The Performance and Competitive Effects of School Autonomy

Damon Clark

University of Florida, National Bureau of Economic Research, and Institute for the Study of Labor

This paper studies a recent British reform that allowed public high schools to opt out of local authority control and become autonomous schools funded directly by the central government. Schools seeking autonomy had only to propose and win a majority vote among current parents. Almost one in three high schools voted on autonomy between 1988 and 1997, and using a version of the regression discontinuity design, I find large achievement gains at schools in which the vote barely won compared to schools in which it barely lost. Despite other reforms that ensured that the British education system was, by international standards, highly competitive, a comparison of schools in the geographic neighborhoods of narrow vote winners and narrow vote losers suggests that these gains did not spill over.

I. Introduction

Market-oriented critics of the U.S. public education system attribute underperformance to the principal-agent problem that arises when voters (principals) are unable to control the activities of school boards (agents). The results, they contend, are school boards captured by vested interests, especially teacher unions, with powers over staffing, discipline, and other matters that should properly be exercised by schools. They argue that this explains why improved "inputs" do not improve public school performance and why public school performance lags private

I thank Ken Chay, Joe Clark, David Figlio, Sarah Hamersma, Graham Hobbs, Mark Hoekstra, Larry Kenny, Jane Leber Herr, David Lee, Paco Martorell, Justin McCrary, Robert McMillan, Imran Rasul, seminar participants at various institutions, and especially David Card for helpful suggestions. I thank Kate Bradford, John Elliott, Richard Howe, and Andrew Ledger for help in constructing the data sets used in the paper.

[*Journal of Political Economy,* 2009, vol. 117, no. 4]

school performance. Among the proposed solutions are policies that give schools more autonomy and policies that exert competitive pressure on schools and school boards (see, e.g., Chubb and Moe 1990).

Whether these kinds of market-based reform will improve public school performance is an open question. Until recently, few such policies had been implemented. Since the mid-1990s, "school choice" reforms have handed some powers to schools (via charter laws) and introduced some competitive pressures into the U.S. system (e.g., via open enrollment and private school vouchers), but these reforms are limited in scope and have been in place for only a short time.[1] As a result, it is perhaps not surprising that several careful studies report conflicting estimates of their early effects, with many researchers arguing that more time will be needed before firmer conclusions can be drawn.[2] Competition between school districts within a metropolitan area is a longer-established source of competitive pressure, but estimates of its impact have produced similarly conflicting findings (Hoxby 2000; Rothstein 2006, 2007).

With these difficulties in mind, this paper evaluates the performance and competitive impacts of school autonomy using a 1988 U.K. reform that allowed public high schools to "opt out" of local schools authority control and become quasi-independent "grant maintained" (GM) schools funded directly by the central government. This was complemented by other reforms, including nationwide open enrollment and the publication of "league tables" of school performance, which ensured that both GM and non-GM schools were operating in an education market that, by international standards, was highly competitive.

There are at least four reasons to think that the GM reform can provide important new evidence on the performance impacts of school autonomy and the competitive effects of this type of school reform. First, schools that became GM acquired a degree of autonomy matched by few charter schools, including complete ownership of all school facilities and complete control of staffing.[3] Second, in order to legitimize what the government (correctly) predicted would be an extremely con-

[1] Only 1.5 percent of public school students are enrolled in charter schools (Hoxby 2004), and at most 15 percent of students in grades 1–12 attend a "chosen" school (Wirt et al. 2004). In many states, school vouchers are limited to particular groups of students or to students from "failing" schools (Kenny 2005).

[2] For example, in the case of charter schools, Hoxby (2004) finds positive effects in some states and negative effects in others. In the case of private school vouchers, there is considerable debate over the size of the effects and the generalizability of the results from the existing demonstration projects; see Howell and Peterson (2002) and Krueger and Zhu (2004) for New York and Rouse (1998) for Milwaukee. Cullen, Jacob, and Levitt (2005, 2006) provide careful evaluations of the "partial" effects of open enrollment in Chicago.

[3] In an analysis of the U.K. reforms, Chubb and Moe (1992) described the GM reform as "truly revolutionary."

troversial policy, schools wishing to become GM had to first win a majority vote among the parents of current students. Third, the reform was large-scale, with almost one-third of high schools holding a GM vote and two-thirds of these winning the vote. These three observations suggest that a comparison of narrow GM vote losers and narrow GM vote winners can provide credible evidence on the causal impact of granting a large measure of autonomy to schools. Fourth, GM schools represented a well-defined source of competitive pressure. Indeed, the Thatcher government predicted that they would improve neighboring school performance through three related channels: first, by pressuring neighboring schools to improve performance via “exit” and “voice” (Hirschman 1970); second, by encouraging other schools to become GM; and third, by encouraging district administrators to become more responsive to non-GM school needs in order to prevent this. By comparing schools in the geographic neighborhoods of narrow vote losers and narrow vote winners, I can provide credible evidence on the compettitive effects of this type of school reform.

In the first part of the paper I estimate the effect of a GM conversion on student achievement in GM schools. I do this using a regression discontinuity design that compares achievement among schools that narrowly won and narrowly lost a GM vote. This strategy reveals dramatic gains in student achievement for schools that converted to GM status—on the order of a one-quarter standard deviation improvement in examination pass rates. These gains emerge as early as the second year after the vote and are stable and persistent for at least 8 years after the vote. Consistent with these performance gains, I estimate that schools converting to GM status enjoyed increases in enrollment and improvements in student quality. The enrollment effects were, however, relatively small, and I estimate that student quality improvements can account for less than one-half of the medium-run performance gains and an even smaller fraction of the short-run performance gains. It is hard to pin down what drove the remaining GM performance effects, although I show that GM conversions were associated with a shake-up of teaching staff that involved increased separations, increased hiring, and a net increase in the number of teachers employed at the school.

In the second part of the paper I estimate the spillovers associated with GM conversions. I use the same discontinuity idea to compare student achievement among schools in the geographic neighborhoods of narrow vote winners and narrow vote losers. Although my estimates are imprecise, my best estimate is that no gains spilled over, and I can rule out spillovers larger than one-half of the own-school effect. These conclusions are robust to a variety of neighbor definitions.

These results point to two conclusions. First, since the overall impact of a GM conversion was large and positive and since these gains came without any increase in overall expenditure, it seems that the GM reform was an effective one. This suggests that other varieties of autonomous school, such as those created by U.S. charter laws, may also be effective. Second, since I find no evidence of any spillovers, my results suggest that the returns to these types of school reforms are most likely to flow through their impacts on the reformed schools, not via spillovers to neighboring schools.

II. The Grant Maintained School Reform

In 1988, the Thatcher government passed legislation giving public secondary schools the option of leaving local education authority (local school district) control and becoming autonomous GM schools. Schools converting to GM status were funded by a new agency of the central government but were owned and managed by the school's governing body, a new 10–15-member entity composed of the head teacher and teacher and parent representatives. As such, control over all staff contracts and ownership of school buildings and grounds were taken from the local school district and given to GM schools.[4] The government also gave GM schools power over admissions, so that students wishing to attend GM schools applied to the school directly. To counter the charge that GM schools would become elitist, the government required them to publish an annual statement of their admissions policies. This had to satisfy certain conditions, the most important of which were that the school could not charge fees or set admissions tests.

The government claimed that schools converting to GM status would receive no additional funds, but in practice, GM schools were funded differently and almost certainly better. On the capital side, GM schools received core funding in proportion to enrollment and additional funding via a nationwide bidding process. Some commentators speculated that in the years following GM conversion, GM capital budgets could have been twice those of non-GM schools (Bush, Coleman, and Glover 1993). The current expenditure premium is easier to estimate since funding for current expenditure was based on an estimate of current expenditure by non-GM schools in the district plus an additional factor (around 15 percent) to compensate for the loss of district services. Most commentators believe that these services could be obtained for less than

[4] This meant that GM schools could, exactly like private schools, dismiss as necessary and pay teaching and nonteaching staff as they wished. Surveys indicate that in practice, schools converting to GM status maintained similar basic pay and conditions but were more likely to make bonus payments and use fixed-term and part-time contracts (Thompson 1992). School governors were free to make alterations to school premises and write contracts with outside organizations for the use of premises but were prevented from selling large quantities of school assets.

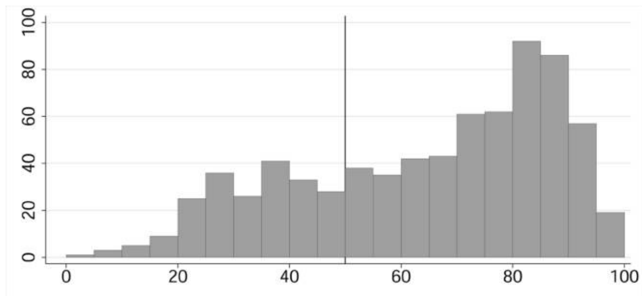


FIG. 1.—Vote counts by vote share. Vote share refers to % yes in the first GM vote held at the school. The sample includes all schools in the main GM sample. $N = 742$.

15 percent of the annual school budget, resulting in a GM current expenditure premium as high as 10–12 percent.

To become GM, a school had to arrange a GM status election. These were secret postal votes of current parents and were organized and monitored by an independent company. With a simple majority in favor of GM status, the school could apply to the government for final approval. Approval was a formality, denied only to the roughly 5 percent of schools that organized a GM election in response to a district decision to close or restructure them. Once converted, the GM legislation made clear that these schools had the power to do "anything which appears to them to be necessary or expedient for the purpose of, or in connection with, the conduct of the school." The legislation made provision for GM schools wishing to convert back to local authority control, although this was extremely involved and no schools attempted to do it. In 1997, when the new Labour government came to power, the GM policy was effectively frozen: existing GM schools kept nearly all their powers, but no new applications were allowed and the GM funding advantage was removed.

Between 1988 and 1997, almost 1,000 secondary schools voted on GM status. As seen in figure 1, the majority of these votes were won (the data are described in more detail below). This is not surprising since the decision to hold a vote would depend in part on the expected outcome. Figure 2 plots the proportion of schools achieving GM status against the share of parents voting in favor of GM status. Vote losers that became GM were those that lost the first election but won a subsequent election (schools wishing to hold subsequent elections had to wait 2 years); vote winners that did not become GM were almost certainly those whose applications the government rejected because local authorities had already marked these schools for closure. This interpretation is consistent with the high vote shares recorded at these schools.

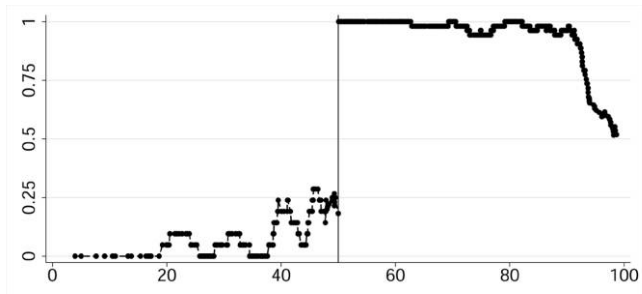


Fig. 2.—Vote share and GM status: smoothed (running mean) fraction of schools that ever acquire GM status. Means are calculated separately above and below the 50 percent threshold using bandwidth 0.1. The sample includes all schools in the main GM sample. $N = 742$.

Although schools could convert to GM status at any time between 1988 and 1997, the majority of votes took place between 1992 and 1995.[5] Anecdotal and survey evidence suggests that head teachers were instrumental in deciding whether a GM vote was held, with the outcome of the vote determined by a campaign heavily influenced by teachers and the local school district (Fitz, Halpin, and Power 1993). Consistent with the existence of a large number of vote losers, anecdotal evidence suggests that much uncertainty surrounded the outcome of these campaigns. Votes were easiest to pass in Conservative-controlled authorities (whose GM policies ranged from hostile to neutral) and hardest in Labour-controlled authorities (universally hostile). Among the full sample of schools "at risk" of becoming GM, linear probability models suggest that a vote was more likely at schools that were (on observables) more successful. Compared with local political control, however, these variables were of secondary importance.[6]

Finally, other school reforms ensured that GM schools operated in an education market that was, by international standards, highly competitive. First, a "nationwide open enrollment" reform meant that parents could apply to any school in the country and could be denied their first-choice school only if that school was oversubscribed. In that case, admissions were based on district-determined criteria, the most impor-

[5] Online App. fig. B2 plots GM ballots by year-term. The sharp increase in the second half of 1992 is consistent with schools awaiting the outcome of the closely fought 1992 general election (held in April) and seeking GM status after the Conservative Party won that election.

[6] See online App. table B1. My inability to find strong correlates of the GM vote decision is consistent with survey evidence pointing to "increased independence" as the most commonly cited reason for seeking GM status among eventual vote winners and "satisfaction with the local schools authority" as the most commonly cited reason why schools did not hold a vote (Bush et al. 1993).

tant of which was usually distance from home to school. Second, from 1992, the government published annual performance tables ("league tables") showing the fraction of a school's grade 11 students passing five or more General Certificate of Secondary Education (GCSE) examinations, standardized examinations taken by all grade 11 students.[7]

III. Empirical Framework

Before I describe the empirical strategy used to evaluate the GM reform process, it is useful to sketch a conceptual framework that will help to understand it. Consider a single school district containing N schools. Suppose that time is discrete and is indexed $t = 0, 1, 2, \ldots, T$, where $t = 0$ is the last prereform period, $t = 1$ is the first postreform period, and $t = T$ is the last year of the observation window.

Consider first the behavior of school districts. At the beginning of period $t \geq 1$, district administrators observe how many schools remain non-GM and decide how much autonomy to grant them. Districts wish to prevent schools from becoming GM, and for reasons explained below, granting non-GM schools more autonomy decreases the probability that they will become GM. However, school autonomy is costly to district administrators, who prefer to make decisions themselves.

Consider next the behavior of schools and assume that school performance Y—for example, test scores—is increasing in autonomy a and school effort e: $Y = g(a, e(a))$.[8] Schools that are already GM at the start of period $t \geq 1$ are fully autonomous ($a = a^{\max}$) and hence decide only how much effort to exert. Effort improves school performance—for example, test scores—but effort is costly. Schools that are still non-GM at the start of period $t \geq 1$ observe how much autonomy they have been granted and decide how much effort to exert and whether or not to become GM (simultaneously). For given effort, non-GM school performance is assumed lower than GM school performance; hence schools have an incentive to become GM. There are, however, costs associated with GM status; hence the GM decision is nontrivial.

Although this sketches a complicated dynamic game, it helps to consider some factors influencing the decision to become GM and the possible effects of a GM conversion. At the start of period 1, when all schools are non-GM, schools will observe autonomy, compute the optimal effort exerted in the non-GM and GM regimes, and choose to become GM if maximized utility in the GM regime is higher than in the non-GM regime. Anticipating this, district administrators choose

[7] Time series of pass rates, enrollment, and the number of operating schools are presented in online App. fig. B1.

[8] I discuss other inputs into the education production function below.

autonomy by trading off the costs of increased autonomy against the decreased probability that schools will become GM.[9] Suppose that one school becomes GM in period 1 and improves performance as a result. Even if autonomy remains unchanged at the start of period 2, other schools might feel pressure to improve performance, causing them to increase effort whether or not they become GM. Since effort may be more productive in the GM regime, the GM regime may now be more attractive. Anticipating this, the district may grant more autonomy in a bid to prevent more GM conversions.

These considerations highlight three important features of the GM reform process. First, districts may have granted more autonomy (more generally, changed behaviors) at the start of period 1, as soon as the reform was announced. As a result, the GM reform may have improved performance before the first school converted to GM. These effects will not be picked up by my analysis. Second, since schools can respond to a GM conversion by becoming GM themselves (I refer to these as "copycat effects"), I cannot interpret spillover effects as purely the result of competition (as distinct from autonomy). Third, the period 2 responses to a period 1 GM conversion will trigger future responses in periods 3, 4, and so on. Since this process may feed back to the school converting to GM status in period 1, I cannot interpret GM performance effects as purely the result of autonomy (as distinct from competition).

A. *Empirical Strategy*

The conceptual framework assumed that schools were identical. In practice, schools differ along many dimensions, and certain types of school may be more likely to hold and win a GM vote (e.g., those with more entrepreneurial head teachers). My empirical approach overcomes this selection problem by focusing on the jump in performance among schools at the 50 percent win threshold. Specifically, I consider variants of the fuzzy regression discontinuity model for school i voting on GM status in year t with outcomes observed in year $t + s$:

$$T_{i,t+s} = \beta_{0s} + \beta_{1s}\mathrm{GM}_{it} + f(V_{it}) + \xi_{i,t+s}, \tag{1}$$

where the dummy variable $\mathrm{WIN}_{it} = I[V_{it} > 50]$ is used as an instrument for the endogenous variable GM_{it} (Angrist and Lavy 1999; Imbens and Lemieux 2008). I use a similar approach to estimate the effects of a GM conversion on neighbor school performance:

$$\overline{T_{j\neq i,t+s}} = \lambda_{0s} + \lambda_{1s}\mathrm{GM}_{it} + g(V_{it}) + \xi_{i,t+s}, \tag{2}$$

[9] Of course, in this kind of multiperiod game, strategies would also depend on expected states and strategies in every period beyond period 2.

where $\overline{T_{j\neq i,t+s}}$ is average performance among neighbor schools, and the dummy variable $\mathrm{WIN}_{it} = I[V_{it} > 50]$ is again used as an instrument for the endogenous variable GM_{it}.

The key identifying assumption underlying this approach is that the functions $f(\cdot)$ and $g(\cdot)$ are continuous through the 50 percent threshold, leaving a vote win as the sole cause of any discontinuity in outcomes. This assumption would be violated if schools could manipulate the outcome of the election (see McCrary 2008), but this is unlikely since parents voted in a secret ballot monitored by an independent agency. In the absence of violations of this assumption and provided that $f(\cdot)$ and $g(\cdot)$ are correctly specified (discussed below), I can interpret estimates of the discontinuity in $T_{i,t+s}$ and $\overline{T_{j\neq i,t+s}}$ at the 50 percent threshold (estimated by β_{1s} and λ_{1s} in eqq. [1] and [2]) as the causal effects of a GM vote win on GM school and neighbor school performance.

B. *Estimation Issues*

Although my empirical strategy appears capable of identifying the causal impacts of the GM reform process, there are some important issues associated with the estimation and interpretation of equations (1) and (2). One issue concerns the empirical specification of the functions $f(\cdot)$ and $g(\cdot)$. As noted above, consistent estimation of β_{1s} and λ_{1s} requires that these functions be specified correctly. I follow the recent applied regression discontinuity literature and proxy for them using low-order polynomials. To verify that my estimates are not overly sensitive to the choice of polynomial, I graph the relationship between outcomes and vote shares (to provide visual evidence of potential specification errors) and experiment with alternative specifications. To help reduce the residual variation in these outcomes, I add a set of predetermined variables to these models. The most important of these is baseline performance, without which the standard errors on the GM effects more than double.

A second issue is how to decompose GM effects into the various channels through which they flow. For example, if GM conversions improve performance, we would like to know what portion of these gains are attributable to class size reductions, improvements in the quality of incoming students, and so on. There is no simple solution to this problem. The estimation strategy described above can be used to estimate GM effects on inputs such as class size, but an estimate of the GM performance effect operating through this channel requires knowledge of the performance effect of class size reductions. It is tempting to estimate the GM performance effect net of this channel by including this input as a covariate in equation (1), but such an estimate is likely to understate these net effects (Angrist and Krueger 1999).[10]

Special issues are raised when the input is student quality. First, student quality is not directly observed. Hence even if measured student quality were included as a covariate, GM conversion may be correlated with changes in unmeasured student quality. Second, when the input is student quality, we can assess its contribution to school performance using an alternative strategy. In particular, we can ask whether a randomly chosen student would benefit from attending a GM school. This does not tell us whether school performance gains driven by GM conversion can be attributed to student selection, since a randomly chosen student could benefit from attending a GM school even if GM conversions serve only to change student composition (via peer effects). Nevertheless, the extent to which students benefit from attending a GM school will provide a useful guide to whether GM effects operate independently of student selection, particularly if high school peer effects are weak.[11]

I use two strategies to estimate the impact of attending a GM school. First, I base a "selection-on-observables" approach on a sample of students who attended a school that voted on GM status. I instrument whether the attended school was GM with whether the attended school won the GM vote (i.e., the regression discontinuity approach), and I control for student as well as school characteristics.[12] The key assumption

[10] For example, suppose that GM performance (Y_{jt}) and the channel variable (X_{jt}) are related via the system

$$Y_{jt} = \beta_0 + \beta_1 \mathrm{GM}_{jt} + \beta_2 X_{jt} + c_j + u_{jt},$$

$$X_{jt} = \delta_0 + \delta_1 \mathrm{GM}_{jt} + \delta_2 c_j + v_{jt},$$

which allows the channel variable to depend on GM status and school effectiveness c_j (assumed fixed over time). Writing the linear projection of c_j on GM_{jt} and X_{jt} as $c_j = r_0 + r_1 \mathrm{GM}_{jt} + r_2 X_{jt} + e_{jt}$, it can be seen that a regression of Y_{jt} on GM_{jt} identifies $\beta_1 + r_1$. We know that

$$r_1 = V(X)\,\mathrm{Cov}\,(\mathrm{GM},\ c) - \mathrm{Cov}\,(\mathrm{GM},\ c) = -\,\mathrm{Cov}\,(\mathrm{GM},\ X)\,\mathrm{Cov}\,(X,\ c)$$

$$= -\,\mathrm{Var}\,(\mathrm{GM})\,\mathrm{Var}\,(c)\delta_1\delta_2 < 0$$

provided that $\delta_1, \delta_2 > 0$ and assuming that GM status is randomly assigned. Intuitively, since the input is increasing in school effectiveness and is higher for GM schools, among a set of schools with $X = x$, the GM schools will be less effective.

[11] The bulk of the empirical literature on peer effects concerns elementary schools. In an important paper relating to high school peer effects, Cullen et al. (2006) show that students who win lotteries to high-achieving schools (defined as those with good peers) score no better than students who lose these lotteries. In the U.K. context, Clark (2007) finds that students admitted to selective high schools score no better than students admitted to nonselective high schools.

[12] A two-step version of this approach could also be implemented. First, regress student-level outcomes on student characteristics and school fixed effects and use the estimated "returns" to these student characteristics to construct adjusted student outcomes. Second, use these adjusted student outcomes as the dependent variable in an instrumental variables

is that student and school characteristics are additively separable and that all relevant student characteristics are included. Since the data sets used here contain detailed information on student socioeconomic status and primary school test scores and since these explain a lot of the across-student performance variation, it is plausible to imagine that they include most if not all of the information on which selection into GM schools could have been based. Second, I base a selection-on-unobservables approach on a sample of students who attended a primary school for which the predicted secondary school voted on GM status. This time I instrument whether the attended school was GM with whether the predicted secondary school won the GM vote (i.e., the regression discontinuity approach), where a student's predicted secondary school is defined as the secondary school in which the largest fraction of the student's primary school classmates enroll. Provided that students did not select into primary school on the basis of the predicted secondary school vote outcome, this strategy will be robust to student selection into GM secondary schools.

A third estimation issue concerns the specification of models designed to assess the competitive effects of GM conversions. One strategy takes vote schools as the unit of analysis and recasts neighbor school outcomes as vote school outcomes. This identifies neighbor school effects operating through all the channels discussed above, including copycat effects. A second strategy takes neighbor schools as the unit of analysis and relates their outcomes to the number of GM schools in their vicinity. Since this second approach conditions on the number of attempts in the neighborhood, it is valid only when GM attempts are independent across schools (i.e., no copycat effects). I present results based on the first of these approaches. Estimates based on the second, as well as estimates of the extent of copycat effects, are available in online Appendix B.

IV. Data

To analyze the impact of the GM reform process, I assemble a new data set of English schools. The data set is built on the Annual School Census, which I have for the period 1975–2003. This census covers all English state-funded secondary schools and, in every year, contains counts of

procedure that instruments whether the attended school was GM with whether the attended school won the GM vote (with controls for GM vote share and attended school characteristics). The two-step version allows for potentially cleaner estimation of the returns to the student characteristics, but the two approaches generate almost identical results. The reason is that the one-step model already includes a large number of school characteristics.

students and teachers.[13] From 1993 it also includes the proportion of a school's students eligible for a free school meal. Because schools open and close, this is an unbalanced panel of roughly 3,500 schools per year. I match this panel to the school performance tables and the Register of Educational Establishments. The School Performance Tables (league tables) were first published in 1992 and contain my primary measure of student achievement, the fraction of grade 11 students who pass five or more GCSE examinations (the "school pass rate"). The Register of Educational Establishments is a database of (open and closed) schools containing school names and addresses. I use the 1999 version of this database to derive geographic information for the schools in this panel.

I analyze the impact of a GM conversion on the schools that converted and on the schools in their neighborhood. The first type of analysis is based on a subset of schools that held a GM vote. Specifically, I analyze a sample of 742 schools that I refer to as the main GM sample. To arrive at this sample I obtained GM vote data on all 950 schools that held a GM vote (from the independent organization that conducted the GM ballots). I then discarded four schools that could not be matched to the school censuses, 78 schools that did not have students in grade 11 (i.e., middle or primary schools), four schools for which I do not have geographic information, and 141 schools that became GM before 1992.[14] I ignore schools that became GM before 1992 because I do not have baseline pass rates for these schools and because, as noted above, GM performance estimates are much more precise when the regression is adjusted for baseline pass rates. Around 90 percent of these schools survive for at least 8 years after the vote (see online fig. B3, similar to survival rates for nonvote schools, not shown), with the probability of survival smooth through the 50 percent win threshold. Appendix table A1 presents descriptive statistics for this sample.

To estimate the impact of attending a GM school, I use student-level data from the Youth Cohort Study (YCS) and the National Pupil Database (NPD). The YCS is a postal survey of 10 percent of grade 11 students in all secondary schools in England and Wales (sampling based on date of birth). Student names and addresses are provided by the

[13] Around 7 percent of students attend a private secondary school. I have no data on private schools; hence they are excluded from my analysis.

[14] Since there is a gap of roughly 12 months between a GM vote and conversion to GM status, the base year is conceptually ambiguous. I define the base year as 1992 for all schools that vote in the second and third terms of 1991 and the first term of 1992 (and similarly across other years). School census base year data will be those collected in January 1992 and school performance base year data those collected in June 1992. This definition will be accurate if treatment begins upon conversion to GM status. It may be slightly "posttreatment" if the treatment begins immediately after the vote. In this case, estimated treatment effects will understate true treatment effects. The first definition generates larger estimation samples, and different definitions produce similar results.

schools attended by these students, and response rates are between 60 and 80 percent. These surveys have been conducted roughly every second year since 1984, and I match the post-1993 surveys to the main GM sample. As seen in Appendix table A1, students in 731 of the 742 schools in the main GM sample responded to the YCS; for both schools and students, YCS match rates are smooth through the 50 percent win threshold (online fig. B3). The NPD, available from 2002, contains demographic and academic achievement information for the population of pupils in English schools. Although this implies that there is no school nonresponse, some schools in the main GM sample do not survive until 2002; hence only 711 schools are matched to the NPD. Not all the students in the NPD have valid primary school test information, a key student characteristic, although the fraction of students with valid test information is smooth through the 50 percent win threshold (online fig. B3).[15] Appendix table A2 provides more information on the YCS and NPD.

V. The Effects of GM Conversion

In this section I estimate the effects of a GM conversion. I focus first on the schools that become GM. I use school-level models to estimate GM impacts on performance and enrollment and student-level models to estimate the impact of attending a GM school. I then examine GM impacts on neighboring schools. In Section VI, I collect and discuss these various findings.

A. The Effects of GM Conversion on GM Schools

1. School Performance

Columns 1–4 of table 1 report estimates of the impact of a GM vote win (the "intention to treat" [ITT] effect) on school pass rates.[16] Column 1 reports estimates from a model without covariates (i.e., the difference in raw pass rates between vote winners and vote losers); column 2 reports estimates from a model that regression adjusts for base pass rate, school type, and vote year-term. The concern that these adjusted pass rates may be correlated with the GM vote share motivates the regression discontinuity specification, which identifies the ITT effect at the 50

[15] Test score information may not be valid because students entered the public school system after the final primary school grade, because students were ill or away during the primary school testing period, because students were exempted from primary school testing, because one or more of the students' scores are missing, or because the students' primary test records cannot be matched to the school census data 5 years later.

[16] To give a sense of the role played by the predetermined variables in these models, online App. table B2 reports all the coefficient estimates from the base year + 2 model.

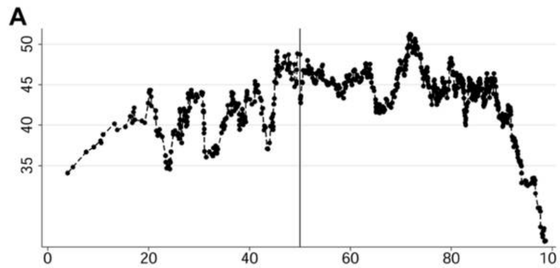
TABLE 1
The Impacts of a GM Conversion on Pass Rates 1–8 Years after the Vote

	RAW-ITT	ADJ-ITT	RD-ITT		2SLS-TE	
			Linear	Lin × Win	Lin × Win	Quad × Win
	(1)	(2)	(3)	(4)	(5)	(6)
	A. Separate Estimates by Year Relative to Base					
Base+1 (N = 729)	3.448	.388	3.221	2.802	3.376	4.499
	(1.579)	(.609)	(1.050)	(1.085)	(1.308)	(2.359)
Base+2 (N = 726)	4.708	1.676	3.587	3.391	4.134	4.529
	(1.588)	(.625)	(1.148)	(1.218)	(1.478)	(2.655)
Base+3 (N = 722)	5.496	2.381	4.616	4.391	5.446	8.228
	(1.585)	(.647)	(1.146)	(1.219)	(1.525)	(2.619)
Base+4 (N = 720)	4.515	1.728	3.365	3.359	4.169	6.971
	(1.576)	(.711)	(1.287)	(1.379)	(1.715)	(2.793)
Base+5 (N = 717)	5.117	2.665	3.531	3.105	3.890	6.824
	(1.594)	(.716)	(1.318)	(1.407)	(1.757)	(2.760)
Base+6 (N = 713)	4.880	2.510	4.506	3.391	4.276	6.946
	(1.578)	(.752)	(1.413)	(1.477)	(1.842)	(2.802)
Base+7 (N = 688)	5.674	3.375	4.292	3.950	5.038	7.306
	(1.578)	(.839)	(1.528)	(1.570)	(1.993)	(3.185)
Base+8 (N = 668)	5.575	3.233	3.941	3.207	4.107	5.528
	(1.620)	(.925)	(1.679)	(1.766)	(2.216)	(3.411)
	B. Pooled Base + 2 through Base + 8: N = 4,954, N(Schools) = 726					
Win	5.207	2.501	4.009	3.578	4.478	6.658
	(1.505)	(.592)	(1.097)	(1.139)	(1.445)	(2.364)

Note.—In panel A, robust standard errors are in parentheses; in panel B, standard errors clustered by school are in parentheses. The sample includes all schools in the main GM sample with relevant pass rate information. Regression adjustment is made for base pass rates, school type, and year-term of vote.

percent threshold. Regression discontinuity estimates of the ITT effect are reported in columns 3 and 4. The column 3 estimates are based on models that specify $f(\cdot)$ as linear; the column 4 estimates are based on models that maintain linearity while allowing for different slopes on either side of the 50 percent threshold. Columns 5 and 6 report estimates of the impact of a GM conversion (the treatment effect). To estimate this treatment effect, the ITT estimates must be scaled up by the probability that a vote win triggers a GM conversion. As noted above, two-stage least squares (2SLS) does this while accounting for the sampling variation introduced in the first stage. The 2SLS estimates reported in column 5 are based on the same specification of $f(\cdot)$ used to generate the ITT estimates reported in column 4; the 2SLS estimates reported in column 6 are based on a more flexible (quadratic) specification of the vote share.

Table 1 points to four conclusions. First, schools that become GM enjoyed large performance improvements in the first few years after the vote. For example, only 2 years after the vote, a GM vote win is estimated to increase performance by around 3 percentage points. This is revealed



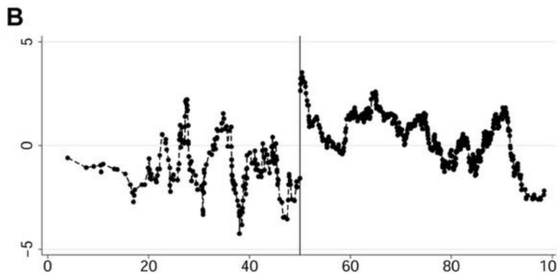


FIG. 3.—*A*, Pass rates in the base year: smoothed (running mean) base pass rates. Means are calculated separately above and below the 50 percent threshold using bandwidth 0.1. The sample includes all schools in the main GM sample. $N = 742$. *B*, The impact of a GM vote win on pass rates 2 years after the vote: smoothed (running mean) regression-adjusted pass rates 2 years after the vote. Means are calculated separately above and below the 50 percent threshold using bandwidth 0.1. Regression adjustment is made for base pass rates, school type, and vote year-term. The sample includes all schools in the main GM sample with relevant pass rate data. $N = 726$.

as a sharp discontinuity in the graph displayed in figure 3*B*, not apparent in the graph of base pass rates displayed in figure 3*A*. Second, the estimates in the table suggest that these effects persist up to 8 years after the vote. Indeed, from the second year onward, the estimates are broadly stable across the observation window. Third, the estimates are robust to the choice of specification. All specifications report positive and significant impacts, with the only consistent difference coming when moving from column 2 to column 3. Visual inspection of figure 3*B* suggests that the reason is that the average difference between winners and losers is slightly smaller than the discontinuity at the 50 percent threshold. Experimenting with a quadratic function for $f(\cdot)$ tends to increase the point estimates, although these are now much less precise.

Given the absence of any time pattern to the estimates reported in panel A of table 1, a natural next step is to pool data across postbase years 2–8 and estimate the model described by the following equation:

$$Y_{jt} = \beta_0 + \beta_1 \mathrm{GM}_j + f(V_j) + t + \varepsilon_{jt},$$

where time dummies (t) are included because the model features multiple observations per school, and the error term allows for residual correlation within schools (handled by clustering standard errors).[17] If GM effects are constant over time, the pooled model will estimate these effects more precisely. This is confirmed by the estimates reported in panel B of table 1, which are comparable to but generally more precise than the separate estimates reported in panel A.[18]

To check the robustness of these results, I estimated variants of the pooled model with hypothetical discontinuities ranging from 20 to 80 percent. I estimated 120 of these models (associated with discontinuities starting at 20 percent and increasing to 80 percent at 0.5-percentage-point intervals), and in each case I recorded the t-ratio associated with the discontinuity. The relationship between these hypothetical discontinuities and the associated t-ratios is plotted in Appendix figure A1. Consistent with a true discontinuity at 50 percent, the t-ratio is maximized at exactly 50 percent and is roughly symmetric around it.[19] I also experimented with the window used to generate these estimates. Specifically, I calculated the estimate and confidence interval obtained when I used a window of x percentage points around the 50 percent cutoff and graph the results from $x = 0$ to $x = 50$ (App. fig. A2). While small windows generate large confidence intervals, the estimates converge on the estimate reported in column 5 of panel B of table 1 for windows as small as 15 percentage points.

2. School Enrollment

Although student achievement is the primary focus of this paper, an analysis of GM impacts on school enrollment can help to shed light on the mechanisms by which any student achievement improvements might have operated. I estimate these enrollment impacts in three steps. First, I estimate impacts on enrollment levels among cohorts already enrolled in secondary school when the GM vote took place (i.e., those in grades 7–11). Second, I estimate impacts on entry enrollment in the years

[17] As a statistical test of the hypothesis that these effects are stable across the observation window, I estimate a more flexible model that allows for a linear trend in these effects. The coefficient on this trend is positive but not statistically different from zero. For example, the estimate (standard error) associated with this trend in the model corresponding to col. 5 of panel B of table 1 is 0.152 (0.137).

[18] Effects estimated over longer periods are less precise because base pass rates capture less of the residual variation in postbase pass rates. The reason is that pass rates, while highly persistent, are not fixed effects. This is reflected in the coefficient on base pass rates in models of pass rates 1 or 2 years after the base year (around 0.9) and is consistent with the results presented by Kane and Staiger (2001).

[19] This symmetry does not reflect the density of the vote share distribution, which has a mean in excess of 50 percent. In other words, this pattern is not found when I perform the same exercise for outcomes that are not affected by GM status.

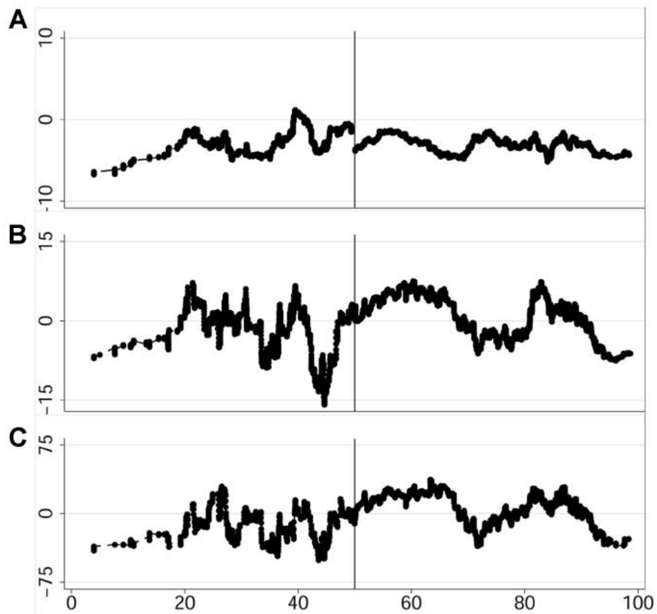


FIG. 4.—The impacts of a GM vote win on enrollment: smoothed (running mean) outcomes. Means are calculated separately above and below the 50 percent threshold using bandwidth 0.1. *A*, The outcome is the change in size of the already-enrolled cohort aged 13 from the base year to base year + 2, when this cohort is in grade 11. The sample includes all schools in the main GM sample with cohort size changes of fewer than 20 students. $N = 705$. *B*, The outcome is regression-adjusted entry enrollment 1–8 years after the base year. Regression adjustment is made for entry enrollment in the base year and the 2 years preceding it, school type, year-term of vote, and year. The sample includes all schools in the main GM sample with relevant entry enrollment data. $N = 5{,}701$, N(schools) $= 739$. *C*, The outcome is regression-adjusted total enrollment 5–8 years after the base year. Regression adjustment is made for total enrollment in the base year and the 2 years preceding it, school type, year-term of vote, and year. The sample includes all schools in the main GM sample with relevant total enrollment data. $N = 2{,}791$, N(schools) $= 719$.

following the vote, defined as enrollment in the first grade of secondary school (usually grade 7). Third, I estimate impacts on total school enrollment and check whether these agree with the total enrollment effects that would be predicted by my estimates of the GM impacts on enrollment among already-enrolled and newly entering cohorts.

I estimate that GM conversions had no impact on the size of already-enrolled cohorts. For example, as seen in figure 4*A*, I find that GM conversion has no impact on the difference between the number of students aged 13 in the base year (in grade 9) and the number of students in grade 11 2 years after the base year. Indeed, the data reveal that among all vote schools, cohort size was typically very stable over these 2 years: for 90 percent of schools, cohort size changed by fewer

TABLE 2
THE IMPACTS OF A GM VOTE WIN/CONVERSION ON ENROLLMENT

	RAW-ITT (1)	ADJ-ITT (2)	RD-ITT: Linear (3)	RD-ITT: Lin × Win (4)	2SLS-TE: Lin × Win (5)	2SLS-TE: Quad × Win (6)
	A. Entry Enrollment: N = 5,701, N(Schools) = 739					
Win/GM	−20.079	3.436	8.839	9.045	11.245	18.108
	(5.086)	(2.106)	(3.843)	(4.235)	(5.241)	(8.626)
	B. Total Enrollment, 1–4 Years after the Vote: N = 2,910, N(Schools) = 739					
Win/GM	−136.442	−15.355	−9.238	−13.132	−14.554	−9.086
	(21.430)	(4.018)	(8.413)	(8.910)	(10.945)	(17.675)
Win/GM × year	18.357	7.242	7.211	7.219	8.277	8.219
	(1.715)	(2.186)	(2.187)	(2.187)	(2.471)	(2.467)
	C. Total Enrollment, ≥ 5 Years after the Vote: N = 2,791, N(Schools) = 719					
Win/GM	−69.645	23.884	35.449	32.963	41.634	54.067
	(22.255)	(10.770)	(19.361)	(20.464)	(25.604)	(40.795)

NOTE.—Standard errors clustered by school are in parentheses. Regression adjustment is made for entry/total enrollment in the base year and the 2 years preceding it, school type, and year-term of vote. The sample includes all voting schools with relevant total enrollment data.

than 10 students; for 65 percent of schools, it changed by fewer than five students. The same patterns are found for the difference between the number of students in grade 7 in the base year and those in grade 11 4 years later, grade 8 in the base year and grade 11 3 three years later, and grade 10 in the base year and grade 11 1 year later (online App. fig. B4). Cohort sizes may have been stable because students did not switch schools, behavior that is relatively uncommon in England.[20] Stable cohort size could, however, mask changes in student composition, a point I return to below.

I estimate that GM conversions increased entry enrollment by about 10 students per year (base enrollment was, on average, 176 students). These estimates are reported in panel A of table 2 and are based on the same pooled model used in panel B of table 1. The corresponding graph is displayed in figure 4*B*. I also estimated effects on enrollment from 1 to 8 years after the base year. Unlike the performance estimates, these are sensitive to observation period and vote share specification (see online App. table B3, panel A). This sensitivity might be driven by the discrete nature of the data (entry enrollment is typically bunched at multiples of 30 students) and the presence of outliers. Despite this caveat, the ITT and 2SLS estimates are uniformly positive. Median re-

[20] Data show that over 95 percent of students finish grade 11 in the school in which they began grade 7 (calculation based on students in grade 11 in 2002 and primary school in 1997; N = 444,376).

gression estimates, which may help deal with the problem of outliers, point to similar effects (not reported).

My estimates of GM impacts on enrollment in already-enrolled grades (zero) and on grades transitioning from primary to secondary school after the vote (around 10 students) suggest that GM conversion would have increased total enrollment in the 4 years following a vote, reaching a maximum of around 50 additional students 5 or more years after the vote (five grades with an additional 10 students per grade). The estimates reported in panel B of table 2, based on models that force this pattern, are consistent with this prediction;[21] a visual impression of these effects is given in figure 4*C*. Estimates of the total enrollment impact 5 years after the vote—around 42 students—correspond to around 5 percent of average base total enrollment (765 students).

3. Student Quality

As noted above, I assess the role of student quality by estimating student-level models of the impact of attending a GM school. I estimate these using the YCS and the NPD, two student-level data sets that contain information on secondary school attended, academic achievement in grade 11, and detailed student background characteristics. I define the treatment in these analyses as attending a GM school for 5 years as opposed to attending a non-GM school for 5 years. Since I observe the school attended only in the final grade, there will be some classification errors in my treatment measure. Since 90 percent of students are observed to attend the same school for 5 years, these errors should be relatively rare.[22]

Estimates based on secondary school samples that control for student characteristics are reported in table 3. Each panel refers to a different sample, and for each sample I report three sets of estimates. The top row reports estimates of the impact of attending a GM school on the probability of passing five or more exams. These estimates facilitate comparison with the estimated effects of GM conversion on school-level pass rates. The middle and bottom rows report estimates of the impact of attending a GM school on total exam points, obtained by assigning points to letter grades (A = 7, B = 6, C = 5, …, G = 1) and summing

[21] Estimates based on separate postbase models are consistent with this pattern (online App. table B3, panel B). If allowance is made for an interaction in the later period, the data cannot reject a zero coefficient on this interaction.

[22] Intuitively, by defining the treatment in this way, I will assign 5 years in a GM school to some students who might have had less and will assign 0 years in a GM school to some students who might have had more. The caveat is that when I instrument GM status with predicted secondary school vote, the classification errors will be in the dependent variable in the first-stage relationship. In that case, I will understate the first-stage relationship and overstate the treatment effects of interest.

TABLE 3
THE IMPACTS OF ATTENDING A GM SCHOOL

	ADJ-ITT (1)	RD-ITT Linear (2)	RD-ITT Lin × Win (3)	2SLS-TE Linear (4)	2SLS-TE Lin × Win (5)	2SLS-TE Quad × Win (6)
	A. SECONDARY SCHOOL SAMPLE					
	1. YCS: N = 15,383, N(Schools) = 718					
Pass ≥ 5	1.912	3.435	4.472	4.277	6.161	7.267
	(1.178)	(2.065)	(2.198)	(2.578)	(3.068)	(5.131)
Exam points	1.338	1.677	1.835	2.088	2.527	2.537
	(.432)	(.763)	(.848)	(.947)	(1.171)	(1.943)
Exam points	1.021	1.109	1.212	1.381	1.672	2.197
	(.392)	(.719)	(.802)	(.890)	(1.102)	(1.852)
	2. YCS, < 5 Years after Vote (1993, 1995, 1997): N = 6,298, N(Schools) = 658					
Pass ≥ 5	1.443	3.545	5.159	4.425	7.186	5.511
	(1.685)	(3.126)	(3.167)	(3.907)	(4.451)	(7.589)
Exam points	1.331	1.882	2.510	2.349	3.497	3.180
	(.554)	(.939)	(.985)	(1.178)	(1.399)	(2.338)
Exam points	1.016	1.596	2.091	1.996	2.917	2.920
	(.505)	(.883)	(.912)	(1.107)	(1.297)	(2.209)
	3. YCS, ≥ 5 Years after Vote (1999, 2001, 2003): N = 9,085, N(Schools) = 656					
Pass ≥ 5	2.313	3.488	4.051	4.329	5.529	8.170
	(1.421)	(2.405)	(2.638)	(2.994)	(3.646)	(6.014)
Exam points	1.349	1.540	1.363	1.911	1.861	2.006
	(.565)	(1.001)	(1.135)	(1.238)	(1.549)	(2.484)
Exam points	1.014	.816	.635	1.013	.867	1.712
	(.513)	(.962)	(1.082)	(1.186)	(1.470)	(2.351)
	4. NPD (2002, 2003): N = 227,755, N(Schools) = 711					
Pass ≥ 5	3.543	3.477	2.810	4.200	3.646	4.351
	(.853)	(1.509)	(1.553)	(1.829)	(2.035)	(3.145)
Exam points	1.851	1.923	1.931	2.323	2.505	3.341
	(.418)	(.752)	(.793)	(.910)	(1.034)	(1.613)
Exam points	1.015	.751	.702	.909	.911	2.204
	(.334)	(.610)	(.649)	(.738)	(.841)	(1.290)
	B. PRIMARY SCHOOL SAMPLE: N(Students) = 213,192, N(Primary Schools) = 3,397, N(Predicted Secondary Schools) = 676, N(Secondary Schools) = 2,958					
Pass ≥ 5	2.183	1.447	.969	2.618	1.907	1.631
	(.575)	(1.079)	(1.095)	(1.954)	(2.151)	(3.502)
Exam points	1.278	.981	.971	1.775	1.912	2.521
	(.266)	(.489)	(.507)	(.886)	(.994)	(1.617)
Exam points	.627	.843	.834	1.557	1.659	2.836
	(.181)	(.326)	(.347)	(.602)	(.689)	(1.089)

NOTE.—In panel A, standard errors are clustered by school. Regression adjustment is made for base pass rate, year, school type, and year-term of vote. In the third row of each panel, the regression is adjusted for student characteristics (see the text for details). The sample in panels A1–A3 includes students in the main GM sample matched to the YCS. The sample in panel A4 includes students with nonmissing data in schools in the main GM sample observed in the NPD. In panel B, estimates in cols. 1–4 are impacts of predicted secondary school vote win on student outcomes, controlling for predicted secondary school vote share and predicted secondary school characteristics (base pass rates, school type, and year-term of vote). Predicted secondary school is based on primary school attended; standard errors are clustered by primary school. Estimates in cols. 5–7 are impacts of attending a GM secondary school using as an instrument whether the predicted secondary school GM vote was won. These models include the same predicted secondary school controls as were included in the col. 1–3 estimates. The sample includes all students who are observed to attend primary schools for which the predicted secondary school was in the main GM sample.

over all exams taken. The middle row reports estimates that do not control for student characteristics (for comparison with the pass five or more estimates); the bottom row reports estimates that control for student characteristics. In the first column of each row I report the adjusted ITT estimates (controlling for school and, where relevant, student characteristics). In subsequent columns I report ITT estimates based on linear and linear-interacted specifications of the vote share and 2SLS estimates based on linear, linear-interacted, and quadratic-interacted specifications of the vote share.

Panel A1 of table 3 reports estimates based on YCS data from 1993 to 2003. The estimates reported in the top row are consistent with, but less precise than, the school-level estimates reported in table 1. This loss of precision is not surprising since YCS samples are much smaller than those underlying the school-level estimates. The estimates reported in the middle row are consistent with those reported in the top row since the estimated impacts of attending a GM school represent around one-fifth of a school standard deviation in both cases. The estimates reported in the bottom row suggest that student characteristics can account for around one-third of the estimates reported in the middle row. A visual impression of this bottom row effect is given by the graph in panel A of figure 5*A*.

In panels A2 and A3 of table 3, I report the same estimates for students taking exams within 5 years of the vote (panel A2) and students taking exams 5 or more years after the vote (panel A3). Since the only enrollment effects were found for cohorts entering secondary school after the vote (taking exams 5 or more years after the vote), observed student quality might play a larger role in this later period. This would also be in line with school-level analyses of GM impacts on school free school meal eligibility rates (online App. table B3, panel C), which suggest small effects within 5 years of the vote (less than half a percentage point off a base of 15 percent) and larger effects 5 or more years after the vote (more than 1 percentage point). Consistent with this conjecture, student characteristics account for around one-half of the estimated impacts of attending a GM school in the later period and around one-fifth of the estimated impacts of attending a GM school in the earlier period. Analyses that use sampling weights to deal with nonresponse in the YCS lead to the same conclusion (online App. table B4, panel A). Analyses based on the NPD also lead to the same conclusion. As seen in panel A4 of table 3, the NPD estimates are more precise than the YCS estimates but are broadly in line with them. Perhaps surprisingly, robustness checks suggest that the addition of the primary school test scores available in the NPD reduces these estimates by only less than a quarter of an exam point (online App. table B4, panel B). Taken together then, the YCS and NPD estimates suggest that attending a GM

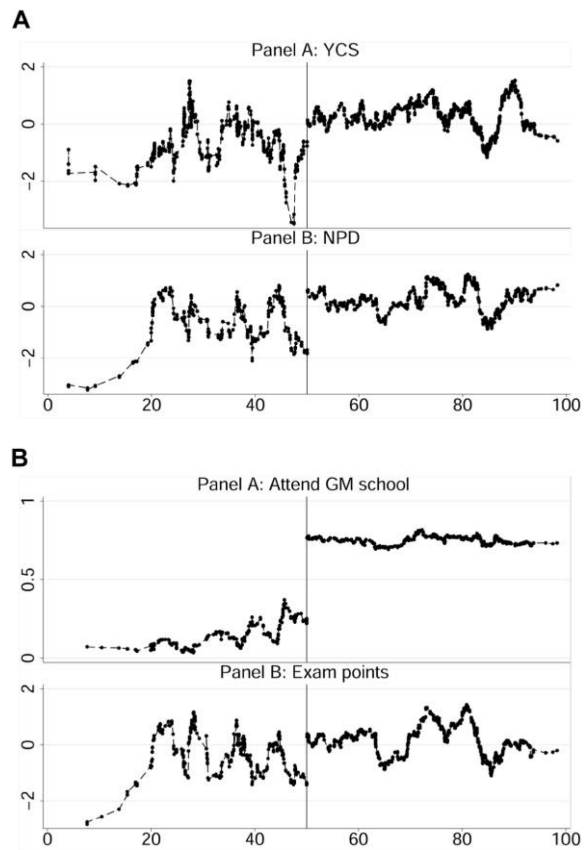


FIG. 5.—*A*, The impact of attending a school that wins a GM vote: smoothed (running mean) regression-adjusted outcomes. Means are calculated separately above and below the 50 percent threshold using bandwidth 0.1. Regression adjustment is made for base pass rates, school type, year-term of vote, and year and student characteristics (see the text for details). In panel A, the sample includes students with nonmissing data in schools in the main GM sample matched to the YCS. $N = 15{,}383$, N(schools) $= 718$. In panel B, the sample includes students with outcomes matched to primary school scores in schools matched to the main GM sample. $N = 227{,}755$, N(schools) $= 711$. *B*, The impact of attending a primary school for which the predicted secondary school wins a GM vote: smoothed (running mean) regression-adjusted outcomes. Means are calculated separately above and below the 50 percent threshold using bandwidth 0.1. In panel A, the outcome is whether or not students attended a GM school. In panel B, the outcome is regression-adjusted total exam points. Regression adjustment is made for predicted secondary school base pass rates, school type, and year-term of vote. The sample includes all students attending primary schools for which the predicted secondary school is in the main GM sample. $N = 213{,}192$, N(primary schools) $= 3{,}397$, N(predicted secondary schools) $= 676$, and N(secondary schools) $= 2{,}958$.

school increased exam performance by around 1 point for students attending 5 or more years after GM conversion and by more than 1 point for students attending within 5 years of GM conversion. A 1-point effect corresponds to a tenth of a school standard deviation in this outcome.

Since GM conversions were associated with changes in observed student characteristics, one might worry that they were also associated with changes in unobserved student characteristics (e.g., parental motivation). To check this, I present estimates based on primary school samples that instrument for whether a student attended a GM school with a dummy variable for whether the student's predicted secondary school won a GM vote. I define a student's predicted secondary school as the one in which the largest fraction of the student's primary school classmates enrolls. As shown by the graph displayed in panel A of figure 5*B*, the probability that students attend a GM school increases sharply as the predicted secondary school vote passes the 50 percent threshold. As shown by the graph displayed in panel B of figure 5*B*, there is a corresponding discontinuity in student outcomes. The associated point estimates, reported in the middle row of table 3, panel B, suggest that attending a GM school increased exam performance by around 1.5 exam points.

Two aspects of these results are worth noting. First, they are robust to the addition of student characteristics (the bottom row of panel B of table 3). This is in sharp contrast with the patterns of estimates reported in panels A3 and A4 of table 3 and is consistent with the assumption that students did not select into primary schools on the basis of the predicted secondary school vote outcome.[23] That assumption is not trivially satisfied because I observe the primary school that students attended in grade 6 in 1996–97 (after the votes), not the primary school that students attended in grade 1 in 1991–92 (before the votes). Additional primary school enrollment analyses confirm that there was no influx of students into those primary schools for which the predicted secondary school vote was won. Second, since these estimates are comparable with the estimates reported in the bottom rows of panels A3 and A4 of table 3, they suggest that those were not severely biased by omitted student characteristics. As such, this analysis adds weight to the earlier conclusion, that attending a GM school increased exam performance by at least 0.1 school standard deviations. In turn, this suggests that at least one-half of the school-level performance improvements could operate net of student composition, with the caveat that the

[23] It is also consistent with the absence of an effect of a predicted secondary school vote win on the behavior of primary schools or parents. This would be a violation of the exclusion restriction that must be satisfied for these effects to be interpreted as purely the result of attending a GM school.

school-level impacts of a GM conversion net of student composition need not be the same as the student-level impacts of attending a GM school (because of peer effects).

B. The Effects of GM Conversion on Neighbor Schools

The above analysis showed that GM conversions had large positive impacts on school performance. With this in mind, it is interesting to ask whether these gains spilled over to neighboring schools. Positive spillovers could have occurred through any of the three channels discussed above: direct competitive pressure (whether the neighbor is GM or non-GM), copycat effects, and effects on district behavior. Negative spillovers could exist if there were no such competitive responses and if GM-driven student quality improvements came at the expense of neighboring schools. I begin the neighbor analysis by estimating models similar to those estimated above, with neighbor school performance as the outcome (cf. eq. [2]) and neighbor schools defined as those in the same district as the voting school.[24] Neighbor school performance is defined as the average pass rate among students in the schools neighboring the voting school. Districts are defined as local schools authorities (local education authorities) in metropolitan areas and county districts in nonmetropolitan areas.[25] There are 528 of these districts, containing an average (mean and median) of seven schools.

The graphs in figure 6 suggest that there were no spillovers associated with GM conversions. Figure 6*A* confirms that there was no discontinuity in neighbor base pass rates; figure 6*B* paints a similar picture for regression-adjusted neighbor pass rates between 5 and 8 years after the base year. I focus on this later period in case spillovers are realized with a lag. Panel A of table 4 reports the corresponding regression estimates and reports regression estimates of impacts on neighbor exam points. The pass rate estimates are close to zero, in line with the impression given by the graphs, and the confidence intervals rule out spillovers larger than one-half of the main effect (around 5 percentage points). The exam points estimates are larger but are not statistically different from zero.

The estimates reported in panels B and C of table 4 suggest that these

[24] Since a school can be the neighbor of two different voting schools, I cluster standard errors at the district level. I cluster standard errors in this way even when the neighborhood definition implies that neighborhoods are nonnested (e.g., distance-based definitions). Experiments with and without clustering suggest that clustering increases standard errors only slightly.

[25] Districts are relatively small units, much smaller than the county-level local education authorities. These seem too large to be single education markets. Consistent with this hypothesis, many large county local education authorities (e.g., Lancashire) have recently been broken up into smaller units (unitary authorities).

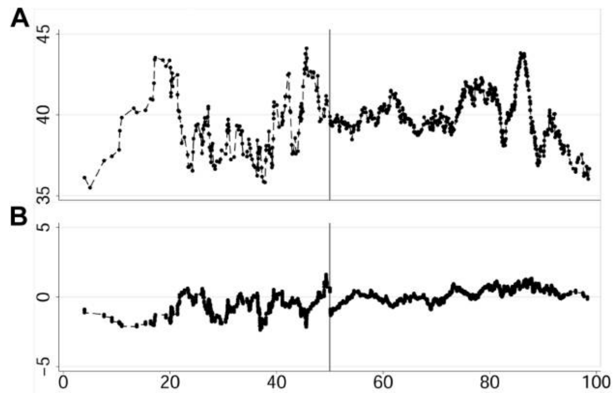


Fig. 6.—The impact of a GM vote win on neighbor school performance. *A*, Neighbor base pass rates; $N = 742$. *B*, Neighbor pass rates, 1998–2003; $N = 2{,}786$ (717 schools). Smoothed (running mean) outcomes; means are calculated separately above and below the 50 percent threshold using bandwidth 0.1. Regression adjustment is made for neighbor base pass rate, school type, and year-term of vote.

results are robust to the definition of the neighborhood. In panel B, I define neighborhoods by distance, weighting neighbor school outcomes according to their distance from the voting school (using triangular weights from 0 to 12 kilometers). A distance-based definition is attractive since open enrollment allows students to cross district lines. In panel C, I define neighborhoods via a primary school dissimilarity index. This can account for school competition not based on distance, for example, between schools with comparable prevote performance.[26] For both sets of estimates, the signs change across the two outcomes (pass rates and exam points) and across the various specifications. Again, the confidence intervals always span zero, and effects larger than one-half of the main GM effects can usually be ruled out.

Since the models estimated above should capture spillovers operating through all the mechanisms outlined earlier—direct effects, copycat effects, and district behavior—they suggest that GM conversions had none of these effects. To check this, I analyzed the impact of GM conversions on the probability that neighboring schools become GM (copycat effects) and analyzed the impact of GM conversions on the performance of non-GM schools (direct effects). My analysis confirmed that there were no copycat effects associated with GM conversions: the relevant graph shows no discontinuity in the relationship between the GM

[26] Unfortunately, I do not have student-level data before 2002; hence the dissimilarity weights are based on posttreatment data. Given the small enrollment effects found above and discussed below, one would expect these to be highly correlated with weights based on pretreatment data.

TABLE 4
The Impacts of a GM Win/Conversion on Neighbor School Performance

	RAW-ITT (1)	ADJ-ITT (2)	RD-ITT: Linear (3)	RD-ITT: Lin × Win (4)	2SLS-TE: Lin × Win (5)	2SLS-TE: Quad × Win (6)
			A. District			
Pass rates (base + 5–8)	2.306	1.108	−.092	.053	.068	−.617
	(1.010)	(.412)	(.778)	(.822)	(1.034)	(1.548)
Exam points	1.738	1.058	.606	.669	.844	.622
	(.419)	(.277)	(.461)	(.504)	(.625)	(.918)
			B. Circle 12 Kilometers—Distance Weighted			
Pass rates (base + 5–8)	2.525	.659	−.451	−.370	−.438	−1.203
	(.853)	(.312)	(.516)	(.541)	(.634)	(1.292)
Exam points	1.598	.630	−.002	.054	.064	−.170
	(.344)	(.188)	(.326)	(.349)	(.403)	(.793)
			C. Circle 12 Kilometers—Dissimilarity Weighted			
Pass rates (base + 5–8)	3.718	.645	−.575	−.309	−.396	−.912
	(.975)	(.375)	(.626)	(.653)	(.826)	(1.305)
Exam points	2.073	.595	−.166	−.051	−.065	−.572
	(.404)	(.237)	(.391)	(.415)	(.518)	(.856)

Note.—Standard errors clustered by district are in parentheses. Outcomes are constructed using information for schools in the relevant neighborhood of the schools in the main GM sample. Regression adjustment is made for base neighbor pass rates, school type, vote year-term, and year. Sample sizes are slightly smaller than the number of schools in the main GM sample because for some of these schools there are no schools in the relevant neighborhood and, hence, no outcomes. Sample sizes in panel A are 2,786 observations on 717 schools in 323 districts in row 1 and 711 schools in 321 districts in row 2, In panels B and C the relevant sample sizes are 2,843/929/327/729/327 and 2,684/688/318/688/318.

vote share and future GM attempts (online App. fig. B5*A*). Since the absence of copycat effects implies that voting was independent across schools, I assessed direct effects by examining whether non-GM school performance was affected by the number of GM schools in the neighborhood. I again took a regression discontinuity approach, this time exploiting the discontinuity between the number of GM schools and the maximum GM vote in the neighborhood of these non-GM schools (the number of neighbor GM schools increases sharply as the maximum vote among neighbors passes 50 percent). The graphs and regression estimates (online App. fig. B5*B*, App. table B5, panel A) provide no evidence of any spillovers associated with GM conversions, although these estimates are less precise than those reported in table 4.

I used two approaches to assess the extent to which these estimates reflect changes in neighboring school student composition. First, I estimated the impact of attending a school in the neighborhood of a GM school, based on analyses parallel to those underlying table 3. These were neither uniformly smaller nor larger than the school-level estimates reported in table 4. Second, I estimated the impact of a GM conversion on district-level performance, exploiting the discontinuity in the relationship between the district maximum vote share and the number of GM schools in the district.[27] If peer effects are "linear in means," district average performance will be invariant to student composition and will deliver an enrollment-weighted average of the performance and competitive effects of GM conversion. The point estimates (online App. fig. B5*C*, App. table B5, panel B) on pass rates are around 1.15, consistent with a 5-percentage-point effect enjoyed by the roughly one-quarter of the students in GM schools and no spillovers. The point estimates on exam points are slightly larger than would be implied by this calculation, although these are not very precise. The confidence intervals on both outcomes are consistent with large own-school performance effects and smaller spillovers in either direction.

VI. Discussion and Interpretation

I estimate that GM conversion increased school pass rates by between 4 and 6 percentage points, roughly 10 percent of the mean pass rate and one-quarter of a standard deviation in base pass rates. These effects are economically significant. For example, a back-of-the-envelope calculation suggests that gains of this magnitude could be worth in excess of $1 million for each cohort of a GM school's students.[28] In this section I discuss what might have driven these gains and how they compare with findings obtained in the related literature.

A. *Were GM Performance Gains Real?*

A natural first question is whether GM gains were real. These gains might not be considered real were they driven by changes in the com-

[27] A more direct approach would be to estimate whether GM conversions reduced enrollment and lowered student quality in neighboring schools (as opposed to nonneighboring and private schools). I estimated the fraction of GM enrollment increases matched by neighboring school enrollment decreases, but these estimates were too imprecise to be useful. The problem is that since GM enrollment effects are small (42 students on a base of 765; see table 2, panel C, col. 5), potential neighbor school enrollment effects are extremely small (up to 42 students on a base of 5,000 students, assuming seven equal-sized neighbor schools). Estimates of GM impacts on indicators of neighbor school student quality (e.g., free school meal eligibility rates) are similarly imprecise.

[28] This calculation assumes a pass rate effect of 5 percentage points, average cohort enrollment of 160 students, a real discount rate of 3 percent, average annual earnings of £22,900 (National Statistics 2005), full-time work from age 18 to 65, an exchange rate of £1 = $1.74, and an estimated return to passing five or more examinations of 20 percent. This is in the middle of effects estimated using the 1958 birth cohort study (National Child Development Study) and the 1970 birth cohort study (British Cohort Study). The NCDS (BCS) estimate of 0.230 (0.147) is conditional on sex, parents' education, parents' social class, and scores on tests taken at ages 7 and 11 (5 and 10). I thank Jo Blanden for these rate of return estimates. High school expenditure conjectures are based on school budget data used in Levacic (2004) and provided by Ros Levacic. Details are available from the author.

position of tested students or by practices in the spirit of teaching to the test. In the U.K. context, these practices could include focusing attention on students close to passing five or more exams or altering the number and mix of courses to maximize a student's chances of passing five or more of them. They could not, however, include some practices studied elsewhere. For example, because students cannot leave school until the end of grade 11 (regardless of age) and because all students are included in the denominator of the school pass rate (whether they are ill, have been suspended, or are classified as being in special education), the fraction of students passing five or more exams is invariant to composition-related practices such as suspending students (Figlio 2006) or reclassifying them as having special educational needs (Jacob 2005). Trivially, it is also invariant to student dropout rates since students cannot drop out until the end of grade 11.[29] In addition, because each exam is specific to a course followed in grades 10 and 11 (e.g., English, history) and covers the material on the predefined course curriculum, the usual interpretation of teaching to the test—test preparation at the expense of curriculum-based activities (Jacob 2005)—is not a natural one in this context.

The evidence suggests that GM gains cannot be attributed to changes in student composition or teaching to the test. First, I showed that changes in student composition are unlikely to explain more than one-half of the medium-run performance effects and even less of the shorter-run performance effects. This is consistent with the enrollment evidence, which suggests that students taking exams in the first few years after a GM conversion were likely those enrolled in the school before the vote took place. Second, I found strong GM performance effects when performance was measured by average points score. This suggests that GM performance effects were not confined to students on the margin of passing five or more exams. Third, additional student-level analysis suggests that GM conversions had a small impact on the number of courses taken and large impacts on performance in mathematics and English (online App. table B6). The course-taking effects are unlikely to account for more than 0.75 of the 2-point exam performance improvement, whereas the effects on performance in mathematics and English are representative of the effects on overall exam perfor-

[29] To be precise, before 1997, students could leave at Easter in their eleventh grade. The law was changed in 1997 (Circular 11/97) so that students had to stay until June of their eleventh grade. In any case, the pass rate denominator is the number of students in the school at the start of the eleventh grade, before any of them reach the compulsory school-leaving age.

mance.[30] This suggests that GM performance improvements were secured in core courses and were not driven by changes in the number of courses pursued.

The evidence also points against three other senses in which these gains may not have been real. First, it is possible that these estimates are local to the 50 percent win threshold (Angrist, Imbens, and Rubin 1996) and that the average effect of a GM conversion was smaller. Although one could construct an explanation consistent with this hypothesis, a more natural assumption is that the vote share is increasing in potential GM gains. Second, it is possible that these estimates reflect deterioration among GM losers rather than improvement among GM winners. It is impossible to rule this out completely, but difference-in-difference comparisons of nonvoters, vote losers, and vote winners show vote losers outperforming nonvoters.[31] Although it is possible that narrow vote losers were especially badly affected by a vote loss, estimates of the impact of a vote win on the subsequent tenure of head teachers is not suggestive of any crises among narrow vote losers.[32] Third, these results are not driven by selection biases induced by differential postvote survival rates. As noted above, the probability of survival is high and smooth through the 50 percent threshold.

B. *What Drove GM Performance Gains?*

If GM performance gains were real, the obvious question is what drove them. The advantages of GM status cited by pro-GM commentators and GM school head teachers fell into three categories: increased resources and the flexibility with which to use them, increased flexibility with regard to teacher deployment, and organizational changes wrought by

[30] Since average points per exam are around 5 (average total points per school 45; average exams taken nine), a 0.15 increase in course taking will generate a 0.75-point increase in total exam points. This is likely to be an upper bound on these effects since the points return to additional courses is likely to be diminishing. In addition, I find that GM school students are slightly less likely to take vocational courses, often considered the easiest to pass. Since students take, on average, nine courses, a 0.35-point improvement in combined mathematics and English performance is consistent with an improvement in overall course performance of $0.35 \times 4.5 = 1.575$ exam points. The school mean (standard deviation) combined mathematics and English exam points is 9.75 (1.55); hence the 0.35 GM effect is again around one-quarter of a school standard deviation.

[31] These difference-in-difference estimates compare 1998–2003 pass rates among schools that won, schools that lost, and schools that did not vote, adjusting for 1992 pass rates and school type. With nonvoters as the omitted category, the estimates (standard errors) associated with vote losers and vote winners are 1.072 (0.660) and 3.436 (0.445), respectively.

[32] If losing schools experienced negative shocks, we might expect tenure to be lower among vote-losing head teachers. I estimate that GM conversion had no statistically significant impact on head teacher tenure (data on head teacher tenure were collected by hand from the U.K. Education Authorities Directories).

TABLE 5
The Impacts of a GM Win/Conversion on Teaching and Nonteaching Staff

	RAW-ITT (1)	ADJ-ITT (2)	RD-ITT Linear (3)	RD-ITT Lin × Win (4)	2SLS-TE Lin × Win (5)	2SLS-TE Quad × Win (6)
	A. Number of Teachers					
Base + 1 ($N = 740$)	−5.632	1.121	1.854	1.827	2.172	2.071
	(1.452)	(.256)	(.462)	(.467)	(.554)	(1.025)
Base + 3 ($N = 727$)	−3.516	2.723	3.150	2.841	3.505	3.703
	(1.490)	(.540)	(.918)	(.984)	(1.178)	(1.993)
Base + 5 ($N = 697$)	−2.088	4.336	4.868	4.561	5.721	5.152
	(1.556)	(.723)	(1.278)	(1.366)	(1.670)	(2.782)
	B. Number of Nonteaching Staff					
Base + 1 ($N = 514$)	.161	.110	1.142	1.216	1.574	1.673
	(.326)	(.217)	(.340)	(.325)	(.457)	(.907)
	C. Separations					
Base + 1 ($N = 661$)	.320	.629	1.233	1.360	1.659	1.995
	(.385)	(.372)	(.629)	(.648)	(.787)	(1.407)
Base + 2 ($N = 652$)	−.133	.189	.257	.401	.502	−.985
	(.352)	(.357)	(.622)	(.719)	(.890)	(1.598)
Base + 3 ($N = 649$)	−1.096	−.442	.114	.976	1.235	2.304
	(.524)	(.425)	(.697)	(.940)	(1.169)	(1.621)
	D. New Hires					
Base + 1 ($N = 661$)	2.278	2.474	3.636	3.866	4.739	4.988
	(.419)	(.436)	(.881)	(.850)	(1.110)	(1.848)
Base + 2 ($N = 652$)	.249	.414	1.275	1.089	1.366	−.173
	(.422)	(.461)	(.751)	(.858)	(1.049)	(1.906)
Base + 3 ($N = 649$)	.005	.318	.794	1.220	1.529	2.144
	(.445)	(.474)	(.692)	(.829)	(1.016)	(1.548)

Note.—Standard errors are clustered by school. Regression adjustment is made for base year outcome, school type, and year-term of vote. Samples in panel A include schools in the main GM sample with nonmissing data; samples in panel B include schools in the main GM sample with base year 1993 or earlier (these data stop in 1994); samples in panels C and D include schools in the main GM sample with base year 1994 or earlier (these data stop in 1997).

GM status. The only observable indicator of school funding is the number of teaching and nonteaching staff employed by the school. The estimates reported in panel A of table 5 reveal that within five schools of converting to GM status, GM schools employed an additional five or six teachers and one additional nonteaching staff member (usually a school bursar to manage the GM school budget). This effect is around 10 percent of base teacher employment, commensurate with estimates of the GM funding advantage. It is hard to know how much of the GM performance effects can be explained by these additional teachers. Without better knowledge of the resource-performance link, I cannot rule out the hypothesis that increased resources explain all the GM performance gains or the hypothesis that the performance effects of GM resources are high because of the freedom with which GM schools were able to use them.[33]

With regard to teacher deployment, GM schools enjoyed additional flexibility in terms of hiring, firing, and paying teachers. For example, GM schools could advertise directly, use recruitment bonuses, fire teachers without following local district procedures, decide which teachers to make redundant, and change the terms of teacher contracts by, for example, linking pay to performance. I quantify some of these changes by using a national teacher database to construct school-level measures of teacher separation (the number of teachers employed last year but not this year) and teacher hiring (the number of teachers employed this year but not last year). As seen in panels C and D of table 5, I estimate that within 3 years of becoming GM, GM schools hired an additional six or seven teachers and separated from an additional three or four teachers (consistent with a net increase of three teachers; cf. the middle row of panel A). Increased hiring and turnover, on top of typical teacher turnover, would have enabled GM head teachers to shape the composition of the teaching body. Since I observe only teacher gender, age, and experience, it is difficult to know how this opportunity was used. Increased hiring and turnover could also have been used to help motivate teachers. Again, I have no data on practices such as performance-based pay, although it is interesting that Lavy (forthcoming) finds that an Israeli teacher incentive program generated GM-sized performance improvements on a comparable outcome.

Organizational change is hardest to measure. As noted above, schools converting to GM status were structured like nonprofit firms, with a new governing body composed of teachers, parents, and the head teacher, who took on a chief executive officer style role and acquired almost full responsibility for school performance. Coupled with the increased accountability requirements facing GM schools—GM schools were required to submit annual performance reports and bid for capital funding—these changes may have made head teachers more performance oriented and made teachers, parents, and students more accepting of head teachers' decisions. To test this claim empirically would require data on the attitudes and effort of teachers, parents, and students. In place of such measures, I assess whether GM performance effects were increasing in GM vote turnout and assess GM effects on student absence. This analysis does not support the hypothesis that GM

[33] In terms of resources, Card and Krueger (1996) estimate that a 10 percent increase in school resources is associated with a 1–2 percent increase in lifetime average earnings. Combined with an estimate of the earnings effect of a pass (28 percent, the central estimate of McIntosh [2002]), this gives a pass rate change of between 3 and 8 percentage points. Heckman, Layne-Farrar, and Todd (1996) argue that the Card and Krueger estimates overstate the impact of resources.

performance effects were increasing in voter turnout, although that is clearly an imperfect measure of support for GM status. The analysis does, however, show that when measured by student absence, defined as the percentage of half days that students miss, GM conversion increases effort (online App. table B7).[34] The absence effects that I estimate—reductions of roughly 0.75 percentage points in the first 4 years after a GM conversion and around 1.25 percentage points 5 or more years after a GM conversion—represent around 10 percent of base absenteeism and around one-tenth of a school standard deviation in absenteeism.[35]

C. *How Do These Estimates Compare to Those Found Elsewhere?*

It is interesting to compare these findings with those obtained for similar school reforms in other countries, particularly U.S. charter laws. Although charter school evaluations have produced mixed findings, few studies have found positive effects on the scale of those found here.[36] The simplest explanation is that the charter school sector is relatively new and extremely heterogeneous: certain types of charter may be more effective than others, and charter schools of all types might be more effective once they are a more permanent part of the educational landscape. This is consistent with evidence showing charter effects varying by type of school and increasing in years of existence. It is also consistent with evidence suggesting larger gains among "conversion" charters, schools that were previously public schools. Indeed, for the schools that most closely resemble GM schools—public high schools converting to charter school status—estimates are among the largest in the charter school literature (Sass 2006). Consistent with my spillover results, these charter studies find mixed evidence on spillovers.[37]

VII. Conclusion

This paper evaluated a U.K. reform that allowed public high schools to "opt out" of local schools authority control by becoming autonomous grant maintained schools. The education secretary that pushed through the GM reform predicted that GM schools would become "beacons of

[34] In other words, the total number of half days missed as a percentage of the total number of half days that all students should have been in school. This will be the same if one student misses 100 half days or if 100 students each miss 1 half day.

[35] Mean (standard deviation) base absenteeism is 11.5 percent (10.5 percent).

[36] Hoxby (2004) and Hoxby and Rockoff (2004) find positive effects of charter schools. Bifulco and Ladd (2006), Sass (2006), and Hanushek et al. (2007) find negative effects.

[37] Estimates from North Carolina (Bifulco and Ladd 2006) and Florida (Sass 2006) show small effects at best; estimates from Arizona (Hoxby 2003) are positive; and those from Michigan (Hoxby 2003; Bettinger 2005) are mixed.

excellence," improving their own performance and pressuring district administrators and other schools to improve theirs (Baker 1993).

My analysis supports the first part of his prediction. Exploiting the requirement that parents vote on GM status, I find that schools that converted to GM status enjoyed large improvements in student achievement, on the order of a one-quarter standard deviation improvement in pass rates on standardized examinations. I estimate that changes in student composition are unlikely to account for more than one-half of the medium-run gains and even less of the short-run gains. In this respect, it is interesting that I find that schools converting to GM status were characterized by both net teacher hiring and increased teacher turnover. These results suggest that autonomy might give school leaders valuable control over the deployment of teachers, with the caveat that GM schools might have used their resource advantage to cream-skim teachers from other schools.

My analysis provides no support for the second part of the former education secretary's prediction. Although my spillover estimates are imprecise, my best estimate is that none of the GM gains spilled over to neighboring schools. This analysis provides support for other policies, such as charter laws, that seek to hand power to schools. Although charter school evaluations uncover much smaller effects, it is plausible to suppose that charter schools will become more effective as the charter sector becomes more firmly established. My analysis does, however, suggest that the returns to these types of reforms are most likely to flow through their impacts on the reformed schools, not via spillovers to neighboring schools.

Appendix A

TABLE A1
Descriptive Statistics for the Main GM Sample

	Observations	Mean	Standard Deviation
GM vote data:			
Base year:			
1992	742	.217	
1993	742	.478	
1994	742	.199	
1995	742	.042	
1996	742	.030	
1997	742	.032	
1998	742	.001	
Number of attempts	742	1.036	.187
1st attempt, number of ballots	742	1.092	.289
Eligible votes (1st attempt, 1st ballot)	742	1,288.730	439.614
% turnout (1st attempt, 1st ballot)	742	61.962	9.452
% yes (1st attempt, 1st ballot)	742	64.540	22.780
GM at first attempt	742	.683	
GM ever	742	.708	
Base year school information:			
Entry enrollment	742	166.183	60.232
Total enrollment	742	764.455	255.492
Pass rate	742	43.232	21.950
Free school meal eligibility rates (schools with base year ≥ 1993)	581	15.421	12.487
Number of teachers (full-time equivalent)	742	52.593	17.219
Number of separations (schools with base year ≤ 1994)	662	6.671	7.019
Number of new hires (schools with base year ≤ 1994)	663	7.394	4.214
Base year neighbor school information:			
Number of schools in same district	742	7.327	2.478
Number of schools within 12 km circle	742	31.100	26.251
Base pass rates among neighbors	742	39.779	10.395
Survival data:			
Survive until base year + 1	742	.978	
Survive until base year + 5	742	.966	
Survive until base year + 8	742	.9	
Matched to YCS data (1993, 1995, 1997, 1999, 2001, 2003):			
Ever observed in YCS	731		
Fraction of students sampled	731	2.672	1.193
Total students sampled	731	25.620	13.928
School mean pass rates	731	59.810	21.821
School mean exam points	731	39.811	9.410
Matched to NPD data (2002, 2003):			
Matched to NPD	711		
Grade 11 students	711	179.389	57.349
Grade 11 students with valid primary school scores	711	.894	.059
School mean pass rates	711	59.753	20.208
School mean exam points	711	45.127	10.278

Note.—Observed in YCS means observed in the YCS sample in 1993, 1995, 1997, 1999, 2001, and 2003.

TABLE A2
STUDENT-LEVEL DATA SETS

	Youth Cohort Survey (YCS)						National Pupil Database		YCS-NPD
Sampling Details	1993 20% (6 Birthdays)	1995 10% (3 Birthdays)	1997 20% (6 Birthdays)	1999 10% (3 Birthdays)	2001 10% (3 Birthdays)	2003 10% (3 Birthdays)	2002 Administrative Data	2003 Administrative Data	2003 Matched YCS/NPD
N(schools)	2,422	1,966	2,626	1,970	1,993	1,789	3,110	3,097	1,788
N(grade 11 students)	15,685	13,278	12,709	11,647	14,687	11,956	531,043	563,622	11,849
N(grade 11 students, valid primary school test scores)							467,405	503,510	10,875
Age (months)	Y	Y	Y	Y	Y	Y	Y	Y	Y
Gender	Y	Y	Y	Y	Y	Y	Y	Y	Y
Ethnicity (5 categories)	Y	Y	Y	Y	Y	Y			Y
Ethnicity (12 categories)							Y	Y	Y
Parental education (2 categories)	Y	Y	Y	Y	Y	Y			Y
Parental education (3 categories)	Y	Y	Y	Y	Y	Y			Y
Housing tenure (5 categories)	Y	Y	Y	Y	Y	Y			Y
Parental employment status (3 categories)	Y	Y	Y	Y	Y	Y			Y

TABLE A2
(*Continued*)

Sampling Details	Youth Cohort Survey (YCS)						National Pupil Database		YCS-NPD
	1993 20% (6 Birthdays)	1995 10% (3 Birthdays)	1997 20% (6 Birthdays)	1999 10% (3 Birthdays)	2001 10% (3 Birthdays)	2003 10% (3 Birthdays)	2002 Administrative Data	2003 Administrative Data	2003 Matched YCS/NPD
Parental socioeconomic group (16 categories)	Y	Y	Y	Y	Y	Y			Y
Eligible for free school meal							Y	Y	Y
First language (5 categories)							Y	Y	Y
Primary school test scores (math, English, science)							Y	Y	Y
Socioeconomic status based on 7-digit post code							Y	Y	Y
% V(exam points) explained by:									
School fixed effects	.324	.318	.353	.338	.302	.32	.205	.199	.325
School fixed effects and YCS characteristics	.451	.447	.45	.443	.412	.432			.428
School fixed effects and test scores							.574	.547	.642
School fixed effects and NPD characteristics							.609	.585	.667
School fixed effects and NPD and YCS characteristics									.685

Note.—Bottom rows report the variance in exam points explained by school fixed effects and various combinations of observed characteristics.

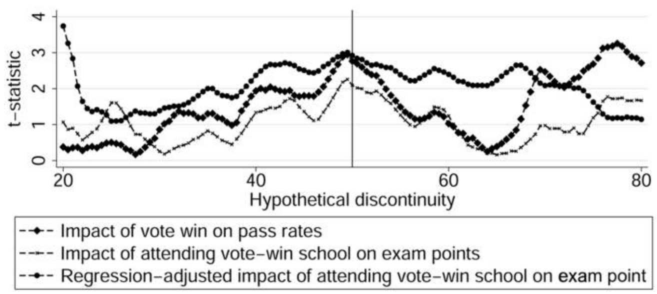


Fig. A1.—Discontinuity checks. The lines represent smoothed (running mean) absolute t-ratios associated with discontinuities from 20 to 80 percent (bandwidth 0.05). The pass rate discontinuities refer to the RD-ITT estimate in column 4 of panel B of table 1; the exam point discontinuities refer to the RD-ITT estimate in column 3 of row 2 of panel A4 of table 3; the exam point (*) discontinuities refer to the ADJ-ITT estimate in column 1 of row 3 of panel A4 in table 3.

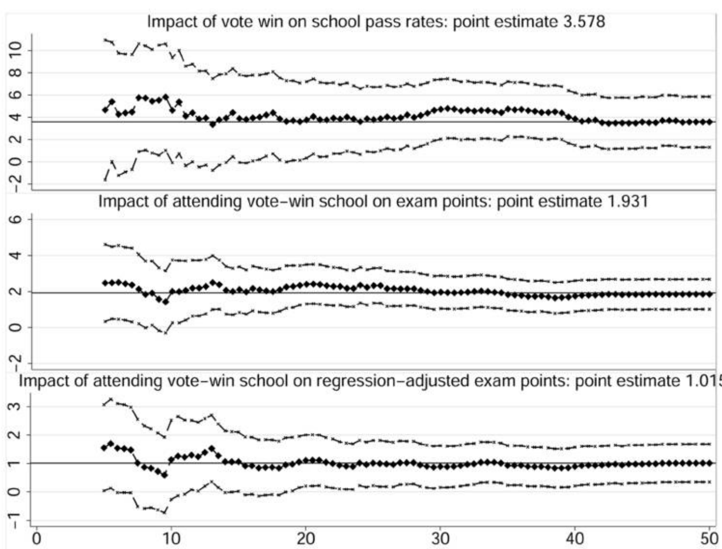


Fig. A2.—Estimates obtained using different windows around the 50 percent threshold. The graphs present estimates (and 95 percent confidence intervals) obtained by estimating models using observation windows of x percentage points around the 50 percent win threshold. The point estimates in the graph titles refer to the estimates obtained using the full range of data. The horizontal lines correspond to these estimates. The estimate in the top graph corresponds to the RD-ITT estimate in column 3 of panel B of table 1; the estimate in the middle graph corresponds to the RD-ITT estimate in column 3 of row 2 of panel A4 of table 3; the estimate in the bottom graph corresponds to the ADJ-ITT estimate in column 1 of row 3 of panel A4 of table 3.

References

Angrist, J., G. W. Imbens, and D. B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *J. American Statis. Assoc.* 91 (434): 448–58.

Angrist, J., and A. B. Krueger. 1999. "Empirical Strategies in Labor Economics." In *Handbook of Labor Economics*, vol. 3, edited by O. Ashenfelter and D. Card. Amsterdam: Elsevier/North-Holland.

Angrist, J., and V. Lavy. 1999. "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *Q.J.E.* 114 (2): 533–75.

Baker, K. 1993. *The Turbulent Years: My Life in Politics.* London: Faber and Faber.

Bettinger, E. 2005. "The Effect of Charter Schools on Charter Students and Public Schools." *Econ. Educ. Rev.* 24 (2): 133–47.

Bifulco, R., and H. Ladd. 2006. "The Impact of Charter Schools on Student Achievement: Evidence from North Carolina." *Educ. Finance and Policy* 1:50–90.

Bush, T., M. Coleman, and D. Glover. 1993. *Managing Autonomous Schools: The Grant Maintained Experience.* London: Chapman.

Card, D., and A. Krueger. 1996. "The Labor Market Effects of School Quality: Theory and Evidence." In *Does Money Matter? The Link between Schools, Student Achievement and Adult Success*, edited by G. Burtless. Washington, DC: Brookings Inst.

Chubb, J. E., and T. M. Moe. 1990. *Politics, Markets and America's Schools.* Washington, DC: Brookings Inst.

———. 1992. *A Lesson in School Reform from Great Britain.* Washington, DC: Brookings Inst.

Clark, D. 2007. "Selective Schools and Academic Achievement." IZA Discussion Paper no. 3182, Inst. Study Labor, Bonn.

Cullen, J., B. Jacob, and S. Levitt. 2005. "The Impact of School Choice on Student Outcomes: An Analysis of the Chicago Public Schools." *J. Public Econ.* 89 (5–6): 729–60.

———. 2006. "The Effect of School Choice on Student Outcomes: Evidence from Randomized Lotteries." *Econometrica* 74 (5): 1191–1230.

Figlio, D. 2006. "Testing, Crime and Punishment." *J. Public Econ.* 90:837–51.

Fitz, J., D. Halpin, and S. Power. 1993. *Grant Maintained Schools: Education in the Market Place.* London: Kogan Page.

Hanushek, E., S. Rivkin, J. Kain, and G. Branch. 2007. "Charter School Quality and Parental Decision Making with School Choice." *J. Public Econ.* 91 (5–6): 823–48.

Heckman, J., A. Layne-Farrar, and P. Todd. 1996. "Does Measured School Quality Really Matter?" In *Does Money Matter? The Link between Schools, Student Achievement and Adult Success*, edited by G. Burtless. Washington, DC: Brookings Inst.

Hirschman, A. 1970. *Exit, Voice and Loyalty: Responses to the Declines in Firms, Organizations and States.* Cambridge, MA: Harvard Univ. Press.

Howell, W., and P. Peterson. 2002. *The Education Gap: Vouchers and Urban Schools.* Washington, DC: Brookings Inst.

Hoxby, C. 2000. "Does Competition among Public Schools Benefit Students and Taxpayers?" *A.E.R.* 90 (5): 1209–38.

———. 2003. "School Choice and School Productivity: Could School Choice Be a Tide That Lifts All Boats?" In *The Economics of School Choice*, edited by C. Hoxby. Chicago: Univ. Chicago Press (for NBER).

———. 2004. "A Straightforward Comparison of Charter Schools and Regular Public Schools in the United States." Manuscript (September), Harvard Univ.

Hoxby, C., and J. Rockoff. 2004. "The Impact of Charter Schools on Student Achievement." Manuscript, Harvard Univ.

Imbens, G. W., and T. Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *J. Econometrics* 142 (2): 615–35.

Jacob, B. A. 2005. "Accountability, Incentives and Behavior: The Impact of High-Stakes Testing in the Chicago Public Schools." *J. Public Econ.* 89:761–96.

Kane, T. J., and D. O. Staiger. 2001. "Improving School Accountability Measures." Working Paper no. 8156, NBER, Cambridge, MA.

Kenny, L. 2005. "The Public Choice of Educational Choice." *Public Choice* 124: 205–22.

Krueger, A., and P. Zhu. 2004. "Another Look at the New York City School Voucher Experiment." *American Behavioral Scientist* 47 (2): 658–98.

Lavy, V. Forthcoming. "Performance Pay and Teachers' Effort, Productivity and Grading Ethics." *A.E.R.*

Levacic, R. 2004. "Competition and the Performance of English Secondary Schools: Further Evidence." *Educ. Econ.* 12 (2): 177–93.

McCrary, J. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *J. Econometrics* 142 (2): 698–714.

McIntosh, S. 2002. "Further Analysis of the Returns to Academic and Vocational Qualifications in the UK." Research Report no. 370 (September), Dept. Educ. and Skills, London.

National Statistics. 2005. "2005 Annual Survey of Earnings and Hours." First Release (November 20), National Statistics, London.

Rothstein, J. 2006. "Good Principals or Good Peers? Parental Valuation of School Characteristics, Tiebout Equilibrium, and the Incentive Effects of Competition among Jurisdictions." *A.E.R.* 96 (4): 1333–50.

———. 2007. "Does Competition among Public Schools Benefit Students and Taxpayers? Comment." *A.E.R.* 97 (5): 2026–37.

Rouse, C. 1998. "Private School Vouchers and Student Achievement: An Evaluation of the Milwaukee Parental Choice Program." *Q.J.E.* 113 (2): 553–602.

Sass, T. 2006. "Charter Schools and Student Achievement in Florida." *Educ. Finance and Policy* 1:91–122.

Thompson, M. 1992. "The Experience of Going Grant Maintained: The Perceptions of AMMA Teacher Representatives." *J. Teacher Development* 1 (3): 133–40.

Wirt, J., S. Choy, P. Rooney, S. Provasnik, A. Sen, and R. Tobin. 2004. *The Condition of Education 2004.* U.S. Dept. Educ., Nat. Center Educ. Statis. Washington, DC: U.S. Government Printing Office.