_i\pi_1 + 2009_i\pi_2 + A_{i,t-1}\beta_1 + X_{it}\beta_2 + \varepsilon_{it}$$

where $A_{i,t-1}\beta_1$ represents lagged measures of reading and math achievement and $X_{it}\beta_2$ represents a vector of observable student background characteristics—such as sex, race/ethnicity, special needs, English learning and free or reduced-price lunch status (descriptive data on these available on request, see also Witte, et al. 2008)—as well as grade fixed effects. The model is again estimated separately for reading and math. As described above, we interpret the π_1 coefficient as the effect of the performance measurement policy on student achievement outcomes. If the models presented in equations (1) and (2) return similar results, then we can have greater confidence in the quality of the estimates. Note that we cannot include student fixed effects in Equation (2), because these results would be inconsistent under typical statistical assumptions.

Table 2 presents the results from estimation of equations (1) and (2). Both the student fixed effects results of (1) and the value-added results from (2) clearly demonstrate no significant

change in student achievement from 2008-09 to 2009-10—prior to implementation of the MPCP performance measurement policy—but a substantial jump in 2010-11, the first year the policy was in place. This pattern is observed in both reading and math. In addition to being statistically significant, the estimated effects are substantively large—nearly two-tenths of a standard deviation across both subjects and specifications. The fact that achievement levels were not statistically different in 2008-09 and 2009-10 provides confidence that the 2010-11 coefficient estimate reflects the effect of the performance measurement policy, as opposed to simply the extension of a preexisting positive trend in MPCP student achievement. Taken as a whole, the results demonstrate that, relative to their prior achievement patterns, MPCP students’ achievement increased substantially in the year that the performance measurement policy was implemented.

[Insert Table 2 about here]

VC.) Difference-in-Differences Analysis

Equations (1) and (2) address the question, *relative to their prior achievement patterns, how did MPCP students’ achievement change after introduction of the performance measurement policy in 2010?* In our second approach, we leverage the fact that MPS students were subject to a similar policy throughout the full time period of our study. Specifically, MPS has been subject to the accountability provisions of No Child Left Behind since 2002 and this policy, which bears resemblance to the performance measurement policy implemented in the MPCP, represented a nearly decade-long status quo in MPS by the time the MPCP performance measurement policy took effect in 2010. Thus, MPS students serve as the comparison group in our second analysis, which addresses the following question: *relative to the achievement of MPS*

students, how did MPCP students' achievement change after introduction of the performance measurement policy in 2010? Our coefficient of interest, δ_1 , in:

$$\text{Eq. 3) } A_{it} = (2010_i * MPCP)\delta_1 + 2010_i\delta_2 + MPCP\delta_3 + X_{it}\beta_1 + \varepsilon_{it}$$

is a difference-in-differences estimate of the private school accountability program; δ_2 is a time trend shared by both sectors, and δ_3 captures average baseline achievement for MPCP students. As in (2), $X_{it}\beta_1$ represents a vector of student background characteristics and grade indicators. Similar to the student fixed effects model in equation (1), difference-in-differences approaches eliminate threats of bias from time-invariant factors—both observed and unobserved—thus providing greater confidence in the causal nature of the estimate of δ_1 .

The second and fourth columns in Table 3 present the results of estimating (3) for reading and math, respectively. Consistent with the interrupted time-series results presented above, the difference-in-differences estimates indicate that—relative to the achievement of MPS students—the achievement of MPCP students increased substantially more from 2009-10 to 2010-11. The difference-in-differences estimates are positive and significant for both math and reading. Although the magnitude of the difference-in-differences estimate for math is similar to that observed in Table 2, the reading estimate is somewhat smaller. This discrepancy is attributable to the fact that the two analyses address wholly different questions and is not problematic; it simply implies that MPS students' reading achievement was higher than the prior reading achievement of MPCP students.

[Insert Table 3 about here]

A potential threat to the validity of difference-in-differences estimates is the presence of differing pre-existing achievement trends in the treatment and control groups, in this case the MPCP and MPS. For evidence that our difference-in-differences analysis is not simply detecting

such trends, consider the first and third columns in Table 3. These columns present estimates of equation (3) in which the 2010 indicators are replaced by 2009 indicators and the sample is changed to student-year observations in 2008-09 and 2009-10. If the results presented above were simply capturing some other extraneous source of growth occurring prior to 2010-11, we would expect to observe a positive and significant estimate for the difference-in-differences terms. Table 3 clearly demonstrates that such results are not present; the difference-in-differences terms are either zero or even slightly negative. Together, the results of the interrupted time-series and difference-in-differences analyses provide strong evidence that implementation of the MPCP performance measurement policy produced substantial increases in the achievement of voucher students attending MPCP schools. Prior to considering the implications of this finding, we explore the possibility that the observed effects are stemming from another factor, specifically selective attrition from MPCP schools.

VD.) Selective Attrition from the MPCP?

Previous work has demonstrated that organizations and institutions—including schools—often respond strategically to the implementation of a performance measurement system (Courty, Kim and Marschke 2011; Hood 2006; Jacob 2005; Jacob and Levitt 2003). One strategy that MPCP schools could potentially employ in response to the new performance measurement requirements involves manipulation of the pool of students who must take the WKCE. In principle, private schools have the ability to expel or counsel students out, and it is possible that they deliberately culled away their lowest performers in anticipation of the new performance measurement regime. Such a response seems somewhat at odds with the fact that most schools depend on voucher students to exist, but it is nevertheless quite possible, especially if schools were concerned about their scores in the first year (2010) under the policy.

A simple test may provide evidence on the role that selective attrition may have played in producing the achievement increases observed in Tables 2 and 3. For this test we estimate equation (3) over an analytic sample where the MPCP group is defined as all voucher students in 2008-09—prior to implementation of the performance measurement policy. In effect, we are estimating an intention-to-treat (ITT) parameter as opposed to the treatment-on-the-treated (TOT) parameters we estimated above. If selective attrition were responsible for the observed achievement increases, then we would expect the ITT estimates to be smaller than the TOT estimates. Table 4 provides the ITT parameter estimates. For both reading and math, the results in Table 4 are strikingly similar to those under the 2010-11 columns in Table 3. Although not fully definitive, this is strong evidence that the achievement gains in voucher schools were not driven by policy-induced systematic exits from the voucher program.¹³

[Insert Table 4 about here]

VI.) Discussion and Conclusion

This paper provides direct empirical evidence for the effect of implementing a performance measurement system on private providers of a public good or service. Specifically, drawing on student outcome measures generated by a study of the Milwaukee Parental Choice Program—the nation’s oldest and largest urban school voucher system—we estimate the effect

¹³ A related possibility is that MPCP schools responded to the impending performance measurement regime not by pushing students away from their schools entirely but by retaining low-performers in earlier grades. That response would represent a sort of “middle ground” between no strategic response to accountability and a targeted attempt to dramatically change the testing population by counseling out poor performers, particularly for the many MPCP schools dependent on each voucher student for survival. We have some anecdotal evidence that, generally, some MPCP schools are hesitant to retain children for reasons similar to their hesitation to label children as special needs, described above. If true, a response that included new retention efforts would represent a shift from past behavior. In our data, only 1 percent of MPCP students were ever retained from 2008 to 2010, compared to 5 percent of students in MPS, and the first year under the policy does not appear to have demonstrated any increase in the probability that a student was retained. In addition, estimates of (2) and (3) controlling for whether a student was retained have little impact on the results above, as we would expect with such low retention probabilities in the first place.

of a performance measurement system requiring private schools to 1) administer standardized tests to all voucher students in grades 3-8 and 10 and 2) submit the results to the state's supervising agency for review, analysis, and public reporting. Results of our analyses demonstrate that the performance measurement system produced gains in student achievement and that the magnitudes of these gains are substantial; they range from about 0.1 to nearly 0.2 standard deviations depending on the subject and analytical approach. The findings have implications for research, policy, and practice.

At a basic level, the results are relevant to public sector efforts to hold private providers of public goods and services accountable. The movement toward third-party provision of public goods shows no sign of abating. To the contrary, the portfolio of public goods and services for which the third-party entities are responsible for delivering continues to grow in both number and complexity. Such growth will increase the challenges that governmental entities face in ensuring that such goods are delivered in a high-quality and efficient manner. To meet these challenges, governmental actors will need to consider a variety of possible accountability mechanisms and work to identify and implement the optimal approach. Historically, it has often been assumed that market forces—primarily choice and competition—would provide the requisite accountability. In many cases this assumption is likely to be correct, but the results presented above demonstrate that this may not always be the case.

Consequently, prior to relying solely on market forces to hold third-party vendors accountable, governmental actors should evaluate whether the relevant goods and services will be provided in a manner meeting the assumptions of perfectly competitive markets, which is the theoretical undergirding of market-based accountability. Several studies, including this one, illustrate that several factors can result in violation of the requisite assumptions. When such

violations occur, it is incumbent upon the public sector to identify and implement policies or procedures designed to correct the relevant violations. In the case of the MPCP—although the market served as the sole accountability mechanism for the first 20 years of the program—it is likely that the assumption of perfect information was violated throughout this period; parents did not possess substantial information about the quality of MPCP schools. Intentionally or not, the performance measurement system had the potential to correct this violation. As a result, it is perhaps not surprising that implementation of the policy resulted in improved achievement outcomes.

Given the preceding discussion, it is clear that the results have implications for the design and implementation of school choice policies. Across the country, legislation has produced a significant expansion in the number and breadth of school choice programs in recent years. Accompanying this expansion have been debates over the mechanisms through which schools participating in these program should be held accountable. One popular position in this debate holds that all accountability should be completely market based, with no formal testing or accountability requirements. This view was expressed succinctly by the spokesman for Louisiana Governor Bobby Jindal when explaining why private schools in the state’s expanded school voucher program would not be subjected to any state-based accountability system. He stated, “Parents are the best accountability measures, not government” (Economist 2012). Our results, coupled with evidence from other studies (e.g., Hastings and Weinstein 2008), suggest that parents are able to best hold schools accountable if they possess the information necessary to

make fully informed schooling decisions.¹⁴ Government is the natural candidate for ensuring that parents have access to such information.

Along with their implications for practice and policy, the results presented above illuminate a number of additional research questions. The findings in this paper were in the context of a large-scale private school voucher program, but there are reasons to expect that the general accountability approach employed in this context—performance measurement coupled with market-based accountability—will be similarly effective in comparable contexts. Specifically, such an accountability approach may be effective in contexts where the public sector does not contract with a single private provider of a good or service but rather approves a set of providers to offer services. Examples of such contexts include publicly subsidized child care or the supplemental educational services provision of the NCLB law. Future research would do well to explore whether the general accountability approach analyzed in this paper is effective in other similar contexts.

This paper clearly demonstrates that implementation of the performance measurement system affecting MPCP schools produced significant gains in student achievement. Less clear are the precise mechanisms through which these gains were generated. Given that the gains are observed in the first year after implementation of the performance measurement policy, it is unlikely that the achievement gains are attributable to student transfers into more effective schools. Seemingly more likely is a situation where MPCP schools anticipate that families will use the measured achievement outcomes in future schooling decisions and thus take actions—such as increased emphasis on the WKCE or better alignment between curriculum and the test—

¹⁴ It should be noted that the competitive pressures that are supposed to be created through school choice policies may not require that all parents have high levels of information. Schneider et al. (1998) find that a small number of informed parents (“marginal choosers”) can create some level of competition amongst schools.

designed to maximize student performance on the WKCE. If accurate, this is suggestive of a scenario where institutional efforts generate short-term quality improvements and longer-term improvements may occur through families' use of the newly provided information to make better schooling decisions (see Hastings and Weinstein 2008). In short, there is the possibility that implementation of a performance measurement system affecting private vendors will induce a long-term cycle of improvement.

Unfortunately, this possibility highlights one limitation of our data. The analyses conducted above were made possible by the fact that an independent study had collected pre-policy outcome measures. But just as that study provided such an advantage, we are likewise limited by its duration, and 2010-11 is the last year of student-level data available to us. We are therefore unable to consider whether implementation of the performance measurement system has a longer-term impact on private school achievement in Milwaukee. Our inability to estimate any long-term impacts, however, does not detract from the primary contribution of this paper: the demonstration that a performance measurement system for third-party private vendors can produce substantial improvements in important outcomes.

References

Belfield, Clive R. 2005. *The Evidence on Education Vouchers: An Application to the Cleveland Scholarship and Tutoring Program*. Paper presented at the Conference on Education and Economic Development.

Bevan, Gwyn, and Christopher Hood. 2006. What's Measured is What Matters: Targets and Gaming in the English Public Health Care System. *Public Administration* 84: 517-538.

Bridge, Gary. 1978. Information Imperfections: The Achilles' Heel of Entitlement Plans. *Social Review*, 3, 504-529.

Brown, Trevor L. and Matthew Potoski. 2003. Managing Contract Performance: A Transaction Costs Approach. *Journal of Policy Analysis and Management* 22: 275-297.

Brown, Trevor L., and Matt Potoski. 2006. Contracting for Management: Assessing Management Capacity Under Alternative Service Delivery Arrangements. *Journal of Policy Analysis and Management* 25: 323-346.

Brown, Trevor L., Matthew Potoski, and David Van Slyke. 2006. Managing Public Service Contracts. *Public Administration Review* 64: 656-668.

Brown, Trevor L., Matthew Potoski, and David M. Van Slyke. 2010. Contracting for Complex Products. *Journal of Public Administration Research and Theory*, 20: i41-i58.

Carnoy, Martin, and Susanna Loeb. 2002. Does External Accountability Affect Student Outcomes? A Cross-State Analysis. *Educational Evaluation and Policy Analysis* 24: 305-331.

Courty, Pascal, Do Han Kim, and Gerald Marschke. 2011. Curbing Cream-Skimming: Evidence on Enrollment Incentives. *Labour Economics* 18: 643-655.

Dee, Thomas S., and Brian Jacob. 2011. The Impact of No Child Left Behind on Student Achievement. *Journal of Policy Analysis and Management* 30: 418-446.

Dias, Janice Johnson and Steven Maynard-Moody. 2007. For-Profit Welfare: Contracts, Conflicts, and the Performance Paradox. *Journal of Public Administration Research and Theory* 17: 189-211.

Dicke, Lisa A. 2002. Ensuring Accountability in Human Services Contracting: Can Stewardship Theory Fill the Bill? *The American Review of Public Administration* 32: 455-470.

Donahue, John D. 1989. *The Privatization Decision: Public Ends, Private Means*. New York: Basic Books.

Economist. 2012. "Governor Jindal extends his reach: Reforms that have transformed New Orleans are applied to the state." April 21 issue, p. 46.

Ewalt, Jo Ann G., and Edward T. Jennings. 2004. Administration, Governance and Policy Tools in Welfare Policy Implementation. *Public Administration Review* 64: 449-462.

Friedman Milton 1962. "The Role of Government in Education" in Milton Friedman and Rose D. Friedman, *Capitalism and Freedom*. Chicago. University of Chicago Press, pp. 85-107 (fortieth anniversary edition, copyright renewed 2002).

Gansler, Jacques S. 2006. Moving toward market-based government: The changing role of government as the provider. In *Competition, choice and incentives in government programs*, by J.M Kamensky and A. Morales, 37-120. Oxford: Rowan and Littlefield.

Greene, Jay P., Paul E. Peterson, and Jiangtao Du. 1999. Effectiveness of School Choice: The Milwaukee Experiment. *Education and Urban Society* 31: 190-213.

Hanushek, Eric A., and Margaret E. Raymond. 2005. Does School Accountability Lead to Improved Student Performance? *Journal of Policy Analysis and Management* 24: 297-327.

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

Heinrich, Carolyn. 2007. False of Fitting Recognition? The Use of High Performance Bonuses in Motivating Organizational Achievements. *Journal of Policy Analysis and Management* 26: 281-304.

Heinrich, Carolyn. 2010. Third-Party Governance under No Child Left Behind: Accountability and Performance Management Challenges. *Journal of Public Administration Research and Theory* 20: i59-i80.

Heinrich, Carolyn J. In press. How Credible is the Evidence and Does it Matter? An Analysis of the Program Assessment Rating Tool. *Public Administration Review*.

Heinrich, Carolyn, and Gerald Marschke. 2010. Incentives and Their Dynamics in Public Sector Performance Management Systems. *Journal of Policy Analysis and Management* 29: 183-208.

Heinrich, Carolyn, and Youseok Choi. 2007. Performance-Based Contracting in Social Welfare Programs. *American Review of Public Administration* 37: 409-435.

Hill, Carolyn J. 2006. Casework Job Design and Client Outcomes in Welfare-to-Work Offices. *Journal of Public Administration Research and Theory* 16: 263-288.

Hood, Christopher. 2006. Gaming in Targetworld: The Targets Approach to Managing British Public Services. *Public Administration Review* 66: 515-521.

Howell, William G., Paul E. Peterson, Patrick J. Wolf, and David E. Campbell. 2006. *The Education Gap: Vouchers and Urban Schools*. Washington, DC: Brookings Institution Press.

Hoxby, Caroline M. 2003. School Choice and School Productivity: Could School Choice Be a Tide that Lifts All Boats? In *Economics of School Choice* pp. 287-341. Caroline M. Hoxby (Ed.). Chicago, IL: University of Chicago Press.

- Jacob, Brian. 2005. Accountability, Incentives and Behavior: Evidence from School Reform in Chicago. *Journal of Public Economics*. 89: 761-796.
- Jacob, Brian and Steven Levitt. 2003. Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating. *Quarterly Journal of Economics*. 118: 843-877.
- Johnston, Jocelyn M., and Barbara S. Romzek. 1999. Contracting and Accountability in State Medicaid Reform: Rhetoric, Theories and Reality. *Public Administration Review* 59: 383-399.
- Kettl, Donald. 1993. *Sharing Power*. Washington, DC: Brookings Institution Press.
- Kettl, Donald. 2002. *The Transformation of Governance: Public Administration for 21st Century America*. Baltimore: Johns Hopkins University Press.
- Kisida, Brian, Laura I. Jensen, and Patrick J. Wolf. 2011. *The Milwaukee Parental Choice Program: Descriptive Report on Participating Schools 2009-2010*. School Choice Demonstration Project Report #27, University of Arkansas.
- Lamarche, Carlos. 2008. Private School Vouchers and Student Achievement: A Fixed Effects Quantile Regression Evaluation. *Labour Economics*. 15: 575-590.
- Lara, Bernardo, Alejandra Mizala, and Andrea Repetto. 2011. The Effectiveness of Private Voucher Education: Evidence from Structural School Switches. *Educational Evaluation and Policy Analysis* 33: 119-137.
- Light, Paul C. 1999. *The True Size of Government*. Washington, DC: Brookings Institution Press.
- Manna, Paul. 2011. *Collision Course: Federal Education Policy Meets State and Local Realities*. Washington, DC: CQ Press.
- Meier, Kenneth J., and Laurence J. O'Toole. 2002. Public Management and Organizational Performance: The Effect of Managerial Quality. *Journal of Policy Analysis and Management* 21: 629-643.
- Meier, Kenneth J., and Laurence J. O'Toole. 2003. Public Management and Educational Performance: The Impact of Managerial Networking. *Public Administration Review* 63: 689-699.
- Metcalf, Kim K., Stephen D. West, Natalie A. Legan, Kelli M. Paul, and William J. Boone. 2003. *Evaluation of the Cleveland Scholarship and Tutoring Program: Summary Report 1998-2002*. Bloomington, IN: Indiana University.
- Milward, H. Brinton, and Keith G. Provan. 2000. Governing the Hollow State. *Journal of Public Administration Research and Theory* 10: 359-379.
- Milwaukee Journal Sentinel*. 2011. Testing Must Remain. April 6. 12A.
- Moynihan, Donald P., and Stephane Lavertu. In press. Does Involvement in Performance Management Routines Encourage Performance Information Use? Evaluating GRPA and PART. *Public Administration Review*.

- Osborne, D. and T. Gaebler. 1992. *Reinventing Government: How the Entrepreneurial Spirit Is Transforming the Public Sector from Schoolhouse to Statehouse, City Hall to the Pentagon*. Reading, MA: Addison-Wesley.
- Peters, B. Guy, and John Pierre. 1998. Governance Without Government? Rethinking Public Administration. *Journal of Public Administration Research and Theory* 8:223-242.
- Prager, Jones and Swati Desai. 1996. Privatizing Local Government Operations. *Public Productivity and Management Review*. 20: 185-203.
- Radin, Beryl A. 2000. The Government Performance and Results Act and the Tradition of Federal Management Reform: Square Pegs in Round Holes? *Journal of Public Administration Research and Theory* 10: 111-135.
- Richards, Erin and Amy Hetzner. 2011. Choice Schools Not Outperforming MPS. *Milwaukee Journal Sentinel*. March 29. 1A, 7A.
- Romzek, Barbara S., and Jocelyn M. Johnston. 2005. State Social Services Contracting: Exploring the Determinants of Effective Contract Accountability. *Public Administration Review*: 436-449.
- Rouse, Cecilia E. 1998. Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program. *Quarterly Journal of Economics*. 113: 553-602.
- Salamon, Lester. 1989. *Beyond Privatization: The Tools of Government Action*. Washington, DC: Urban Institute Press.
- Savas, E.S. 1987. *Privatization: The Key to Better Government*. Chatham, NJ: Chatham House.
- Savas, E.S. 2000. *Privatization and Public-Private Partnerships*. Chatham, NJ: Chatham House.
- Schneider, Mark, Paul Teske, Melissa Marshall, and Christine Roch. 1998. Shopping for Schools: In the Land of the Blind, The One-Eyed Parent May Be Enough. *American Journal of Political Science* 42: 769-793.
- Stalebrink, Odd J. 2009. National Performance Mandates and Intergovernmental Collaboration: An Examination of the Program Assessment Rating Tool (PART). *The American Review of Public Administration* 39: 619-639.
- Stoker, Gerry. 1998. Governance as Theory: Five Propositions. *International Social Science Journal*. 155: 17-28.
- Van Slyke, David M. 2003. The Mythology of Privatization in Contracting for Social Services. *Public Administration Review*. 63: 296-315.
- Van Slyke, David M. 2007. Agents or Stewards: Using Theory to Understand the Government-Nonprofit Social Service Contracting Relationship. *Journal of Public Administration Research and Theory* 17: 157-187.

Witte, John F. 2000. *The Market Approach to Education; An Analysis of America's First Voucher Program*. Princeton, NJ: Princeton University Press.

Witte, John F., Patrick J. Wolf, Joshua M. Cowen, David J. Fleming, and Juanitia Lucas-McLean. 2008. *MPCP Longitudinal Educational Growth Study Baseline Report*. University of Arkansas Education Working Paper Archive.

Wolf, Patrick, Brian Kisida, Babette Guttmann, Lou Rizzo, and Nada Eissa. 2011. The Evaluation of the DC Opportunity Scholarship Program: A Summary of Experimental Impacts. In *School Choice and School Improvement: Research in State, District and Community Contexts*, by Mark Berends, Marissa Cannata and Ellen Goldring. Cambridge, MA: Harvard Education Press.

Yang, Kaifeng, and Gary VanLandingham. In press. How Hollow Can We Go? A Case Study of Florida's Efforts to Outsource Oversight of Privatized Welfare Services. *The American Review of Public Administration*.

Yang, Kaifeng, Jun Yi Hsieh, and Tzung Shiun Li. 2009. Contracting Capacity and Perceived Contracting Performance: Nonlinear Effects and the Role of Time. *Public Administration Review* 69: 681-696.

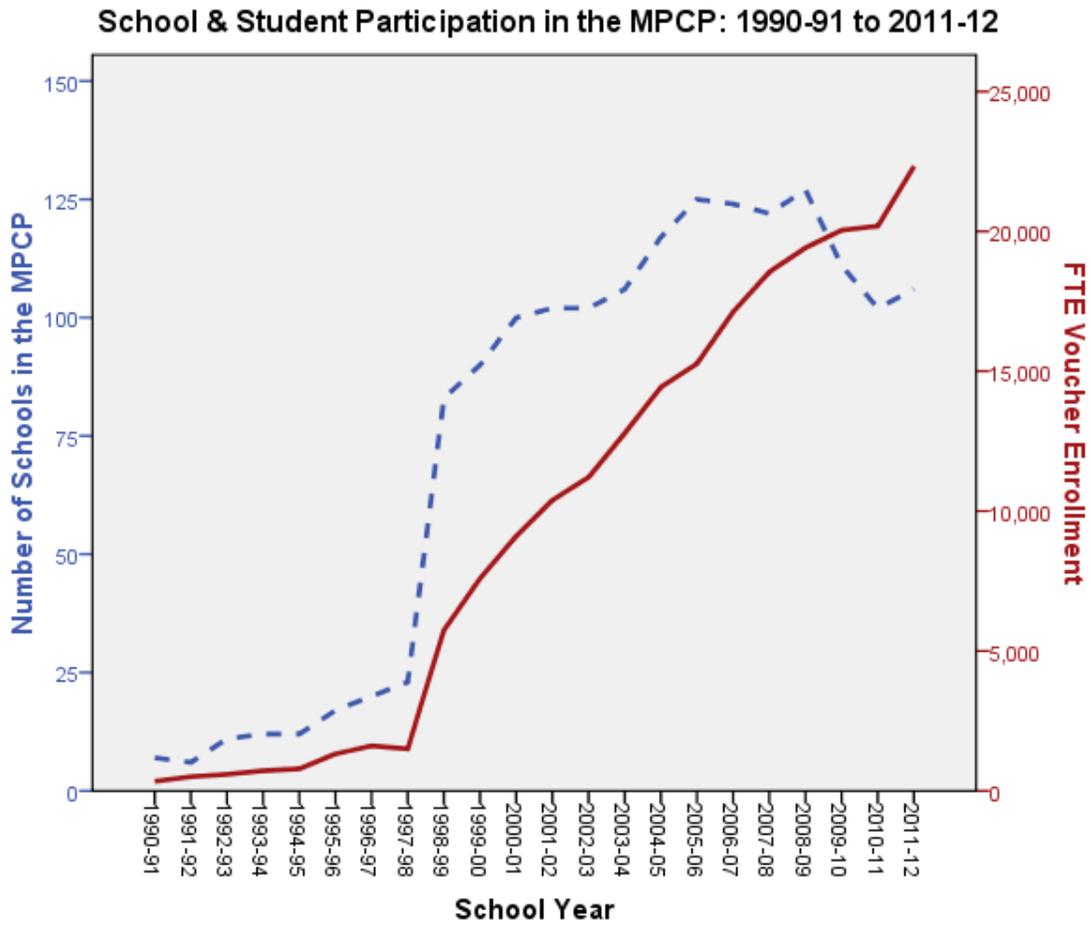
Tables and Figures

Table 1. School and student participation in the MPCP: 1990-91 to 2011-12

Year	Number of Schools	FTE Enrollment
1990-91	7	337
1991-92	6	504
1992-93	11	591
1993-94	12	718
1994-95	12	786
1995-96	17	1,320
1996-97	20	1,606
1997-98	23	1,501
1998-99	83	5,740
1999-00	90	7,596
2000-01	100	9,104
2001-02	102	10,391
2002-03	102	11,209
2003-04	106	12,788
2004-05	117	14,427
2005-06	125	15,274
2006-07	124	17,126
2007-08	122	18,550
2008-09	127	19,414
2009-10	111	20,042
2010-11	102	20,189
2011-12	106	22,328

Source: Wisconsin Department of Public Instruction

Figure 1:



Source: Wisconsin Department of Public Instruction

Table 1: Summary of MPCP Accountability Changes, 2006-2011

Effective Date	Description
Fall 2006	Program expanded; Expansion authorizes a new independent five-year evaluation comparing a “representative” panel of voucher students to a “comparable” panel of public school students (2005 Wisconsin Act 125); Evaluators provided access to selected panelists in each school; individual panelists may opt out of study.
Fall 2009	Schools required to pay additional participation fees and new schools required to obtain accreditation via Marquette University’s Institute for Transformation of Learning (ITL) (2009 Wisconsin Act 28)
Fall 2010	<ul style="list-style-type: none"> • Schools required to test all students in grades 3-8 and 10 with state WKCE examination in reading and math. • Schools required to test all students in grades 4, 8 and 10 with state WKCE examination in English/language arts, writing, science, and social studies. • Scores submitted by school to state; • Schools must develop and report individual academic standards for each subject; • All schools teachers required to be accredited • Yearly instructional hour minimums (1,050 for elementary and 1,137 hours for junior high school) set for all MPCP schools.
Fall 2010-Spring 2011	<ul style="list-style-type: none"> • Scores reported for each private school based on Fall 2010 results.

Figure 2. Mean MPCP Achievement (standardized by MPS mean): by Subject

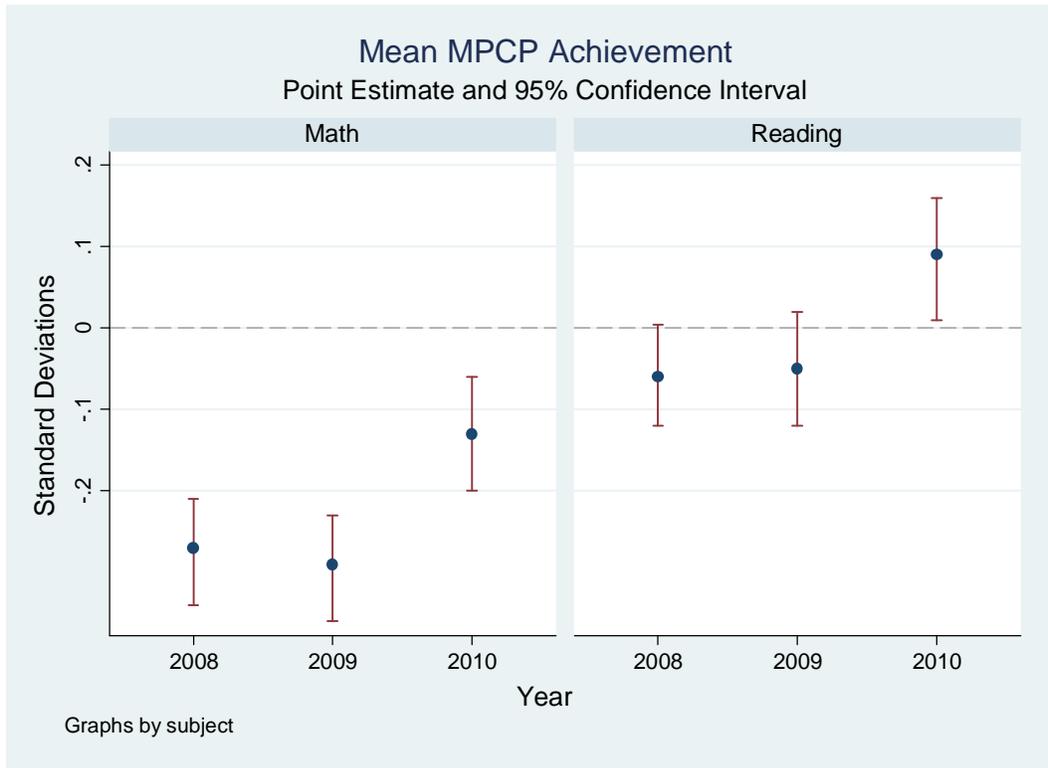


Figure 3: Standardized Mean MPCP-MPS Achievement Differences: by Subject

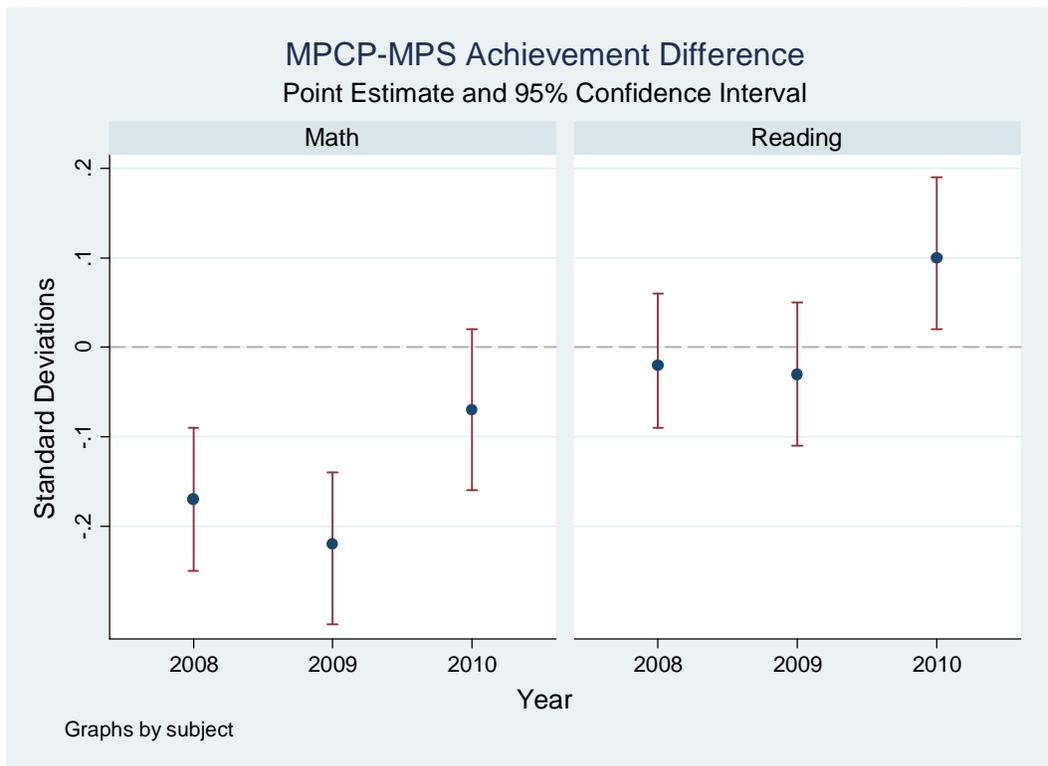


Table 2: Coefficients and Standard Errors from Estimation of Interrupted Time-Series Models of Student Achievement

VARIABLES	Reading		Math	
	Student FE	Student VA	Student FE	Student VA
2010 dummy	0.171*** (0.030)	0.187*** (0.036)	0.178*** (0.026)	0.218*** (0.036)
2009 dummy	-0.002 (0.026)	-0.023 (0.036)	-0.026 (0.023)	0.008 (0.037)
Reading t-1	—	0.566*** (0.035)	—	0.187*** (0.026)
Math t-1	—	0.227*** (0.029)	—	0.646*** (0.026)
Constant	-0.062*** (0.015)	0.364* (0.187)	-0.281*** (0.013)	0.497*** (0.159)
N (Student-Year)	2,005	1,478	2,002	1,475
R²	0.038	0.590	0.059	0.637

*** p<0.01, ** p<0.05, * p<0.10, robust clustered standard errors by student in parenthesis (per Rouse 1998; hypothesis tests are also robust to school-clustered errors). Student value-added models also include grade fixed effects and student demographics (race, gender, English language learners, free/reduced lunch and special needs indicators). Descriptive statistics for covariates available on request.

Table 3: Coefficients and Standard Errors from Estimation of Difference-in-Differences Models of Student Achievement

VARIABLES	Reading		Math	
	2009-10	2010-11	2009-10	2010-11
MPCP*Year	0.010 (0.036)	0.105*** (0.039)	-0.063* (0.035)	0.173*** (0.039)
Year	0.025 (0.019)	0.041** (0.019)	0.046** (0.019)	-0.023 (0.019)
MPCP	-0.202*** (0.038)	-0.196*** (0.041)	-0.354*** (0.040)	-0.424*** (0.039)
Constant	0.621*** (0.061)	0.797*** (0.098)	0.538*** (0.065)	0.831*** (0.123)
N (Student-Year)	6,501	5,551	6,506	5,543
R²	0.259	0.286	0.227	0.247

*** p<0.01, ** p<0.05, * p<0.10, robust clustered standard errors by student in parenthesis (per Rouse 1998; hypothesis tests are also robust to school-clustered errors). Models also include grade fixed effects and student demographics (race, gender, English language learners, free/reduced lunch and special needs indicators). Descriptive statistics for covariates available on request.

Table 4: Coefficients and Standard Errors from Estimation of Difference-in-Differences Models of Student Achievement—Intention-to-Treat Estimates

VARIABLES	Reading	Math
In MPCP*2010	0.103*** (0.036)	0.166*** (0.037)
2010 dummy	0.019 (0.018)	-0.042** (0.019)
In MPCP	-0.220*** (0.032)	-0.407*** (0.033)
Constant	0.267 (0.253)	0.465*** (0.153)
N (Student-Year)	6,981	6,969
R²	0.268	0.231

*** p<0.01, ** p<0.05, * p<0.10, robust clustered standard errors by student in parenthesis (per Rouse 1998; hypothesis tests are also robust to school-clustered errors). Models also include grade fixed effects and student demographics (race, gender, English language learners, free/reduced lunch and special needs indicators). Descriptive statistics for covariates available on request.