Privatizing Education In Philadelphia: Are Educational Management Organizations Improving Student Achievement?

Martha Abele Mac Iver and Douglas J. Mac Iver
Johns Hopkins University

Paper prepared for presentation at the annual meeting of the American Political Science Association, Philadelphia, PA. August 31 - September 3, 2006

DRAFT: Not for citation

The authors would like to thank Vaughan Byrnes, Suzanne Blanc, Allen Ruby, and Robert Balfanz for their help and comments on this article. The research reported here was supported by the Research on Learning and Education (ROLE) Program at the National Science Foundation, grant number 0411796.
Abstract

This paper reports initial findings from a study of middle grades educational reforms in Philadelphia. We use multilevel change models to analyze the impact on student mathematics and reading achievement of privatization through the use of educational management organizations (EMOs), taking account of the structural reforms (creation of new K-8 schools to replace selected middle schools) occurring simultaneously within the district. Overall, the longitudinal mathematics and reading achievement gains for students in EMO-managed schools were not larger than those for students in schools managed by the district. While students in long-established K-8 schools generally outgained students in middle schools, newly-established K-8s did not consistently outperform middle schools. There are hopeful signs, however, that broader systemic reforms -- such as district-wide increases in the quality and coherence of curriculum and professional development -- are bearing fruit: The 2004 and 2005 cohorts of eighth-graders showed math achievement gains across the middle grades that were much larger than those shown by the 2003 cohort. This finding was broad based, occurring in middle schools and K-8 schools, in privately-managed and district-managed schools.
Privatization has become a popular “arrow” in the school reform quiver, particularly since Chubb and Moe (1990) launched a veritable revolution by linking higher student achievement to lower levels of bureaucratic organization in schools. There was considerable methodological debate over the Chubb and Moe study (e.g., Sukstorf, Wells, & Crain, 1993), and though proponents of educational privatization can cite some additional empirical support from studies of academic achievement to support their case (e.g., Peterson, 1999), much of their advocacy is based more on philosophical confidence in market mechanisms and decentralization as principles rather than on convincing evidence of improved outcomes for children. Privatization advocates like Paul Hill and his associates (1997a, 1997b, 1997c, 1997d; 1998) and Finn (1991) describe how schools are constrained and burdened by bureaucratic requirements and practices that hinder students from making academic progress and offer privatization solutions with little if any solid evidence that they result in achievement gains. Similarly, many opponents of privatization rely more on rhetoric than hard evidence in their arguments (e.g., Bracey, 2003; Saltman, 2005).

The movement for privatization in education is philosophically related to two related trends: site-based management (SBM) and decentralization. Proponents of SBM and decentralized school districts argue that the people who are in the best position to create and sustain high-performing learning environments are the people who are closest to teaching and learning: classroom teachers, school administrators, and parents. The theory of site-based management is that the individuals who work in, run, and send their children to school will develop the most effective and lasting strategies for improvement if they are free from district constraints but held accountable for high standards (Bryk et al., 1998; Hill & Celio, 1998; Hill, Campbell, & Harvey, 2000; Mohrman & Wohlstetter, 1994; Odden & Hill, 1997; Ouchi, 2003; Wohlstetter, Mohrman, & Roberson, 1997). For example, Odden & Hill (1997) argued that the success of site-based management hinges on districts granting decision-making power to teachers and school leaders over their day-to-day operations, allocation of financial and human resources, curriculum, instructional methods, professional development, and teacher hiring and evaluation while holding the teachers and leaders to high performance standards and creating a variety of support systems to help them succeed. Authors advocating this approach typically point to high-performing organizations in other fields than education that nimbly and effectively use decentralized authority and decision-making to produce high-quality, cost-effective products and services. They also cite a number of schools and a few districts that have had some success with site-based management (Bryk et al., 1998; Palmaffy, 1998; Wohlstetter, Mohrman, & Robertson, 1997). But much of the evidence supporting site-based management is based on experiences at elementary schools. Even some of decentralization’s strongest proponents concede that site-based management has had less of a positive effect on large high schools and middle schools (e.g., Palmaffy, 1998; Wohlstetter, Mohrman, & Robertson, 1997). Overall, the research on the effectiveness of site-based management in improving student achievement is at best mixed (Leithwood & Menzies, 1998; Murphy & Beck, 1995). Providing greater educational choice for families is another resounding theme in the privatization debate, though not the focus of our current privatization study. As Chester Finn (1993) has put it, “The main argument for justice in the choice debate is to make it possible for the poor to do what the rich already do.” Besides expanding choices within the public school
system, recent education policy debates have emphasized the expansion of charter schools and the provision of vouchers for use at private schools as avenues for educational choice. Despite some positive effects of charter school expansion on parental satisfaction and teacher empowerment (Hess, Maranto, & Milliman, 2001; Hess, Maranto, & Milliman, 2000), there is not yet compelling evidence that charter schools do a better job educating students (Miron & Nelson, 2002; Murphy & Shiffman, 2002; Hess & Maranto, 2000; Wells et al., 1998). Based on evidence from Milwaukee, Cleveland, and Edgewood, Texas, Hess (2002) also argues that vouchers offering school choice to some families have not made the districts more productive or efficient, as the theoretical case for privatization had claimed. Again, as Belfield and Levin (2005) point out, the evidence produced by those on opposite sides of the debate has an uncanny, extremely high correlation with the positions espoused by the authors.

Our focus in the current study, the privatization of instructional delivery and educational management, has mixed reviews when one examines the research studies to date. Murphy and colleagues (1998) emphasized the blend of positive and negative indicators in both the available evidence and the theoretical arguments, concluding that “contracting out is neither a magic wand nor an inherently flawed policy mechanism” (p. 85). Several studies note how the public controversy surrounding privatization efforts has affected outcomes of those efforts (Doughty, 1997; Hunter, 1995; Leak & Williams, 1997; Walsh, 1995). Some independent evaluations have noted such achievements as lower student-teacher ratios, increased teacher autonomy, parental empowerment, increased attendance, greater student access to computers, and better-maintained facilities in schools under private management (GAO, 1996; Edwards, 1997; Peeler and Parham, 1994). While preliminary results from testing in a couple of privatization experiments were cautiously positive or mixed (Blasik & Hodges, 1997; Maryland Higher Education Commission, 1997; Jennings & Maruyama, 2000; Mac Iver & Stringfield, 2000), most studies have noted the lack of evidence of student achievement gains under privatization models (Ascher, 1996; Ascher, Fruchter, & Berne, 1996; GAO, 1996; Richards, 1996; Ligas, 1997; Ligas, 1998). Though the number of privately managed public schools has grown dramatically over the past few years, there were still just over 400 such schools nationwide during the 2002-03 school year, and that recent study comparing student outcomes in privately managed and traditional schools in six cities found mixed results (GAO, 2003).

It is important to distinguish between for-profit and not for-profit Education Management Organizations (EMOs) (Levin, 2002). Several studies of the achievement effects of management by the for-profit EMO, Edison Schools, have now emerged, with mixed results. The most recent study, conducted by RAND (Gill et al, 2006), emphasizes the importance of analyses over an extended period of time, since studies of Edison Schools in early implementation stages (e.g., Miron & Applegate, 2000; AFT, 1998, 2000, 2003; Dryden, 2004) did not find the same positive achievement effects that the RAND study began to find after 4 or 5 years of implementation. The RAND study also notes that not all the positive Edison effects are statistically significant, and that sample sizes for analyses of 4 and 5 years of implementation are small.

Additional longitudinal studies of educational privatization, including for-profit EMOs, are certainly needed. Our study examines the achievement effects of recent privatization efforts in Philadelphia, where a “state takeover” of the School District of Philadelphia in December 2001 led to the appointment of a new School Reform Commission (SRC) that embraced
privatization and mandated restructuring as reform strategies (Maranto, 2005; Useem & Balfanz, 2002). The state and city negotiated increased expenditures for the district as part of a “friendly takeover,” with the state contributing an additional $75 million and the city promising to contribute an additional $45 million. The SRC Among the district’s 264 schools:

- 45 schools were designated to be run by private entities, both non-profit and for-profit groups
- 21 schools were “restructured,” singled out to be in a separate subdistrict and to implement mandated reforms
- 4 schools were instructed to become self-governing public charter schools
- 16 low-performing schools that had shown sustained improvement were given additional funds to continue their successful reforms (but were not placed under a new governance structure)
- The remaining 178 schools experienced little change in governance.

In July 2002 the SRC hired the former head of Chicago’s school system, Paul Vallas, to serve as the district’s new CEO. Vallas made it clear that he was in charge of all the schools, serving notice that he would cancel contracts for education management organizations (EMOs) and charter school managers whose schools performed. He led the district to implement ambitious and comprehensive reforms that emphasize coherent districtwide instructional programs in reading and mathematics (and later science) as well as corresponding professional development and school-based coaching (Neild, Useem, Travers, & Lesnick, 2003). The Vallas administration simultaneously launched a “Campaign for Human Capital” designed to improve teacher quality by increasing the number of qualified applicants for the district’s teaching vacancies (through expanded incentives, outreach, and marketing, and through reforms in the hiring process). The reform period was thus characterized by decisive action at the district office level as well as by privatized management of some district schools.

Our study focuses on the effect of this privatization effort on students in the middle grades (grades 5-8), an important and often neglected analytical gradespan. Though it is high school reform that tops the agenda of the National Governors Association, the roots of the dropout crisis can be traced back to sixth grade, when students’ marks, attendance, and behavior falter (Balfanz & Herzog, 2006). Because middle grades reform has lagged far behind reforms at the elementary level, a focus on this crucial educational stage before high school is important from a policy perspective.

A map of the middle grades terrain in Philadelphia is an important starting point as we describe our study. Of the 99 Philadelphia schools serving eighth grade students in 2002-03, a total of 26 were under EMO governance. More than half (14) of these were run by Edison, Victory, Chancellor Beacon, Foundations, Universal, and Temple University managed two schools each. There was a charter school and one run by the University of Pennsylvania. Chancellor Beacon lost its contract with the district the following year. In 2002-03, about 30 percent of the 8th grade students with test scores were attending the 26 middle grades schools run by an EMO.

The introduction of EMO-run schools in Philadelphia coincided with a movement to increase the number of K-8 schools (reducing the number of students attending middle schools).
Several Philadelphia schools began converting to K-8s in the late 1990s. By 2002-03, there were eight new K-8s (six of which had eighth grade by 2000 or 2001). Almost 30 more Philadelphia schools have been in the process of converting to K-8s since 2003.

Both the privatization and K-8 conversion reforms were undertaken with the expectation that student achievement would improve under these new structures. The analyses undertaken in this report seek to test the underlying hypotheses of these reform efforts. Do students at schools managed by EMOs make significantly greater achievement improvement during the middle grades than students at other schools? Do students at the newly converted K-8s make significantly greater improvement than students in the remaining middle schools? And what are the interaction effects between governance and gradespan structures?

We believe there are theoretical reasons to doubt that there will be significant positive effects of these reforms, especially in their early years. Given the limited experience that Philadelphia’s EMOs had with urban high-poverty middle grades schools, we do not expect the initial achievement gains in the privately managed middle grades schools to be significantly greater than in district-managed schools. Neither do we expect that mere conversion to a K-8 structure will result in higher achievement gains. These reforms do not directly influence the manipulable variables most closely associated with student achievement: teacher quality and the quality and coherence of curriculum, instruction, and professional development.

Data and Descriptive Analyses

Our analyses focus on the both the reading and mathematics portions of the state-administered standardized achievement tests. Students take the state-administered Pennsylvania System of School Assessment (PSSA) just before entering the middle grades (in the spring of 5th grade) and at the end of the middle grades (in the spring of eighth grade).

Median state percentile scores for all Philadelphia eighth-graders taking the PSSA mathematics test remained relatively stable from 1995 to 2000 and then began to increase. The greatest increase occurred between 2003 and 2004, the year that CEO Vallas mandated a system-wide implementation of a completely specified core curriculum, pacing guides and curriculum-based assessments every six weeks, together with teacher coaches and increased time for professional development.

The schools that were assigned to EMO management, beginning in fall 2002, had significant pre-existing differences from those not assigned to EMOs. First, the EMO schools were all high-poverty schools. Second, even when we compare EMO schools just to other high-poverty schools, there remain some pre-existing differences between the EMO and district-managed schools (Table 1). The average poverty rate was significantly higher in Edison schools and non-significantly higher in other EMO schools than in district-managed high-poverty schools. Average fifth grade PSSA scores in math were significantly lower at schools that would become EMO schools than at the high-poverty schools that were not assigned to an EMO. Schools that would become EMO schools had significantly higher percentages of non-certified teachers and somewhat (but not significantly) lower teacher retention rates the year before they were assigned to EMO status. We control for these prior differences in the analyses described
below. There were also large prior differences among school types (middle schools, established K-8 schools, and new K-8 schools) that we summarize in Table 2 below.

The analyses that follow include longitudinal data from the first three cohorts of students to reach eighth grade in schools managed by EMOs (fifth graders in 1999-00, 2000-01, and 2001-02, who reached eighth grade in 2002-03, 2003-04, and 2004-05). The first of these cohorts began fifth grade during the last year of the Children Achieving initiative under David Hornbeck, had sixth and seventh grades during the leadership transition years of 2000-01 and 2001-02, and were the first to experience eighth grade instruction in an EMO school. The second of these cohorts began fifth grade in the leadership transition year of 2000-01, and had seventh and eighth grade instruction in schools that were either governed by EMOs or the district. The third of these cohorts began fifth-grade in the last year prior to privatization and had all their middle grades instruction in schools that were either governed by EMOs or the district.

Modeling Student Achievement Growth

We use multilevel change models (Raudenbush & Bryk, 2002; Seltzer, Choi, & Thum, 2003) to estimate the impact of the EMO and new K-8 reforms on students’ achievement growth during the middle grades, since students are nested within schools. Student level data files were provided by the School District of Philadelphia. School level data files were built using publicly available school level variables as well as some variables constructed from teacher level files provided by the School District of Philadelphia. With just two time points per student on the state tests, we model initial status and total growth rather than a full “growth curve.” In estimating a 3-level growth model, we specify a within-student model, a between-student model, and a school level model. There were two records for each student on each test, fifth grade and eighth grade mathematics and reading scale scores on the Pennsylvania System of School Assessment (PSSA).

At level 1, within students, we model students’ achievement scores as a function of grade: \( Y = P_0 + P_1(EIGHTH) + \epsilon \). \( EIGHTH \) is a dummy variable, coded “0” if the achievement score is from a student’s fifth-grade year, and a “1” if the score is from a students’ eighth-grade year. Thus, the coefficient for the intercept (\( P_0 \)) represents students’ prior achievement in the spring of fifth grade and the slope coefficient for grade (\( P_1 \)) represents students’ cumulative achievement growth between the spring of fifth and the spring of eighth grade.

At level 2, the between-student model, we take account of differences in prior achievement between students as they enter the middle grades and model differences in achievement growth during the middle grades that are associated with characteristics and experiences that vary between students who attend the same school. Dummy variables were constructed for demographic characteristics (gender and ethnicity) and special educational status (special education and English language learners), with the named characteristic coded as 1. Individual level data on student’s socioeconomic status was not available from district records.
and so students were assigned the poverty level of their 5th grade school as a surrogate measure of family poverty level. In modeling student achievement growth, we also included a variable (SAMESCH) distinguishing those who had 3 years (grades 6 to 8) at the same school (coded 1) from those who did not (coded 0). We specified the following between-students model:

\[
P_0 = B_{00} + B_{01}(\% \text{ LOW INCOME STUDENTS IN 5TH GRADE SCHOOL}) + B_{02}(\text{SPECIAL EDUCATION}) + B_{03}(\text{ENGLISH LANGUAGE LEARNER}) + B_{04}(\text{ASIAN}) + B_{05}(\text{HISPANIC}) + B_{06}(\text{CAUCASIAN}) + B_{07}(\text{FEMALE}) + R_0
\]

\[
P_1 = B_{10} + B_{11}(\text{FEMALE}) + B_{12}(\text{ATTENDED SAME SCHOOL FOR 6TH to 8TH}) + B_{13}(\text{ASIAN}) + B_{14}(\text{HISPANIC}) + B_{15}(\text{CAUCASIAN}) + B_{16}(\text{SPECIAL EDUCATION}) + B_{17}(\text{ENGLISH LANGUAGE LEARNER})
\]

At Level 3, the school level, we tested for interactive effects of the management and grade span interventions on students’ achievement growth in mathematics (Table 3) or reading (Table 4), controlling for differences between schools that already existed (in 01-02) prior to the assignment of schools to be run by an EMO or not (prior percent of non-certified teachers), as well as school poverty level (% of students eligible for free or reduced price lunch) and average achievement of each school’s incoming sixth-grade cohort (average math or reading NCE score on the PSSA math at the end of fifth grade). Each school’s management (EDISON or OTHER EMO) and grade span (OLD K8 or NEW K8) were coded as dummy variables (with named type equal to 1, and unnamed types – district-managed schools and middle schools -- equal to 0) and a set of product variables (such as EDISON × OLDK8) denoted each school’s cross-classification. We specified the following between-schools model:

\[
B_{00} = G_{000} + U_{00}
\]
\[
B_{01} = G_{010} + U_{01}
\]
\[
B_{02} = G_{020} + U_{02}
\]
\[
B_{03} = G_{030} + U_{03}
\]
\[
B_{04} = G_{040} + U_{04}
\]
\[
B_{05} = G_{050} + U_{05}
\]
\[
B_{06} = G_{060} + U_{06}
\]
\[
B_{07} = G_{070} + U_{07}
\]
\[
B_{10} = G_{100} + G_{101}(\text{NEW K-8}) + G_{102}(\text{OLD K-8}) + G_{103}(\text{EDISON}) + G_{104}(\text{OTHER EMO}) + G_{105}(\text{EDISON} \times \text{NEW K-8}) + G_{106}(\text{EDISON} \times \text{OLD K-8}) + G_{107}(\text{OTHER EMO} \times \text{OLD K-8}) + G_{108}(\text{SCHL’S % FRL}) + G_{109}(\text{SCHL’S PRIOR % OF NON-CERTIFIED TCHRS}) + G_{1010}(\text{AVG INCOMING MATH OR READING NCE @ END OF 5TH GRADE}) + U_{10}
\]
\[
B_{11} = G_{110} + U_{11}
\]
\[
B_{12} = G_{120} + U_{12}
\]
\[
B_{13} = G_{130} + U_{13}
\]
\[
B_{14} = G_{140} + U_{14}
\]
\[
B_{15} = G_{150} + U_{15}
\]
\[
B_{16} = G_{160} + U_{16}
\]
\[
B_{17} = G_{170} + U_{17}
\]

Figures 1, 2, & 3 show the adjusted mean math gains in each type of school for the first three cohorts of students to complete 8th-grade after privatization was implemented and Table 3
shows the HLM estimates for these cohorts. Omnibus testing of the product term coefficients indicate there were significant Management X Gradespan interaction effects (p < .01) on students’ math gains in each cohort (2003 cohort $\chi^2(3) = 61.2$, 2004 cohort $\chi^2(3) = 35.3$, 2005 cohort $\chi^2(3) = 16.5$). The nature of these effects can be seen in the figures and in coefficients from the last seven rows in the table.

[Table 3 and Figure 1 about here]

Consider the EDISON and OTHEREMO coefficients in Table 3. These coefficients compare the 3-year math achievement gains for eighth-graders in middle schools that have been under Edison or other EMO management for one year (2003 cohort), two years (2004 cohort) or 3 years (2005 cohort) with those in district-managed middle schools. These gains are depicted in Figure 1. Gains for eighth-graders in Edison-managed middle schools were not significantly different than in district-managed middle schools. However, eighth-graders in district-managed middle schools outgained those in other EMO-managed middle schools by 26.7 points (p < .001) in the 2003 cohort, by 26.9 points (p < .1) in the 2004 cohort, and by 33.8 points in the 2005 cohort (p < .05).

The OLDK8 coefficients indicate that eighth-graders in district-managed older K-8 schools outgained those in district-managed middle schools by about 33 points in the 2003 and 2004 cohorts and by 60 points in the 2004 cohort (see Figure 2). The EDISON × OLDK8 coefficients indicate that the small number of eighth-graders in Edison-managed old K-8 schools (e.g., n = 35 in Cohort 2003) achieved 3-year gains that were significantly higher than those in district-managed old K-8 schools in the 2003 cohort but not in the 2004 and 2005 cohorts. In contrast, the OTHEREMO × OLD K8 coefficients indicate that the small number of eighth-graders in non-Edison EMO-managed old K8 schools gained significantly less than did eighth-graders in district-managed old K8 schools in two of the three cohorts. The gains in older K-8 schools in each cohort are depicted in Figure 2. (The gains in middle schools are also shown in order to make visual comparisons possible.)

The NEWK8 coefficients indicate eighth-graders in newly-established district-managed K-8s began to outperform those in district-managed middle schools in the second cohort but that this performance advantage did not approach statistical significance until the third cohort. The EDISON × NEW K8 coefficients indicate that 8th-graders in Edison-managed new K-8 schools did not significantly outgain those in district-managed New K8s in any cohort. At the beginning of the privatization reforms, two of the K-5 schools being gradually converted into K-8s were assigned to be managed by Temple University. These two schools did not have any 8th-graders until their fifth-graders from Spring 2002 reached 8th-grade in Spring 2005. Therefore, the model does not include an OTHEREMO × NEWK8 product variable until the 2005 cohort. As both the Figure 3 and Table 3 indicate, the math achievement gains attained by the 63 8th-graders in these schools were significantly higher than those in district-managed and Edison-managed New K-8s.

[Figures 2 and 3 about here]

We present comparable analyses of reading score gains in Table 4 and Figures 4 to 6. The EDISON and OTHEREMO coefficients in Table 4 compare the 3-year reading achievement
gains for eighth-graders in middle schools that have been under Edison or other EMO management for one year (2003 cohort), two years (2004 cohort) or 3 years (2005 cohort) with those in district-managed middle schools. These gains are depicted in Figure 4. Except in the first cohort (in which there was a significantly negative impact of Edison management), gains for eighth-graders in Edison-managed middle schools were not significantly different than in district-managed middle schools. Eighth-graders in district-managed middle schools outgained those in other EMO-managed middle schools by 31.3 points (p < .01) in the 2003 cohort, and by 17.6 points (p < .05) in the 2004 cohort, and by 40.4 points in the 2005 cohort (not statistically significant because of large variation in this cohort among the middle schools managed by other EMOs.)

As shown in Figure 5, if one compares the reading gains achieved by students in the older K-8 schools, students in the one Edison-managed older K-8 school outgained those in district-managed older K-8 schools. Reading gains were significantly higher at district-managed established K-8 schools (compared to district-managed middle schools) only in the 2003 cohort as well. The reading gain results for established K-8 schools managed by other EMOs varied widely by cohort (marginally higher than district-managed Old K8s in the 2003 cohort, significantly lower than district-managed Old K8s in the 2004 cohort, and not significantly different in the 2005 cohort).

The NEWK8 coefficients in Table 4 indicate that reading gains at the newly established K8 schools managed by the district were not significantly different from those at district-managed middle schools. Students at the new K8s managed by Edison marginally outgained those at district-managed new K-8s in reading (p < .10) in the 2003 cohort, but did not perform significantly better in later cohorts.

Privatization has been an expensive experiment in Philadelphia. So far, (through Spring 2005) this experiment has not paid off by producing consistently better achievement gains in the privatized schools than in the district-managed schools. Non-Edison EMO-managed schools have actually performed worse than district-managed schools during this period. (However, two Temple University-managed new K-8s achieved promising results with their first small cohort of 8th graders.) The results in Edison schools have been generally positive but not significantly nor consistently different from those in district-managed schools. In the first cohort encountering privatization, the 35 Edison 8th-graders from an older K-8 school outgained the 553 8th graders in district-managed older K-8 schools in math and reading. But Edison did not significantly outperform the district when managing large middle schools and there was no significant Edison advantage in K-8 schools in the second or third cohorts.

The early results from the K-8 conversion experiment are also mixed. While students in long-established K-8’s generally outperformed students in middle schools, students in newly-converted K-8s did not always significantly outgain those in middle schools nor match the gains found in older K-8 schools. This suggests that Philadelphia’s attempt to replicate the achievement success often found in the older generation of K-8 schools by creating a new generation of such schools has not yet been entirely successful.
What is even more notable in these data than comparisons between Edison- and district-managed schools and between K-8 schools and middle schools are comparisons between the 2003 cohort of students and the 2004 and 2005 cohorts. The 2004 and 2005 cohorts of eighth-graders showed math gains that were much larger than those shown by the 2003 cohort of eighth-graders. This finding was broad based, occurring in middle schools and K-8 schools, in Edison schools and district-managed schools. This finding suggests that the systemic reforms of the Vallas era that have impacted all Philadelphia schools regardless of grade span or management (such as the adoption in all schools of a strong, NSF-supported math curricula accompanied by ongoing professional development and coaching for teachers, with regular benchmark testing of students) may be having very positive effects. The pattern was not as pronounced in reading, however, since reading gains for the 2005 cohort were more similar to those of the 2003 cohort than those of the 2004 cohort.

Turning now to the demographic predictors in our analyses, we find it interesting that student gender was not a significant predictor of baseline (5th grade) math score in any cohort, but was a significant predictor of baseline reading achievement for the 2004 and 2005 cohorts (girls finished the elementary grades with better reading skills than boys). Other demographic variables were stronger and more consistent predictors of student achievement at the end of the elementary grades. While demographic variables had few significant effects on student growth during the middle grades in the 2003 cohort, there were several significant effects in the 2004 and 2005 cohorts (for example, the strong gains in these cohorts were even stronger for females and Asians.) This suggests that some subgroups benefited more from the systemic changes than did others.

The analyses also show that students who remain in the same school throughout the middle grades consistently learn more during this critical period of their educational career; attending the same school for 6th, 7th, and 8th grade was a strong predictor of growth in both math and reading achievement across all cohorts. This finding may indicate that, other things being equal, students benefit from the greater stability and coherence in relationships and instructional programs that they encounter when they stay in the same school. Of course, this predictor may also be a proxy for other disruptive events in the lives of some young adolescents’ (such as a changing one’s school due to residential relocations caused by economic, marital or other types of distress in the family unit.)

Discussion

Analyses of the first three years of EMO management in Philadelphia middle grades schools have not yet yielded evidence of a positive effect for this reform strategy. We believe that the political decision to replace district management with private management as a strategy for solving the problem of low achievement among middle grades students reflected a fuzzy understanding of the kind of reforms necessary to improve instruction and help students make achievement gains. Private management itself does not directly affect the quality of classroom instruction, particularly when private educational management providers have little experience with middle grades curriculum and meeting high poverty adolescent students’ needs. In this case, all of the Non-Edison EMOs chose to follow the district curriculum and pacing guide and Edison had a similar curriculum and pacing guide, and so there is not as much variation on that
factor in Philadelphia as there could be in other districts. It is possible that EMOs may in the long run attract higher quality teachers, who will be able to help raise student achievement significantly more than teachers in district-administered schools. In the short run, however, the percentage of certified teachers schools remained significantly lower at EMO schools than non-EMO schools (as it had been prior to the assignment of schools to EMO providers). As other measures of teacher quality become available for analysis, it is possible that we will find evidence of higher quality teachers in EMO schools over time, but thus far this has not occurred.

Some may argue that the absence of a positive impact of EMOs on student achievement in Philadelphia is due to incomplete implementation of a privatization model. As Bulkley, Mundell, and Riffer (2004) point out, the Philadelphia “diverse provider model” has not fully meet the criteria set forth by Hill, Campbell, and Harvey (2000). Both teachers and principals remained employees of the school district, and providers had to initially honor union contracts. Principals had to answer to both the EMO provider and to the district. This situation was in part necessitated by contractual and political realities in a city with strong and respected unions and by the district’s non-assignable legal responsibilities (e.g., to assure student safety and the provision of special education services). Except in a rare district where all such issues can be completely resolved, it will be impossible for private providers to have total control over the public schools they manage. Special education liability issues are particularly hard to resolve given the huge numbers of special education students in big city districts and the court-ordered monitoring of services in many of these districts. Yet, any defensible diverse provider model will require that each provider serve its fair share of special needs students. (It would not be acceptable or legal, for example, to tell special education students that they are welcome only at district-managed schools.) So far, politically- and legally-acceptable district contracts with EMOs -- as they have been implemented in Philadelphia and in other large urban districts -- have not yet been shown to be a bet that has paid off.

On the other hand, providing more access to a K-8 structure for students who would ordinarily have attended large middle schools could be a promising development. Our analyses show the importance of attending the same school over the middle grades, and the absence of a transition year associated with attendance at middle school after elementary school may similarly help to explain the higher achievement of students in some of the new K-8 schools (even though these schools, like the middle schools, had significantly higher numbers of non-certified teachers than the established K-8 schools). The simple fact of significantly fewer students in grades 6 through 8 at these new K-8s (compared with middle schools) may reduce the middle school behavioral “chaos” factor so that it is possible for more learning to occur in these schools, and more individual attention can be paid to students struggling to master academic subject matter. Preliminary analyses indicate that the incidence of poor behavior grades is lower at new K8s than at middle schools (Balfanz & Herzog, 2006), and that controlling for grade size reduces the impact of the K-8 effect (indicating it is a probable component of that effect) (Byrnes, 2005). But conversion to K-8 will require additional leadership qualities on the part of the elementary principals involved, who will have a steep learning curve to assure instructional leadership for higher level academic content and adolescent student needs not previously part of their responsibility. Recruiting teachers with sufficient content knowledge, under the new NCLB requirements, and assuring adequate professional development opportunities for these teachers who do not have the same type of “team” environment possible in the middle school context, is
another challenge for these new K-8s. Creating space for three more grade levels at scores of
elementary schools may also prove daunting. Further analyses on subsequent years’ data, as the
numbers of newly converted K-8 schools increase, will be important to determine whether the
promising trends for New K-8 schools in our analyses become stronger and more consistent over
time.

Large 5th to 8th grade achievement gains in the eighth grade cohorts of 2004 and 2005 are
among the most interesting findings in these analyses. It appears that these gains may be due to
increased coherence and coordination of curricula and professional development, increased focus
on student outcomes, and increased resources for low-performing schools (which may be more
sustained in math than in reading). Greater centralization, writ large, may have also played a
role in increasing student achievement. Further analyses using cohort years and looking at
changing interventions across all schools and between different categories of schools will be
necessary to confirm or disconfirm these hypotheses. On the surface, it appears that greater
instructional coherence, along with increased resources from the state, especially compared with
decentralization and paucity of funding under the Children Achieving reforms, has contributed to
student success in Philadelphia. These findings echo those from New York District #2 (Elmore &
Burney, 1997; D’Amico, Harwell, Stein, & van den Heuvel, 2000) and San Diego (Darling-
Hammond et al., 2002; Hightower, 2002) and other studies of district central offices (e.g.,
Snipes, Dolittle, & Herlily, 2002; Hightower, Knapp, Marsh, & McLaughlin, 2002; see review in
Mac Iver and Farley, 2003). In many ways, lessons learned from some of the comprehensive
school reform models (e.g., Mac Iver & Balfanz, 2000; Herlily & Kemple, 2003) have been
scaled up to the district level in Philadelphia, with notable initial success. Whether such gains
will continue, and urban students will actually begin to close the achievement gap with their
more advantaged counterparts, graduating from high school and entering the productive work
force in higher numbers, remains to be seen.
References


<www.aft.org/research/edisonschools>.

<www.aft.org/research/edisonproject>.


<table>
<thead>
<tr>
<th>School Characteristics</th>
<th>Edison Schools (n=14)</th>
<th>Other EMO Schools (n=12)</th>
<th>Non-EMO Schools (n=26)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average 5th grade Incoming PSSA NCE Math Score</td>
<td>23.9*</td>
<td>23.6*</td>
<td>29.0</td>
</tr>
<tr>
<td>Average 5th grade Incoming PSSA NCE Reading Score</td>
<td>26.0*</td>
<td>26.2*</td>
<td>28.8</td>
</tr>
<tr>
<td>% Low Income Students (2001-02)</td>
<td>88.6*</td>
<td>82.9</td>
<td>80.5</td>
</tr>
<tr>
<td>% Non-certified Teachers (2001-02)</td>
<td>23.8*</td>
<td>18.5*</td>
<td>10.9</td>
</tr>
<tr>
<td>% of Teachers Returning in 01-02 (from prior year)</td>
<td>82.3</td>
<td>84.1</td>
<td>86.2</td>
</tr>
</tbody>
</table>

* Significantly different from non-EMO schools at p<.05.
Table 2
Pre-Existing Differences (Spring 2002) between School Types

<table>
<thead>
<tr>
<th>School characteristics</th>
<th>Middle Schools (n=28)</th>
<th>Old K-8 Schools (n=16)</th>
<th>New K-8 Schools (n=8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average 5th grade Incoming PSSA NCE Math Score</td>
<td>25.0</td>
<td>28.8*</td>
<td>26.4</td>
</tr>
<tr>
<td>Average 5th grade Incoming PSSA NCE Reading Score</td>
<td>26.7</td>
<td>28.8</td>
<td>27.3</td>
</tr>
<tr>
<td>% Low Income Students</td>
<td>83.7</td>
<td>80.0</td>
<td>89.0*</td>
</tr>
<tr>
<td>% Non-certified Teachers</td>
<td>18.5</td>
<td>9.9*</td>
<td>20.4</td>
</tr>
<tr>
<td>% of Teachers Returning in 01-02 (from prior year)</td>
<td>84.6</td>
<td>85.6</td>
<td>83.3</td>
</tr>
</tbody>
</table>

* Significantly different from middle schools at p<.05
Because of schools that were transitioning into becoming K-8s, there were actually more than 100 schools serving middle grades students, but many had not yet added 8th grade. Only those high poverty schools with 8th graders having PSSA math test scores were used in the following analyses. Alternative schools for students with behavioral problems were also excluded from the analyses since school level data were not available from the district on these schools.

Although two of the K-5 schools being gradually converted into K-8s were assigned to a Non-Edison EMO, these schools did not have any 8th-graders until their fifth-graders from Spring 2002 reached 8th-grade in Spring 2005. Therefore, the model does not include an OTHEREMO × NEWK8 product variable until the 2005 cohort.

When a chi-square test indicated that the residual parameter variance associated with certain effect in the model was not significant, we simplified the model by setting the residual parameter variance to zero. For example, in the math achievement analyses of the 2003 cohort, the residual variance was set to zero for these parameters: the effects of FEMALE on P0 (math scale score in Spring of 5th grade) and the effects of FEMALE, ENG LANG LEARNER, ASIAN, HISPANIC, WHITE, and ATTENDED SAME SCHL FOR 6TH to 8th on P1 (math scale score gain between Spring of 5th Grade and Spring of 8th Grade).
<table>
<thead>
<tr>
<th>Fixed Effect</th>
<th>Coeff (se)</th>
<th>Signif</th>
<th>Coeff (se)</th>
<th>Signif</th>
<th>Coeff (se)</th>
<th>Signif</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model for P0 (math scale score in 5th grade)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1104.7 (4.8) ***</td>
<td>1117.1 (5.2) ***</td>
<td>1114.9 (5.3) ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>-2.9 (3.2)</td>
<td>3.0 (3.4)</td>
<td>-2.7 (3.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>84.2 (13.1) ***</td>
<td>106.6 (14.1) ***</td>
<td>100.5 (13.1) ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>18.2 (5.7) **</td>
<td>19.1 (7.0) **</td>
<td>21.2 (6.6) ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>62.0 (8.9) ***</td>
<td>51.0 (10.6) ***</td>
<td>77.3 (13.9) ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Special Educ</td>
<td>-72.1 (8.5) ***</td>
<td>-31.3 (8.2) ***</td>
<td>-115.6 (11.4) ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eng Lang Learner</td>
<td>-52.0 (15.8) **</td>
<td>-58.4 (16.0) ***</td>
<td>-68.8 (10.3) ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Low Inc (5th Grade Sch)</td>
<td>-0.6 (0.2) ***</td>
<td>-1.3 (0.3) ***</td>
<td>-0.3 (0.2) **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Model for P1 (math scale score gain)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>32.1 (4.9) ***</td>
<td>75.1 (10.8) ***</td>
<td>83.6 (13.7) ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attended same school</td>
<td>17.2 (3.7) ***</td>
<td>21.9 (4.7) ***</td>
<td>20.1 (6.3) **</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>1.2 (3.1)</td>
<td>11.5 (3.4) ***</td>
<td>11.3 (4.7) *</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>11.9 (9.6)</td>
<td>23.3 (8.6) **</td>
<td>47.2 (13.1) ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>-5.0 (7.3)</td>
<td>2.3 (6.8)</td>
<td>1.6 (5.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>-18.0 (8.5) *</td>
<td>-18.0 (10.0) †</td>
<td>-7.5 (12.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Special Educ</td>
<td>6.0 (6.8)</td>
<td>-15.4 (4.9) **</td>
<td>-2.0 (13.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eng Lang Learner</td>
<td>-5.4 (10.8)</td>
<td>-32.8 (15.9) *</td>
<td>28.5 (15.4) †</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg. Incoming Read NCE</td>
<td>-1.7 (0.7) *</td>
<td>-0.6 (1.0)</td>
<td>0.0 (1.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schl's prior % non-cert</td>
<td>-0.3 (0.3)</td>
<td>0.8 (0.6)</td>
<td>0.2 (0.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schl's %FRL</td>
<td>-2.5 (0.8) **</td>
<td>-0.3 (0.9)</td>
<td>-0.6 (1.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Edison</td>
<td>-14.3 (9.3)</td>
<td>19.3 (18.4)</td>
<td>23.6 (20.2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other EMO</td>
<td>-26.7 (7.2) ***</td>
<td>-26.9 (13.7) †</td>
<td>-33.8 (14.5) *</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old K8</td>
<td>32.8 (9.4) ***</td>
<td>60.0 (14.2) ***</td>
<td>33.8 (21.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Edison X Old K8</td>
<td>52.1 (15.6) **</td>
<td>4.3 (21.1)</td>
<td>11.2 (26.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other EMO x Old K8</td>
<td>-26.5 (10.4) *</td>
<td>-81.0 (16.5) ***</td>
<td>-22.4 (22.2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New K8</td>
<td>-1.8 (16.6)</td>
<td>25.9 (23.9)</td>
<td>32.6 (19.0) †</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Edison x New K8</td>
<td>27.2 (21.8)</td>
<td>5.7 (42.7)</td>
<td>-15.8 (35.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other EMO x New K8</td>
<td>NA</td>
<td>NA</td>
<td>163.4 (61.3) **</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

† p<.1  *p<.05  **p<.01  ***p<.001
### Table 4

<table>
<thead>
<tr>
<th>Fixed Effect</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff (se)</td>
<td>Signif</td>
<td>Coeff (se)</td>
</tr>
<tr>
<td><strong>Model for P0 (reading scale score in 5th grade)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1098.4 (3.3) ***</td>
<td></td>
<td>1102.5 (4.8) ***</td>
</tr>
<tr>
<td>Female</td>
<td>3.6 (3.9)</td>
<td></td>
<td>31.7 (4.8) ***</td>
</tr>
<tr>
<td>Asian</td>
<td>51.6 (13.4) ***</td>
<td></td>
<td>67.3 (12.3) ***</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-11.7 (5.6) *</td>
<td></td>
<td>-2.9 (8.8)</td>
</tr>
<tr>
<td>Caucasian</td>
<td>67.4 (10.0) ***</td>
<td></td>
<td>36.6 (10.1) **</td>
</tr>
<tr>
<td>Special Educ</td>
<td>-65.1 (11.5) ***</td>
<td></td>
<td>-38.1 (10.4) **</td>
</tr>
<tr>
<td>Eng Lang Learner</td>
<td>-50.3 (17.2) **</td>
<td></td>
<td>-98.1 (15.2) ***</td>
</tr>
<tr>
<td>% Low Inc (5th Grade Sch)</td>
<td>-0.8 (0.3) **</td>
<td></td>
<td>-1.6 (0.3) ***</td>
</tr>
<tr>
<td><strong>Model for P1 (reading scale score gain)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>46.9 (9.8) ***</td>
<td></td>
<td>94.6 (9.3) ***</td>
</tr>
<tr>
<td>Attended same school</td>
<td>18.2 (4.9) ***</td>
<td></td>
<td>20.6 (6.5) **</td>
</tr>
<tr>
<td>Female</td>
<td>29.7 (4.3) ***</td>
<td></td>
<td>22.1 (4.2) ***</td>
</tr>
<tr>
<td>Asian</td>
<td>25.0 (8.2) **</td>
<td></td>
<td>43.0 (10.0) ***</td>
</tr>
<tr>
<td>Hispanic</td>
<td>13.1 (8.9)</td>
<td></td>
<td>12.0 (9.9)</td>
</tr>
<tr>
<td>Caucasian</td>
<td>-16.6 (7.7) *</td>
<td></td>
<td>4.2 (9.7)</td>
</tr>
<tr>
<td>Special Educ</td>
<td>-78.2 (9.6) ***</td>
<td></td>
<td>-64.9 (9.1) ***</td>
</tr>
<tr>
<td>Eng Lang Learner</td>
<td>0.9 (9.7)</td>
<td></td>
<td>-34.5 (17.6) *</td>
</tr>
<tr>
<td>Avg. Incoming Read NCE</td>
<td>2.3 (1.5)</td>
<td></td>
<td>5.6 (1.1) ***</td>
</tr>
<tr>
<td>Schl's prior % non-cert</td>
<td>-0.5 (0.5)</td>
<td></td>
<td>-0.7 (0.5)</td>
</tr>
<tr>
<td>Schl's %FRL</td>
<td>-0.1 (1.2)</td>
<td></td>
<td>-2.2 (1.3) †</td>
</tr>
<tr>
<td>Edison</td>
<td>-50.2 (14.6) **</td>
<td></td>
<td>-8.1 (16.8) **</td>
</tr>
<tr>
<td>Other EMO</td>
<td>-31.3 (11.3) ***</td>
<td></td>
<td>-17.6 (8.2) *</td>
</tr>
<tr>
<td>Old K8</td>
<td>30.2 (14.2) **</td>
<td></td>
<td>8.9 (15.5)</td>
</tr>
<tr>
<td>Edison X Old K8</td>
<td>93.6 (18.9) ***</td>
<td></td>
<td>40.2 (23.3) †</td>
</tr>
<tr>
<td>Other EMO x Old K8</td>
<td>28.6 (15.0) †</td>
<td></td>
<td>-35.2 (14.9) *</td>
</tr>
<tr>
<td>New K8</td>
<td>21.5 (17.4) †</td>
<td></td>
<td>22.1 (18.8)</td>
</tr>
<tr>
<td>Edison x New K8</td>
<td>39.4 (20.0) †</td>
<td></td>
<td>-16.9 (29.3)</td>
</tr>
<tr>
<td>Other EMO x New K8</td>
<td>NA</td>
<td></td>
<td>NA</td>
</tr>
</tbody>
</table>

* † p<.1  
**p<.05  
***p<.01  
****p<.001
Figure 1. Adjusted Mean Math Gains In Middle Schools for Three Student Cohorts
(the number of 8th-graders served is shown above each bar)
Figure 2. Adjusted Mean Math Gains in Older K-8 Schools Compared to Those in Middle Schools (the number of 8th-graders served in Older K-8 Schools is shown above each bar)
Figure 3. Adjusted Mean Math Gains in Newly-Converted K-8 Schools Compared to Those in Middle Schools
(the number of 8th-graders served in New K-8 Schools is shown above each bar)
Figure 4. Adjusted Mean Reading Gains In Middle Schools for Three Student Cohorts (the number of 8th-graders served is shown above each bar)
Figure 5. Adjusted Mean Reading Gains in Older K-8 Schools
Compared to Those in Middle Schools
(the number of 8th-graders served in Older K-8 Schools is shown above each bar)
Figure 6. Adjusted Mean Reading Gains in Newly-Converted K-8 Schools Compared to Those in Middle Schools (the number of 8th-graders served in New K-8 Schools is shown above each bar)