Exploring the Correlates of Academic Success in Pennsylvania Charter Schools[1]

Christopher Nelson
RAND Corporation

Gary Miron
The Evaluation Center
Western Michigan University

Abstract

To date, most of the scholarly and policy debate regarding charter schools has focused on two questions: (a) whether charter schools are using their autonomy to engage in innovative practices and (b) whether students in charter schools, taken as a group, perform better or worse than similar students in noncharter public schools. For the most part, however, scholars have done relatively little to bring these two sets of questions together in order to assess whether some uses of charter school autonomy are more academically productive than others. This study seeks to identify factors that distinguish academically successful charter schools from others. Using a unique data set on Pennsylvania charter schools, the study tests explanations about the correlates of academic success. Using what might best be termed "pseudo-growth curve analysis," the study finds that charter schools with higher degrees of perceived accountability produce stronger score growth. Similarly, charter schools with higher degrees of teacher mission commitment and leadership stability produce stronger growth rates in reading and math. Schools with higher degrees of classroom autonomy appear to have lower growth rates, perhaps reflecting recent research on the importance of shared professional culture in teaching and learning. Finally, parent volunteerism appears to be negatively associated with score growth, though it is not clear whether this is simply a proxy for poor governance. While conclusions are limited by the properties of available test score data and a small number of cases, the findings from this study provide a useful early foray into an important policy and school-design question.

[1] Funding for this study was provided by a Field Initiated National Activities Program grant awarded to The Evaluation Center by the Public Charter Schools Program, U.S. Department of Education. The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the sponsor or the researchers' host institutions.

Introduction

Over the course of the past decade, the charter school movement has evolved from the germ of an idea to one of the most widely discussed education reforms. Recent estimates indicate that nearly 3,000 charter schools are providing an education to more than 700,000 children across the country (U.S. Department of Education, 2004). Charter school proponents have long argued that the idea behind the movement is not only to provide new educational opportunities for charter school students, but also to serve as a field laboratory in which practitioners can prove the worth of new educational ideas. In short, charter schools are to be public education's "R&D" sector.

Previous studies indicate that there is considerable variation in the curricular, pedagogical, and organizational approaches these schools take and in the quality of the outcomes they produce. What is less certain is whether some operationalizations of the charter school concept are more effective than others. Without a better understanding of the mechanisms and processes within the "black box" of the charter concept, it is difficult to know whether certain elements in the charter school R&D "portfolio" are more worthy of scale-up than others and how a fully scaled charter school system might perform (Gill, Timpane, Ross, & Brewer, 2001).

This study represents an early foray into the charter school black box by way of an attempt to identify factors that distinguish academically successful charter schools from others. The focus of the study is on Pennsylvania charter schools open as of the 2001-02 academic year. With 77 charter schools operating at that time, this state offers enough cases to test some rudimentary hypotheses about the correlates of academic success. Previous evaluations (Meister & Schuh, 2000; Miron, Nelson & Risley, 2002) have demonstrated that charter schools in the commonwealth represent considerable variation in approach.

The article begins with a brief review of what is known about the academic productivity of charter schools. While the research is beginning to provide a picture – though a mixed one – about the efficacy of charter schools relative to noncharter schools, very little is known about why some charter schools are more academically productive than others. Next, the article sketches a model from which hypotheses about the correlates of academic success are drawn and discusses data sources, measures, and analytical methods. Finally, the key findings from the study as well as their implications are presented and discussed.

Previous Literature

In spite of the charter school reform's prevalence in educational policy debates, there remains a dearth of clear evidence about the schools' academic productivity. Citing the newness of the reform, many researchers have declined to address the achievement question altogether. Other studies have been hampered by data limitations, including small numbers of charter schools, limited data on student characteristics, and so on.

To date, most of the recent literature has focused on charter schools' aggregate academic performance relative to comparable noncharter public schools. Few studies have even attempted to provide systematic explanations for *why* some charter schools are more successful than others.

Studies Estimating Overall Academic Impact

Two groups, one from RAND and the other from Western Michigan University, reviewed existing literature on charter schools' academic productivity. While using different methods, both concluded that the evidence is mixed. Gill, Timpane, Ross, and Brewer (2001) of RAND restricted their review to three statewide studies (Arizona, Michigan, and Texas) that were judged to be of high analytical quality and based on high quality data. An examination of three years of achievement data in Arizona (Solmon, Paark, & Garcia, 2001) found that, when compared with students remaining in noncharter public schools, students who attended charter schools for two to three years made higher gains in reading and similar to slightly higher gains in math. Interestingly, the Arizona study found that the effects of one year of charter school attendance were negative, but that "the positive effect of charter schools outweighed the negative effect of moving" (Solmon, Paark, & Garcia, 2001, p. 20).[2] At the other end of the spectrum, the RAND team adduced an early study of Michigan charter schools (Bettinger, 1999) that found charters to be lagging behind in fourth grade scores and on par with comparable noncharter public schools in the seventh grade. The authors caution, however, that the Michigan findings came early in the state's reform experiment and might not reflect the potential of charter schools. Finally, the RAND team reviewed findings from a study of Texas charter schools (Gronberg & Jansen, 2001) that found that while the overall effect was negative, schools targeting at-risk students produced higher gains than similar noncharter public schools. Both the Arizona and Michigan studies suggest that older charter schools tend to outperform newer ones.

A parallel review of the research by Miron and Nelson (2004) approached the literature somewhat differently, reviewing studies with a wider range of technical quality but assigning quality weights in the final synthesis. This synthesis gave highest quality rankings to the studies included in the RAND analysis, but also included 14 additional studies.[3] For instance, this review included positive findings from Connecticut's charter reform (Miron & Horn, 2002), a mixed-to-negative reanalysis of California data (Rogosa, 2002), and mixed-to-negative findings from North Carolina (Noblit & Corbett, 2001). Miron and Nelson's conclusions were similar to those of the RAND team," . . . with studies of some states suggesting positive impacts, studies from other states suggesting negative impacts, and some providing evidence of both positive and negative of impacts" (Miron & Nelson, 2004, p. 173).[4]

Since the appearance of these two reviews, a number of new studies have come out. A RAND study of California charter schools found that, on average, charter schools produce similar levels of academic achievement to comparable noncharter public schools (Zimmer et al., 2003). Hanushek, Kain, and Rivkin (2002) found no significant charter effect relative to conventional schools in Texas. Bilfulco and Ladd (2003) found small negative effects for charter schools in North Carolina, and Crew and Anderson (2003) found that charter schools in Florida were not performing

[2] While recognizing the study as an importance advance, Nelson and Hollenbeck (2001) raised concerns about the match between the Arizona studies sampling strategy and the interpretation of the regression coefficients, the effects of specification error, and the authors' benefit-cost estimates.

[3] Some of the included studies were not available in time for inclusion in the RAND analysis. However, others were excluded by the RAND team on quality grounds.

[4] Interestingly, adding the lower-quality studies to the mix did little to change the overall picture.

as well as noncharter public schools. Eric Bettinger (2002) updated his earlier study of Michigan charter schools. As before, he found that charter schools in Michigan have lower scores than matched noncharter public schools and they are gaining less over time. These findings are in line with the findings from the two state evaluations of charter schools (Kleine, Scott, & White, 2000; Horn & Miron, 2000; Miron & Nelson, 2002) and a study conducted by The Upjohn Institute for Employment Research (Eberts & Hollenbeck, 2002). Similarly, Solmon and Goldschmidt (2004) released a paper that confirms Solmon's earlier achievement findings from Arizona (see above) and added findings on retention and school switching. Finally, in a preliminary study of student achievement in Wisconsin, Dickman et al. (2003) found that charter schools had higher scores than similar noncharter public schools in 2001-02.

Variation Among Charter Schools

Instead of prescribing a set of particular educational approaches, charter school laws seek to free local educators to try a wide variety of approaches. Thus, we might expect considerable variation in performance among charter schools. A small number of studies provide evidence on such variations.

Loveless' (2002) ten-state study of achievement in charter schools provides evidence of variation *between* states. In four states, charter schools were performing significantly lower (i.e., Massachusetts, Michigan, Minnesota, and Texas), in one state charter schools were performing slightly higher (i.e., Colorado), and in the remaining five states (Arizona, California, Florida, Pennsylvania, and Wisconsin) charter schools were performing slightly lower than their comparison groups. Loveless also found that urban charter schools performed better than rural or suburban charter schools, large schools performed better than smaller schools, and older charter schools performed better than those with less experience.

Other studies have documented large *within-state* differences among charter schools. RAND's study of California charter schools (Zimmer et al., 2003), for instance, found differences according to start-up vs. conversion status and between classroom- vs. non-classroom-based schools. Likewise, Gronberg and Jansen (2001) found differences between "at-risk" and other charter schools.[5] Even in poor performing states like Michigan, a number of charter schools were performing much better than their comparison districts. Likewise, in Connecticut, where the results were generally positive, a few schools showed gains far lower than their host districts (Miron & Horn, 2002).[6] A few studies have sought to explore such within-state performance differentials.

However, while studies such as these are starting to describe these differences among charter schools, to our knowledge, none has provided an *explanation* for such differences. This study seeks to help advance the discussion about the mechanisms that drive differences in charter school performance through an analysis that replaces discussion of differences in categorical distinctions

[5] The findings from Hanushek, Kain, and Rivkin (2002) differ from Gronberg and Jansen (2001) since they found no positive effects for disadvantaged students.

[6] Several of the poor performing charter schools in Connecticut were closed or chose not to seek renewal of their charters, which lifted the overall aggregate results in the state.

with analysis of school-level organizational characteristics.[7] We turn next to an identification of hypotheses about the correlates of success in charter schools that drives the empirical analysis presented later in the article.

Identifying Potential Correlates of Success

Variation in performance among charter schools creates fruitful opportunities to explore whether certain uses of charter school autonomy are more productive than others. Having identified these variations, the next step is to identify which attributes of charter schools and their contexts might be systematically related to variations in academic productivity. We call these factors the "correlates of success." This section of the article seeks to identify such correlates by means of a conceptual analysis of the charter school program theory and on empirical literature that illuminates key elements of that theory.

Identifying potential correlates of success in charter schools is difficult given that there have been surprisingly few attempts to articulate the mechanisms through which charter schools are supposed to work. Existing efforts, moreover, tend to be cast at a high level of generality and are rarely linked to the larger literature on school productivity. Indeed, Gill et al. (2001) bemoan the dearth of studies that attempt to get inside the "black box" of the charter and voucher program theories. While a comprehensive attempt to link the charter theory to the larger literature is beyond the scope of this study, we hope that even this modest attempt will move the discussion forward.

State-Level Correlates

Rather than prescribing specific actions, the charter school theory of action prescribes a set of structural changes that seek simultaneously to expand the choice set made available to school actors and to transform the matrix of incentives and constraints faced by these actors as they seek to craft specific school-level designs out of this choice set. First, charter schools are exempt from a wide variety of rules that apply to other public schools. While varying from state to state, these exemptions typically cover teacher certification, curricula, collective bargaining, and the school calendar. Deregulation is supposed to free up previously underutilized creativity and increase the sense of ownership and stewardship among school staff (see, e.g., Miron & Nelson, 2002; Nelson & Miron, 2004). Moreover, by driving key decisions down to the school level, deregulation might allow enlightened school leaders to craft more coherent and focused educational offerings than is possible under the kaleidoscope of cross-cutting mandates faced by most public schools.

The grant of autonomy to charter schools, however, is not a blank check. Indeed, the other side of the oft-discussed "autonomy-accountability bargain" is that the schools are subject to increased scrutiny and incentives to perform. One such incentive to perform is the charter document itself, which grants the schools the authority to operate as public institutions on the condition that they meet goals enumerated in the charter.[8]

---

[7] As Przeworski and Tueune (1970) stated the matter, comparative analysis is the process of "changing proper names into variables."

[8] In most states, charter schools are also accountable for meeting state performance standards. More recently, they have become subject to standards under the No Child Left Behind Act.

Third, since charter schools are schools of choice, and since money follows the students, charter schools that fail to attract and retain students will, in theory, go out of business. Also, choice initiates a sorting process in which "consumers" are matched with providers whose services best match consumers' educational preferences. The end result, in theory, is that more consumers will end up in schools that match their tastes than under a more centralized assignment system (see, e.g., Chubb & Moe, 1990).

For the most part, these structural changes in the school environment are matters of state law and should be expected to vary state-to-state. For instance, states vary in the degree to which exemptions from public school regulations are automatic and the range of such exemptions. Similarly, state policy usually sets out parameters governing eligibility for and selection into charter schools (e.g., random lottery with sibling exceptions). Finally, performance accountability standards, and the extent to which they apply to charter schools, are determined by state policy.

Given that this study examines data from a single state, we shall spend little time focusing on such structural features of the charter school program theory. Nevertheless, while decisions about deregulation and choice are usually made at the state level, previous studies have found that there is often considerable variation in the extent to which accountability is actually implemented. Miron, Nelson, and Risley (2002), for instance, found considerable variation in the degree of oversight exercised by authorizing agencies in Pennsylvania. Similarly, Miron and Horn (2002) found variability in the clarity and measurability of performance standards identified in charters for Connecticut charter schools. Thus,

> *H_1: Perceived accountability*. Other things equal, charter schools subject to more stringent accountability pressures will be more academically productive than other charter schools.

School-Level Correlates

It is important to note that the policy-level correlates of success identified above are quite distant from the classroom and the teaching and learning processes. Thus, to fully understand why some charter schools are more academically productive than others, we must identify potential within-school correlates of success. We draw upon two sources in identifying such correlates. First, most state laws include policy goals and instruments that relate to internal school operations, including governance, parental and community involvement, and teacher professionalism (see, e.g., Miron & Nelson, 2002). Second, we have sought to draw, where appropriate, from the broader literature on school effectiveness about types of practices that are likely to leverage improvements in school productivity. Throughout the discussion, we note how many of these school-level mechanisms might interact with the structural changes described above (deregulation, accountability, and choice).

Teachers. As noted above, the charter concept advocates granting greater discretion over teaching and learning to teachers and other school-level personnel. Wohlstetter, Smyer, and Mohrman (1994) cautioned, however, that decentralizing reforms are unlikely to work unless policymakers provide mechanisms through which newly autonomous teachers and school staff can acquire the skills and capacity to use their autonomy in productive ways. As the primary deliverers of instruction, teachers are the transmission belt that links school designs with educational

outcomes.[9] Thus, we should expect charter schools whose teachers have greater skills to produce better outcomes than other charter schools.

> H_2: *Teacher human capital*. Other things equal, charter schools with higher levels of human capital will be more academically productive than other charter schools.

Advocates of charter schools, however, point out that teacher skills are unlikely to have a positive impact unless there are also changes to the social organization of teaching. Echoing calls from other segments of the school reform literature, many charter school statutes include enhanced teacher professionalism and autonomy as a policy goal (see, e.g., Nelson & Miron, 2004). Autonomous and professional teachers, on this view, are more likely to be able to effectively diagnose and respond to the needs of individual students. More generally, calls for professionalism are based on an understanding of teaching as a nonrepetitive, nonroutine task (Rowan, Raudenbush, & Cheong, 1993). What works for one group of students might not work well for other groups of students. Organic management structures that empower teachers, therefore, will better allow them to assess individual student needs and customize educational interventions designed to address those needs.

It is important to emphasize, however, that teacher professional autonomy does not necessarily imply the freedom to be left alone in the classroom. Indeed, schools have long been criticized as loosely coupled "organized anarchies" in which there is too little coordination among educational activities (see, e.g., Weick, 1976). A shared professional culture is one way to break down this traditional isolation through a shared understanding of educational problems, greater sharing of information on promising practices, and enhanced focus on student learning (see, e.g., Marks & Louis, 1997; Louis, Marks, and Kruse, 1996). In short, the impact of teacher autonomy is more likely to be positive only if teachers are linked by a shared professional culture.

> H_3: *Teacher professional autonomy*. Charter schools with more professional *and* autonomous teachers will be more academically productive than other charter schools.

Teacher professionalism and autonomy might also engage teachers in responsibilities beyond their specialized work, overburdening teachers with a sense of accountability (Howard & Schneider, 1987) and consume teachers' energy and distract them from instructional matters (Bryk, Easton, Kerbow, Rollow, & Sebring, 1993).[10] Due to the small size of many charter schools, teachers in charter schools might have more responsibilities that are not directly related to teaching and learning. Thus,

> H_4: *Teacher academic focus*. Charter schools whose teachers are not burdened with excessive nonacademic duties will be more academically productive than other charter schools.

---

[9] See Sanders & Rivers (1996), Wright, Horn, & Sanders, (1997), and Wenglinsky (2002) for examples of studies linking teacher attributes with student achievement.

[10] Much of this discussion is informed by an unpublished literature review compiled by Catherine Awsumb Nelson.

Before leaving the subject of teacher impacts in charter schools, it is necessary to note that the impact of teacher capacity likely depends upon curriculum. In particular, schools with curricula that are less scripted should show a stronger link between teacher skill and achievement than schools with highly-scripted curricula.

> *H_5: Interactions between teacher skill and curriculum*. Charter schools with less scripted curricula will be more dependent on a highly skilled teaching force than schools with highly scripted curriculum.

Indeed, some charter school directors note that their curricula were chosen to be "teacher proof."

Organizational/mission coherence. As noted above, a key component of the charter school design is the charter itself. This document elucidates the school's specific education goals and, in theory, provides a clear basis on which to hold it accountable. In extreme cases, schools might be closed for failing to live up to the goals set forth in their charters. Alternatively, schools that fail to live up to their promises might fail to attract and retain customers and, eventually, go out of business.

More than a vehicle for accountability, however, the goals and values articulated in the charter document can provide a tool for attracting like-minded families and teachers. Choice works not only through the pressures of competition, but also through a Tiebout-like sorting process in which families seek to find schools that best match their educational preferences (see, e.g., Hoxby, 2001). When sorting occurs on the basis of race and other suspect categories, its consequences can be legally suspect. However, when sorting takes place on the basis of educational preferences, it can allow schools to spend less time "selling" programs and approaches to its stakeholders and more time implementing them. Moreover, the focus that preference homogeneity brings can allow schools to capitalize on economies of specialization. Thus,

> *H_6: Attitudinal congruence*. Other things equal, schools that attract families with more homogeneous educational preferences will be more academically productive than other schools.

Coherent school missions, however, require more than clear statements in documents. They also require buy-in from a core group of stakeholders. Such buy-in, in turn, can be threatened by high rates of turnover.

> *H_7: Staff and student turnover*. Other things equal, schools with lower rates of staff and student turnover will be more academically productive than other schools.

Indeed, several previous studies of charter schools identify teacher and student turnover as threats to charter school missions (see, e.g., Miron, Nelson & Risley, 2002). However, we know of no study that has linked turnover to academic productivity in charter schools.

Parent and community involvement. Another common intermediate goal in charter school statutes is enhanced parent and community involvement in schools. Indeed, it is hoped that enhanced choice, along with increased site-based autonomy, will leave parents and community

organizations more invested in local schools and more willing to contribute time and other resources to their operation. Although the literature lumps a wide range of activities, behaviors, and values under the heading "parent involvement," they are united by the premise that greater parent involvement with the school enhances the home learning environment and the kinds of expectations communicated to students (particularly by making them more consistent with what is happening at school) (SEDL, 2002; Epstein, 1992,1995; Baker and Soden, 1997; Coleman, 1987, 1991). This enhancement of the home learning environment is expected to benefit student achievement both through coproduction (more productive and better aligned learning activities go on in the home) and normative support (students hear the same messages about learning at home and school). Unfortunately, the empirical literature on parental involvement is often tinged with advocacy and provides few clear guideposts about the direction and magnitude of any causal impact.[11] Nonetheless, the following hypothesis seems reasonable:

> H_8: *Parental involvement*. Other things equal, schools with higher rates of parental involvement will be more academically productive.

Similar arguments also apply to community groups. Thus,

> H_9: *Support from community groups*. Other things equal, schools with higher rates of parental involvement will be more academically productive.

In summary, the charter school program theory features three core structural policy changes that seek to alter the environment in which public schools operate. These include deregulation, performance and contractual accountability, and the introduction of choice. However, the program theory is less clear about how these structural changes might lead to changes in curriculum, instruction, and other mechanisms with a more direct impact on achievement. Here, the charter program theory resembles an artist's pallette that includes a number of preexisting reforms. These include teacher professional autonomy, parent-community involvement, and mission coherence. We hypothesize that variation in charter school academic productivity will depend on which elements are drawn from the pallette and how they are combined on the "canvas" of the school.

Data and Methods

This section describes the sample of charter schools against which we tested hypotheses about the correlates of academic success, our measure of academic productivity, the variables used to represent these hypotheses, and the estimation methods used.

Sampling and Case Selection

Ideally, an empirical investigation of the correlates of success in charter schools would involve schools from several states. This would allow the analysis to consider the influence of both

[11] Once again, this section is informed by an unpublished literature review by Catherine Awsumb Nelson.

state-level policy variables and variations in school design. The present study represents a first foray into the area through the examination of the correlates of success in Pennsylvania charter schools.

Pennsylvania was selected for several reasons. First, the state has a considerable number of charter schools, creating enough variation for an empirical test of the correlates of success. Second, while all state charter school laws have unique features, the Pennsylvania law is regarded by advocates as neither particularly strong nor weak (the Center for Education Reform gives the law a "B"). Thus, the law represents the middle of the road in charter school policy design. Finally, the authors' five-year evaluation of the state's charter school law provides a wealth of qualitative insights with which to supplement the largely quantitative analysis presented here.

Measuring Student Achievement Gains

This study seeks to describe and explain variations in charter schools' success in leveraging achievement gains in their students. Developing a measure of achievement gains proved to be a particular challenge in Pennsylvania. The absence of a widely used commercial assessment in Pennsylvania charter schools forced us to rely on the state's criterion-referenced assessment, the Pennsylvania System of School Assessment (PSSA). Administered to 5th, 8th, and 11th graders in reading and math,[12] the assessment provides no way to track individual students (or even cohorts of students) from one year to the next. Thus, growth or decline in a given school might be due to changes in student composition, as well as value added by the school.[13]

In order to separate score changes due to changes in student composition from those due to school effects, we regressed PSSA scores from noncharter public schools against (a) percent of students eligible for free or reduced-price lunch, (b) percent of students in various racial categories (Black, Asian, Hispanic, and Indian), (c) percent of students with individualized education plans, (d) urbanicity, (e) PSSA participation rate, and (f) school enrollment.[14] Separate regressions were estimated for each year charter schools have been open in the state (1997-98 to 2001-02). Estimates from the regressions were used to generate residualized scores for each charter school and year. Thus, changes in the residuals provide an estimate of growth or decline in achievement levels net of changes in student composition. These "filtered" PSSA scores comprise the dependent variable in the remaining analyses (see Miron, Nelson, & Risley, 2002; Nelson & Applegate, 2002).

It is necessary to note at the outset that while the filtered scores seek to approximate estimates of value added, they are subject to considerable weaknesses. Indeed, we were unable to obtain reliable data on mother's education, family structure, and other demographic factors known to be

---

[12] The PSSA also includes a writing assessment. However, writing scores were omitted from the analysis because of several years' missing data for both charter and noncharter schools.

[13] Attempts to track cohorts of students from 5th to 8th and 8th to 11th grades are compromised by considerable student turnover in many of the schools.

[14] Achievement scores were PSSA scaled scores. Data on free and reduced-price lunch enrollment, race, PSSA participation, and enrollment were taken from the state's School Profiles system (available on-line). Remaining variables were obtained from the National Center for Education Statistics' Common Core of Data. The urbanicity variable was entered as discrete dummy variables in order to allow for nonlinearity in the relationship to achievement. Estimates were weighted by total school enrollment. In total, the variables accounted for approximately 75 percent of school-to-school variance in each of the 5 years.

correlated with achievement. Moreover, even the best set of demographic measures would miss score change due to unobservables such as motivation and the value families place on education.

Measuring Potential Correlates

As noted above, the sample of charter schools examined in this study come from a single state. Thus, it was not possible to test propositions about policy-level correlates. Data for school-level correlates were derived from state administrative data and data from original surveys administered to charter school teachers and parents. Administrative data were obtained from the state education department's Web site and were, where necessary, supplemented with data from the National Center for Education Statistics' (NCES) Common Core of Data. Surveys were administered to all charter school teachers and a cluster sample[15] of parents during the 1998-99, 1999-00, and 2001-02 academic years. Response rates on the teacher survey were typically 73 to 86 percent, while responses on the parent surveys were considerably lower: 44 to 51 percent.

The relatively low response rate on the parent survey, combined with the fact that there were no surveys administered during the 2000-01 academic year, left a preponderance of missing values on many variables representing correlates of success. The combination of a large proportion of missing values and small number of schools in the sample (37) made the use of standard imputing methods suspect. Instead, we opted for the more straightforward approach of taking the within-school median[16] value of each correlate as a time-invariant factor. Thus, the analysis can say little about the impact of *changes* in correlates on score growth.[17] Table 1 summarizes the hypotheses presented in the previous section and provides information on measures and corresponding data sources. The remainder of this section briefly discusses each of the measures identified in the table.

Table 1: Hypotheses, Variables, and Data Sources

Hypothesis	*Variables*	*Data Source*
H_1: Accountability	Perceived accountability (mean)	Teacher/staff survey
H_2: Teacher human capital	Mean teacher experience (years)	Teacher/staff survey
	Percent teachers with graduate degree	Teacher/staff survey
	Pupil-teacher ratio	Administrative data
H_3: Teacher professional autonomy	Perceived classroom autonomy (mean)	Teacher/staff survey
	Perceived schoolwide influence (mean)	Teacher/staff survey

[15] School was the unit of clustering in the sample.

[16] The median was chosen to reduce the influence of outliers in some schools' time series.

[17] We considered including slope estimates of change in correlates as elements in the model. However, the small number of cases restricted our ability to add right-hand side variables to the analysis.

Hypothesis	*Variables*	*Data Source*
	Perceived mission commitment among other teachers (mean)	Teacher/staff survey
H_4: Teacher academic focus	I have too many noninstructional duties (mean)	Teacher/staff survey
H_5: Skill-curriculum interaction	No measure of curriculum	N.A.
H_6: Attitudinal congruence	Index of attitudinal incongruence (mean)	Teacher/staff and parent survey
H_7: Turnover	Student instability (entries plus exits as a proportion of total enrollment)	Administrative data
	Teacher stability (average years at school divided by number of years that school is open)	Teacher/staff survey
	Current administrator present at school founding	Teacher/staff survey
H_8: Parental involvement	Teachers' perceptions	Teacher/staff survey
H_9: Support from community groups	School receives support or assistance from nonprofit organization	School documents and interviews
	School contracts with EMO	School documents

Accountability. An ideal measure of accountability would rely on an independent and objective assessment of degree and quality of oversight and performance incentives. Such an examination, however, was beyond the scope of this study. Instead, we rely on teachers' perceptions of the extent to which their school is accountable for performance, as expressed on the teacher/staff survey. Specifically, the question asks staff to respond to the statement, "Teachers and school leaders are accountable for student performance" on a 5-point Likert scale (1 = strongly disagree; 5 = strongly agree). While there is no way to determine whether these perceptions are grounded in reality, the strength of this measure lies in the fact that accountability requirements are not likely to have an impact unless they are perceived as sufficiently stringent by school staff.

Teacher human capital. We used two measures of teacher human capital (teaching skill): average teacher experience (in years) and the percentage of teachers holding graduate degrees. As previous research has demonstrated, such measures are generally not strongly related to student achievement. Unfortunately, we were unable to obtain reliable data on more direct measures of teacher cognitive ability and pedagogical skill (see, e.g., Wenglinsky, 2002). Finally, the pupil-teacher ratio (the number of pupils divided by the number of teachers in school *i*) provides a rough measure of student access to teachers.

Teacher professional autonomy. As noted above, the literature on teacher empowerment draws a distinction between two types of teacher autonomy (see, e.g., Ingersoll, 1996). Classroom

autonomy is measured by a survey item that asked teachers to rate their agreement with the following statement: "Teachers are autonomous in their classrooms." Schoolwide influence is measured by a similarly structured item that asks teachers to respond to the claim that "Teachers are able to influence the steering and direction of the school." While professional culture is a rich and multi-faceted concept, available survey items allowed us only to ask teachers about the extent to which they agreed with the statement, "Teachers are committed to the mission of this school." All three items allowed teachers to judge the statement to be "false," "partly true," or "true."[18]

Teacher academic focus. Teacher academic focus, while very important in discussions of teacher professionalism, is difficult to measure. An examination of studies of charter schools – particularly charter school start-up – suggests that one particularly salient threat to an academic focus is the existence of too many nonacademic duties. Thus, teachers were asked to rate the extent to which they agree with the statement, "Teachers have many noninstructional duties" (5-point Likert scale).

Skills-curriculum interaction. As noted above, the extent to which variations in teacher skill influences instruction and student performance will likely depend on the extent to which the curriculum used at the school is scripted and "teacher-proof." While we were unable to obtain consistent and reliable indicators of curriculum and instruction, the evidence presented below exploits variation between math and reading curricula to approximate the effect.

Preference homogeneity. As noted above, one of the consequences of choice might be schools whose stakeholders are more focused around a core set of attitudes and beliefs about education than others. While our surveys were not explicitly designed to test this proposition, we were able to construct an index of dissimilarity in teacher versus parent attitudes on a number of core educational issues, including standards, accountability, and parental influence. First, we calculated the mean difference between parent and teacher responses to a common set of survey items. Next, those deviation scores were factor analyzed in order to create a weighted index of "attitudinal incongruence." Higher values on the index indicate increasing disagreement between parents and teachers.

Turnover. We included measures of both student, teacher, and administrator turnover in our analysis. The measure of student stability is the Commonwealth of Pennsylvania's measure of the percentage of students for whom school *i* was a new school during year *t*, plus the percentage of students that left the school during the school year.[19] The measure of teacher stability was calculated by dividing the average teacher experience at the school (in years) by the number of years the school had actually operated. Data for this indicator were based on teacher surveys. Based on surveys, school documents, and interviews we were also able to determine whether or not the administrator was at the school from the time it was started, which comprised our indicator of administrator turnover.

Parent and community involvement. We included several measures of parent and community involvement in charter schools. Parent involvement was measured by an item on the parent survey asking how many hours the family volunteers to the school during a typical month. Another item

[18] A 5-point Likert scale would have been preferable. However, space constraints on the survey instrument precluded it.

[19] Both entries and exits were included in the index, since both contribute to instability in the student population.

asked teachers to respond to the statement, "Parents can influence instruction in this school" on a 5-point Likert scale. We also sought to assess the degree to which schools received institutional and organizational support from the community. Thus, we constructed a dummy variable coded 1 if the school was formally sponsored by a nonprofit organization, 0 otherwise. Data for this indicator came from examination of school documents and interviews with key school staff.

A related indicator assessed whether the school was run by or received services from an educational management organization (EMO). While this is not "community support" in the conventional sense of the term, EMOs can provide an important source of logistical support for charter schools. As with the nonprofit variable above, this was dummy-coded 1 if the school was run by an EMO, 0 otherwise.

Student composition. As noted above, a number of measures of student composition were partialed out of the test score data prior to the correlates analysis. However, the charter school survey afforded additional opportunities to control for student demographics. One such measure assessed the proportion of students in charter school *i* who had attended a nonpublic school before attending the charter school.

Data Analysis

The analysis of correlates of academic success in charter schools focused on explaining variations in growth rates on the filtered achievement scores described above. As noted above, growth in the filtered scores indicates that the charter school in question posted score gains in excess of what would be expected due to observable change in student compositional alone. However, given that these data are not truly longitudinal, it is perhaps more appropriate to refer to the analysis as a pseudo growth curve analysis.

Taking these (albeit imperfect) filtered scores as the response variable, we executed a classic multilevel growth curve analysis (see, e.g., Bryk & Raudenbush, 1992). Specifically, the level 1 model regressed the filtered achievement score for a given school *i* during year *t* ($SCORE_{it}$) against a time counter ($TIME_t$) and an independently and identically distributed Gaussian disturbance term, ε (see Equation 1).

$$SCORE_{it} = \pi_0 + \pi_1 TIME_t + \varepsilon_{it} \quad (1)$$

The $TIME_t$ variable was scaled so that the value zero represents the first year of each school's operation. Thus, π_0 provides the predicted starting value for the growth curves, while π_1 is the annual growth rate thereafter.

Before seeking to explain variations in starting points and in growth rates, we estimated a simple unconditional two-level model in order to determine whether there was sufficient variance in starting values (π_0) and growth rates (π_1) to support an analysis of the correlates of success. This involved modeling these parameters as a function of a fixed effect, β, and a random effect, *u* (see Equations 2 and 3).

$$\pi_0 = \beta_{00} + \upsilon_0 \quad (2)$$

$$\pi_1 = \beta_{10} + \upsilon_1 \quad (3)$$

Substituting and rearranging terms yields the following model, which was estimated using mixed models theory (Bryk & Raudenbush, 1992).

$$SCORE_{it} = \beta_{00} + \beta_{10} TIME_t + \upsilon_0 + \upsilon_1 + \varepsilon_{it} \quad (4)$$

To test the assumption of linear growth we also reestimated Equation 1 as a second-order polynomial with a squared "time" term. The squared coefficients were not discernibly different from zero, suggesting that it is safe to proceed on the assumption that growth is linear.

In the linear models the typical school gained approximately 13 points per year net of changes in demographic and geographic factors. This is a modest gain given that the range of the PSSA is approximately 600 points. The model also shows that a typical school began approximately 49 points behind its comparison group. Overall, then, Pennsylvania charter schools appear to be modestly successful in generating improvements in achievement, as measured by these imperfect filtered scores (see also Miron, Nelson, & Risley, 2002; Nelson, 2004).

Examination of the variance components estimated from Equation 4 suggests that there is statistically discernible variance in both starting values and growth rates – though more so in the latter than the former. Given the larger variance components associated with the slopes and in order to conserve degrees of freedom, we relied mainly on a model in which only the slopes are conditional on the covariates.[20] In the slopes-only models, Equation 3 was extended to include *k* covariates (*C*) measuring the correlates of success discussed above (see Equation 5).

$$\pi_1 = \beta_{10} + \sum_{j=1}^{k} \beta_j C + \upsilon_1 \quad (5)$$

Similarly, in the fully conditional model Equation 2 was also extended to include the same set of covariates (see Equation 6).

$$\pi_0 = \beta_{00} + \sum_{j=1}^{k} \beta_j C + \upsilon_0 \quad (6)$$

For the slopes-only model, substituting into the level 1 model results in interactions between the "time" variable and each of the *k* covariates (Equation 7).

$$SCORE_{it} = \beta_{00} + \beta_{10} TIME + \beta_{10k} \sum_{j=1}^{k} (C * TIME) + \upsilon_1 + \varepsilon_{it} \quad (7)$$

Similar substitutions for the fully conditional model yields the same time-covariate interactions and a vector of unconditioned covariates (Equation 8).

[20] Estimates from the variance components analysis are available from the first author.

$$SCORE_{it} = \beta_{00} + \beta_{10}TIME + \beta_{0k}\sum_{j=1}^{k} C + \beta_{1k}\sum_{j=1}^{k}(C * TIME) + \upsilon_0 + \upsilon_1 + \varepsilon_{it} \quad (8)$$

All growth curve models were estimated using generalized estimating equations (GEE) theory with robust standard errors to account for clustering by school.[21] All covariates were centered so that the intercept may be interpreted as the predicted starting value for the typical charter school.

The paucity of cases severely limited the number of covariates that could be included in the model. We employed two strategies for selecting variables. First, we sought to include mostly variables that are causally "closest" to the teaching and learning process. Thus, for instance, funding levels might be a correlate of success in charter schools. Yet, it is likely that its influence would be manifested in other measures. Second, we carefully and judiciously eliminated variables that yielded small coefficient estimates. Mindful of the dangers of omitted variable bias, the decision to remove a variable was preceded by careful examination of multicollinearity with other potential covariates and sensitivity tests to determine empirically the impact of omitting the variable. As noted above, the paucity of cases (along with findings from the variance components analysis) also drove the decision to focus primarily on variations in slopes only.

Findings

This section summarizes key findings from the statistical analysis.[22] Descriptive statistics on all variables is provided in Appendix A. Table 2, in turn, provides estimates from the slopes-only model (Equation 7). Findings on each of the hypotheses are presented below. Separate models were estimated for math, reading, and for math and reading combined. The proportion of residual variance explained by the slopes-only model was 44 percent. Interestingly, that proportion only rose to 50 percent when we moved from the slopes-only to the fully conditional model. Thus, not much explanatory power was sacrificed by the decision to emphasize the slopes-only model.

Accountability

As noted above, most structural factors related to the charter school design (e.g., choice, deregulation, accountability) are constant within state. Given that the data for this study come exclusively from one state, we cannot test the impact of these factors. An exception is accountability, which can vary depending upon the characteristics of the chartering agency and the school's local political climate.

Across all models, increases in perceived accountability were associated with increases in rates of improvement. However, the relationship was statistically discernible only for math. There, a 1 point increase in the 5-point Likert scale was associated with approximately a 50 point per year increase in achievement. However, we must place this apparently high effect size estimate in

[21] The G-matrix of within-school errors was assumed to be exchangeable. Unrestricted estimates of the G-matrix were impossible given the paucity of cases.

[22] Standard regression diagnostics evaluated the plausibility of the standard Gauss-Markov regression assumptions, given the deviations already accounted for in the GEE models.

Table 2. Conditional Slopes HLM Estimates (Weighted GEE Estimates; Standard errors in Parentheses)

	Math & Reading		*Math Only*		*Reading Only*	
Variable	*I*	*II*	*III*	*IV*	*V*	*VI*
Intercept	-40.7***	-39.7***	-47.9***	-47.0***	-32.4**	-31.2**
	(11.7)	(13.4)	(11.6)	(11.9)	(14.2)	(14.4)
Time	5.4	1.4	0.5	-6.7	9.4	6.2
	(10.2)	(13.4)	(10.6)	(13.7)	(10.5)	(13.1)
Accountability and Oversight						
Perceived accountability	30.9	35.5	50.9**	54.4**	6.3	12.9
	(30)	(23.4)	(26.4)	(22.8)	(26.4)	(25.0)
Teacher Human Capital						
Mean teacher experience (years)	1.7	1.3	-0.7	-1.0	3.9*	3.5
	(2.2)	(2.4)	(2.0)	(2.2)	(2.3)	(2.4)
Percent teachers with graduate degree	-0.8**	-0.6	-0.6	-0.3	-1.1***	-0.9**
	(0.4)	(0.4)	(2.0)	(0.4)	(0.4)	(0.4)
Pupil-teacher ratio	0.7	-0.4	0.5	-0.8	1.1	0.2
	(1.4)	(1.5)	(1.4)	(1.6)	(1.3)	(1.5)
Teacher Professional Autonomy						
Perceived classroom autonomy	-114***	-96.6***	-98.7***	-84.6***	-123***	-101***
	(34.5)	(32.3)	(35.5)	(33.7)	(33.3)	(32.7)
Perceived schoolwide influence	-19.1	-22.1	-35.4*	-38.0	-9.9	-14.6
	(19.7)	(24.2)	(20.7)	(26.4)	(20.7)	(23.9)
Perceived mission commitment among teachers	89.4**	89.1*	85.4**	96.3**	102**	92.3*
	(41.0)	(48.3)	(40.7)	(49.3)	(43.1)	(48.8)
Teacher Academic Focus						
Perceived volume of noninstructional duties	37.3**	24.6	44.2**	28.3	29.1*	20.5*
	(17.6)	(17.0)	(18.5)	(17.9)	(16.8)	(16.2)
Preference Heterogeneity						
Attitudinal incongruence	-12.2	-6.9	-12.1	-6.2	-11.3	-7.4
	(10.9)	(9.0)	(11.2)	(9.1)	(10.9)	(9.3)
Turnover						
Student instability	-0.4**	-0.6***	-0.5***	-0.7***	-0.4**	-0.6***
	(0.2)	(0.3)	(0.2)	(0.2)	(0.2)	(0.2)
Teacher stability	-96.0***	-36.4	-79.4*	-15.2	-106***	-55.3
	(39.0)	(47.5)	(42.4)	(58.0)	(37.2)	(53.0)
Administrator present at founding	23.0**	34.5***	30.1***	41.8***	19.6**	29.2***
	(10.4)	(10.3)	(11.2)	(12.1)	(9.7)	(11.3)
Parent-Community Involvement						
Perceived parental influence in school (judged by teachers)	-7.3	-34.3	-12.1	-44.7	0.9	-24.4
	(28.2)	(29.3)	(28.7)	(29.9)	(29.1)	(29.1)
Hours volunteered in typical month	-17.8	-21.2*	-9.2	-12.3	-21.1*	-24.5**
	(11.3)	(11.8)	(11.2)	(11.7)	(11.5)	(12.2)
School supported by nonprofit	--	-25.0**	--	-24.0*	--	-25.2**
		(11.7)		(13.9)		(13.0)
School contracts with EMO	--	8.7	--	14.5	--	6.7
		(16.4)		(16.0)		(15.2)
Conversion school	--	-30.6**	--	-33.8**	--	
		(13.7)		(14.9)		
Student Composition						
% previously in nonpublic school	2.0***	1.4**	1.7***	0.9	2.2***	1.6***
	(0.6)	(0.7)	(0.6)	(0.7)	(0.5)	(0.6)

Note that the Intercept indicates average level at the initial observation

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

perspective. For the average charter school we had just 3 years of data, and the PSSA has a range of approximately 600 points and a standard deviation of approximately 100 points. Also, the standard deviation of the accountability variable was just 0.5. Thus, this is equivalent to saying that a 2 standard deviation increase in accountability is associated with a 50-point increase in growth rate. Similar caveats hold for interpretation of most other coefficient point estimates.

The data provide no clear explanations for the larger effect size for math scores. However, it is possible that the more procedural and objective nature of the math component of the PSSA is more amenable to teaching to the test than the reading examination.

Teacher Human Capital

For the most part, the teacher human capital variables did not perform well. As expected, the relationship between mean years of teacher experience and growth rates was generally positive, but significant only for reading scores. Here, a 1 year increase in average experience level was associated with just a 4-point increase in growth rate – a small gain on a test with a standard deviation of approximately 100 points. The fact that the relationship was significant for reading but not for math might indicate the sort of skill-curriculum interaction described above. In most schools, reading curricula are less scripted than math curricula. Thus, it is possible that the reading-math difference is a proxy for this sort of interaction. This, however, is only speculation, since we have no reliable measures of curriculum and instruction in the schools.

Somewhat surprisingly, the percentage of teachers with graduate degrees was negatively related to achievement growth rates. While the relationship was statistically discernible in both the reading and in the pooled reading-math models, the point estimates were extremely small. One possible explanation for the unexpectedly negative relationship might lie in the state's certification requirements for charter school teachers. Specifically, teachers with graduate degrees might be more likely than others to be among the 25 percent of teachers in each school that are allowed to teach without certification. However, examination of certification data suggests that, if anything, charter school teachers with graduate degrees are more likely to be certified than other teachers. Thus, this finding remains puzzling.

The final measure of teacher human capital was pupil-teacher ratio. This is less a measure of teacher skill than of the extent to which students are likely to be exposed to individualized attention from teachers. Here, the point estimates were extremely small and statistically indiscernible from zero. It is worth noting, however, that the measures are only distant proxies for human capital.

Teacher Professional Autonomy

The findings on teacher professional autonomy are among the most interesting in the analysis. First, there was a strong and significant negative relationship between teachers' perceptions of classroom autonomy and achievement score growth. The finding was consistent across models and specification, with point estimates of approximately 100. As with the accountability findings, the scaling of the autonomy variable makes interpretation of the point estimate somewhat difficult.

As before, the magnitude of the coefficient is largely explained by the fact that the variable lies on a 3-point scale and has a relatively low standard deviation.

These findings are consistent with claims that classroom-level autonomy can lead to loose coupling in schools. Indeed, students of teacher professional culture have argued that teacher autonomy must be coupled with common professional norms in order to offset the centripetal tendencies that come with teacher autonomy. As Lee, Dedrick, and Smith (1991, p. 194) observed, "[g]roup consensus [around an organizational culture]. . . reduces the uncertainty and ambiguity of roles – factors that make general functioning difficult." Culture is a particularly important mechanism of coordination among actors in organizations that grant extensive autonomy to individuals and have open and flexible organizational structures (Miller, 1992). It is perhaps not surprising, therefore, that the coefficient on the mission commitment variable is positively related to academic productivity in charter schools. Once again, this finding is consistent across models and specifications.

Preference Heterogeneity

Along with teacher professional culture, another form of normative coherence in charter schools might be congruence among stakeholder attitudes. As noted above, choice might create a sorting process that leaves schools with relatively homogeneous clusters of like-minded individuals. As expected, our measure of attitudinal incongruence is negatively related to academic productivity. However, none of the coefficients was statistically discernible. Thus, at most we can say that the data are suggestive, but not conclusive, on this point.

Student Turnover and Teacher Stability

It is often noted that excessive turnover in school personnel and stakeholders can hinder attempts to nurture a shared culture and sense of community. Bryk and Schneider (2002), for instance, write of the need to develop relationships of trust among members of the school community and provide compelling evidence of the link between such trust and academic productivity. Trust, in turn, often takes time to develop – a process that can be short-circuited by high student and staff turnover.

As expected, turnover among students was negatively related to academic productivity in our charter school data. Interestingly, teacher stability appears to be negatively related to performance. This is a surprising finding, but may be explained by the fact that many of the schools were very new and high turnover may suggest that "functional attrition" was taking place. Effective school administrators might have been using their new autonomy to remove teachers who did not buy into the mission or who were not proving effective in the classroom. However, the coefficient becomes nonsignificant in models that include dummy variables for conversion status, EMO involvement, and nonprofit support, suggesting that the attrition variable might be a proxy for factors related to these other variables.

Teacher Academic Focus

One perennial concern about high-involvement models of school organization is that involving teachers in decision making might overwhelm them with duties that interfere with their

core teaching functions (Smylie, Lazarus, & Brownlee-Conyers, 1996, p. 194). Our findings seem to suggest the opposite, since increases in the extent to which teachers report having many noninstructional duties is positively related to academic productivity. One possible explanation for these unexpected findings is that charter school teachers who report noninstructional duties also find that these activities provide them with an enhanced sense of school stewardship and ownership (Goodlad, 1996), both of which have been suggested as contributors to academic productivity. Teachers who are stewards of the school will move the school beyond loose coupling (Cusick, 1992; March & Olsen, 1976; Weick, 1976, 1984) at one extreme and "schools as top-down bureaucracies" (Conley & Cooper, 1991; Sergiovanni & Moore, 1989) at the other.[23] The stewardship might correct the decoupling of the organizational structure and learning activity in the school (e.g., Meyer & Rowan, 1992; Shulman, 1987) and lead to teacher empowerment and higher academic achievement for students (Marks & Louis, 1997).[24]

Parent-Community Involvement

Most charter school statutes list parent-community involvement as an intermediate program goal. Perhaps surprisingly, parent involvement (as judged by teachers) is negatively related to academic productivity, calling the presumptions found in these statutes into question. This variable, however, tells us nothing about the *quality* of such involvement. Thus, there is no way to know whether high parent involvement is actually a proxy for dysfunctional governance.

Community support for charter schools can also come from formal organizations. Thus, we expected charter schools with the backing of nonprofits to be more productive than other schools. However, the coefficient estimates suggest quite the opposite. One possible explanation for this finding is that nonprofit support is endogenous and that nonprofits tend to support schools that serve students and neighborhoods that are disadvantaged in ways not fully measured by the demographic variables used to generate the filtered scores.

While perhaps not "community" involvement *per se,* EMO involvement in charter schools is often justified in similar ways – that they can provide needed managerial and technical support. Indeed, the coefficient on the EMO variable is positive but not significant. However, an ever-present confound in attempts to estimate EMO/non-EMO differences is positive selection not otherwise measured in the model. Thus, while there is perhaps some suggestive evidence of a positive EMO effect in Pennsylvania, the findings are far from conclusive.

Student Composition

Finally, we included an item from the charter school parent survey on previous school attendance to test whether schools with high proportions of former nonpublic school students might enjoy advantages over other schools. As expected, schools with higher proportions of such students do appear to have slightly higher growth rates.

---

[23] Interestingly, teachers in smaller schools were more likely to note noninstructional burdens than teachers in larger schools. Spearman's rho for the "duties" variable and school enrollment was -0.29 ($p < 0.001$).

[24] Thanks to Prof. Jianping Shen of Western Michigan University for suggesting this line of argument.

Alternative Estimations

In interpreting the findings above, we were concerned that the findings might be biased by two factors. First, recall that we concluded, based on a variance components analysis, that there was more variance in slopes than in starting points (intercepts). Thus, we focused our attention on modeling slopes rather than intercepts. However, by forcing a common intercept estimate on all schools, our models might be exaggerating differences in growth rates. To test this, we also estimated a fully conditional model in which both slopes and intercepts were conditioned on the covariates (see Equation 8). These models had twice as many variables as the models presented above. As a consequence of the reduction in degrees of freedom, the number of statistically discernible relationships dropped considerably. However, the direction of the relationships was largely the same as in the slopes-only model. For instance, the coefficient on the mission commitment variable remains near 100 and is statistically discernible at the 0.10 level. Similarly, the teacher autonomy variable is near -100 and remains statistically discernible at the 0.01 level. Only one intercept-time interactions (related to the teacher schoolwide influence) was discernibly different from zero and suggested that schools with high levels of perceived schoolwide teacher influence tend to have higher starting values and negative growth rates over time. Overall, findings from the fully conditional model leave us reasonably confident that our findings are not overly influenced by the decision to treat the intercept as fixed across cases.

Second, given the paucity of data, we were concerned that the findings could be overly sensitive to a small number of outlying observations. To test this, we reestimated the models using robust regression with Huber weights and biweights. Once again, the findings remained substantially the same. Estimates from both the fully conditional GEE model and the robust regression models are available from the authors.

Discussion

Charter schools are commonly touted as public education's R&D sector. Given the autonomy and flexibility afforded the schools, one goal of this reform is to provide a field laboratory in which entrepreneurial educators might try out ideas not possible in traditional public schools. Yet, surprisingly little is known about whether certain charter school practices are more academically productive than others. This study represents an initial foray into this sort of analysis based on data from Pennsylvania charter schools.

Key Findings and Policy Implications

At the heart of the charter school program theory lies the oft-noted "autonomy-accountability bargain," whereby schools receive enhanced autonomy in exchange for more accountability for outcomes. While the single-state sample used in this study precluded a thorough investigation of accountability mechanisms, our data do provide some evidence of the importance of accountability as a driver of charter school quality. These findings, however, are largely restricted to math outcomes and are based on *perceptions* of accountability mechanisms, not direct observation of those mechanisms themselves. Without knowledge of the accountability mechanisms, it is difficult to

discern clear policy implications from these findings without a clearer understanding of what types of policy mechanisms are most likely to engender *perceptions* of accountability in charter schools. Also, as has been noted elsewhere (see, e.g., Miron & Nelson, 2002; Miron, Nelson, & Risley, 2002), charter school accountability is an elusive concept with multiple meanings and manifestation. Thus, further exploration is required to assess the policy design implications of our findings.

Our findings on autonomy are particularly interesting. Teacher autonomy, it appears, is negatively related to school productivity. Thus, our findings are consistent with the tradition of research and commentary on schools as loosely coupled systems (see, e.g., Weick, 1976). Indeed, later researchers working on teacher professional community have noted that teacher autonomy must be counterbalanced with strong professional norms, culture, and other informal coordinative mechanisms (see, e.g., Marks & Louis, 1997; Louis, Marks, & Kruse, 1996). Thus, it appears that teacher autonomy must be approached with some caution by school designers. As with other professions, autonomy must be coupled with strong professional norms and a sense of mission (see, e.g., Perrow, 1986; Miller, 1992).

One threat to professional culture is turnover and instability. Previous studies have noted that this is often a problem in charter schools, especially during the start-up process. Our findings bear this out for student turnover but not for teacher turnover. While student turnover appears to have a modest impact on school productivity, teacher stability was negatively related to score growth. Conversely, schools with stable leadership appear to do better than other schools. These findings might suggest that relationships among school leaders and families is more important than between leaders and teachers. More likely, the teacher findings might reflect "functional" attrition in which teachers who do not agree with the schools' missions are replaced by those who do.

Finally, while the charter school policy design identifies parent and community involvement as important performance drivers, our data provide some suggestion that parent involvement might be counterproductive. But the findings are weak and, moreover, might be proxies for weaknesses in school governance. Nonetheless, it appears that school designs should approach this issue with some caution.

Limitations

There are a number of limitations in this study, many of which have been mentioned earlier in the article. We begin with limitations specific to the study of Pennsylvania charter schools and move on to limitations that might apply to studies of the correlates of charter school academic success in any state.

Measuring achievement growth. First, the achievement measures available for Pennsylvania charter schools do not allow scholars to track students (or even cohorts of students) over time. This study used regression residuals to generate "pseudogrowth" curves that seek to measure the amount of score growth above and beyond what would be expected through changes in student composition alone. However, we were unable to find usable measures of family structure and other variables known to be correlated with academic achievement. Thus, our achievement growth curves are likely contaminated by unmeasured variance in student composition.

Missing values and questions about temporal order. Second, as noted above, there was a high proportion of missing values on most of the survey-based items. This was due both to

nonresponse at the item level and more seriously to the fact that the charter school surveys were not administered during the 2000-01 academic year. Given the high proportion of missing values and small number of nonmissing cases, we were uncomfortable using regression-based or similar imputation algorithms. Instead, we used the school mean as time-invariant level 2 covariates in the HLM models. This conservative approach, while requiring fewer assumptions about missing values can be predicted based on non-missing values. This approach reduces the variance on the covariates and likely creates downward bias on the standard error estimates. Perhaps more importantly, the missing values restrict our ability to determine temporal order – that is, whether changes in the putative correlates of academic success preceded changes in achievement.

Generalizability. Third, while Pennsylvania had some 77 charter schools in operation during the last year for which data are available, we had usable data (i.e., two or more years of test data, nonmissing survey and administrative data) on just 36 of those schools. Thus, we must cautious in making inferences from our sample to the larger population of Pennsylvania charter schools. Similarly, the fact that all sampled charter schools came from one state requires caution in drawing inferences about schools in other states.

Variation in policy variables. Fourth, by virtue of the fact that all of the charter schools examined come from one state, we cannot test explanations related to state-level policy variables. Comparisons across states, however, raise thorny issues related to the comparability of test scores across states. Some scholars (see, e.g., Loveless, 2002; Nelson & Applegate, 2002) have proposed work-arounds that can be applied to the analysis of charter schools.[25] However, these solutions are all based on generous assumptions. Nevertheless, unless scholars can identify sufficiently large and representative cross-state samples of charter schools using a common commercial test (e.g., SAT9), it seems likely that future research will have to rely on these imperfect approaches.

Measures of curriculum and instruction. Finally, as noted above, this study lacks good measures of curriculum and instruction in charter schools. Thus, most of the correlates of success examined here relate to school organization. While school organization bulks large in the charter school program theory, it is still somewhat causally distant from the activities that most directly impact student achievement. Absent better knowledge of classroom practice, our understanding of charter school academic productivity will remain incomplete.

Future Research

In spite of the limitations noted above, we believe that this study represents a useful exploration of an important, but previously understudied, topic. Future research should be able to address some, but perhaps not all, of these problems.

First, limitations related to achievement growth could be addressed by studies that focus on charter schools in states with testing regimes that allow scholars to track individual students over time (e.g., North Carolina, Texas). Alternatively, scholars could study the correlates of academic success in cities with (a) a large number of charter schools and (b) a testing regime that meets the specifications above. Unfortunately, the testing properties that make assessment of score growth

[25] McLaughlin & Drori (2000) proposed a technique for doing cross-state achievement analysis that uses NAEP scores as moderator test. However, the fact that charter schools have not, in the past, participated in the NAEP limits the utility of this approach in studies of charter schools (see Nelson & Applegate, 2002).

difficult in Pennsylvania also exist in a large number of other states and cities. Thus, until states make considerable progress in implementing No Child Left Behind's consecutive grade testing requirements, scholars seeking to study the correlates of academic success in charter schools are likely to face similar problems.

Second, as noted throughout the article, our understanding of academic productivity in charter schools is limited by the dearth of reliable data on curricula and instructional practices. Indeed, the structural and organizational elements of the charter school program theory by themselves do not produce student learning. Rather, they seek to improve the contextual conditions under which effective teaching and learning can take place. Without better knowledge of curriculum and instruction in charter schools, however, we will not be able to subject this premise to rigorous examination.

Unfortunately, it is extremely difficult to develop reliable large sample indicators of curriculum and instruction (see, e.g., Hamilton et al., 2003; Spillane & Zeuli, 1999). An alternative approach would be to blend the type of quantitative approach taken in this study with in-depth ethnographic research on the mechanisms and processes that might translate the charter theory's structural and organizational elements into teaching and learning and, ultimately, student learning. This approach would blend the quantitative approach's strengths in drawing clear and rigorous comparisons with the qualitative approach's strength in providing thick description of processes and mechanisms (see, e.g., Maxwell, 2004).

Finally, future research should seek to include comparisons across states. As noted above, such comparisons are hindered by difficulties in identifying or developing common achievement metrics. Thus, it appears that researchers have two choices. First, they can rely on the work-around approaches noted above, including use of regression residuals as in Loveless's (2002) work. Alternatively, scholars could employ the sort of mixed method comparative strategies used by students on comparative politics and public policy. Such an approach might begin by estimating separate statistical models of student achievement across a number of states. Having derived estimates of the impact of various inputs and processes on achievement, scholars would use qualitative-comparative approaches to test explanations based on state level factors. In essence, this approach mimics multilevel statistical modeling approaches (such as hierarchical linear models) by using "higher" level factors (in this case measured through qualitative comparisons) to explain state-by-state variations in the impact of school-level predictors on student achievement. While we know of no systematic attempt to use this approach in combining qualitative and quantitative techniques in the education literature, it has been used in the field of comparative politics and public policy (Ragin, 1987; Rueschemeyer, Stephens, & Stephens, 1992).

Conclusion

Further examination of the correlates of success in charter schools might help bring clarity to the confusing and muddled debate on charter school innovation. The identification of innovative practices would be more reliable and informative if it could be linked to outcomes. In the absence of such linkages, past researchers have been forced to judge innovation in terms of whether charter

school processes are sufficiently different from those of noncharter schools.[26] This method of identifying innovations, however, is at odds with the charter concept's preference for accountability for results over accountability for inputs and processes. Moreover, the likelihood of diffusion would be increased if assessments of innovation were based on solid research that can link these practices to outcomes. By providing an initial and exploratory assessment of the linkages between charter school educational practices and achievement outcomes, we hope to have provided a sound, if preliminary, contribution to this debate.

[26] The Pennsylvania Charter Appeals Board, for instance, defines innovation as any practice that is not offered by surrounding noncharter public schools. See Miron and Nelson (2000) for a discussion of these issues.

Appendix A: Descriptive Statistics

Variable	*Mean*	*Std. Dev*	*Min.*	*Max.*
Filtered achievement score (pooled)	-38.2	90.2	-300.0	379.2
Filtered achievement score (math)	-44.8	93.8	-342.0	396.7
Filtered achievement score (reading)	-30.7	97.1	-343.1	427.7
Pupil-teacher ratio	20.0	23.0	2.6	286.5
% Students previously in nonpublic school	27.3	17.8	0.0	90.9
Teachers are committed to the mission of the school (3-point scale)	2.6	0.3	1.8	3.0
Teachers are able to influence the steering and direction of the school (3-point scale)	2.3	0.4	1.3	3.0
Teachers are autonomous and creative in their classrooms (3-point scale)	2.6	0.3	1.7	3.0
Teachers have many noninstructional duties (5-point scale)	3.1	0.6	1.0	4.4
Average teacher experience (years)	5.5	3.0	1.0	18.1
Percentage of teachers with graduate degrees	26.5	20.3	0.0	100.0
Parents can influence instruction (3-point scale, as perceived by teachers)	2.2	0.3	1.2	3.0
Number of hours spent volunteering during average month (as perceived by parents)	2.0	0.7	1.0	4.0
School supported by nonprofit (dummy variable)	0.3	0.5	0.0	1.0
Schools run by EMO (dummy variable)	0.2	0.4	0.0	1.0
Conversion school (dummy variable)	13.0	0.3	0.0	1.0
Student stability (index)	27.8	50.0	0.0	125.8
Teacher stability (years at school/years school is open)	0.8	0.2	0.3	1.5
Current administrator present at school founding	0.7	0.5	0.0	1.0
Attitudinal congruence among teachers and parents*	0.0	1.0	-2.6	2.0
Perceived accountability (5-point scale)	3.9	0.4	2.3	4.9

* Index compares parent and teacher attitudes on accountability, standards, and parental influence. Teacher-parent deviation scores were factor analyzed; factor loadings were used to generate weighted summary index.

References

Baker, A.J.L., & Soden, L.M. (1997). *Parent involvement in children's education: A critical assessment of the knowledge base*. Paper presented at the annual meeting of the American Educational Research Association, Chicago. ED407127.

Bettinger, E. P. (1999). *The effect of charter schools on charter students and public schools*. New York: National Center for the Study of Privatization in Education, Occasional Paper No. 4.

Bettinger, E. P. (2002, November). *The effect of charter schools on charter students and public schools.* Cleveland: Case Western Reserve University.

Bilfulco, R., & Ladd, H. (2003, November) *The impacts of charter schools on student achievement: Evidence from North Carolina*. Paper prepared for the Association for Public Policy Analysis and Management Meeting, Washington, DC.

Bryk, A. S., & Schneider, B. L. (2002). *Trust in schools: A core resource for improvement*. New York: Russell Sage Foundation.

Bryk, A.S., Easton, J. Q., Kerbow, D., Rollow, G., & Sebring, P.A. (1993). *A view from the schools: The state of school reform in Chicago.* University of Chicago, Consortium on Chicago School Reform Research.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods.* Newbury Park, CA: Sage Publications.

Chubb J., & Moe, T. (1990). *Politics, markets, and American schools.* Washington, DC: The Brookings Institution.

Coleman, J. (1987). Families and schools. *Educational Researcher*, *16*, 32-38.

Coleman, J. S. (1991). *Parental involvement in education*. Washington, DC: U.S. Dept. of Education, Office of Educational Research and Improvement.

Conley, S., & Cooper, B. (1991). *The school as a work environment*. Boston: Allyn & Bacon.

Crew, R.E., & Anderson, M. R. (2003). Accountability and performance in charter schools in Florida: A theory-based evaluation. *The American Journal of Evaluation, 24*(2), 189-212.

Cusick, P. *(1992). The educational system: Its nature and logic.* New York: McGraw-Hill.

Dickman, A., Van Dunk, E., Witte, J., Schlomer, P., & Weimer, D. (2003). *Charter schools in Wisconsin: Assessing form and performance.* Paper presented at the American Political Science Association annual meeting, Philadelphia.

Eberts, R. W., & Hollenbeck, K. M. (2002). *Impact of charter school attendance on student achievement in Michigan*. Staff Working Paper No. 02-080. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Epstein, J. L. (1992). School and family partnerships. In M. Alkin (Ed.), *Encyclopedia of educational research* (Sixth Edition). New York: Haworth Press.

Epstein, J. L. (1995). School/family/community partnerships: Caring for the children we share. *Phi Delta Kappan*, *76*, 701-712.

Gill, B. P., Timpane, P. M., Ross, K. E., & Brewer, D. J. (2001). *Rhetoric versus reality: What we know and what we need to know about vouchers and charter schools*. Santa Monica, CA: RAND.

Goodlad, J. I. (1996). Sustaining and extending educational renewal. *Phi Delta Kappan*. *78*, 228-34.

Gronberg, T., & Jansen, D. W. (2001) *Navigating newly chartered waters: An analysis of Texas charter school performance*. Austin: Texas Public Policy Foundation.

Hamilton, L., McCaffrey, D., Stecher, B., Klein, S., Robyn, A., & Bugliari, D. (2003). Studying large-scale reforms of instructional practice: An example from mathematics and science. *Educational Evaluation and Policy Analysis, 25,* 1-29.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2002). *The impact of charter schools on academic achievement*. Retrieve on-line: http://edpro.stanford.edu/eah/down.htm, March 24, 2004.

Horn, J., & Miron, G. (2000). An evaluation of the Michigan charter school initiative: Performance, accountability, and impact. Kalamazoo: The Evaluation Center, Western Michigan University.

Hoxby, C. M. (2001). *School choice and school productivity (or, could school choice be a tide that lifts all boats?)*. Paper prepared for National Bureau of Economic Research Conference on the Economics of School Choice, Islamorada, FL.

Ingersoll, R. (1996). Teachers' decision-making power and school conflict. Sociology of Education, 69, 159-176.

Kleine, R., Scott, C., & White, R. (2000). *Issues in Michigan's public school academy initiative phase II.* Lansing, MI: Public Sector Consultants.

Lee, V., Dedrick, R., Smith, J. (1991). The effect of the social organization of schools on teachers' efficacy and satisfaction. *Sociology of Education 64*, 190-208).

Louis, K.S., Marks, H.M., & Kruse, S. (1996). Teachers' professional community in restructuring schools. *American Educational Research Journal, 33*(4), 757-798.

Loveless, T. (2002, September). *The 2002 Brown Center report on American education: How well are American students learning?* Washington, DC: Brookings Institution.

March, J. G., & Olsen, J. P. (1976). *Ambiguity and choice in organizations*. Bergen, Norway: Universitetforlaget.

Marks, H., & Louis, K.S. (1997). Does teacher empowerment affect the classroom? The implications of teacher empowerment for instructional practice and student academic performance. *Educational Evaluation and Policy Analysis, 19,* 245-275.

Maxwell, J. (2004). Causal explanation, qualitative research, and scientific inquiry in education. *Education Researcher, 33*(2), 3-11.

McLaughlin, D., & Drori, G. (2000). *School-level correlates of academic achievement: Student assessment scores in SASS public schools.* Washington, DC: National Center for Education Statistics.

Meister, G., & Schuh, A. (2000). *A report on Philadelphia charter schools: Facing challenges, forging solutions*. Philadelphia: Drexel University/FOUNDATIONS.

Meyer, J. W., & Rowan, B. (1992). The structure of educational organizations. In J. W. Meyer & W. R. Scott (Eds.) *Organizational environment: Ritual and rationality* (pp. 71-97). Newbury Park, CA: Sage.

Miller, G. (1992). *Managerial dilemmas: The political economy of hierarchy*. New York: Cambridge University Press.

Miron, G., & Horn, J. (2002). *Evaluation of Connecticut charter schools and charter school initiative: Final report.* Kalamazoo: The Evaluation Center, Western Michigan University.

Miron, G., & Nelson, C. (2000). Autonomy in exchange for accountability: An initial study of Pennsylvania charter schools. Kalamazoo: The Evaluation Center, Western Michigan University.

Miron, G., & Nelson, C. (2002). *What's public about charter schools? Lessons learned about choice and accountability*. Thousand Oaks, CA: Corwin Press.

Miron, G., & Nelson, C. (2004). Student achievement in charter schools: What we know and why we know so little. In K. Bulkley & P. Wohlstetter (eds.). *Taking account of charter schools: What's happened and what's next?* New York: Teachers College Press.

Miron, G., Nelson, C., & Risley, J. (2002). *Strengthening Pennsylvania's charter school reform: Findings from the statewide evaluation and discussion of relevant policy issues*. Harrisburg, PA: Pennsylvania Department of Education.

Nelson, C. (2004). *The academic productivity of charter schools: A methodological discussion and evidence from Pennsylvania*. Kalamazoo: The Evaluation Center, Western Michigan University.

Nelson, C., & Applegate, B. (2002). *Developing a cross-state comparable method of measuring student achievement in charter schools.* Kalamazoo: The Evaluation Center, Western Michigan University.

Nelson, C., & Hollenbeck, K. (2001). *Does charter school attendance improve test scores? Comments and reactions on the Arizona achievement study*. Staff Working Paper 01-70. Kalamazoo: W.E. Upjohn Institute for Employment Research.

Nelson, C., & Miron, G. (2004). Professional opportunities for teachers: A view from inside charter schools. In K. Bulkley & P. Wohlstetter (Eds.). *Taking account of charter schools: What's happened and what's next?* New York: Teachers College Press.

Noblit G., & Corbett, D. (2001). *North Carolina charter school evaluation report*. Raleigh, NC: Department of Public Instruction.

Perrow, C. (1986). *Complex organizations: A critical essay*. Glenview, IL: Scott, Foresman.

Przeworski, A., & Teune, H. (1970). *The logic of comparative social inquiry*. New York: Wiley.

Ragin, C. (1987). *The comparative method: Moving beyond qualitative and quantitative strategies*. Berkeley: University of California Press.

Rogosa, D. (2002). *A further examination of student progress in charter schools using California API*. Palo Alto, CA: Stanford University, Department of Statistics. [Online]. http://www-stat.stanford.edu/~rag/api/charter.pdf. Retrieved March 24, 2004.

Rowan, B., Raudenbush, S., & Cheong, Y. (1993). Teacher as a nonroutine task: Implications for the management of schools. *Educational Administration Quarterly, 29*, 479-500.

Rueschmeyer, D., Stephens, E., & Stephens, J. (1992). *Capitalist development and democracy*. Chicago: University of Chicago Press.

Sanders, W., & Rivers, J. (1996). *Cumulative and residual effects of teachers on future student academic achievement.* Knoxville: University of Tennessee Value-Added Research and Assessment Center.

Southwest Educational Development Laboratory (2002). *Emerging issues in school, family, and community connections: Annual synthesis 2001.* Austin: SEDL.

Sergiovanni, T. J., & Moore, J. H. (Eds.) (1989). *Schooling for tomorrow: Directing reforms to issues that count*. Boston: Allyn and Bacon.

Shulman, L. S. (1987). Paradigms and research programs in the study of teaching: A contemporary perspective. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed.). New York: Macmillan.

Smylie, M., Lazarus, V., & Brownlee-Conyers, J. (1996). Instructional outcomes of school-based participative decision making. *Educational Evaluation and Policy Analysis*, *18*(3), 181-198.

Solmon, L., & Goldschmidt, P. (2004). *Comparison of traditional public schools and charter schools on retention, school switching, and achievement growth.* Policy Report # 192. Phoenix, AZ: Goldwater Institute.

Solmon, L., Paark, K., & Garcia, D. (2001). *Does charter school attendance improve test scores? The Arizona results*. Phoeniz, AZ: Goldwater Institute.

Spillane, J., & Zeuli, J. (1999). Reform and teaching: Exploring patterns of practice in the context of national and state mathematics reforms. *Educational Evaluation and Policy Analysis, 21,* 1-27.

U.S. Department of Education. (2004). Charter school facts. <www.charterschools.org> (online: retrieved March 24, 2004.

Weick, K. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly, 21*, 1-19.

Wenglinsky, H. (2002). The link between teacher classroom practices and student academic performance. *Education Policy Analysis Archives, 10*(*12*).

Wilson, J.Q. (1989). *Bureaucracy*: *What government agencies do and why they do it.* New York: Basic Books.

Wohlstetter, P., Smyer, R., & Mohrman, S. (1994). New boundaries for school-based management: The high involvement model. *Educational Evaluation and Policy Anlaysis, 16*, 268-286.

Wright, S., Horn, S., & Sanders, W. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education, 11*(1), 57-67.

Zimmer, R., Buddin, R., Chau, D., Daley, G., Gill, B., Guarino, C., et al. (2003). *Charter school operations and performance: Evidence from California*. Santa Monica, CA: RAND.

About the Authors:

Christopher Nelson
RAND Corporation
201 N. Craig St., Suite 202
Pittsburgh, PA 15213-1516
Telephone: 412-683-2300
E-mail. cnelson@rand.org

Christopher Nelson is Political Scientist at the RAND Corporation, where he is also Associate Director of Research for the Pittsburgh Office. Previously, he was Senior Research Associate at The Evaluation Center, Western Michigan University.

Gary Miron
The Evaluation Center
4405 Ellsworth Hall
Western Michigan University
Kalamazoo, MI 49008-5237
Telephone: 269-387-5895
Email: gary.miron@wmich.edu

Gary Miron is the Chief of Staff at The Evaluation Center and Professor of Education at Western Michigan University.