DOES GREATER AUTONOMY IMPROVE SCHOOL PERFORMANCE? EVIDENCE FROM A REGRESSION DISCONTINUITY ANALYSIS IN CHICAGO

Matthew P. Steinberg

Graduate School of Education University of Pennsylvania Philadelphia, PA 19104 steima@gse.upenn.edu

Abstract

School districts throughout the United States are increasingly providing greater autonomy to local public (non-charter) school principals. In 2005-06, Chicago Public Schools initiated the Autonomous Management and Performance Schools program, granting academic, programmatic, and operational freedoms to select principals. This paper provides evidence on how school leaders used their new autonomy and its impact on school performance. Findings suggest that principals were more likely to exercise autonomy over the school budget and curricular/instructional strategies than over professional development and the school's calendar/schedule. Utilizing regression discontinuity methods, I find that receipt of greater autonomy had no statistically significant impact on a school's average math or reading achievement after two years of autonomy. I do find evidence that autonomy positively affected reading proficiency rates at the end of the second year of autonomy. These findings are particularly relevant for policy makers considering the provision of greater school-based autonomy in their local school districts.

1. INTRODUCTION

In an era of educational accountability, school leaders face increasing pressure to align their efforts with the educational needs of students and to improve school performance. To achieve these goals, policy makers have paid increasing attention to schools' organizational capacity and the institutional relationship between schools and large, complex district bureaucracies. Recognizing that decision makers at the school level are most aware of and potentially most able to efficiently respond to schools' organizational needs, school districts throughout the United States have recently begun granting more autonomy to local public schools. Over the past decade, decentralization of decision-making authority to local schools has unfolded in some of the nation's largest school districts.¹

Decentralization of decision making as an education reform strategy, however, has a long history in the United States. In response to the growth in state funding in the 1970s, school-based (site-based) management (SBM) gained traction in the 1980s among education reformers who believed greater schoolbased autonomy could improve school performance (Chubb and Moe 1990). SBM was often implemented by appointing a council at each school, made up of teachers, parents, and community members, that was given some responsibility for budget, personnel, and curriculum; in practice, however, local school councils tended to have advisory rather than decision-making authority (Wohlstetter and Buffett 1992; Mohrman and Wohlstetter 1994). By the early 1980s, approximately 60 percent of school districts with enrollments of at least 50,000 students had experimented with some form of SBM (Van Langen and Dekkers 2001), but there is little evidence that these structural reforms had any impact on school performance (Fullan 1993). Around the same time, researchers began to attend to the relationship between school achievement and how schools and school districts are structured and make decisions, with much evidence pointing to the principal's central role in school governance (Purkey and Smith 1983).

In the most recent iteration of decentralization, the locus of decisionmaking resides with principals rather than local school councils. The extent of autonomy has also expanded to include teacher professional development, school climate, safety strategies, and the school's calendar and schedule, alongside curriculum and budget. SBM programs, by contrast, generally limited control to only one or two areas of decision-making authority among budget, personnel, and curriculum decisions (Wohlstetter and Mohrman 1996).

School-based autonomy has been implemented by school districts in Boston, Chicago, Houston, New York City, Oakland, Seattle, St. Paul, and San Francisco, with officials in Clark County, Nevada, and Los Angeles school districts planning to provide greater school-based autonomy to local public schools (Ouchi and Goldschmidt 2009).

Moreover, whereas SBM historically gave local school councils budgetary authority, evidence suggests this authority was nominal at best—upward of 95 percent of the school budget was often determined before dollars were allocated to the school (Wohlstetter and Mohrman 1996).

More recently, federal education policy has reflected the shift toward decentralization of decision-making authority to principals. Under the 2010 Race to the Top (RTTT) competition, the U.S. Department of Education gave priority to states that not only expanded the number of authorized charter schools but also instituted a broader reform strategy enabling local school districts to operate autonomous public, non-charter schools. Under RTTT, autonomous public schools have greater authority to define their curriculum, hire and replace staff, organize their school day or year, and control their budget (USDOE 2009).²

Giving principals greater autonomy moves the locus of decision making to the organizational level where schools provide educational services, the goal of which is to improve student learning and achievement. Scholars argue that school leaders (in most cases, the principal) must control the key mechanisms of an educational governance system (Hansen and Roza 2005). These mechanisms include decisions pertaining to school budgets, personnel and staffing, curriculum and instruction, and general operational and administrative conditions (such as the organization of the school day and calendar) (Ouchi 2006). Of these four mechanisms, Ouchi (2004) notes that budgetary control is the crucial element and that "true decentralization" requires that budgetary control be given to each school (p. 22). The theory of comparative advantage provides one justification for school-based autonomy (Hill 2004; Hansen and Roza 2005). Because educational production (student learning) and school-based inputs to the educational production function (curriculum, instruction, and classroom management) occur at the school level, school-level professionals are strategically positioned to organize instruction and respond to students' individual learning needs (Hansen and Roza 2005). Indeed, many organizational theory scholars believe that large bureaucracies-state agencies with at least 3,000 employees—can realize improvements in performance only through decentralized decision making (Ouchi and Goldschmidt 2009). On the other hand, some scholars argue that one of the main disadvantages of decentralization is that some communities are unable and unequipped to handle the complex problems associated with local management. Indeed, these scholars note that large public agencies play a fundamentally important role in recruiting human capital (De Vries 2000). In the context of urban schools,

^{2.} The charter schools and autonomous (innovative) schools portion of the RTTT competition represented approximately 10 percent of points available for states to earn.

this amounts to the district assuming responsibility for recruiting and hiring qualified teachers and skilled school staff. The major concern regarding greater autonomy is the extent to which an unequal distribution of school resources exists. In districts where resource quality, such as the human capital capacity of teachers and school principals, varies across schools, the decentralization of decision-making authority may exacerbate the unequal distribution of schooling outcomes.

Whether school-level professionals are better positioned than central administrators to organize local schools for improvement depends on a number of factors. Among them is the capacity of local school leaders to more effectively identify areas for improvement (among the school-based inputs to educational production) and make changes to school organization in ways that lead to improvements in instruction and student achievement. If, however, the costs (both informational and transitional) to principals and teachers associated with the identification and implementation of organizational change are large, we might not expect improvements in school performance. Indeed, when an institution embarks on new innovations requiring the leader to engage new knowledge and skills, such as those required of principals under decentralization, transition difficulties may produce an implementation dip-a decline in organizational performance as leaders encounter an innovation that requires new skills and new understandings of organizational management (Fullan 2001; Herold and Fedor 2008). As a consequence, greater autonomy may lead to short-term declines in school performance as principals learn how to best use their new organizational and managerial options. The timing of organizational change is also critical. If principals have little time to translate their new autonomy into meaningful changes to how schools operate, it is unlikely that we would see any significant changes in performance. If principals become more adept over time at incorporating new leadership responsibilities into their management of school operations, however, schools may realize longerterm performance improvements. The extent to which short-term outcomes are impacted by principals' new decision-making authority is an empirical question addressed in the context of this paper.

As local school districts across the United States are increasingly encouraged to take innovative approaches to school reform strategies, we need a clearer understanding of greater school-based autonomy's potential to improve school performance. Despite the emphasis on reform strategies that decentralize decision-making authority, little evidence exists on the impact that greater local autonomy (among non-charter schools) has on school performance. In a multi-district study of decentralization, Ouchi and Goldschmidt (2009) show that the provision of school autonomy is correlated with student achievement through the school's student–teacher ratio (which they refer to as Total Student

Load or TSL); however, the authors do not provide evidence of a direct effect of school-based autonomy on student performance.3 Lottery-based estimates from Boston find that students attending pilot schools, which have decisionmaking authority over school budgets, curriculum, and scheduling but remain part of the Boston Public School district, realized mixed achievement results in math and reading (Abdulkadiroglu et al. 2011). In the international context, a recent cross-national comparison of decentralized decision-making authority finds that autonomy reforms are related to improvements in student achievement in countries with greater levels of economic development, although adversely affecting achievement in less-developed countries with lower levels of economic development (Hanushek, Link, and Woessmann 2011). Following a 1988 reform in the United Kingdom that gave public secondary schools the option to opt out of local school district control and become autonomous schools funded directly by the central government, Clark (2009) finds that these newly autonomous secondary schools realized large achievement gains in the pass rates of eleventh-grade students.

In this paper, I explore a recent policy change in the Chicago Public Schools (CPS) that gave principals greater autonomy among a select group of schools. I first describe the nature of the autonomy granted to CPS elementary school principals and explore the choices school leaders made that reflect their preferences about areas for autonomy. My findings suggest that school leaders were more likely to exercise greater autonomy over their budget and curricular and instructional strategies than over professional development and the school's calendar and schedule. In addition, higher-achieving autonomous schools were less likely to use their newly provided autonomy, selecting, on average, one less autonomy option (of the ten options made available to them) than did lower-achieving autonomous schools. I then explore the impact of greater school-based autonomy on math and reading performance among CPS elementary schools. Using regression discontinuity methods, I find that, for schools at the discontinuity margin, there is no statistically significant evidence that the receipt of autonomy differentially impacted math or reading achievement after two years of autonomy. However, when attention is focused on the extensive margin of school performance-the share of students meeting state-determined proficiency standards in Illinois-evidence suggests that autonomous schools increased reading proficiency rates after two years of autonomy.

^{3.} Indeed, experimental evidence elsewhere has shown the robust relationship between class size and student achievement (Krueger 1999). And though Ouchi and Goldschmidt (2009) note that their conception of TSL is different from class size, they admit that "the two are related" (p. 34).

2. SCHOOL AUTONOMY IN CHICAGO

In the 2005–06 school year, CPS initiated the Autonomous Management and Performance Schools (AMPS) program. This program granted certain academic, programmatic, and operational freedoms to select schools (Cohort 1). CPS granted AMPS schools increased decision-making authority in four areas: (1) budget; (2) curriculum, instruction, and assessment; (3) calendar and schedule; and (4) professional development. These four domains, which Hansen and Roza (2005) describe as the key areas in which schools should be provided greater autonomy, reflect the definition of autonomous schools under RTTT. Indeed, some of the key policy makers overseeing the RTTT competition—in particular, Secretary of Education Duncan—were instrumental in creating the AMPS initiative in Chicago. Of the 576 regular CPS schools (486 elementary and 90 high schools), 73 elementary and 10 high schools were granted Cohort 1 AMPS status.⁴ For the purposes of this paper, I focus exclusively on the elementary schools.

CPS elementary schools were granted Cohort 1 AMPS status primarily based on the schools' prior academic achievement.⁵ According to Dr. Barbara Eason-Watkins, who was the CPS Chief Education Officer, the district sorted schools based on their academic performance from the 2003-04 school year (school growth over time was not considered), with additional consideration given to the soundness of the school's fiscal operations.⁶ In very few cases, consideration was given to schools for which the district had confidence in the principal's leadership.7 The measure of academic performance used by CPS to select schools was a composite of the 2004 Illinois Standards Achievement Test (ISAT). A school's 2004 ISAT composite results were based on the share of a school's students who were proficient on the math, reading, and science portions of the 2004 ISAT.⁸ Though CPS did not use an explicit threshold to determine AMPS status, CPS policy makers indicated through informal discussions with the author that they followed a heuristic approach-ISAT composite proficiency rates between 55 and 60 percent—when selecting schools. My empirical identification of the achievement threshold described later in

The total number of CPS schools for 2005–06 excludes 49 charter or nontraditional schools (e.g., small high schools and alternative learning centers).

Subsequent AMPS cohorts submitted applications to the district and were selected based on factors such as the principal's leadership and the school's potential for academic improvement, but there was no application process for Cohort 1 schools (Sartain et al. 2009).

A school's fiscal operations were evaluated based on the district's ratings of its financial management. For a list of the fiscal operation indicators considered, see Elmore, Grossman, and King 2006.

Author's communication with Dr. Barbara Eason-Watkins, former CPS Chief Education Officer (8 March 2011).

^{8.} For the 2004 ISAT, students in grades 3, 5, and 8 were tested in reading and mathematics, and students in grades 4 and 7 were tested in science.

detail confirms an achievement threshold of 55 percent of students proficient on the 2004 ISAT composite measure.

Schools were notified of their AMPS status in the summer before the 2005–06 school year (approximately June 2005). According to AMPS area officers, schools and school leaders were not notified prior to receipt of AMPS status that the opportunity to acquire greater school-based autonomy would become available to them in the 2005–06 school year.⁹ Cohort 1 AMPS schools were then given two years of autonomy before their AMPS status was to be re-evaluated by the school district.

The Chief Executive Officer of CPS, Arne Duncan, encouraged the district to play a minimal role in managing the newly autonomous schools. Duncan encouraged CPS to give these schools the "freedom to innovate" and emphasized that the "best thing (the district) can do is get out of their way" (Elmore, Grossman, and King 2006, p. 9). In this context, should policy makers expect that greater autonomy will improve school performance? Or, should district leaders' offer of greater autonomy be viewed as a gamble that school leaders take upon acceptance of increased decision-making authority? Indeed, Melissa Megliola-Zaikos, the CPS district official principally responsible for supporting the AMPS schools, recognized the inherent gamble schools were taking to participate in this initiative. She noted that the district was "asking schools to take new risks with us. We did not want anyone to fear being penalized if performance went down as a result" (Elmore, Grossman, and King 2006, p. 10).

Before the provision of greater autonomy in Chicago, schools had very clear expectations of district requirements concerning the organization and management of school operations. For example, the district set and maintained the school calendar, with specified days for staff professional development. The type of professional development and new teacher induction was mandated by the school district. If a principal wished to fund additional staff development or an innovative after-school program from the existing school budget, he or she was required to seek district approval to reallocate financial resources. The principal's actions were largely circumscribed by district rules and oversight. Under the AMPS initiative, however, participating schools and principals were given the opportunity to choose not only which (of ten) autonomies to take advantage of but also how to use them. This new decision-making authority represented a deviation from principals' and school leaders' prior experience with the management and operation of their schools. Indeed, principals now had the option to make substantive changes to school operations and management. They could alter the curriculum, reallocate resources, change

^{9.} Author's communication with Melissa Megliola-Zaikos, Chief Area Officer, AMPS (1 April 2011).

instructional practices, and choose professional development models for new teachers.

3. DATA AND SAMPLE

Data for this paper are from CPS student-level administrative files and publicly available CPS school files.¹⁰ The student-level data include a rich set of demographic characteristics, including a student's race, gender, grade, and birth date (which I use to create a continuous measure of a student's age). Information is also available about whether a student receives free or reducedprice lunch, has an individualized education plan (IEP), or is limited English proficient (LEP).¹¹ I link students to their schools both within and across years using student and school identifiers.

Student-level test scores in math and reading are available for the 2005–06 and 2006–07 school years. The reading and mathematics scale scores from the ISAT are available for elementary school students in grades 3–8. The ISAT is the exam that Illinois uses to measure a school's progress toward meeting adequate yearly progress benchmarks for purposes of school performance evaluation under the No Child Left Behind (NCLB) Act.

The provision of school autonomy in Chicago is a school-level intervention. As a result, it is important to consider school composition in the empirical analysis. The most important school-level variable is school achievement in the 2003–04 academic year—the primary criterion upon which schools were granted AMPS status in 2005–06. A school's prior achievement is measured by the share of students meeting or exceeding Illinois state standards on the 2004 ISAT composite measure. Additional school-level characteristics include total school enrollment, gender composition (percentage of male and female students), racial composition (percentage of African American, Hispanic, white, Native American, Asian, and multi-race students), the proportion of students identified as LEP, the proportion of students with an IEP, and the proportion of students receiving free or reduced-price school lunch.

To identify which schools were granted Cohort 1 AMPS status in the 2005–06 school year, the AMPS office provided data on both the extensive and intensive margins of school autonomy in Chicago. The extensive margin characterizes whether a CPS elementary school was granted AMPS status in the 2005–06 school year; this is an indicator variable that takes on the value of one for schools offered AMPS status and zero otherwise. The intensive margin

CPS student-level administrative files provided to the author by Chicago Public Schools and the Consortium on Chicago School Research (CCSR); publicly available CPS school files downloaded by the author from www.cps.edu.

A student with an IEP received special education services for the academic school year. A student is identified as LEP if enrolled in a bilingual education program during the academic school year.

School Characteristic	Full Sample ($n = 450$)	AMPS $(n = 73)$	Non-AMPS $(n = 377)$
Enrollment, n	646.4	684.0	639.1
	(316.7)	(370.7)	(305.2)
Male, %	50.9	49.9	51.1
	(2.6)	(2.9)	(2.5)
African American, %	53.8	28.1	58.8
	(43.3)	(35.7)	(42.9)
Hispanic, %	32.4	39.9	30.9
	(36.4)	(35.8)	(36.4)
White, %	8.2	21.2	5.7
	(15.7)	(22.5)	(12.5)
Asian, %	2.8	5.5	2.3
	(7.8)	(9.8)	(7.2)
Other Race, %	2.8	5.2	2.3
	(3.2)	(3.9)	(2.9)
IEP, %	10.7	9.6	10.9
	(5.0)	(3.6)	(5.2)
LEP, %	12.4	15.7	11.8
	(15.1)	(15.6)	(14.9)
FRPL, %	79.2	60.7	82.7
	(19.5)	(26.6)	(15.4)

Table 1. Analytic Sample: School-Level Characteristics

Notes: Means (standard deviations) shown for the 2005–06 school year. Other Race is the share of students identified as either Native American or Mixed Race. *IEP* is the proportion of a school's students in receipt of an individualized education plan; *LEP* is the proportion of a school's students identified as limited English proficient; and *FRPL* is the proportion of a school's students receiving free/reduced price lunch.

captures the extent to which AMPS schools took up autonomy, summarized by the number of autonomies (out of ten offered) that schools chose. Table 1 summarizes the school-level characteristics of the sample.

Of the 486 non-charter regular CPS elementary schools in 2005–06, information for 450 schools is available. In most cases, schools excluded from the sample did not have test score information for the 2004 ISAT composite measure, the prior achievement variable used to grant schools AMPS status. As a result, all analyses are done for the sample of 450 schools, which include all 73 AMPS and 377 non-AMPS elementary schools.

4. HOW AUTONOMY UNFOLDED IN CHICAGO

Across the four areas in which AMPS principals were granted greater decisionmaking authority—budget; curriculum, instruction, and assessment; calendar and schedule; and professional development—principals had the option to choose any of ten specific autonomies. Table 2 describes these autonomies

Table 2. AMPS Autonomies

Autonomy	Number of Schools (Percent of 73)
New Teacher Induction Program. Option allowed schools to opt out of the district's teacher induction program, design their own program, and receive \$800 per new teacher.	32 (44%)
School Calendar. Option allowed schools to organize their school year calendar, with flexibility in determining the opening and closing of the school year and dates of professional development days.	32 (44%)
Out of Area. Option allowed schools to opt out of their Area Instructional Office (AIO). While schools lost access to area staff, coaches, and professional development, they were not required to hold walkthroughs with the area team and the principal did not have to attend area principal meetings.	34 (47%)
Restructured Day Calendar. Option allowed schools to choose their own quarterly professional development days while non-AMPS schools were required to follow a schedule of half-days, known as restructured days.	38 (52%)
Self-Directed Status. Option granted schools responsibility for school operations and maintenance, such as contracting vendors directly instead of going through the district and area office.	39 (53%)
Curriculum . Option allowed schools to opt out of district curriculum initiatives and to use more innovative and untraditional programs, for which the AMPS office would provide additional financial support for new curriculum materials or additional professional development/training around the curriculum.	47 (64%)
After-School Counts. Option provided schools with greater flexibility in the hours, course content, and budget allocation for after-school programming (AMPS schools received no additional funding but could re-allocate existing funds to after-school programming).	51 (70%)
School Improvement Plan. Option allowed schools to manage the planning process (all Illinois schools are required to submit a plan to the state board of education) without area oversight and review while completing a shorter document.	54 (74%)
Attendance Plan. Option enabled schools to design their own plan to improve/increase student attendance instead of following the district's plan.	55 (75%)
Budget Transfers. Option allowed schools to transfer money across line items without having their budgets reviewed by or requiring approval from the local area office.	60 (82%)

Notes: Author's calculations from data provided by AMPS office. Of the 73 Cohort 1 AMPS schools, five schools did not choose any of the ten autonomies. Source: Elmore, Grossman, and King 2006; author's conversation with Anthony Jelinek, 12 April

2011.

	Budgetary Control	Curriculum, Instruction Assessment	Calendar & Schedule	Professional Development
Budgetary Control	61 (0.84)	58 (0.79)	37 (0.51)	45 (0.62)
Curriculum, Instruction Assessment	-	64 (0.88)	40 (0.55)	47 (0.64)
Calendar & Schedule	-	-	40 (0.55)	32 (0.44)
Professional Development	-	-	-	50 (0.69)

Table 3. Joint Distribution Matrix of AMPS Decision Areas

Notes: Cells contain the total number of schools that chose the decision areas, with the share of 73 schools in parentheses. The *Budgetary Control* area includes budget transfers and self-directed status; the *Curriculum*, *Instruction*, and Assessment area includes curriculum, after-school counts, school improvement plan, and attendance plan; the *Calendar & Schedule* area includes school calendar and restructured day calendar options; and the *Professional Development* area includes new teacher induction program and the out of area options. See table 2 for a description of each of the ten autonomies.

and summarizes their distribution across the 73 elementary schools offered Cohort 1 AMPS status.

On average, Cohort 1 AMPS schools chose approximately six autonomies.¹² Moreover, 63 percent of AMPS schools (46 of 73) chose at least six autonomies, indicating that most schools offered greater autonomy through this district initiative actively engaged their new decision-making authority. Of the ten options, AMPS schools most frequently (82 percent) chose the *Budget Transfers* option, granting schools the ability to transfer money across line items without having their budgets reviewed by or requiring approval from the local area office. Ouchi (2004) describes budgetary control as "crucial to decentralization" (p. 22), and the fact that this option was chosen most often demonstrates that school leaders granted AMPS status also viewed budgetary control as the most critical component in the provision of autonomy over school operations.

To further explore the mechanisms through which school autonomy unfolded in Chicago, I explore the distribution of autonomy choices made by AMPS schools. Specifically, I collapse the ten autonomy choices into one of the four decision-making areas in which schools were granted greater decisionmaking authority.¹³ Table 3 shows the distribution of school choices across each

^{12.} Of the 68 schools that chose at least one autonomy, two schools selected one autonomy, five schools selected two autonomies, five schools selected three autonomies, seven schools selected four autonomies, three selected five autonomies, seven selected six autonomies, nine selected seven autonomies, eight selected eight autonomies, seventeen selected nine autonomies, and five selected all ten autonomies.

^{13.} The budgetary control area includes the following autonomy options (described in table 1): budget transfers and self-directed status; the curriculum, instruction, and assessment area includes curriculum, after-school counts, school improvement plan, and attendance plan; the calendar and

of these four decision areas, as well as how principals jointly exercised their autonomy choice across the four decision areas.

The vast majority of AMPS schools chose the Budgetary Control and the Curriculum, Instruction, and Assessment areas. In particular, 84 percent of all AMPS schools chose greater autonomy over budget-related decisions, and 88 percent of all AMPS schools chose to have greater autonomy and decision-making authority around the type of curriculum utilized in the classroom and the manner in which instruction occurred. These two decision areas independently represented a much larger share of the autonomies chosen by schools. Just over half of all AMPS schools chose the option to alter the school and day calendar (55 percent), whereas approximately two-thirds (69 percent) of all AMPS schools chose to exercise greater autonomy over professional development. The joint distributions provide insight into the mix of strategies that schools exercised in their decision to undertake school-based autonomy. For example, schools that chose budgetary authority were far more likely to also choose greater autonomy over curriculum, instruction, and assessment. In particular, 95 percent of schools that selected among the Budgetary Control area also selected among the Curriculum, Instruction, and Assessment decision area.¹⁴ The conditional probability of selecting from the Calendar and Schedule and Professional Development areas was much lower. Specifically, conditional on a school selecting among the Budgetary Control area, 61 and 74 percent of schools chose the Calendar and Schedule and Professional Development areas, respectively. Therefore, evidence suggests that most autonomous schools in Chicago chose to exercise their autonomy in areas that allowed them more flexibility in budgetary concerns while simultaneously enabling them to have more options over how instruction took place during the school day.

Although a principal's individual preferences for areas of autonomy likely influenced the options they chose, principals also faced potential constraints when determining the extent of autonomy take-up. In an effort to promote organizational change and innovation in schools, the district gave schools two years of autonomy before reevaluating their autonomous status. However, principals might have limited the extent of their school's participation in autonomy as a response to the extent of teacher buy-in necessary to initiate and successfully implement organizational change. Moreover, school-level autonomy was also accompanied by a reduction in district supports, some of which principals might have viewed as essential for limiting the risk of performance

schedule area includes school calendar and restructure day calendar; and professional development area includes new teacher induction program and the out of area option.

^{14.} I use the rule for conditional probabilities and the joint proportions from table 3 to calculate the conditional proportions. For example, P(Curriculum, Instruction and Assessment | Budgetary Control) = P(Curriculum, Instruction and Assessment ∩ Budgetary Control)/P(Budgetary Control). All calculations are made considering the total number of schools (73) granted AMPS status.

declines following organizational change. For example, schools that chose to implement their own attendance plans lost access to the district-level resources that work with schools to reduce truancy.¹⁵ Additionally, structural factors at the school level, including the characteristics of students the school served, may have also influenced the extent of autonomy take-up. In particular, did lower-achieving (among newly autonomous) schools have more to gain from greater autonomy? Did principals in schools with larger student enrollments face greater barriers to implementing organizational change? Table 4 summarizes a series of regressions that explore the influence of such structural factors on autonomy take-up across AMPS schools.

Evidence suggests that school size, based on student enrollment, had no influence on the number of autonomies principals selected, or on whether a principal selected any given autonomy. However, higher-achieving schools, on average, selected one less autonomy (of the ten options available) than did lower achieving schools, suggesting that, indeed, lower-achieving schools may have had more to gain from organizational change. In particular, lowerachieving schools were more likely to make changes to the school calendar, attendance plan, school improvement plan, and budget transfers.

5. DID GREATER AUTONOMY IMPACT SCHOOL ACHIEVEMENT?

To explore whether CPS elementary schools benefited from greater autonomy and realized achievement gains, I use a regression discontinuity approach. As previously discussed, Cohort 1 AMPS status was based primarily on a school's prior year achievement. Though no explicit achievement threshold was used to determine AMPS status, the reliance on prior school achievement as the primary selection criterion created a discontinuity in the likelihood of receiving AMPS status. I estimate this point of discontinuity (what I call the "achievement threshold") in the probability of receipt of AMPS status. I then use this achievement threshold as an exogenous source of variation to estimate the impact of school autonomy in an instrumental variables framework. The fundamental assumption underlying the regression discontinuity approach is that the achievement threshold provides exogenous variation in the treatment (e.g., AMPS selection) and that unobserved characteristics of schools vary continuously with a school's achievement in the 2004 school year, which is the observable characteristic that determines AMPS status (Imbens and Lemieux 2008). Moreover, receiving AMPS status should be the sole reason for any discontinuity in outcomes. I proceed by describing the strategy for estimating the achievement threshold and then turn to formalizing and estimating a model of the impact of AMPS status on school achievement. I

^{15.} Author's conversation with Anthony Jelinek, 12 April 2011.

Table 4. Structural Factors Related to Autonomy Take-Up

					Depe	endent Varia	ble				
	(1)	(2)	(3)	(4)	(2)	(9)	(2)	(8)	(6)	(10)	(11)
	New Teacher			Restructured			After-	School			
School	Induction	School	Out of	Day	Self-Directed	Curriculum	School	Improvement	Attendance	Budget	Total
Characteristic	Program	Calendar	Area	Calendar	Status		Counts	Plan	Plan	Transfers	Autonomies
Enrollment	0.0001	0.0002	0.0001	0.0001	0.0000	-0.0000	-0.0001	-0.0001	0.0001	0.0001	0.0005
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0001)	(0.0011)
Male (%)	0.0048	-0.0155	-0.0210	-0.0306	-0.0321	0.0052	-0.0185	-0.0107	-0.0325*	-0.0518**	-0.2026*
	(0.0026)	(0.0170)	(0.0242)	(0.0218)	(0.0213)	(0.0249)	(0.0203)	(0.0213)	(0.0164)	(0.0212)	(0.1215)
African American (%)	0.189*	0.0803	0.0907	0.0725	0.01388	0.0952	0.1297	0.0529	0.0877	0.0344	0.8468
	(0.0992)	(0.1148)	(0.1229)	(0.1314)	(0.1204)	(0.0999)	(0.1138)	(0.0969)	(0.0831)	(0.0956)	(0.6736)
Hispanic (%)	0.191*	0.0817	0.0936	0.0769	0.0166	0.0992	0.1347	0.0552	0.0958	0.0404	0.8848
	(0.1002)	(0.1155)	(0.1238)	(0.1322)	(0.1216)	(0.1007)	(0.1142)	(0.0975)	(0.0832)	(0.0954)	(0.6769)
White (%)	0.1966*	0.0802	0.0870	0.0728	0.02147	0.0963	0.1389	0.0582	0.1004	0.0465	0.8983
	(0.1004)	(0.1146)	(0.1230)	(0.1316)	(0.1218)	(0.1005)	(0.1142)	(0.0979)	(0.0833)	(0.0961)	(0.6746)
Asian (%)	0.1956*	0.0913	0.0960	0.0828	0.0221	0.1044	0.1383	0.0611	0.1014	0.0427	0.9357
	(0.1012)	(0.1149)	(0.1250)	(0.1322)	(0.1223)	(0.1001)	(0.1149)	(0.0970)	(0.0824)	(0.0945)	(0.6703)

Other Race (%)	0.1836^{*}	0.0865	0.0836	0.0938	0.0327	0.0878	0.1347	0.0755	0.0873	0.0298	0.8954
	(0.1052)	(0.1199)	(0.1294)	(0.1352)	(0.1252)	(0.1001)	(0.1096)	(0.0978)	(0.0869)	(0.1007)	(0.6776)
IEP (%)	-0.0033 (0.0258)	0.0043 (0.0208)	0.0110 (0.0243)	-0.0088 (0.0217)	-0.0081 (0.0213)	-0.0089 (0.0220)	-0.0042 (0.0208)	0.0067 (0.0164)	0.0018 (0.0143)	-0.0011 (0.0124)	-0.0105 (0.1086)
LEP (%)	-0.0047 (0.0081)	0.0113 (0.0069)	0.0016 (0.0075)	0.0069 (0.0074)	0.0008 (0.0079)	0.0086 (0.0068)	0.0015 (0.0061)	0.0097* (0.0057)	-0.0006 (0.0037)	-0.0106** (0.0042)	0.0245 (0.0390)
FRPL (%)	0.0036 (0.0055)	-0.0084 (0.0056)	-0.0038 (0.0055)	-0.0060 (0.0049)	0.0006 (0.0053)	-0.0067 (0.0045)	-0.0017 (0.0042)	-0.0049 (0.0043)	-0.0032 (0.0034)	0.0029 (0.0034)	-0.0275 (0.0348)
Prior School Achievement	0.0399 (0.1115)	-0.2389** (0.0971)	-0.0384 (0.1140)	-0.0765 (0.0942)	-0.1661 (0.1018)	0.0686 (0.0949)	-0.1305 (0.0913)	-0.1786** (0.0828)	-0.1713** (0.0667)	-0.1512** (0.0632)	-1.04* (0.5374)
Sample Size (Schools)	73	73	73	73	73	73	73	73	73	73	73
R ²	0.0757	0.3167	0.1191	0.2624	0.1317	0.2469	0.2346	0.3427	0.5401	0.2984	0.3830

vith robust standard errors) in columns (1)-(9) are from linear probability models estimating the relationship between school characteristics and the	Ig each of the ten autonomy options; coefficients in column (10) represent the relationship between school characteristics and the number of total	0) selected in the 2005–06 school year. Prior School Achievement is the 2005 ISAT Composite score (the share of students in grades 3–8 who met or	e standards on math, reading, and science portions of the 2005 ISAT during the 2004-05 school year), and has been standardized within the sample	Coefficients statistically significant at the *10% and **5% levels.
Notes: Coefficients (with robust stands	probability of choosing each of the ter	autonomies (out of 10) selected in the	exceeded Illinois state standards on m	of 73 AMPS schools. Coefficients statis

address the assumptions underlying this empirical approach in the context of the analysis.

Identification of the Achievement Threshold

Following the applied literature on structural breaks (Card, Mas, and Rothstein 2008; Bertrand, Hanna, and Mullainathan 2010), I estimate the achievement threshold—the discontinuity point in the probability of receipt of AMPS status. Equation 1 models the receipt of autonomy for CPS elementary schools in the 2005–06 school year:

$$AMPS_s = \theta + \beta(Threshold_s) + g(Achievement_{s,2004}) + X_s\zeta + \mu_s.$$
(1)

In equation 1, *AMPS*_s represents actual receipt of school-based autonomy for school *s* beginning in the 2005–06 school year; *Achievement*_{*s*,2004} is the percentage of students in school *s* meeting or exceeding Illinois state standards on the 2004 ISAT composite measure; X_s is a vector of school-level characteristics for the 2005–06 school year, and includes total school enrollment, gender composition (percentage of male and female students), racial composition (percentage of African American, Hispanic, white, Native American, Asian, and multi-race students), the proportion of students identified as LEP, the proportion of students with an IEP and the proportion of students receiving free or reduced-price school lunch; and μ_s is a random error. The variable *Threshold*_s is the achievement threshold to be estimated from the data, and a dummy variable whereby: *Threshold*_s = I[*Achievement*_{*s*,2004} > η], where η is the threshold value. The function g(.) is a smooth function modeled with higher-order polynomials.

To obtain candidate values for η , I follow Card, Mas, and Rothstein (2008), who use an R²-based search method for locating the threshold value. I first estimate equation 1 by ignoring the school-level covariates and approximating *g*(.) by a constant function and iterating over values in the range {40, ϕ } for *Thesholds* in discrete one percentage point increments. Equation 1 becomes:

$$AMPS_s = \theta + \beta (Theshold_s) + \mu_s \text{ for } 40 \le Achievement_{s,2004} \le \phi.$$
 (2)

I set $\phi = 75$ and select the value of *Achievement*_{*s*,2004} that maximizes the R² of equation 2. I then estimate the following variants of equation 1: (a) approximating *g*(.) with a linear function with no school-level controls; (b) approximating *g*(.) with a linear function and controlling for school-level covariates; (c) approximating *g*(.) with a quadratic function with no school-level controls; and (d) approximating *g*(.) with a quadratic function and controlling for school-level controls; and (d)



Figure 1. AMPS Receipt by School Achievement. *Notes:* Each dot indicates the rate of AMPS receipt for schools at 1 percentage point intervals of school achievement. *School Achievement* is the running variable and represents the share of students meeting or exceeding state standards on the 2004 ISAT composite measure. The solid lines are locally weighted regressions of AMPS status on the running variable. The vertical line represents the estimated achievement threshold—55 percent of elementary students meeting or exceeding state standards. Of the five schools granted AMPS status that chose zero autonomy options, their values on the running variable were {56.8, 60.2, 74.5, 79.1, 90.9}.

covariates. I then include interactions of *Theshold*_s with g (*Achievement*_{s,2004}), and equation 1 becomes:

$$AMPS_{s} = \theta + \beta (Threshold_{s}) + g(Achievement_{s, 2004}) + \gamma (Threshold_{s})^{*}g(Achievement_{s, 2004}) + X_{s}`\zeta + \mu_{s}, \quad \text{for } 40 \le Achievement_{s, 2004} \le \phi.$$
(3)

I estimate variants of equation 3 that include linear and quadratic polynomials for g(.), both with and without school-level controls in the range {40, 75} for *Achievement*_{s,2004}. For all variants of equations 1–3, the R² reaches a maximum at 55 (e.g., 55 percent of students meeting or exceeding Illinois state standards on the 2004 ISAT composite measure). This point is henceforth considered the achievement threshold and used as an instrumental variable in the subsequent regression discontinuity analysis.¹⁶ Figure 1 shows the discontinuity in the probability of receiving AMPS status as a function of the 2004 ISAT composite measure at the estimated 55 percent achievement threshold.

^{16.} As previously discussed, the empirical identification of the achievement threshold confirms what I learned through informal discussions with CPS leaders in the AMPS office. Although CPS did not use an explicit threshold to determine AMPS status, the policy makers indicated that the rule of thumb followed when selecting schools was between 55 and 60 percent of students proficient on the 2004 ISAT composite measure.

Fundamental to this approach is the assumption that an achievement threshold actually exists. In Card, Mas, and Rothstein (2008), it is not clear, a priori, that a discontinuity exists, raising the concern that they might locate a discontinuity point even if one is not present. Like Bertrand, Hanna, and Mullainathan (2010), however, who explore affirmative action policies in India, the authors know that a threshold exists—there is an entrance exam score below which no additional students can gain university admission because all enrollment seats are filled. In my case, there is a composite school performance score below which no school was eligible for AMPS status. In particular, any elementary school scoring below 40 percent on the 2004 ISAT composite measure was ineligible to receive greater autonomy through the AMPS initiative.¹⁷ As can be seen from figure 1, only four (of 73) AMPS schools had 2004 ISAT composite performance below 55 percent.¹⁸

Another concern is the use of average school performance to determine participation in the AMPS initiative as well as to estimate the impact of autonomy on subsequent school achievement outcomes. In particular, test scores have been shown to be a noisy measure of school performance due to transitory error shocks, providing a misleading indicator of a school's true year-to-year performance (Kane and Staiger 2002). In such cases, Chay, McEwan, and Urquiola (2005) point out that mean reversion due to idiosyncratic testing noise may lead to misleading estimates of policy effects (for example, in the context of difference-in-differences strategies). Chay, McEwan, and Urquiola (2005) show that regression discontinuity designs can help overcome the potential biases induced by mean reversion when test-based ratings are used to assign program participation. I turn now to formalizing the regression discontinuity strategy used to estimate the impact of autonomy on school outcomes.

Regression Discontinuity Framework

Interest centers on the impact that greater school-based autonomy has on school performance in Chicago. The relationship between school autonomy and performance may be captured in the following student-level education production function:

$$Y_{ist} = \alpha + \theta (Autonomy_{is, 2006}) + Z_{ist} \Gamma + X_{st} \zeta + \upsilon_{ist}.$$
(4)

Schools scoring 40 percent or less on the 2004 ISAT composite measure were designated by CPS as "Schools on Probation" or "Schools of Challenge." Neither type of school was eligible to receive AMPS status (Elmore, Grossman, and King 2006).

^{18.} These four AMPS schools had 2004 ISAT composite performance of 46.0, 46.2, 46.3, and 52.5 percent.

In equation 4, Y_{ist} is an outcome for student *i* in school *s* at the end of school year *t* and *Autonomy*_{is,2006} is a variable indicating whether student *i* attended a school in 2005–06 that received AMPS status. The variables Z and X are student- and school-level covariates, respectively, and v_{ist} is a random error term.

The key estimation issue concerning equation 4 is that the provision of school-based autonomy (e.g., receipt of AMPS status) is likely correlated with unobserved characteristics of the school. Specifically, even after conditioning on student and school characteristics, the conditional mean zero assumption— $E(v_{ist}|Z_{ist}, X_{st}, Autonomy_{is,2006})$ —is likely violated. For example, unobserved characteristics of schools, such as school culture and other idiosyncrasies, are likely positively correlated with the provision of school autonomy.¹⁹ Therefore, to the extent that unobserved school heterogeneity is positively correlated with student outcomes, the endogeneity of school autonomy status will induce an upward bias into the estimate of the impact of AMPS status and the receipt of school-based autonomy on a range of student (and school) outcomes.

To overcome the endogeneity of autonomous school status and to consistently estimate the effect of the receipt of school-based autonomy (e.g., AMPS status), I use a fuzzy regression discontinuity approach. Because the provision of autonomy is a school-level intervention, I aggregate all student-level variables to the school level, and, following Lee and Lemieux (2010), use a local linear regression specification. In a two-stage least squares framework, I first predict receipt of school-based autonomy (AMPS status) for the 2005–06 school year, using the exogeneity of the estimated achievement threshold (55 percent), and then estimate the effect of predicted receipt of autonomy on school achievement outcomes. The local linear regression equations may be specified as:

$$AMPS_{s,2006} = \alpha + \gamma (Threshold_s) + \lambda (ISATdist_s^*Below_s) + \zeta (ISATdist_s^*Threshold_s) + X_{st}\Gamma + \varepsilon_{s,2006},$$
(5)

and

$$Y_{st} = \theta + \beta (\widehat{AMPS}_{s,2006}) + \pi (ISATdist_s^*Below_s) + \delta (ISATdist_s^*Threshold_s) + X_{st}\phi + \mu_{st}.$$
(6)

For equation 5, $AMPS_{s,2006}$ represents actual receipt of school-based autonomy for the 2005–06 school year for school *s*; *Threshold*_s is an indicator

^{19.} Formally, $cov(Autonomy_{is,2006}, v_{ist}) > 0$.

variable that equals one for schools at or above the achievement threshold (55 percent), and zero otherwise; *Belows* is an indicator variable that equals one for schools below the achievement threshold, and zero otherwise; *ISATdists* is the distance between the school's share of students meeting or exceeding state standards on the 2004 ISAT composite measure and the achievement threshold; **X** is a vector of school-level characteristics for the 2005–06 school year, including total school enrollment, gender composition (percentage of male and female students), racial composition (percentage of African American, Hispanic, white, Native American, Asian, and multi-race students), the proportion of students identified as LEP, the proportion of students with an IEP, and the proportion of students receiving free or reduced-price school lunch; and $\varepsilon_{s,2006}$ is a random error. The estimated magnitude of the discontinuity in the probability of receiving autonomy ($\hat{\gamma}$), from equation 1, is approximately 30 percentage points, and highly statistically significant.²⁰

I use the estimated value of $AMPS_{s,2006}$ from equation 5 to estimate the impact of autonomy on school performance. The parameter β in equation 6 estimates the difference in school outcomes at the achievement threshold. Given that I have access to student-level test score data, I use these data in the following way. I first regress the student-level math and reading test score outcomes separately on a vector of student characteristics—gender, race, lunch status, bilingual status, special education status, grade, and age—to partial out the idiosyncratic effect of student-level heterogeneity on student performance. I next recover the residuals from these regressions and aggregate the residuals to the school level. These aggregated residuals become the school-level mean performance measure (Y_{st}) for school *s* in year *t*.²¹ Because CPS did not reevaluate a school's Cohort 1 AMPS status until after two full school years with autonomy, I estimate both one- and two-year effects of autonomy for the *t* school years 2005–06 and 2006–07, respectively.

The range of the running variable over which to estimate the local linear regressions is of particular import. The choice of bandwidth involves considering the bias-precision tradeoff. A larger bandwidth will incorporate more school observations, generating more precise estimates. The cost of a wider bandwidth is the potential bias of the treatment effect, however, as the linear specification is less likely to be accurate further away from the achievement threshold. Lee and Lemieux (2010) suggest two approaches for selecting an optimal bandwidth that balances this bias-precision tradeoff. The first is a rule of thumb

^{20.} The estimated discontinuity, from equation 1 without school covariates, is $\hat{\gamma} = .313$ (with an associated *F*-statistic of 23.74); with the inclusion of school covariates, $\hat{\gamma} = .347$ (with an associated *F*-statistic of 25.38).

^{21.} I also estimate models at the student level without aggregation to the school level, clustering the standard errors at the school level to account for within-school correlation. The coefficients estimates are very similar to the school-aggregated results.

selection formula, and the second, following Imbens and Kalyanaraman (2009), is a data-dependent method for choosing the optimal bandwidth.²² The optimal bandwidth generated by these two methods is approximately 7 percentage points. Therefore, for estimation purposes, the preferred range of the running variable is between 48 and 62 percentage points on the 2004 ISAT composite measure.

As discussed, the first assumption underlying the "fuzzy" regression discontinuity approach used in this paper is that the achievement threshold provides exogenous variation in the treatment status (e.g., AMPS selection). The first stage regression results indicate that, at the achievement threshold, the probability of receiving AMPS status increases by approximately 30 percentage points, and is highly statistically significant.

The second assumption concerns schools' ability to manipulate the forcing variable—2004 school achievement—to become eligible for AMPS status. According to Lee and Lemieux (2010), "this is probably the most important question to ask when assessing whether a particular application should be analyzed as an RD [regression discontinuity] design" (p. 292). More specifically, this assumption generates the local random assignment around the achievement threshold. Although this assumption cannot be tested directly, Lee and Lemieux (2010) recommend assessing the continuity of the distribution of the assignment variable. Figure 2 shows the frequency distribution of the assignment variable. Because schools and principals were unaware before summer 2005 that the opportunity for greater school-based autonomy would become available, it is implausible that schools would have had the ability to manipulate their 2004 test score results to create such an opportunity for their schools. Figure 2 confirms this, showing no evidence of a discontinuity at the achievement threshold.²³

The third assumption underlying the empirical approach is that unobserved characteristics of schools vary continuously with a school's achievement in the 2004 school year—the observable characteristic that determines

^{22.} The rule of thumb bandwidth selection formula in kernel density estimation is: $0.9^{\circ} \hat{\sigma}^{\circ} N^{-1/5}$, where $\hat{\sigma}$ is an estimate of the variability in the running variable and N is the sample size. Using the interquartile range of the running variable (since 2004 ISAT composite is right-skewed), the rule of thumb bandwidth = $(0.9)^{\circ} (IQR_{2004ISAT} = Q_3 - Q_1 = 60.4 - 33.5 = 26.9)^{\circ} (450^{-1/5}) = 7.13$. The method suggested by Imbens and Kalyanaraman (2009) yields an optimal bandwidth of 6.93. Considering the solutions generated by both approaches, I utilize an optimal bandwidth in the range [48, 62] on the 2004 ISAT running variable.

^{23.} A density discontinuity test confirms the absence of a jump in the distribution of the running variable at the achievement threshold. The test is conducted by first partitioning the assignment variable into 1 percentage point intervals (bins) and computing the frequencies within each interval. Local linear regressions are then estimated using the frequency count as the dependent variable. The estimated jump in the frequency at the achievement threshold is not statistically significant (for the full sample, the jump is -0.19 [p-value = .640]; for the sample with preferred bandwidth of 7 percentage points, the jump is -0.748 [p-value = .374]).



Figure 2. Density of the Forcing Variable (School Achievement). *Notes:* Each dot indicates the frequency count, at one percentage point intervals, of school achievement. *School Achievement* is the running variable and represents the share of students meeting or exceeding state standards on the 2004 ISAT composite. The solid lines are locally weighted regressions. The vertical line represents the estimated achievement threshold—55 percent of elementary students meeting or exceeding state standards.

AMPS status (Imbens and Lemieux 2008; Lee and Lemieux 2010). As Lee and Lemieux (2010) note: "As in a randomized experiment, the distribution of observed baseline covariates should not change discontinuously at the threshold. . . . This is akin to the tests performed to empirically assess whether the randomization was carried out properly in randomized experiments" (pp. 292, 296). Though I am unable to empirically verify this assumption for unobserved characteristics of schools, I do explore whether the main school covariates vary continuously around the achievement threshold. Figure 3 confirms that the distribution of the share of male students, the share of students by race and ethnicity, the share of students who receive free or reduced-price lunch, and the share of students classified as LEP varies continuously at the 55 percent achievement threshold point. There is some evidence that the share of students who have an IEP and the size of the school (enrollment) vary discontinuously at the threshold. This indicates, for example, that schools with larger enrollments may have been given some priority, at the achievement threshold, for AMPS status. This does not invalidate the fuzzy regression discontinuity design, however. As long as schools cannot precisely manipulate their position vis-à-vis the assignment variable, the treatment near the threshold remains as if randomized (Lee and Lemieux 2010). Like experimental settings where there is some covariate imbalance across treatment and control groups, conditioning on the variables that are discontinuous at the threshold achieves balance on these covariates.

6. DISCUSSION OF IMPACTS ON SCHOOL ACHIEVEMENT

Of primary interest is whether greater autonomy improves school performance. I explore two margins of school performance. The intensive margin considers average school performance based on standardized values of student-level ISAT math and reading scale scores. The extensive margin considers a school's proficiency rate-the share of students who meet or exceed Illinois state proficiency standards in math and reading-as the outcome measure. Column 1 of table 5 offers preliminary evidence on whether autonomy improves school performance by reporting ordinary least squares (OLS) estimates of the relationship between AMPS status and school math and reading achievement after the first (2005-06) and second (2006-07) years of autonomy. On average, AMPS elementary schools perform better than non-AMPS schools. After the first year of autonomy, AMPS schools performed 0.23 standard deviations better than non-AMPS schools in math and 0.22 standard deviations better in reading. Moreover, there is a 22 and 21 percentage point difference in the proficiency rate in math and reading, respectively, between AMPS and non-AMPS schools. These performance differences persist after the first year of autonomy.

As is evidenced in table 1, however, AMPS schools are quite different from non-AMPS schools. Indeed, AMPS schools teach fewer students from economically disadvantaged circumstances—60.7 percent of AMPS students receive free or reduced-price lunch, compared with 82.7 percent in non-AMPS schools—and serve a smaller share of minority students—68.0 percent of students in AMPS schools are African American or Hispanic, compared with 89.7 percent of students in non-AMPS schools. Once observable differences in school characteristics are accounted for (see column 2, table 5), autonomy no longer significantly predicts a school's average math and reading achievement, while achievement differences in math and reading proficiency rates persist (except when focused on reading proficiency rates for the full sample of schools in year 1).

Based on the evidence from the OLS regressions, conditional on school characteristics, greater autonomy does not impact average school achievement (the *intensive* margin of school performance), but continues to have an effect on reading and math proficiency rates (the *extensive* margin of school performance). For the OLS regression estimates to provide a valid causal effect of autonomy, however, no differences, observable or unobservable, must remain between AMPS and non-AMPS schools once we condition on school characteristics, such as the share of students receiving free lunch and the share of minority students. But there are a number of reasons why low-performing schools (those in the bottom quartile of the performance distribution) are different relative to high-performing elementary schools. Factors such as the



Figure 3. Distribution of School Covariates. *Notes:* The dependent variables are school-level characteristics. The dots in each panel indicate the average of the dependent variable for schools at 1 percentage point intervals (bins) of school achievement. *School Achievement* is the running variable and represents the share of students meeting or exceeding state standards on the 2004 ISAT composite. The solid lines are locally weighted regressions of dependent variable on the running variable. The vertical line represents the estimated achievement threshold—55 percent of elementary students meeting or exceeding state standards.

extent of neighborhood resources, parental involvement, and teacher quality likely differ systematically across low- and high-achieving schools. Therefore, aside from indicating that AMPS schools perform better than non-AMPS schools, the OLS estimates provide little guidance on how the provision of autonomy directly impacts school performance. To gain traction on the causal effect of autonomy, I turn to the instrumental variable estimates in the context of the regression discontinuity design.

Regression Discontinuity Estimates of Achievement Discontinuity

Figure 4 presents visual evidence of the magnitude of the discontinuity in math and reading outcomes at the achievement threshold. The figure illustrates the distribution of the outcomes—math and reading results after the first (2005–06) and second (2006–07) years of autonomy—with the running variable, the 2004 ISAT composite measure, on the horizontal axis. Across the full sample of elementary schools, panels A–D of figure 4 suggest that there was no difference in the average math or reading performance on the intensive margin, at the 55 percent achievement discontinuity threshold. Though there appears to be no difference in proficiency rates after the first year of autonomy (panels E and F of figure 4), panels G and H suggest that math and reading proficiency rates in year 2 increased at the achievement threshold, with a more pronounced shift in reading.

To further explore the magnitude of the discontinuity, columns 4-5 of table 5 provide instrumental variable estimates of the impact of school-based autonomy in Chicago on school math and reading performance using the preferred bandwidth specification. The year 1 and year 2 results confirm the visual evidence presented in figure 4; school-based autonomy did not have a statistically significant effect on reading or math on the intensive margin of school performance. Although the impact on math achievement is -0.15 standard deviations after the first year, there is no conclusive evidence of a statistically significant impact. After the second year of autonomy, the impact estimates for math and reading are small and not statistically different from zero.

Therefore, based on a school's average math and reading performance, the empirical evidence suggests AMPS schools did not realize relative performance improvements in the two years of autonomy. However, for the purposes of school sanctions under NCLB legislation, schools are evaluated on the share of their students proficient in math and reading. It is plausible, then, that schools used their greater autonomy to develop strategies to target students at the proficiency benchmark, focusing on improving the school's proficiency rate by allocating effort to students at the margin of proficiency while reducing effort given to very high- and very low-achieving students. In fact, evidence

Table 5. Estimates of the Impact of Autonomy on School Achievement Outcomes

	(1)	(2)	(3)	(4)	(5)	(9)	(7)	(8)
I						Alter	native Bandwi	dths
Outcome	SIO	SIO	OLS (Preferred RD BW)	Main RD-IV Specification	Main RD-IV Specification	2004 ISAT: 43–67%	2004 ISAT: 40–70%	2004 ISAT: 35-75%
Year 1 (2005-06) Math Achievement	0.231*** (0.0328)	-0.016 (0.0230)	0.053 (0.0349)	-0.224 (0.2953)	-0.146 (0.1780)	-0.010 (0.1272)	-0.057 (0.1464)	-0.095 (0.1136)
Reading Achievement	0.217*** (0.0304)	-0.025 (0.0185)	0.047* (0.0238)	-0.091 (0.2109)	-0.033 (0.1285)	0.046 (0.1135)	0.017 (0.1268)	0.0463 (0.0886)
School Proficiency (Math)	22.15*** (1.279)	1.39* (0.749)	3.24** (1.391)	4.82 (8.662)	2.12 (7.710)	1.26 (7.11)	-1.29 (7.635)	-0.128 (5.256)
School Proficiency (Reading)	20.88*** (1.409)	1.12 (0.689)	2.65* (1.444)	17.12 (12.376)	15.28 (9.869)	9.99 (8.101)	7.08 (7.915)	5.92 (5.429)
Year 2 (2006–07) Math Achievement	0.221*** (0.0316)	-0.007 (0.0236)	0.046 (0.0377)	-0.031 (0.2439)	-0.024 (0.1809)	0.0112 (0.1414)	-0.041 (0.1656)	-0.010 (0.1217)

Reading	0.202***	-0.026	0.027	-0.054	0.014	0.087	0.100	0.120
Achievement	(0.0292)	(0.0203)	(0.0259)	(0.2093)	(0.1317)	(0.1208)	(0.1325)	(0.1006)
School Proficiency	20.18***	1.25*	3.15**	16.06	9.91	3.69	2.72	1.645
(Math)	(1.173)	(0.759)	(1.570)	(11.893)	(8.554)	(7.377)	(8.087)	(5.635)
School Proficiency	20.51***	1.43*	2.90*	23.36	18.73*	16.75*	15.99*	12.00*
(Reading)	(1.310)	(0.752)	(1.555)	(14.894)	(10.154)	(9.301)	(9.695)	(6.414)
School Characteristics		×	×		×	×	×	×
Sample Size (Schools)	450	450	111	111	111	185	221	287

its (with robust standard errors) for Math and Reading Achievement outcomes are in standard deviation units. For Year 1, the outcome variable for Math	ement is the standardized 2006 ISAT math (reading) score (mean zero and unit variance across the sample of tested students in grades 3–8); for Year 2, the	e for Math (Reading) Achievement is the standardized 2007 ISAT math (reading) score. The Math and Reading Achievement outcomes have been residualized	from a regression of math or reading achievement on the following student-level characteristics: gender, race, grade level (age as of September 1 of the	e/reduced price lunch status, IEP status and LEP status) and aggregated (averaged) to the school level. Coefficients (with robust standard errors) for School	omes are in percentage points; School Proficiency represents the share of a school's students who met or exceeded Illinois state standards on the ISAT math	School characteristics include controls for enrollment, gender, the percent of African American, Hispanic, white, Asian and other race (Native American and	ents, the proportion of students identified as LEP in receipt of an IEP and receiving free/reduced-price lunch. For models (1)–(3), the running variable (2004	is included in the set of school covariates. Coefficients are statistically significant at the *10%, **5%, and ***1% levels.
Notes: Coefficients (with robu	(Reading) Achievement is the	outcome variable for Math (Re	(e.g., residuals from a regree	school year), free/reduced pr	Proficiency outcomes are in p	or reading test. School chara	multi-race) students, the prop	ISAT composite) is included in



Panel G: School Proficiency: Math (Year 2)

Panel H: School Proficiency: Reading (Year 2)



Figure 4. Distribution of School Achievement. *Notes:* For panels A–D, the dots indicate the average of the dependent variable (Math or Reading achievement) for schools at 1 percentage point intervals (bins) of school achievement. For panels E–H, the dots indicate the proportion of students proficient in math or reading for schools at 1 percentage point intervals (bins) of school achievement. School *Achievement* is the running variable and represents the share of students meeting or exceeding state standards on the 2004 ISAT composite measure. The dependent variables in panels A–D are regression-adjusted average school performance (net of student-specific characteristics) measured in standard deviation units. The dependent variables in panels E–H are the proportion of students are locally weighted regressions of the dependent variable on the running variable. The vertical line represents the estimated achievement threshold—55 percent of elementary students meeting or exceeding or exceeding state standards.

from Chicago suggests that after the introduction of NCLB, math and reading test scores improved for students in the middle of the achievement distribution, near the proficiency benchmark, but not among the lowest-performing students; indeed, for students in the bottom 20 percent of the performance distribution, there is evidence of lower-than-expected performance in math (Neal and Schanzenbach 2010).

Looking at the extensive margin of school performance, the regression discontinuity estimates indicate a large positive effect of autonomy on reading proficiency rates. I am unable to reject the null of no effect after year 1, but there is evidence of a statistically significant impact in reading proficiency, on the order of 18 percentage points, after year 2. There is no statistically significant evidence, however, of any impact on math proficiency rates in either year. Though there were no significant changes in average math or reading after two years of autonomy, this evidence suggests that simply looking at school proficiency rates as the only measure of performance masks real differences in average school performance, differences that more adversely impact the lowest- and highest-achieving students than students in the middle of the achievement distribution. Although student reading performance, on average, did not benefit from attendance at more autonomous schools, it appears that students at the proficiency benchmark likely did benefit.

The specific mechanisms through which students in the middle of the achievement distribution benefitted can't be determined from this analysis. But it is clear that autonomous schools may have used their autonomy for more targeted instruction, or perhaps even allocated resources to enhance the potential that students at the margin of proficiency would meet achievement standards. Indeed, AMPS schools chose approximately two more autonomies at the achievement threshold.²⁴ Moreover, each of the ten autonomy choices exhibits a statistically significant increase in the probability that it was used by schools; that is, each of the ten choices is significantly and positively related to the achievement threshold. Four autonomy choices-After School Counts, School Improvement Plan, Attendance Plan, and Budget Transfers-stand out. The probability of a school choosing each of these four autonomy options increases by approximately 25 percentage points at the achievement threshold. This increase is 6-10 percentage points greater than the increase for the other six autonomy options. I cannot attribute the effect of autonomy to these options, but this evidence provides more insight into the choices schools made at the achievement threshold. Indeed, each of these options is part of either

^{24.} In an analysis not reported here, I explored the extent to which each of the ten autonomies jump discontinuously at the achievement threshold. In particular, I estimated the probability that an AMPS school chose each of the ten autonomies at the achievement threshold. Results from this analysis are available from the author upon request.

the *Curriculum, Instruction, and Assessment* or *Budgetary Control* decision areas, which influence the type of instruction in the classroom and the allocation of resources in the school.

Though the empirical strategy limits my ability to look out past two years, the positive effect on reading proficiency rates after two years suggests that as principals and school leaders are given time to learn how to use their resources and teachers become accustomed to new organizational changes, school performance can improve over time. As a result, the practice of implementing autonomy strategies appears to require both learning on behalf of school leaders and teachers and time to update how they use their new autonomy. Finally, an important caveat pertains to the generalizability of these results to the full distribution of CPS elementary schools. In particular, the impact estimates represent a localized effect (e.g., local average treatment effect, or LATE) and capture the impact of autonomy in Chicago for schools at the 55 percent achievement threshold. As a consequence, this LATE estimate does not generalize to schools across the achievement distribution and therefore does not provide for an estimate of the causal effect of autonomy in Chicago for schools with much higher prior achievement (for example, 75 percent of students meeting proficiency on the 2004 ISAT composite measure).

Robustness of Regression Discontinuity Estimates

To explore the robustness of the findings to bandwidth choice, columns 6–8 of table 5 and figure 5 present regression discontinuity estimates from alternative bandwidths. The findings on the impact of school autonomy are robust across a range of bandwidth choice. Indeed, there is no evidence of a statistically significant effect of autonomy on average reading or math achievement after two years of autonomy. The only exception is for math achievement in years 1 and 2 when attention is paid to the full sample of schools. Though the impact of autonomy on average achievement appears negative, the inclusion of schools farthest away from the achievement threshold likely biases the impact estimates, generating misleading conclusions about the impact of autonomy. The inpact of autonomy on reading proficiency rates in year 2 is robust to alternative specifications (at the 10 percent level of significance), while there remains no evidence of a statistically significant autonomy effect on math proficiency in years 1 and 2 or on reading proficiency in year 1.

7. CONCLUSION

Over the past decade, some of the nation's largest school districts have given greater autonomy to local public (non-charter) school principals. The trend toward greater school-based autonomy as a district-level plan to improve



Figure 5. Regression Discontinuity Estimates by Bandwidth Selection. *Notes:* The figures summarize the regression discontinuity estimates (with associated 95 percent confidence intervals) across a range of bandwidth choices. All estimates are generated from local linear regressions with controls for school characteristics and robust standard errors. Bandwidths (BW) are in percentage points. For BW = 2, there are 35 schools; for BW = 5, there are 87 schools; for BW = 7, there are 111 schools; for BW = 10, there are 156 schools; for BW = 12, there are 185 schools; for BW = 15, there are 221 schools; for BW = 20, there are 287 schools; for BW = 25, there are 348 schools; the full sample includes all 450 schools.

educational performance has found support at the national level. Indeed, some of the chief architects of Chicago's AMPS initiative, including Secretary of Education Arne Duncan, modeled federal policy after Chicago's experience with school autonomy. Specifically, the 2010 Race to the Top competition awarded states for designing innovative school reform strategies that let local school districts provide greater school-based autonomy. Up to this point, however, little evidence existed on how this emerging district-level reform strategy has unfolded. This paper offers initial evidence on how school leaders utilized greater decision-making authority in Chicago—with attention to the mechanisms through which this autonomy operated—and whether autonomy substantively impacted school achievement.

I have described how schools and school leaders responded to the offer of autonomy. The majority of AMPS schools (63 percent) chose at least six of the ten autonomy options; budgetary control and greater autonomy over curriculum, instruction, and assessment were the most desired autonomy areas. I find that elementary schools in Chicago that received AMPS status did not experience statistically significant changes in average math or reading performance relative to non-AMPS schools. There do appear to be statistically significant differences in reading proficiency rates after two years of autonomy, however, suggesting that schools require time to efficiently implement autonomy, and may be targeting resources to marginal students. As a result, policy makers should not expect immediate performance gains from the provision of autonomy but rather, if such policies are pursued, should allow time for schools and school leaders to learn how best to implement their new autonomy. Policy makers should be aware of the potential adverse impact on the lowest- and highest-achieving students along the performance distribution if resources are focused on improving proficiency rates, rather than average school performance.

Compare these findings to a similar study of decentralization of decisionmaking authority in the Boston Public Schools (Abdulkadiroglu et al. 2011), where the authors explored the impact of pilot schools on student achievement. The pilot schools determined their own budgets, staffing, curricula and scheduling, and their teachers were Boston Teachers Union members, very similar to the AMPS initiative in CPS. Lottery-based estimates of the effect of pilot schools for students in grades 3-8 were mixed; the authors found positive effects on elementary (grades 3-4) reading but not math, and no effect on reading among middle school students (grades 6–8) but a negative effect on math performance. In CPS, the effect of autonomy on average math achievement after one year is imprecisely estimated, although the magnitude of the negative effect is very similar to the estimated effect of attending a pilot school among sixth to eighth graders in Boston. Moreover, though I cannot reject zero impacts of autonomy on reading achievement, I do find large and marginally statistically significant effects on school proficiency rates in reading. Clark (2009) also assesses the impact of greater autonomy on pass rates, similar

to my consideration of proficiency rates. He looks at public high schools in the U.K. between 1988 and 1997, and finds large, immediate, and persistent effects on pass rates—the fraction of eleventh-grade students who passed five or more General Certificate of Secondary Education exams, standardized exams taken by all eleventh graders in the U.K. When I explore proficiency rates among a sample of students in grades 3–8, I similarly find large effects on reading proficiency, but only after two years of autonomy.

Ultimately, the findings from Chicago are short-term in nature. If schools do require more time to incorporate their new autonomy into changes in school organization, then in the long term autonomy might increase not only proficiency rates but also benefit all students along the distribution of student achievement. This, of course, is an empirical question, and presents a robust area for future research. Moreover, an assessment of school performance should extend beyond how students perform on standardized test scores and incorporate the full range of educational and social experiences that may result from greater school-level autonomy. For example, does greater autonomy improve student safety and establish a school climate that is more conducive to learning? Does autonomy offer teachers more control over their instructional practice inside the classroom and professional collaboration opportunities with other teachers outside the classroom? Understanding how greater autonomy affects these (and other) areas would provide a more nuanced understanding of the impact that decentralized decision-making has on school organization and performance.

The author thanks Kerwin Charles, Stephen Raudenbush, and Ofer Malamud for helpful discussions and suggestions on earlier drafts of this article. The author benefitted from the suggestions of two anonymous referees and comments from seminar and conference participants at the University of Chicago, the School of Public Affairs at CUNY-Baruch College, the Association for Public Policy Analysis and Management, and the American Educational Research Association. The author thanks Jonathan Wallace for editorial assistance and the Chicago Public Schools district and the Consortium on Chicago School Research for providing access to the data for this article. The author acknowledges funding from the Institute of Education Sciences.

REFERENCES

Abdulkadiroglu, Atila, Joshua D. Angrist, Susan M. Dynarski, Thomas J. Kane, and Parag A. Pathak. 2011. Accountability and flexibility in public schools: Evidence from Boston's charters and pilots. *Quarterly Journal of Economics* 126(2): 699–748. doi:10.1093/qje/qjr017

Bertrand, Marianne, Rema Hanna, and Sendhil Mullainathan. 2010. Affirmative action in education: Evidence from engineering college admissions in India. *Journal of Public Economics* 94(1–2): 16–29. doi:10.1016/j.jpubeco.2009.11.003

Card, David, Alexandre Mas, and Jesse Rothstein. 2008. Tipping and the dynamics of segregation. *Quarterly Journal of Economics* 123(1): 177–218. doi:10.1162/ gjec.2008.123.1.177

Chay, Kenneth Y., Patrick J. McEwan, and Miguel Urquiola. 2005. The central role of noise in evaluating interventions that use test scores to rank schools. *American Economic Review* 95(4): 1237–1258. doi:10.1257/0002828054825529

Chubb, John E., and Terry M. Moe. 1990. *Politics, markets and America's schools*. Washington, DC: The Brookings Institution.

Clark, Damon. 2009. The performance and competitive effects of school autonomy. *Journal of Political Economy* 117(4): 745–783. doi:10.1086/605604

De Vries, Michiels S. 2000. The rise and fall of decentralization: A comparative analysis of arguments and practices in European countries. *European Journal of Political Research* 38(2): 193–224. doi:10.1111/1475-6765.00532

Elmore, Richard, Allen S. Grossman, and Caroline King. 2006. Managing the Chicago public schools. Public Education Leadership Project at Harvard University, Report No. PEL-033.

Fullan, Michael. 1993. Coordinating school and district development in restructuring. In *Restructuring schooling: Learning from ongoing efforts*, edited by Joseph Murphy and Philip Hallinger, pp. 143–164. Newbury Park, CA: Corwin Press.

Fullan, Michael. 2001. Leading in a culture of change. San Francisco, CA: Jossey-Bass.

Hansen, Janet S., and Marguerite Roza. 2005. Decentralized decisionmaking for schools: New promise for an old idea? RAND Corporation Occasional Paper No. OP-153-EDU.

Hanushek, Eric A., Susanne Link, and Ludger Woessmann. 2011. Does school autonomy make sense everywhere? Panel estimates from PISA. NBER Working Paper No. 17591.

Herold, David M., and Donald G. Fedor. 2008. *Change the way you lead change: Leader-ship strategies that really work*. Stanford, CA: Stanford University Press.

Hill, Paul. T. 2004. Recovering from an accident: Repairing governance with comparative advantage. In *Who's in charge here: The tangled web of school governance and policy,* edited by Noel Epstein, pp. 75–103. Denver, CO: Education Commission of the States, and Washington, DC: Brookings Institution Press.

Imbens, Guido W., and Karthik Kalyanaraman. 2009. Optimal bandwidth choice for the regression discontinuity estimator. NBER Working Paper No. 14726.

Imbens, Guido W., and Thomas Lemieux. 2008. Regression discontinuity designs: A guide to practice. *Journal of Econometrics* 142(2): 615–635. doi:10.1016/j.jeconom.2007 .05.001

Kane, Thomas J., and Douglas O. Staiger. 2002. The promise and pitfalls of using imprecise school accountability measures. *Journal of Economic Perspectives* 16(4): 91–114. doi:10.1257/089533002320950993

Krueger, Alan. 1999. Experimental estimates of educational production functions. *Quarterly Journal of Economics* 114(2): 497–532. doi:10.1162/003355399556052

Lee, David S., and Thomas Lemieux. 2010. Regression discontinuity designs in economics. *Journal of Economic Literature* 48(2): 281–355. doi:10.1257/jel.48.2.281

Mohrman, Susan Albers, Priscilla Wohlstetter, eds. 1994. *School-based management: Organizing for high performance*. San Francisco, CA: Jossey-Bass.

Neal, Derek, and Diane Schanzenbach. 2010. Left behind by design: Proficiency counts and test-based accountability. *Review of Economics and Statistics* 92(2): 263–283. doi:10.1162/rest.2010.12318

Ouchi, William G. 2004. Academic freedom. Education Next 4(1): 21-25.

Ouchi, William G. 2006. Power to the principals: Decentralization in three large school districts. *Organization Science* 17(2): 298–307. doi:10.1287/orsc.1050.0172

Ouchi, William G., and Pete Goldschmidt. 2009. A national study of school district decentralization and student performance. Unpublished paper, University of California-Los Angeles.

Purkey, Stewart C., and Marshall S. Smith. 1983. Effective schools: A review. *Elementary School Journal* 83(4): 426–452. doi:10.1086/461325

Sartain, Lauren, Raymond McGhee, Jr., Lauren Cassidy, Maria I. Abasi, Viki M. Young, Susan E. Sporte, and Patrick M. Shields. 2009. *High school reform in Chicago public schools: Autonomous management and performance schools.* Menlo Park, CA: SRI International.

U.S. Department of Education (USDOE). 2009. *Race to the top program: Executive Summary*. Available http://www2.ed.gov/programs/racetothetop/executive-summary.pdf. Accessed 10 February 2013.

Van Langen, Annemarie, and Hetty Dekkers. 2001. Decentralization and combating educational exclusion. *Comparative Education* 37(3): 367–384. doi:10.1080/ 03050060120067839

Wohlstetter, Priscilla, and Thomas Buffett. 1992. Decentralizing dollars under school-based management: Have policies changed? *Educational Policy* 6(1): 280–294. doi:10.1177/0895904892006001003

Wohlstetter, Priscilla, and Susan A. Mohrman. 1996. Assessment of school-based management. Available www2.ed.gov/pubs/SER/SchBasedMgmt/index.html. doi:Accessed 10 February 2013.