
Martha Abele Mac Iver and Douglas J. Mac Iver¹
Johns Hopkins University

Abstract
This article 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 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 achievement gains for students in EMO-managed schools were not larger than those for students in schools managed by the district. Non-Edison EMO schools actually performed worse than district-managed schools. With the exception of one older K-8 school in one cohort, Edison schools did not significantly outperform district-managed counterparts. Students in long-established K-8 schools generally outgained students in middle schools, but gains were not as large in newly-established K-8 schools. Across all types of schools, the second cohort of students obtained greater gains than did the first.

The dropout problem in urban school systems, where at least half of all ninth graders leave high school without a diploma (Balfanz & Legters, 2001, 2004; Mac Iver, Farley, & Wayman, 2003; Swanson, 2004), is a primary factor in urban poverty, crime, and the continued decline of American cities (Land & Legters, 2002; Wald & Losen, 2003). High school reform now tops the policy agenda of the National Governors’ Association, which held a national summit on the issue in February 2005. Policymakers using “crisis discourse” (Harvey & Housman, 2004, p. 10) have emphasized the need for more rigorous standards and the importance of assessments in addressing high school reform, but as Harvey and Housman (2004, p. 3) point out, “little guidance has emerged on how to bring about desired changes.”

High school reform by itself will not be enough to solve the dropout problem because the problem has its roots in the middle and elementary grades. Dropout outcomes can be predicted as early as 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. During the middle grades, students must make a successful transition from elementary understandings to more sophisticated academic competencies to increase their chances of high school graduation and college completion (Ingles, Owings, Kaufman, Alt, & Chen, 2002; Moses & Cobb, 2001; Pelavin & Kane, 1990; U.S. Department of Education, 1997; National Research Council, 2001; RAND Mathematics Study Panel, 2003). Results from the National Assessment of Educational Progress (NAEP) show significantly worse student achievement at inner city middle schools with high minority populations from low-income families (Braswell et al., 2001; O’Sullivan, Lauko, Grigg, Qian, & Zhang, 2003; U.S. Department of...
Education, 2003). These results illustrate the findings from numerous studies that show the unequal opportunities to learn (Blank & Langsen, 1999; Campbell & Silver, 1999; Raudenbush, Fotiu, & Cheong, 1998) and the fewer supports and resources available to students in these schools (Flanagan & Grissmer, 2002; Mullis et al., 2000), including differences in teacher quality. Teachers at high-poverty secondary schools are generally not as well prepared for their positions (National Center for Education Statistics, 1999), and as a result provide less effective instruction (Grossman, 1990).

The problems described above have prompted considerable research on urban district reform (e.g., Bryk, Sebring, Kerbrow, Rollow, & Easton, 1998; D’Amico, Harwell, Stein, & van den Heuvel, 2001; Hightower, 2002; Hightower, Knapp, Marsh, & McLaughlin, 2002; Hill, Campbell, & Harvey, 2000; Quinn, Stewart, & Nowakowski, 1993; Resnick & Harwell, 2000; Ross, 2001; Snipes, Dolittle, & Herlihy, 2002; Westat, 2001). Our study contributes to this research literature by analyzing longitudinal student outcome data in one urban district that has undergone a variety of educational reforms over the past decade.

Background

Researchers, policymakers, and district leaders from across the nation have been watching Philadelphia ever since February 1995 when Superintendent David Hornbeck used $50 million from the Annenberg Challenge (and $100 million in local matching funds) to launch “Children Achieving.” This systemic reform initiative relied on standards, decentralization, and accountability reforms while vigorously pursuing fairer, more adequate funding for the district’s schools and mandating that all schools be divided into smaller units (SLCs) that would provide more personalized environments for teaching and learning (e.g., Christman, 2001; Wong & Alkis, 1999; Wong & Brown, 1998; Wong & Sunderman, 2000). Significant achievement gains were made during the Children’s Achieving initiative (Corcoran & Christman, 2002; Tighe, Wang, & Foley, 2002), including gains in access to challenging math and science courses and achievement in these content areas noted in studies of the Urban Systemic Initiative (USI) in Philadelphia (Kim, Crasco, Leavitt, & Karantonis, 2002). But these gains were not big enough or fast enough to shield the initiative from financial and political problems. These, in turn, led to the threatened state takeover and the drastic budget cuts, which prompted Hornbeck’s resignation and the abrupt end of this initiative after 5 1/3 years.

The state takeover became a reality in December 2001 when Philadelphia’s Board of Education was replaced by a School Reform Commission (SRC) with a majority of members appointed by the Governor Mark Schweiker. 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 embraced privatization and mandated restructuring as reform strategies (Useem & Balfanz, 2002). Among the district’s 264 schools:
Forty-five were designated to be run by private entities, both nonprofit and for-profit groups;

Twenty-one were “restructured,” singled out to be in a separate subdistrict and to implement mandated reforms;

Four were instructed to become self-governing public charter schools;

Sixteen 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 Paul Vallas, former head of Chicago’s school system, as the district’s new chief executive officer (CEO). Vallas has led the district to implement ambitious and comprehensive reforms that emphasize a coherent instructional program that provides strong instructional guidance through the systemwide implementation of a completely specified core curriculum, pacing guides, and curriculum-based assessments every 6 weeks. The Vallas administration also strengthened the support system for new teachers by hiring coaches to mentor them and by providing paid summer training for all new teachers. In addition, the administration increased support in mathematics and literacy by setting aside a half-day during school hours twice a month for professional development during the rollout year of the new curricula. Schools are also required to have school-based content leaders with some release time to assist in implementing the new curricula (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). Vallas also 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 poorly. In addition, providers were welcome to use the district’s new core curriculum in the schools that they managed.

Literature on Educational Governance

Chubb and Moe (1990) linked higher student achievement to lower levels of bureaucratic organization in schools, setting off a veritable privatization revolution in education despite the methodological critiques of their analysis (e.g., Sukstorf, Wells, & Crain, 1993). Others have proposed downsizing or dramatically scaling back district functions (Effron & Concannon, 1995; Finn, 1991; Hill, 1997; Keedy, 1994; Parsley, 1991; Scambio & Graber, 1991) or greatly reducing the ratio of central office administrators to students in big school districts (Ornstein, 1989). Restructuring, school-based management (SBM), and charter school development have similarly become the cornerstones of reform efforts for more than a decade, despite the mixed evidence of effectiveness in improving student achievement (e.g., Murphy & Beck, 1995; Murphy & Shiffman, 2002).
Despite calls for the end of school districts, there is generally broad agreement about two functions they can and should fulfill: setting high performance standards, and developing and enforcing an accountability framework (Archbald, 1998; Odden & Hill, 1997; Palmaffy, 1998; Wong & Alkins, 1999). There is no consensus, however, on the role the school district should play in helping schools meet these standards. There are those who argue that districts should do little more than set standards, enforce accountability, and ensure adequate funding while giving each school much freedom in deciding how to best improve performance and in choosing external partners to help them. (The Hornbeck administration was an exemplar of this approach.) There are those who believe that district leaders should strongly and actively enforce common approaches for district schools. (The Vallas administration is an exemplar of this approach.) There are still others who believe that neither school nor district leaders are capable of doing what needs to be done and that private companies, universities, community organizations, and charter school boards should be the primary educational management organizations. (Former Pennsylvania Governor Mark Schweiker was a firm believer in this approach. When the state took control of the district, his initial plan called for much more privatization than was eventually implemented. See Useem & Balfanz, 2002).

Proponents of site-based management 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, operate, 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 et al., 2000; Mohrman & Wohlstetter, 1994, Odden & Hill, 1997; Ouchi, 2003; Wohlstetter, Mohrman, & Roberson, 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 (Palmaffy, 1998; Wohlstetter et al., 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). Research on Philadelphia’s attempt to promote decentralized decision making about curriculum under the Children Achieving initiative found that many school leaders needed guidance in selecting appropriate curriculum for standards-based instruction and many Philadelphia teachers needed more concrete help with unit and lesson planning to address the standards developed under the initiative (Corcoran & Christman, 2002; Foley, 2001; Simon, Foley, & Passantino, 1998; Wong & Brown, 1998).

Research on the effects of privatizing instructional delivery is also decidedly mixed. The most recent study comparing student outcomes in privately managed and traditional schools in six cities found student achievement higher in one type in some cities and the other type in others (General Accounting Office [GAO], 2003). Previous studies similarly found mixed evidence of improved achievement (Ascher, 1996; Edwards, 1997; GAO, 1996; Mac Iver & Stringfield, 2000; Peeler & Parham, 1994; Williams & Leak, 1995).
School Governance and Gradespan Reform Efforts in the Middle Grades

While our larger study includes analyses of the impact of various curricular and organizational reforms and teacher/administrator quality variables on student achievement over the past decade, this article focuses on the more recent governance and gradespan reform efforts. As we noted, privatization began in Philadelphia in fall 2002. Of the 99 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% of the eighth grade students with math test scores were attending the 26 middle grades schools run by an EMO. In 2003–04, EMOs ran 24 middle grades schools, with about 27% of eighth grade students.

Advocates who emphasized the beneficial effects of K-8 schools over 6–8 (middle) schools helped fuel the movement toward creating more K-8 schools in Philadelphia. 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? And do students at the newly converted K-8s make significantly greater improvement than students in the remaining middle schools?

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 grade 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. This structural reform does 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 mathematics portion of the state-administered standardized achievement tests because they are part of a larger study of the impact of systemic reforms in mathematics funded by the National Science Foundation. Students take the state-administered Pennsylvania System of School Assessment (PSSA) just before entering the middle grades (in the spring of fifth 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 systemwide implementation of a completely specified core curriculum, pacing guides, and curriculum-based assessments every 6 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 preexisting 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 preexisting differences between the EMO and district-managed schools (Table 1). The average poverty rate was significantly higher in Edison schools and nonsignificantly 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 noncertified 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 following analyses. 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.

The analyses that follow include longitudinal data from the first and second cohorts of students to reach eighth grade in schools managed by EMOs (fifth graders in 1999–2000, who reached eighth grade in 2002–03, and fifth graders in 2000–01 who reached eighth grade in 2003–04). 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.

Table 1. Preexisting Differences between Schools Chosen for EMO Management and Other Schools

<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 PSSA NCE Math Score</td>
<td>23.9*</td>
<td>23.6*</td>
<td>29.0</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>% Noncertified 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 < 0.05

Table 2. Preexisting 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 PSSA NCE Math Score</td>
<td>25.0</td>
<td>28.8*</td>
<td>26.4</td>
</tr>
<tr>
<td>% Low-Income Students</td>
<td>83.7</td>
<td>80.0</td>
<td>89.0*</td>
</tr>
<tr>
<td>% Noncertified 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 < 0.05
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 at high-poverty schools 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 three-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, a fifth grade and an eighth grade mathematics scale score on the Pennsylvania System of School Assessment (PSSA). For each cohort there were roughly 4,800 students in high-poverty schools, and just over half were enrolled in EMO schools.

At level 1, within students, we model students’ achievement scores as a function of grade:

\[ Y = P_0 + P_1(EIGHTH) + E \]

where 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 were not available from district records, and so students were assigned the poverty level of their fifth 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:

\[
P_0 = B00 + B01(\text{% LOW INCOME STUDENTS IN } 5^{TH} \text{ GRADE SCHOOL}) + B02(\text{SPECIAL EDUCATION}) + B03(\text{ENGLISH LANGUAGE LEARNER}) + B04(\text{ASIAN}) + B05(\text{HISPANIC}) + B06(\text{CAUCASIAN}) + B07(\text{FEMALE}) + R0
\]

\[
P_1 = B10 + B11(\text{FEMALE}) + B12(\text{ATTENDED SAME SCHOOL FOR } 6^{TH} \text{ to } 8^{TH}) + B13(\text{ASIAN}) + B14(\text{HISPANIC}) + B15(\text{CAUCASIAN}) + B16(\text{SPECIAL EDUCATION}) + B17(\text{ENGLISH LANGUAGE LEARNER})
\]

At level 3, the school level, we tested for interactive effects of the management and grade span interventions on students’ mathematics achievement growth, con
trolling for differences between schools that already existed (in 2001–02) prior to the assignment of schools to be run by an EMO or not (prior percent of noncertified 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 NCE score on the PSSA math at the end of fifth grade). Each school’s management (EDISON or OTHER EMO) and grades 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 × OLD K8) denoted each school’s cross-classification. Other school characteristics were measured as percentages, except for a school’s average incoming math achievement. We specified the following between-schools model:

\[
B_{10} = G_{100} + G_{101}(\text{NEW K8}) + G_{102}(\text{OLD K8}) + G_{103}(\text{EDISON}) + G_{104}(\text{OTHER EMO}) + G_{105}(\text{EDISON} \times \text{NEW K8}) + G_{106}(\text{EDISON} \times \text{OLD K8}) + G_{107}(\text{OTHER EMO} \times \text{OLD K8}) + G_{108}(\text{SCHL’s % FRL}) + G_{109}(\text{SCHL’s PRIOR % OF NONCERTIFIED TCHR}) + G_{1010}(\text{AVG INCOMING MATH NCE @ ENDOF 5TH GRADE}) + U_{10}
\]

\[
B_{11} = G_{110} + U_{11}
\]
\[
B_{12} = G_{120} + U_{12}
\]
\[
B_{13} = G_{130} + U_{13}
\]
\[
B_{14} = G_{140} + U_{14}
\]
\[
B_{15} = G_{150} + U_{15}
\]
\[
B_{16} = G_{160} + U_{16}
\]
\[
B_{17} = G_{170} + U_{17}
\]
Figure 1 shows the adjusted mean math gains in each type of school for the first two cohorts of students to complete eighth 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 \((p < 0.01)\) on students’ math gains in each cohort (2003 cohort \(\chi^2 (3) = 61.2\), 2004 cohort \(\chi^2 (3) = 35.3\)). The nature of these effects can be seen both in the figure and in coefficients from the last seven rows in the table.

Consider the \textit{EDISON} and \textit{OTHER EMO} 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 1 year (2003 cohort) or 2 years (2004 cohort) with those in district-managed middle schools. Gains for eighth graders in Edison-managed middle schools were not significantly different than eighth graders 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 \((p < .01)\) in the 2003 cohort and by 26.9 points \((p < 0.1)\) in the 2004 cohort. The \textit{OLD K8} coefficients indicate that eighth-graders in district-managed older K-8 schools significantly outgained those in district-managed middle schools by about 33 points (2003 cohort) and 60 points (2004 cohort) across their three years in the middle grades. The \textit{EDISON} \texttimes \textit{OLD K8} coefficients indicate that eighth graders in the one Edison-managed old K-8 school 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 cohort. In contrast, the \textit{OTHEREMO} \texttimes \textit{OLDK8} coefficients indicate that eighth graders in the one non-Edison EMO-managed old K-8 school gained significantly less than did eighth graders in district-managed old K-8 schools in both cohorts. The \textit{NEW K8} coefficients indicate eighth graders in newly established district-managed K-8s did not significantly outperform those in district-managed middle schools. Finally, the \textit{EDISON} \texttimes \textit{NEW K8} coefficients indicate that eighth-graders in Edison-managed
new K-8 schools did not significantly outgain those in district-managed New K8s.  

Privatization has been an expensive experiment in Philadelphia. So far (through spring 2004), this experiment has not paid off by producing consistently better math achievement gains in the privatized schools than in the district-managed schools. The non-Edison EMO schools actually have performed worse than district-managed schools during this period. The results in Edison schools have been similar to those in district-managed schools. While it is true that the 35 Edison eighth-graders who attended one older K-8 school significantly outgained their counterparts in district-managed K-8 schools in the first cohort encountering privatization, Edison did not outperform the district when managing large middle schools and there was no significant Edison advantage in K-8 schools in the second cohort.  

The early results from the K-8 conversion experiment are mixed. While students in long-established K-8’s generally outperformed students in middle schools, students in newly converted K-8s did not significantly outgain those in middle schools nor match the gains found in older K-8 schools. Although there were some hopeful signs in the student achievement gains at new K-8 schools, 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. Analyses with subsequent years’ data, in a larger number of new K-8 schools, will be needed to ascertain whether this bet eventually pays off.

Table 3. Modeling Prior Math Achievement and Math Achievement Growth: HLM Estimates

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>2003 Cohort</th>
<th>2004 Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1,104.7*** (4.8)</td>
<td>1,117.1*** (5.2)</td>
</tr>
<tr>
<td>\textit{FEMALE}</td>
<td>-2.9 (3.2)</td>
<td>3.0 (3.4)</td>
</tr>
<tr>
<td>\textit{ASIAN}</td>
<td>84.2*** (13.1)</td>
<td>106.6*** (14.1)</td>
</tr>
<tr>
<td>\textit{HISPANIC}</td>
<td>18.2** (5.7)</td>
<td>19.1** (7.0)</td>
</tr>
<tr>
<td>\textit{CAUCASIAN}</td>
<td>62.0*** (8.9)</td>
<td>51.0*** (10.6)</td>
</tr>
<tr>
<td>\textit{SPECIAL EDUC}</td>
<td>-72.1*** (8.5)</td>
<td>-31.3*** (8.2)</td>
</tr>
<tr>
<td>\textit{ENG LANG LEARNER}</td>
<td>-52.0** (15.8)</td>
<td>-58.4*** (16.0)</td>
</tr>
<tr>
<td>% \textit{LOW INC (5TH GRADE SCHL)}</td>
<td>-0.6*** (0.2)</td>
<td>-1.3*** (0.3)</td>
</tr>
</tbody>
</table>

Model for P1 (math scale score gain between Spring of 5th Grade and Spring of 8th Grade)

<table>
<thead>
<tr>
<th>Fixed effect</th>
<th>2003 Cohort</th>
<th>2004 Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>75.1*** (10.8)</td>
<td></td>
</tr>
<tr>
<td>\textit{ATTENDED SAME SCHL FOR 6TH to 8TH}</td>
<td>17.2** (3.7)</td>
<td>21.9*** (4.7)</td>
</tr>
<tr>
<td>\textit{FEMALE}</td>
<td>1.2 (3.1)</td>
<td>11.5*** (5.4)</td>
</tr>
<tr>
<td>\textit{ASIAN}</td>
<td>11.9 (9.6)</td>
<td>23.3** (8.6)</td>
</tr>
<tr>
<td>\textit{HISPANIC}</td>
<td>-5.0 (7.3)</td>
<td>2.3 (6.8)</td>
</tr>
<tr>
<td>\textit{CAUCASIAN}</td>
<td>-18.0* (8.5)</td>
<td>-18.0* (10.0)</td>
</tr>
<tr>
<td>\textit{SPECIAL EDUC}</td>
<td>6.0 (6.8)</td>
<td>-15.4** (4.9)</td>
</tr>
<tr>
<td>\textit{ENG LANG LEARNER}</td>
<td>-5.4 (10.8)</td>
<td>-32.8* (15.9)</td>
</tr>
<tr>
<td>\textit{AVG INCOMING MATH NCE. (END OF 5TH)}</td>
<td>-1.7* (0.7)</td>
<td>-0.6 (1.0)</td>
</tr>
<tr>
<td>\textit{SCHL’S PRIOR % OF NON-CERTIFIED TCHRS}</td>
<td>-0.3 (0.3)</td>
<td>-0.8 (0.6)</td>
</tr>
<tr>
<td>\textit{SCHL’S %FRL}</td>
<td>-2.5* (0.8)</td>
<td>-0.3 (0.9)</td>
</tr>
<tr>
<td>\textit{EDISON}</td>
<td>-14.3 (9.3)</td>
<td>19.3 (18.4)</td>
</tr>
<tr>
<td>\textit{OTHEREMO}</td>
<td>-26.7** (7.2)</td>
<td>-26.9† (13.7)</td>
</tr>
<tr>
<td>\textit{OLDK8}</td>
<td>32.8** (9.4)</td>
<td>60.0*** (14.2)</td>
</tr>
<tr>
<td>\textit{EDISON × OLDK8}</td>
<td>32.1** (15.6)</td>
<td>4.3 (21.1)</td>
</tr>
<tr>
<td>\textit{OTHEREMO × OLDK8}</td>
<td>-26.7* (10.4)</td>
<td>-81.0*** (16.5)</td>
</tr>
<tr>
<td>\textit{NEWK8}</td>
<td>-1.8 (16.6)</td>
<td>25.9 (23.9)</td>
</tr>
<tr>
<td>\textit{EDISON × NEWK8}</td>
<td>27.2 (21.8)</td>
<td>5.7 (42.7)</td>
</tr>
</tbody>
</table>

Note: Each predictor variable, except for those listed in bold, has been centered around its grand mean.

\[p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001\]
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 two cohorts of students considered here. The 2004 cohort of eighth graders showed 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 that impacted all Philadelphia schools regardless of grade span or management (such as the adoption in all district schools and all Edison 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 beginning to have very positive effects.

Turning now to the demographic predictors in our analyses, we find it interesting that student gender was not a significant predictor of baseline (fifth grade) math score in either cohort, while the other demographic variables were strong predictors of student achievement at the end of the elementary grades. While demographic variables had few significant effects on student growth in the 2003 cohort, there were many significant effects in the 2004 cohort (the strong gains in this cohort were even stronger for females and Asians, but somewhat weaker for Caucasians, special education students, and English Language Learners.) This suggests that some subgroups benefited more from the systemic changes than did others.

Discussion

Analyses of the first 2 years of EMO management in Philadelphia middle grades schools have not yet yielded strong 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, most of the EMOs chose to follow the district curriculum or had a similar 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 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.

It is possible that the absence of a positive impact of EMOs on student achievement was 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 the potential liability of the district for provision of special education services and to assure student safety. Until the special education liability issue is resolved, it will be impossible for private providers to have complete control over schools that include students with special education needs, and therefore impossible to implement a diverse provider model that serves the same types of students in both privately managed and district-managed schools. Research studies focused on this issue will continue to be faced with the dilemma of choosing between having comparable groups of schools, and having a more faithful implementation of a diverse provider model. District contracts with EMOs, as they were implemented in Philadelphia (and will probably be implemented 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 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 noncertified 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 dramatically 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 K-8s 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 number of newly converted K-8 schools increases, will be important to determine whether the promising findings in our analyses become stronger and more consistent over time.

Large achievement gains in eighth grade scores and fifth to eighth grade gains between the eighth grade cohorts of 2003 and 2004 are one of the most interesting findings in these analyses. It appears that these gains may be due to increased coherence and coordination of curricula, increased focus on student outcomes, and increased resources for low-performing schools. 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 No. 2 (Elmore & Burney, 1997) and San Diego (Darling-Hammond, Hightower, Husbands, LaFors, & Young, 2002; Hightower, 2002) and other studies of district central offices (e.g., Hightower et al., 2002; Snipes et al., 2002; see review in Mac Iver & Farley, 2003). In many ways, lessons learned from some of the comprehensive school reform models (e.g., Mac Iver & Balfanz, 2000; Herlihy & 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.

Notes

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

2 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 eighth 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.

3 None of the newly converted K-8s were assigned to a Non-Edison EMO. Therefore, the model does not include a $OTHEREMO \times NEWK8$ product variable.

4 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. In 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 fifth 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 fifth grade and spring of eighth grade). In the 2004 cohort, the residual parameter variance was set to zero for the effects of $FEMALE$, $HISPANIC$, $WHITE$ AND $ENG\ LANG\ LEARNER$ on $P0$ and for the effects of $FEMALE$, $ASIAN$, $HISPANIC$, $WHITE$, and $ATTENDED\ SAME\ SCHL\ FOR\ 6th\ to\ 8th$ on $P1$.

5 Several schools changed EMO status from 2002–03 to 2003–04. For the 2003–04 cohort, schools that were returned to the district were coded according to whether they were ever under EMO management.

About the Authors

Martha Abele Mac Iver is Associate Research Scientist at the Center for Social Organization of Schools at Johns Hopkins University, Baltimore, Maryland. Her current research includes evaluations of whole school reform programs and analyses of systemic change in urban school districts. Her recent articles have appeared in Urban Education, Journal of Vocational Education Research, and Educational Evaluation and Policy Analysis.

Douglas J. Mac Iver is Principal Research Scientist and Co-Director of the Center for Social Organization of Schools at Johns Hopkins University, Baltimore, Maryland. He directs the
Talent Development Middle Grades Program and conducts research on the impact of reform efforts on middle grades student achievement. His recent articles have appeared in the *Journal of Research in Mathematics Education*, *Journal of Curriculum and Instruction*, and *Research in Middle Level Education Quarterly*.

**References**


