THE INTRODUCTION OF ACADEMY SCHOOLS TO ENGLAND'S EDUCATION

Andrew Eyles
University College London
and London School of Economics

Stephen Machin
London School of Economics

Abstract
This paper studies the origins of what has become one of the most radical and encompassing programmes of school reform seen in the recent past in advanced countries—the introduction of academy schools to English education. Academies are independent state funded schools that are allowed to run in an autonomous manner outside of local authority control. Almost all academies are conversions from already existent state schools and so are school takeovers that enable more autonomy in operation than was permitted in their predecessor state. Studying the first round of conversions that took place in the 2000s, where poorly performing schools were converted to academies, a focus is placed on legacy enrolled pupils who were already attending the school prior to conversion. The impact on end of secondary school pupil performance is shown to be positive and significant. Performance improvements are stronger for pupils in urban academies and for those converting from schools that gained relatively more autonomy as a result of conversion. (JEL: I20, I21, I28)

1. Introduction

The English academy schools programme is turning out to be one of the most radical and encompassing programmes of school reform seen in a developed country. Unlike traditional community schools, academies are autonomous, state-funded schools that are managed and run outside of local authority control. In almost all cases, they are conversions of pre-existing schools that inherit pupils already enrolled in the school, but enjoy significantly more autonomy in operation than was permitted in their

The editor in charge of this paper was M. Daniele Paserman.

Acknowledgments: This is a very substantively revised and overhauled version of a paper based only on school-level data previously circulated as "Changing School Autonomy: Academy Schools and Their Introduction to England's Education" (Machin and Vernoit 2011). It should be viewed as the current version of that paper. We would like to thank the Department for Education for providing access to the pupil-level data used in the paper. We received many helpful comments and suggestions from the Editor, Daniele Paserman, five referees, and participants in a large number of seminars and conferences. This work was supported by the Economic and Social Research Council grant ES/M010341/1.

E-mail: a.eyles@lse.ac.uk (Eyles); s.j.machin@lse.ac.uk (Machin)

Journal of the European Economic Association 2018 0(0):1–40 DOI: 10.1093/jeea/jvy021

predecessor state.[1] At the time of writing, nearly 65% of England's secondary schools and a further 15% of primary schools had become academies.[2] The vast majority became academies after a change of government in May 2010 when legislation—the 2010 Academies Act—greatly widened the remit of the programme.[3]

The genesis of the English academies programme is what is studied in this paper. The programme was initiated under the 1997–2010 labour government when strong concerns were being expressed that schools in particular local authorities (usually those serving disadvantaged urban neighbourhoods) were not delivering pupils a good enough education. A widespread recognition emerged that something needed to be done to improve standards in schools where it had been said that "teachers had lost control of the corridors". The proposed solution was to replace an existing school with a new type of state school managed by a private team of independent co-sponsors. The sponsors of the new academy school delegate the management of the school to a largely self-appointed board of governors who have responsibility for employing all academy staff, agreeing levels of pay and conditions of service and deciding on the policies for staffing structure, career development, discipline, and performance management.

This paper studies the causal impact of academy school conversion on pupil performance. This line of enquiry answers two key policy questions. Did the programme have the desired effect on the population it was targeted at? And is the programme likely to benefit similar pupils in the future? As the discussion has already made clear, it was pupils in disadvantaged schools that formed the target population. This selection is accounted for in the research design by using a combination of differences-in-differences and instrumental variables. Namely, it compares outcomes of pupils enrolled in academy schools to outcomes of those enrolled in a group of comparison schools—a set of state schools that go on to become sponsored academies after the sample period ends. Potentially endogenous sorting into academies is circumvented by using enrolment in the academy before conversion—legacy enrolment—as an instrument for actual attendance. The approach has similarities to that taken in Abdulkadiroglu et al. (2016), who study charter school takeovers in New Orleans. They limit their study to what they term "grandfathered" pupils; that is, those who passively enrol in a charter, by virtue of already being enrolled in the school prior to the takeover.

1. They are different from most US charter schools which are typically, though not always, set up from scratch. A closer comparison to the typical charter school in England is free schools, recent additions that are brand new schools (often set up by parent or community groups). A closer US comparison to academies are "in-district" charters where an already existent public school is converted to a charter as a school takeover—these are less commonplace than US charters as a whole, but there are places where conversions of public schools to charters have taken place (like Boston and New Orleans—see Abdulkadiroglu et al. 2016).

2. In England, secondary schooling takes place from ages 11 to 16 and primary schooling from ages 5 to 11.

3. Prior to the Act only secondary schools could become academies and to convert they were required to sign up a sponsor. Afterwards, primary schools were permitted to become academies, free schools were introduced and a sponsor was no longer required for conversion to take place. See Eyles, Hupkau, and Machin (2016) for more details.

Although this study informs the current policy debate in England, it also complements two often overlapping strands of research in the economics of education. The first focuses on how the type of school one attends influence test scores, whereas the second focuses on increasing the amount of autonomy in previously centralised education systems. The former is exemplified by the US literature on catholic school attendance (Evans and Schwab 1995; Neal 1997; Altonji, Elder, and Taber 2005) and charter school attendance (see Epple, Romano, and Zimmer 2016). Examples of the latter include cross country evidence on the contribution of greater school autonomy to international test score differences (Hanushek and Woessmann 2011, 2015; OCED 2011) and the effects of private school voucher programs (Epple, Romano, and Urquiola 2017).

Focusing on pupils who enrolled in academies prior to conversion, the results suggest that academy schools considerably raise the achievement of their attendees. The preferred estimate is that pupils who attend an academy gain on average 0.12 of a standard deviation (σ) higher end of school test scores relative to otherwise similar pupils who attend traditional schools. This effect increases over time, with pupils who attend for four years reaping gains of 0.28σ. Suggestive evidence is also presented to show that improvements are concentrated amongst those schools that gained the most autonomy after conversion. Mirroring findings on charter schools, improvements appear stronger in schools in urban areas. Alongside performance improvements, there is evidence that schools change their intake upon conversion; in particular, incoming cohorts of students have higher baseline test scores after conversion. This legitimises the adopted research design, which uses pupil level data, and explicitly accounts for such compositional changes. Although the changes are sizeable, empirical tests also show that the estimated performance improvements do not seem to come about because performance effects and changes in peer composition are related.

2. Context

2.1. Academy Schools

Academy schools were first introduced in the early 2000s. With hindsight, their introduction can be viewed as a landmark in the history of education in England.[4] Academies are now the predominant school type in the English secondary sector, but are not without controversy. The almost evangelical fervour for academisation shown by its advocates has been countered by an equal lack of enthusiasm by detractors.[5] Lord

4. It is only England, and not in the other nations of the United Kingdom (Northern Ireland, Scotland, and Wales) who run their own devolved education systems, where academies have been introduced. In the OECD's Programme for International Student Assessment (PISA) data, this has resulted in England becoming the highest ranked country in school autonomy over resource allocation in the 2012 PISA—see Eyles, Hupkau and Machin (2016) for more detail on this aspect of academisation of English schools, and the policy context more generally.

5. For example, the anti-academies alliance (see the website at http://antiacademies.org.uk).

Adonis—the key player in setting up the labour academies programme—eloquently describes this in his 2012 book (Adonis 2012).

The first clutch of academies opened in the school year beginning in September 2002. Academies are independent, nonselective, state-funded schools that fall outside the control of local authorities. In most cases, they are conversions of already existing predecessor schools. The first tranche of academies that are studied in this paper are managed by a private team of independent co-sponsors. The sponsors of the academy school delegate the management of the school to a largely self-appointed board of governors, which has responsibility for employing all academy staff, agreeing levels of pay and conditions of service and deciding on policies for staffing structure, career development, discipline, and performance management. Since the introduction of the Academies Act 2010, the programme's remit has changed. Converter academies—good and outstanding schools that gain academy status without a sponsor—now dominate the English educational landscape.[6] It is important to note that all the academies that studied in this paper, both in the treatment and control group, are sponsored academies that—in the main analysis—were approved for opening prior to this change.[7]

2.2. Secondary School Types in England and Academy Introductions

The English secondary education system is composed of seven school types: independent schools, academy schools, city technology colleges (CTCs), voluntary aided schools, foundation schools, voluntary controlled schools, and community schools. Each school type is characterised by a unique set of features regarding their autonomy and governance. This is shown in Table 1. In the table, the different school types are ordered by the amount of autonomy that their governing body/management body has, ranging from those with the most (private fee-paying independent schools that operate outside of the state sector) to those with the least (community schools that are largely operated under the remit of local authority control).

In the time period under study, the main impetus of the programme was to replace failing schools with academies by moving away from the conventional school type that had populated the English secondary sector in the past.[8] The path to establishing an academy school in a local authority involved a number of steps. The key feature was the need to sign up a sponsor, who worked with the local authority (LA) where the school operates, and to complete a formal expression of interest (this made the

6. Post-2010 the academies programme was also extended to cover primary schools as well (see Eyles, Machin, and McNally 2017).

7. The latest opening date, for the control schools, is September 2010, which coincides with the first openings of converter academies. A focus is placed on those that follow the sponsor route as these are underperforming schools gaining academy status via the same route as those in the treatment sample. Converters voluntarily gain academy status and are not comparable to the schools that are studied.

8. There were some other cases, for example, where schools that already had more autonomy than a typical state community school became an academy, or as a means for fee-charging independent schools to broaden their intake of pupils by becoming academies (Department for Children, Schools and Families 2007), but as the numbers discussed below will show, these were the exception rather than the norm.

TABLE 1. Characteristics of autonomy and governance in English secondary schools.

	Non-LA Admission Authority	Maintained by Non-LA body	Not obliged to follow National Curriculum	Fee charging
Registered independent school[a]	✓	✓	✓	✓
Academy[b]	✓	✓	✓	×
City technology college[c]	✓	✓	✓	×
Voluntary-aided[d]	✓	×	×	×
Foundation[e]	✓	×	×	×
Voluntary-controlled[f]	×	×	×	×
Community[g]	×	×	×	×

a. Registered independent schools are independent of the local authority (LA), and are fee-charging.

b. Academy schools (prior to 2010/2011): all ability independent specialist schools, which do not charge fees, and are not maintained by the local authority; established by sponsors from business, faith, HE institutions or voluntary groups, working in partnership with central government. Sponsors and the DfE provide the capital costs for the Academy. Running costs are met by the DfE in accordance with the number of pupils, at a similar level to that provided by local authorities for maintained schools serving similar catchment areas.

c. City Technology Colleges: all ability independent schools, which do not charge fees, and are not maintained by the local education authority. Their curriculum has a particular focus on science and technology education (see West and Bailey 2013). They were established by sponsors from business, faith or voluntary groups. Sponsors and the DfE provided the capital costs for the CTC. Running costs are met by the DfE in accordance with the number of pupils, at a similar level to that provided by local authorities for maintained schools serving similar catchment areas.

d. Voluntary-aided schools are maintained by the local authority. The foundation (generally religious) appoints most of the governing body. The governing body is responsible for admissions and employing the school staff. Land at voluntary-aided schools is usually owned by trustees, although the local authority often owns any playing field land (Department for Schools, Children and Families 2008).

e. Foundation (formerly grant-maintained) schools are maintained by the local authority. The governing body is responsible for admissions, employing the school staff, and either the foundation or the governing body owns the school's land and buildings (Department for Schools, Children and Families 2008).

f. Voluntary-controlled schools are maintained by the local authority. These are mostly religious schools where the local authority continues to be the admission authority. Land at voluntary-controlled schools is usually owned by trustees (Department for Schools, Children and Families 2008).

g. Community schools are maintained by the local authority. The local authority is responsible for admissions, employing the school staff, and it also owns the school's land and buildings.

case that an academy in the proposed area was both needed and feasible). The phase is completed when the LA and sponsor send the expression of interest to the Secretary of State for Education for his or her ministerial approval. After approval the process moves on to the feasibility stage and beyond that to actual conversion of the already existing school to an academy.

Figure 1 shows the evolution of school types between the school years 2001/2002 and 2009/2010. The change in the composition of schools is modest relative to the vast post-2010 expansion of the programme. In 2001/2002 there were no academies; by 2009/2010, 203 academies were in operation. By 2008/2009—the final school year in the sample to be analysed—there were 133 academies open and operating. These had a gradual introduction, with the first three opening in the 2002/2003 school year, and then

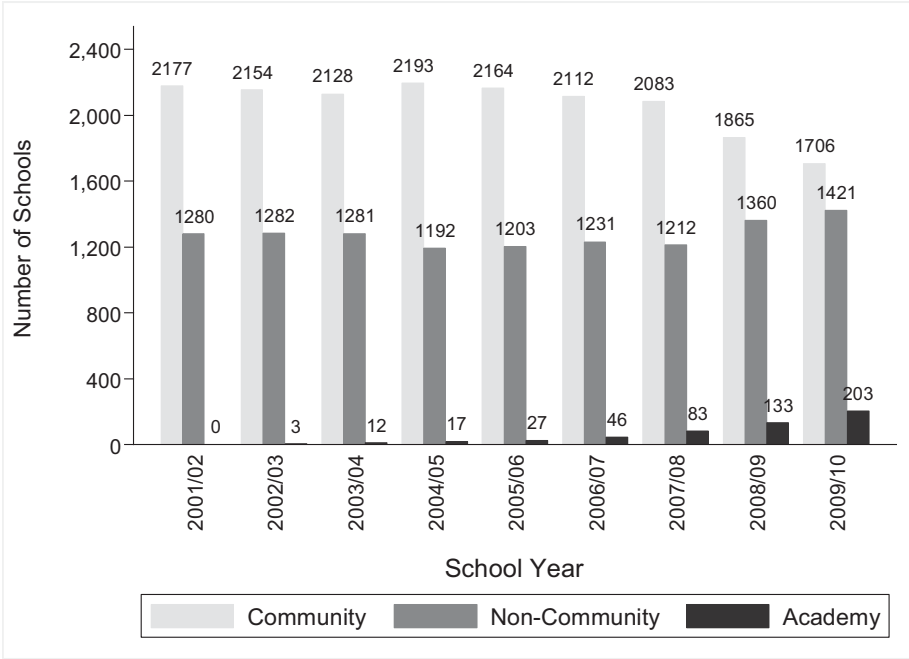

FIGURE 1. Secondary school types in England, 2001/2002 to 2009/2010. Author's calculations using Department for Education data. Includes middle schools. Excludes special schools. Noncommunity schools are city technology colleges, foundation schools, voluntary aided schools, and voluntary controlled schools. See Table 1 for more detail on these school types.

in the subsequent school years as follows: 2003/2004: 9; 2004/2005: 5; 2005/2006: 10; 2006/2007: 19; 2007/2008: 37; 2008/2009: 50.

Table A.1 shows that (at least) one school from every secondary school type converted to become an academy prior to 2008/2009. Because of the research design to be implemented, the focus is limited to schools that convert from an already-existing school; furthermore, the analysis is based upon schools that enrol pupils at age 11 and have students sit their final compulsory schooling exams at the school (this corresponds to the conventional secondary school in England). The final sample consists of 94 treatment schools drawn from the seven cohorts of schools opening prior to the 2008/2009 school year and 114 control schools opening in the 2009/2010 and 2010/2011 school years.[9] These latter two cohorts consist solely of sponsored academies that were approved to become academies prior to the new regime that arose after the 2010 Academies Act.[10]

9. The two main discrepancies between the sample the 133 and 94 are the removal of newly built academies, of which there were 12, and of 12 conversions from City technology colleges. The rationale for this latter omission is given below.

10. For inclusion in the analysis, the approval of "future" academies had to have taken place before May 2010, when the government changed and the new coalition introduced the Academies Act.

2.3. Related Literature

Although there is a sizeable body of research on the impact of different schooling systems on pupil performance, there are fewer studies that look at what happens when the type of school attended by pupils changes. One study that looks at schools changing status in England is Clark (2009). He looks at what happened when schools became grant-maintained (GM)—a school type that enjoyed substantial operational autonomy.[11] He utilises a regression-discontinuity design based on the fact that the decision to change status was decided by parental vote. As narrow GM vote winners experienced a significant improvement in pupil performance (of about a quarter of a standard deviation) compared to the narrow GM vote losers, his results suggest that increased school autonomy can bring about performance improvements.

GM schools were introduced in the late 1980s and conversion to GM status involved little turnover in management; indeed, the process was voluntary and often instigated by the school's governors. The granting of greater autonomy to already successful schools contrasts with the initial academies programme, where managerial changes were imposed on schools deemed to be struggling. In this respect, the United States work on charter schools is more relevant to the analysis undertaken in this paper.[12]

Initial findings from the literature on charters, based upon quasi-experimental research designs, produced mixed to negative results. For instance, Betts et al. (2006) find that charters perform roughly at a similar level to public schools in the 16 charters they study in San Diego whereas two studies carried out by CREDO (2009, 2013) find little average effects when looking at charters across 16 and 27 states.

Concerns with nonrandom selection into charters subsequently led researchers to begin to look at lottery based estimates of the effect of charter attendance. These studies exploit the fact that some schools use lotteries to allocate places when the school is oversubscribed. The vast majority of these papers find substantial positive test score gains for pupils "lotteried" in to charters relative to those "lotteried" out (see Hoxby, Murarka, and Kang 2009; Angrist et al. 2010; Abdulkadiroglu et al. 2011; Dobbie and Fryer 2011; Angrist et al. 2013; Dobbie and Fryer 2013 for studies of test score gains; and Dobbie and Fryer 2014; Angrist et al. 2016 for evidence of students' longer-run outcomes, including college attendance).

An exception to the above is Gleason et al. (2010) who use lottery estimates from 36 charters across 15 states and find little evidence of improvements in pupil performance on average. However, they do find performance improvements for disadvantaged children (defined as those on free school meals). Similarly, Angrist, Pathak, and Walters (2013) find that when splitting their Massachusetts sample between urban and nonurban charters, gains are positive in urban schools but negative for nonurban charters. As the majority of the lottery studies are based upon charters serving disadvantaged children in urban areas—such as New York and Boston—these latter studies shed light on seemingly disparate findings between lottery and nonlottery based studies.

11. GM schools were renamed as foundation schools (see Table 1) in the Schools Act of 1998.

12. Epple, Romano, and Zimmer (2016) provide an in-depth and up-to-date survey of the work on charter schools.

Charters differ from academies in two important dimensions; first, charters are often newly built or set up schools; and second, applications to charters tend not to be co-ordinated with applications to other local schools.[13] Some recent studies appear pertinent to the case of English academies in these two dimensions. A small number of US studies have looked at conversions of already existing public schools to charters (as in the study of school takeovers in Boston and New Orleans by Abdulkadiroglu et al. 2016), as well as the introduction of practices used in charters to US public schools (as in Houston schools studied by Fryer 2014). These report substantial improvements in test scores in those setting due to the use of methods of "best practice". Alongside these, Abdulkadiroglu et al. (2017) report lottery estimates for charters in Denver where, contrary to usual practice, places at charters and public schools are allocated using a common assignment scheme. In a school choice setting similar to the one studied here, they find positive effects of charters on performance.

On academies themselves, there remains little rigorous work. Very early work by Machin and Wilson (2008) looked at differences in pupil performance between a small sample of the first academy schools and a matched group of schools, finding modest, statistically insignificant, relative improvements. A PwC Report (2008) reported higher percentage point increases in the results of academies compared to the national average (which is not a good comparison since academies are well below average performers in their predecessor state), whereas a National Audit Office (2010) report on the labour academies looked at their performance compared to a selected group of maintained schools, with similar pupil intakes and performance to the academies pre-treatment, finding a significant improvement in pupil performance in the academies. There is also some largely descriptive, noncausal school-level empirical work in the education field.[14]

3. Data and Research Designs

3.1. Data

The main data source is the National Pupil Database (NPD).[15] The NPD is centrally collected census data containing pupil and school characteristics combined with the

13. While academies can set their own admissions criteria so long as it accords with legislative guidelines, applications to state schools are co-ordinated at the local authority level. Compliance with local authority co-ordination of admission arrangements is part of an academy's funding agreement.

14. See, for example, Gorard (2014) or West and Bailey (2013).

15. The use of pupil-level data throughout and a heavily refined research design are the key innovations compared to the version of this paper circulated earlier (Machin and Vernoit 2011). Of course, use of pupil-level data (which the earlier version did not have full access to) makes the analysis more appropriate in that the right level of treatment is the effect of schools on the pupils that attend them compared to schools they would otherwise have attended. Put another way, changing pupil composition due to academy conversion because the demand for places alters compared to the predecessor school can render school-level estimates biased.

annual National Curriculum key stage attainment data at the pupil level. The Pupil Level Annual Census data (PLASC) contains information on characteristics of all pupils in the English maintained sector. This has been collected three times per year (January, May, and September) from the 2001/2002 school year onwards (though pupils can be traced back to earlier years of the key stage attainment data via their unique id). For this paper, only use the year-on-year January collection is used because this collection is the most available and consistent over time.

In England, compulsory education is organised around four key stages for years of schooling from ages 5 to 16. These are key stage 1 (in grades 1 and 2) and key stage 2 (grades 3–6) in primary school; and key stage 3 (grades 7–9) and key stage 4 (grades 10 and 11) in secondary school. In studying academy conversion impacts, the two outcomes of interest are pupil intake and pupil performance. To study intake for pupils enrolling in secondary school in grade 7, the first grade of secondary school, the focus is on the key stage test exams (KS2) that pupils take at the end of primary school (aged 10/11 at the end of grade 6) before they make the transition to secondary school. To study performance in grade 11, the final year of compulsory secondary schooling, the key stage 4 (KS4) examinations that pupils take at the end of compulsory schooling (aged 15/16 at the end of grade 11) are studied. These school leaving exams are known as GCSEs (standing for the General Certificate of Secondary Education).

The impact of academy conversion needs to be analysed at the pupil-level. This is because the underlying composition of students attending schools may change over time and, indeed, it turns out that pupil intake does change post-conversion. It is important to devise an empirical strategy that is not contaminated by the changing quality of post-conversion enrolees. A causal effect of academy attendance on pupil performance is therefore identified by focusing on pupils who were already enrolled in an academy pre-conversion. These pupils are referred to as being legacy enrolled. Because they had been enrolled in the school prior to conversion this avoids the endogeneity of the post-conversion enrolment decision that would contaminate estimates obtained from also looking at newly enrolled students.[16]

3.2. Comparison Schools

The research design combines difference-in-differences with instrumental variables. Before going into specific details, first the comparison schools are defined. Table 2 compares pre-treatment characteristics of academy schools and other types of maintained English secondary schools. The 106 academies (who have both a grade 7 intake and grade 11 exam takers) very clearly have significantly different pupil characteristics and levels of pupil performance than other state maintained secondary schools.

16. One further practical issue concerns the definition of schools that convert to academies. There are a small number of examples where multiple predecessor schools combine to create a single academy school. Where this occurs, create one hypothetical pre-academy school is created (see a fuller discussion in the Data Appendix). This adopts hypothetical characteristics that are a weighted-average of the characteristics of the merged schools.

TABLE 2. Pre-conversion characteristics and tests of balancing.

	End of primary test score (KS2) (mean)	End of secondary test score (KS4) (mean)	Proportion getting 5 or more A*-C GCSEs or equivalents (mean)	Proportion male	Proportion White	Proportion eligible for free school meals	Number of Schools
A: All Schools							
City technology college	74.786	57.804	0.934	0.487	0.968	0.095	2
Voluntary aided	66.763	43.323	0.578	0.505	0.798	0.126	502
Foundation	65.516	43.340	0.573	0.522	0.85	0.092	470
Voluntary controlled	66.827	43.515	0.579	0.510	0.876	0.077	96
Community	61.983	38.312	0.460	0.503	0.828	0.153	1933
Academies (pre-conversion)	57.230	31.689	0.316	0.536	0.725	0.250	106
B: Academy Schools							
Current academies (treatment group)	55.408	29.619	0.267	0.536	0.730	0.264	94
Future academies (control group)	56.476	30.912	0.285	0.515	0.812	0.232	114
Difference	−1.068	−1.293	−0.018	0.021	−0.082**	0.032*	
	(0.796)	(0.834)	(0.018)	(0.015)	(0.040)	(0.019)	

Notes: Standard errors clustered at school level reported in parentheses. Both panels refer to characteristics in the 2001/2002 school year. Panel A is maintained schools in England, which do not convert to academies prior to, or in, the academic year 2008/2009. All variables with the exception of KS2 point score, which refers to characteristics of the incoming 2001/2002 cohort, that is, incoming pupils in the school year 2001/2002, before any academies had opened, refer to those in their final school grade. *Significant at 10%; **significant at 5%.

The fact that these schools show higher signs of disadvantage and record lower achievement in school leaving tests is not surprising as labour's academy programme was aimed at poorly performing schools. Thus, a naive comparison between academy schools and all other state-maintained schools is likely to suffer from significant selection bias. There is one exception here, as the 12 conversions from City Technology Colleges (CTCs) were already highly autonomous schools that were performing well. These are therefore omitted and the treatment group defined as the 94 new academies that converted from the four groups of state maintained schools: community, voluntary controlled, foundation and voluntary aided schools.[17] In addition to selection on observables, a related issue is that schools that go on to become academies may have common unobservable characteristics (e.g., they have a type of school ethos that is more in line with the academy model). Finally there is also scope for mean reversion, as academies were badly performing schools in their predecessor state.

Panel B of Table 2 shows the pre-treatment characteristics of both the 94 schools that become academies in the sample period and 114 schools that become academies later after the study sample period ends. In contrast with the top panel, the characteristics of these two sets of schools appear balanced in the pre-treatment period; that is, for most of the variables considered (the exceptions being the proportion White for a 5% significance level and free school meal eligibility at the 10% level), one cannot reject the null hypothesis that the 94 academies that convert in the sample period and the 114 academies that convert in the following two school years have, on average, the same sets of characteristics.[18] This partially legitimises the use of pupils attending future converters as a control group in the D-i-D setting. It is further legitimised in the empirical findings described in what follows where there is no sign of differential pre-conversion trends between treatment and control schools in test scores, thus allaying concerns of mean reversion.

3.3. Modelling Approach

To first study the issue of changing pupil composition post-conversion, a brief analysis of the impact of academy school conversion on pupil intake is first presented. For some of this analysis, intake is measured in terms of ability composition by the end

17. In fact, some commentators have identified CTCs as the precursors of academies (see West and Bailey 2013). Almost all CTCs took up the opportunity to become academies when it arose with the introduction of academy schools. They were already highly autonomous schools already, being able to not fully follow the national curriculum, to run their own admissions, and not being maintained by the local authority. One can argue that the autonomy gains they experienced from academy conversion were negligible, unlike for the state maintained schools that converted studied in this paper. The working version of this paper (Eyles and Machin 2015) contains results that include city technology colleges—the results are largely unchanged by their omission.

18. A test of joint significance was carried out by collapsing the data to school level and running a probit model of treatment status on all of the variables. The null hypothesis that the variables were jointly insignificantly in predicting treatment status could not be rejected (the p-value testing joint insignificance was 0.153).

of primary school standardised KS2 average points score[19] of pupils who enrol into grade 7, the first year of secondary school. Alongside this ethnicity, free school meal status, and gender of the incoming cohort are also considered.

For the main analysis—the impact of academies on pupil performance—the outcome of interest is the KS4 performance of pupils, measured for the majority of the analysis as the standardised best 8 exams points score of individual grade 11 students.[20] How robust the findings are to the use of different measures of pupil performance is also considered.

In terms of the timing of academy conversion, an academy is designated as starting for the first whole school year when it has academy status. One can define c as the number of academic years before or after conversion. In the intake analysis, the first treated individuals ($c = 0$) are those entering the academy in the September that it opens for business. For the performance sample, $c = 0$ refers to those sitting their GCSEs in the school in the following May/June, that is, those that sit their exams in the school's first academic year as an academy. Limiting the sample to pupils in schools that either convert or are set to convert after the sample period enables implementation of the treatment-control comparison across conversion cohorts that is described in what follows.

3.4. Research Design

The main empirical question of interest is the impact of becoming an academy on end of secondary school examination performance. To clarify the research design it is useful to first introduce some notation. Let t denote academic year, which runs from 2001/2002–2008/2009, $g(i, t)$ denote the grade in which pupil i is enrolled in year t (this takes values 7–11 for secondary school pupils), and let $j(i, t)$ denote the school in which pupil i is enrolled in year t. The year of conversion for school j can be defined as $\mathrm{CY_j}$. Finally, conversion cohorts are sets of schools—S_t—that convert in the same academic year t.

The focus on legacy enrolled pupils in the main analysis is initially justified by showing that the pupil composition of academy schools changed post conversion. This involved looking at treatment-control differences in several intake measures for the population of grade 7 pupils entering schools over the 2001/2002–2008/2009 period. In the following equation, for an intake measure for pupil i, who enrolled in grade 7 at school j in year t, the key parameter of interest is the differences-in-differences coefficient δ_1:

$$\mathrm{Intake}_{it} = \alpha_j + \delta_1 \mathrm{Academy}_{it} + T_t + u_{1it} \quad (1)$$

19. This is calculated by totalling (for each pupil) their raw scores in English, Maths, and Science, then averaging across the three before standardising to have mean zero and standard deviation one.

20. The precise measures used for KS2 and KS4 are described in detail in the Data Appendix, together with additional performance results for a range of different KS4 measures.

In (1), α_j and T_t, respectively, denote school and year fixed effects and u_{1it} is an error term. The key independent variable in the regression is defined as follows:

$$\text{Academy}_{it} = \begin{cases} 1 & \text{if } t \geq \text{CY}_{j(i,t)} \\ 0 & \text{if } t < \text{CY}_{j(i,t)} \end{cases} \quad (2)$$

The estimates of (1) reported in what follows show that pupil intake did change after academisation. This change in composition means it would be misleading to study pupil performance effects for the children newly enrolling post-conversion. This is dealt with by using legacy enrolment as an instrument for academy attendance. The approach has similarities to that taken in Abdulkadiroglu et al. (2016), who study school takeovers in New Orleans, referring to pupils who stay in a converting school as "grand-fathered" pupils.

In the main pupil performance analysis, the focus is placed on grade 11 pupils in the school years 2001/2002–2008/2009. For these pupils the legacy enrolment instrument—Z_{it}—is defined as

$$Z_{it} = \begin{cases} 1 & \text{if } j\,(i, t-1) \in S_t \text{ and } 11 > g\,(i, t-1) \geq 7 \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

In other words, pupils are legacy enrolled if they are enrolled in an academy in the year prior to conversion and are not in their final year of compulsory schooling (grade 11).

The empirical research design to study performance effects can now be operationalised by means of the following value added equation, where KS4 is end of secondary school examinations performance and KS2 is end of primary school performance for pupil i in year t

$$\text{KS4}_{it} = \beta_j + \theta_1 \text{Academy}_{it} + \pi_1 \text{X}_{it} + \varphi_1 \text{KS2}_{it} + T_t + v_{1it} \quad (4)$$

In (2), X is a vector of control variables and v_1 is an error term. Importantly, β_j is a school fixed effect measured at the date of legacy enrolment, rather than the date when KS4 exams are sat. As there are multiple conversion cohorts—from 2002/2003–2008/2009—equation (4) is estimated separately for each conversion cohort, each time using pupils in schools that convert out of sample—in 2009/2010 and 2010/2011—as a control group. Estimates are pooled together across conversion cohorts in the presentation of the results.[21]

An ordinary least squares estimate of θ_1 from (4) may not reflect a causal estimate if individuals sort into academies post-conversion in a nonrandom way. Selection into

21. As the estimates are pooled, the same pupils are used multiple times as controls. In almost all cases, multiple observations of control pupils occur within the same school. Standard errors are therefore clustered at school level.

and out of treatment is accounted for as follows:

(i) From the point of conversion onwards, focus is placed on a fixed set of pupils who are legacy enrolled in the school; therefore, for conversion cohort S_t, the focus is on grade 11 pupils sitting their exams in schools $j \in S_t$ before academic year t (who form the "before" group of pupils in the difference-in-differences analysis) alongside pupils that are legacy enrolled in these schools in $t-1$ (who form the "after" group of pupils in the difference-in-differences analysis). Pupils enrolling in schools $j \in S_t$ after, or in, year t are removed from the sample.

(ii) To make the treatment and control groups as comparable as possible, the same restrictions applied to the treatment group are applied to the control group; accordingly, when considering conversion cohort S_t, the control group comprises grade 11 pupils who sit their exams in control schools prior to t, and those enrolled in a control school in $t-1$, but who are not in their final year of compulsory schooling.[22] The composition of the treatment and control groups is shown in Table A.2.

(iii) The legacy enrolment variable—Z_{it}—is used as an instrument for Academy_{it}, to estimate a local average treatment effect (LATE). Because pupils selecting into schools after the point of conversion are removed from the sample, the LATE estimate corrects for the fact that not all legacy enrolled pupils remain in the academy until grade 11.

Formally, the local average treatment effect is obtained from estimates of the following two reduced forms:

$$\text{Academy}_{it} = \beta_j + \theta_2 \text{Z}_{it} + \pi_2 \text{X}_{it} + \varphi_2 \text{KS2}_{it} + T_t + \text{v}_{2it} \tag{5}$$

$$\text{KS4}_{it} = \beta_j + \theta_3 Z_{it} + \pi_3 X_{it} + \varphi_3 \text{KS2}_{it} + T_t + \text{v}_{3it} \tag{6}$$

In the first stage, equation (5), estimates of θ_2 measure the proportion of the legacy enrolled pupils that stay in the academy and take KS4 exams there. Equation (6) is the reduced form regression of KS4 on the instrument. The instrumental variable (IV) LATE estimate is then the ratio of the reduced form coefficient to the first stage coefficient, θ_3/θ_2.

The specifications adopted so far impose an average post-conversion effect across all post-conversion years. A more flexible specification estimates separate treatment effects for pre- and post-conversion years in an event study setting. The IV setting already described can be extended to an event study framework where separate

22. The rationale for restricting the control group in a similar fashion to the treatment group is to avoid conflating estimates with the effect of mobility on test scores. If we were to take as a control group all grade 11 students in 2009/2010 and 2010/2011 converters then, for conversion cohort S_t, control group pupils observed after t would be more mobile than legacy enrolled pupils observed in the same academic years. This restriction also allows harmonisation of the school fixed effects—β_j—across treatment and control schools. For conversion cohort S_t, β_j corresponds to the $t-1$ school for those sitting their grade 11 exams after, or in, t. For those sitting their KS4 exams in, or prior to, $t-1$, β_j corresponds to the grade 11 school.

TABLE 3. Pupil intake, various measures, enrolled in Grade 7, 2001/2002 to 2008/2009.

	End of primary KS2 test score (1)	Free school meals (2)	White ethnicity (3)	Male (4)
Enrols in Academy in Grade 7	0.093***	−0.033***	−0.015	−0.008
	(0.023)	(0.009)	(0.014)	(0.010)
School Fixed Effects	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Sample Size	1,321,157	1,321,157	1,321,157	1,321,157
Number of Treatment and Control Schools	208	208	208	208

Notes: Robust standard errors (clustered at the school level) are reported in parentheses. The outcome in column (1) (the End of Primary KS2 Test Score) is calculated by totalling (for each pupil) their raw scores in English, Math, and Science. It is then averaged across the three before standardising to have mean zero and standard deviation one—see the Data Appendix for precise definitions. ***Significant at 1%.

estimates are obtained for each of the four post-conversion years ($c = 0$ to $c = 3$) using four separate instruments, which are equivalent to dummies for the enrolment grade of a pupil (7–10) in the pre-conversion year. The event study structural form comparable to equation (4) becomes

$$\text{KS4}_{it} = \beta_j + \sum_{c=-4}^{3} \theta_{4c} \text{Academy}_{it} \times 1\left[\text{CY}_{j(i,t)} = t - c\right] + \varphi_4 \text{KS2}_{it} + \pi_4 \text{X}_{it} + T_t + \text{v}_{4it} \quad (7)$$

In (5), $1[\text{CY}_{j(\text{i,t})} = t - c]$ is an indicator for whether pupil i takes their KS4 exams in a treatment school c years before/after conversion, with conversion taking place at $c = 0$. Therefore, event study estimates of four pre-conversion θ_4s (from $c = -4$ to $c = -1$) and four conversion year and post-conversion θ_4s (from $c = 0$ to $c = 3$) can be obtained. The former are informative about differential pre-conversion trends. All time periods 5 or more years before conversion comprise the omitted category that is set to zero.

4. Empirical Results

4.1. Academies and Pupil Intake

Table 3 shows results from the analysis of changing composition for grade 7 enrolees at the start of secondary school. The table reports differences-in-differences estimates based on equation (1), with the following four dependent variables: the end of primary school KS2 test score, and dummy variable indicators of free school meal status, being of White ethnic origin, and being male. In each case, and in all that follows, standard errors are clustered at the school level.

The estimated coefficients in Table 4 show that academies, post-conversion, did alter their intake in a number of dimensions. They were less likely to admit free school meal eligible pupils, and they admitted pupils with significantly higher KS2 scores. Column (1) shows that, on average, pupils enrolling in an academy at grade 7 have a KS2 mean points score that is 0.09σ higher than those attending schools yet to attain academy status. Column (2) shows a 3.3 percentage point fall in the number of free school meal pupils entering academies post conversion. By contrast, the gender and ethnic composition of their intake appear unchanged by a school becoming an academy.[23]

These results are shown simply to make clear the need to study the legacy enrolment cohorts in the main pupil performance analysis. They indicate that the composition of newly enrolled children, beginning their secondary school years, did change differentially in treatment versus control schools before and after academisation. Hence, for the pupil performance analysis that comes next, to avoid biases from changing composition it is necessary to focus on legacy enrolled pupils.

4.2. Academies and Pupil Performance

The first set of results from the analysis of the main question of interest—the impact of academies on pupil performance—are reported in Table 4. It shows OLS, reduced form and IV estimates of the impact of academy conversion on end of secondary school Key Stage 4 pupil performance for grade 11 children. Columns (1)–(3) show estimates of the impact of academy conversion on pupil performance from specifications without control variables. Columns (4)–(6) show estimates from value added specifications that net out end of primary school pupil performance and other pupil characteristics. Columns (1)–(3) show that being in an academy school increases pupil's KS4 test scores by a statistically significant 0.12σ to 0.20σ. Adding the prior achievement measure (KS2) and control variables in columns (4)–(6) reduces this by a very small amount to a range of 0.12σ to 0.18σ, with all estimates remaining strongly significant.

The interpretation of the legacy enrolment estimate in column (5) is of a 0.11σ higher KS4 score for children enrolled in a pre-conversion school as compared to children enrolled in control schools in the same school years. The IV estimate in column (6) corrects for the fact that not all legacy enrolled children sat their KS4 examinations in the school. In fact the vast majority—93.8%[24]—did as the highly significant first stage at the bottom of the table shows. Because of the high rate of

23. It is noteworthy that, while academies gain freedom to handle their own admissions, they remain subject to the same statutory rules as other state schools and operate under a common admissions regime. In particular, unless they are oversubscribed they must admit all children who apply and—in the case of oversubscription—cannot discriminate on any of the outcomes studied. These intake results are therefore indicative of a post-conversion change in preferences of the local community, rather than a change in recruitment practices of the schools.

24. The implied degree of pupil mobility in the secondary school years from this 93.8% (or 6.2% moving) lines up well with pupil mobility numbers for English schools described in Machin, Telhaj, and Wilson (2006).

TABLE 4. Pupil performance, end of secondary school test scores, Grade 11, 2001/2002 to 2008/2009.

	Standardised end of secondary KS4 test scores					
	OLS (1)	Legacy enrolment (2)	IV (3)	OLS (4)	Legacy enrolment (5)	IV (6)
Takes KS4 in Academy	0.195***	0.111***	0.118***	0.182***	0.108***	0.115***
	(0.029)	(0.028)	(0.030)	(0.029)	(0.029)	(0.031)
Standardised End of Primary KS2 Test Score				0.599***	0.599***	0.599***
				(0.005)	(0.005)	(0.005)
Control Variables	No	No	No	Yes	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	1,263,751	1,263,751	1,263,751	1,263,751	1,263,751	1,263,751
Number of Treatment and Control Schools	208	208	208	208	208	208
First Stage Coefficient on Legacy Enrolment			0.938***			0.938***
			(0.003)			(0.003)

Notes: Robust standard errors (clustered at the school level) are reported in parentheses. Control variables included are dummies for whether the pupil is male, the pupil's ethnicity group, and whether they are eligible for free school meals, entered together with end of primary school KS2 test scores and a dummy variable for pupils for whom KS2 data is unavailable. The dependent variable is the standardised best 8 examinations point score of the pupil—see the Data Appendix for precise definitions. ***Significant at 1%.

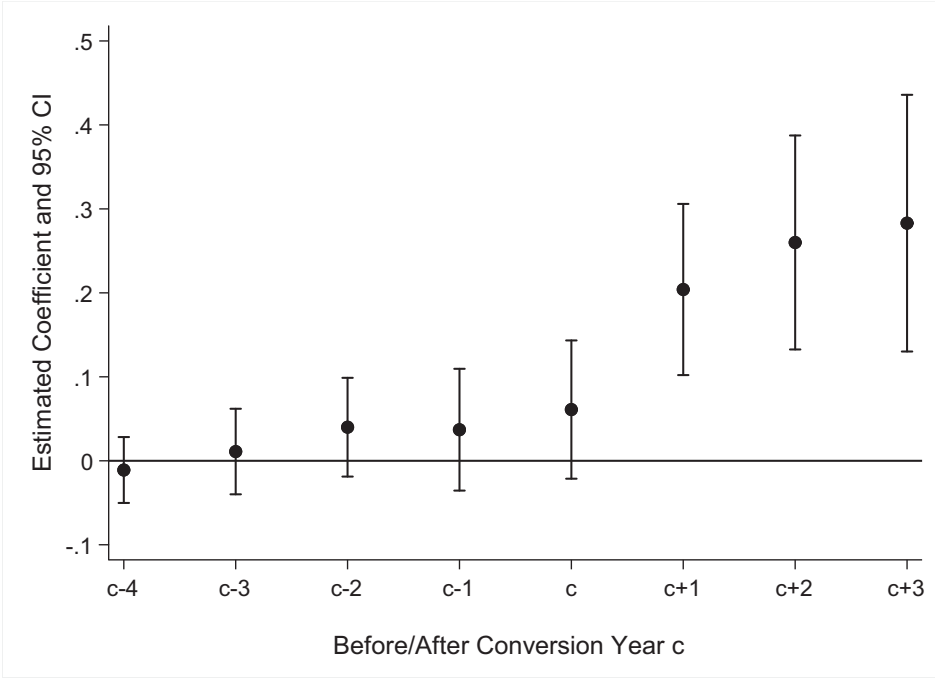

FIGURE 2. Event study instrumental variable estimates of pupil performance and academy conversion, end of secondary school test scores, grade 11 pupils. Event study estimates, from specification in column (1) of Table A.3. The outcome measure is the best 8 capped point score.

compliance, the IV estimate rises only a touch compared to the reduced form estimate, increasing to 0.12σ. This is the preferred baseline average causal estimate of academy conversion.

Aside from the fact that pupil achievement is significantly higher on average for pupils attending schools that converted to an academy, a further point stands out from the results shown in the table—the estimates are similar regardless of estimation method and the set of control variables used. This reflects two aspects of the data; first, the treatment and control pupils are well balanced in terms of covariates, including end of primary school KS2 test scores; and second, there is a high rate of compliance for legacy enrolled students. Because of this, the reduced form and IV estimates broadly align with each other.

Figure 2 plots the event study D-i-D estimates from the IV specification including control variables.[25] Importantly, there are seen to be no discernible pre-treatment trends in the outcome variable. However, there is a significant positive, and rising over time, impact after conversion. Conversion year test scores are 0.06σ higher (though

25. The full set of event study estimates are shown in Table A.3.

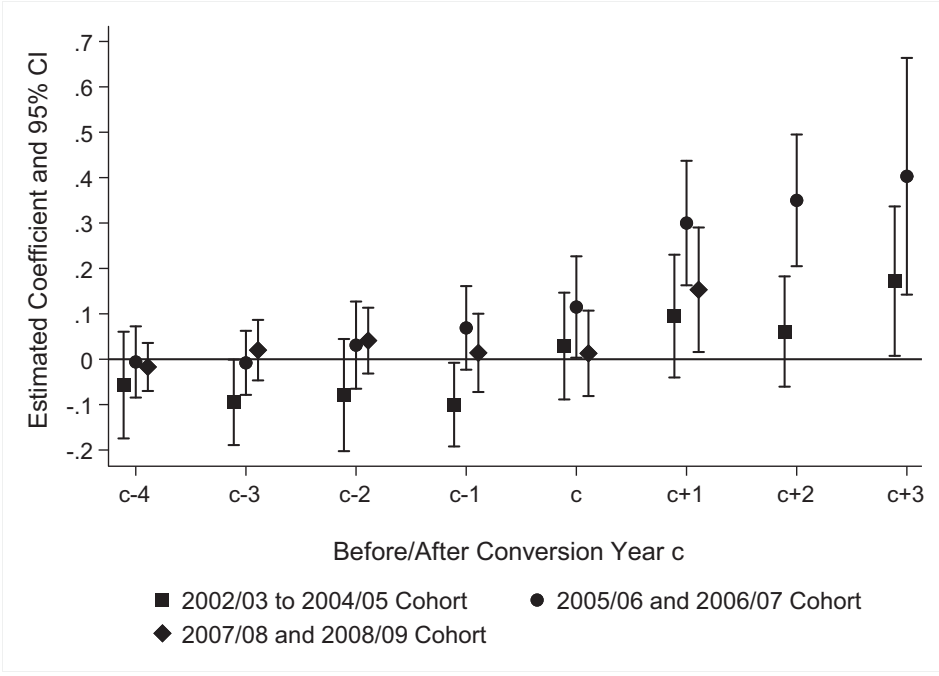

FIGURE 3. Event study instrumental variable estimates of pupil performance and academy conversion, end of secondary school test scores, grade 11 pupils, by groups of academy conversion cohorts. Event study estimates, from cohort specific specifications comparable to column (1) of Table A.3. The outcome measure is the best 8 capped point score.

statistically insignificant) at $c = 0$, the conversion year, and rise to (a statistically significant) 0.28σ four years post-conversion.

The results reported previously are pooled versions of difference-in-difference estimates from different cohorts of academy conversions occurring in the school years 2002/2003 through 2008/2009. Figure 3 plots IV estimates from the models separately by conversion cohort.[26] It is very clear that a null hypothesis of the same average effects across cohorts is not rejected by the data. The gradually rising positive performance effects are seen across the three cohort groups of conversions shown in the figure. Furthermore, the lack of differential pre-treatment trends for all cohorts is highly supportive of the research design that is implemented.

For pupils that attend academies four years after conversion, these impacts of academisation are quite large. To contextualise this, it is worth comparing them with

26. The breakdown by cohort is: 11 conversions in school years 2002/2003 to 2004/2005 (3 from 2002/2003, 6 from 2003/2004, and 2 from 2004/2005), 21 conversions from school years 2005/2006 and 2006/2007 (7 from 2005/2006 and 14 from 2006/2007) and 62 conversions from 2007/2008 and 2008/2009 (25 from 2007/2008 and 37 from 2008/2009). In each case they are compared to the control group of 114 schools that convert after the study sample period ends.

some of the results found in the US charter school literature. Dobbie and Fryer (2013) exploit lottery admission in New York City charters and find gains of $0.13\sigma/0.05\sigma$ in math/ELA tests scores for middle school aged students. The research designs most similar to our own (Fryer 2014; Abdulkadiroglu et al. 2016)—in that they either inject charter practices into pre-existing schools or focus on pupils "grandfathered" into takeover schools—report estimates of between 0.15σ and 0.32σ for math middle and high school students and insignificant to 0.39σ for ELA students. Although these correspond to a slightly different age range, and a broader measure of achievement—to reflect to the nature of KS4 exams—is used, the results fall in line with the high achieving charter school findings.

4.3. Heterogeneous Effects

Although there are positive estimates of performance effects for pupils attending academies, it is possible that substantial heterogeneity is obscured by a focus on average effects. For instance, the charter school literature tends to find positive effects for disadvantaged pupils in urban charters, but little (and sometimes even negative) effects in nonurban settings. Similarly, as noted earlier in the paper, treatment intensity varies in the setting studied in this paper: schools that become academies from community schools gain much greater freedom than those converting from religious schools or foundation schools.

Table 5 shows IV estimates that allow the treatment effect to vary by predecessor school type and whether the school is in an urban area. The estimates in columns (1)–(4) show that although the average effect for academy attendance is positive, effects appear larger for pupils who are pre-enrolled in community schools and those who attend urban academies. The sizeable effect of 0.14σ for community converters and 0.11σ for urban academies contrast with the insignificant coefficients for pupils in nonurban schools and in noncommunity predecessors.

4.4. Robustness: Pre-Trends, Choice of Control Group, and Outcome Measures

Various robustness checks were also undertaken with an intention of testing the sensitivity of the main results to various assumptions. These checks are motivated by the following possible concerns; there might be school specific pre-trends in outcomes; results may depend on the choice of control group; and results may be sensitive to the KS4 outcome used to measure end of secondary school performance.

To address the first concern, specifications that control for a set of school specific pre-conversion linear trends were also estimated. To aid in the precision of these estimates, data going back to 1997/1998 for which there are data on KS4 scores, but no pupil level covariates, was added. For brevity's sake, only the result of the main estimate is reported here: the equivalent of the IV estimate in column (6) of Table 4. Adding school specific trends shifts the main estimate by a small amount up to 0.13σ (0.03), which remains statistically significant. Thus it does not appear to be the case that school specific pre-trends in test scores explain the results.

TABLE 5. Pupil performance, end of secondary school test scores, Grade 11, 2001/2002 to 2008/2009, heterogeneous effects.

	Standardised end of secondary KS4 test scores			
	Pupils in community predecessor school IV (1)	Pupils in noncommunity predecessor school IV (2)	Pupils in urban schools IV (3)	Pupils in nonurban schools IV (4)
Takes KS4 in Academy	0.140***	0.069	0.113***	0.061
	(0.034)	(0.064)	(0.034)	(0.048)
Standardised End of Primary KS2 Test Score	0.603***	0.582**	0.591**	0.625***
	(0.005)	(0.012)	(0.006)	(0.009)
Control Variables	Yes	Yes	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Sample Size	981,249	282,502	974,434	289,317
Number of Treatment and Control Schools	158	50	170	38
First Stage Coefficient on Legacy Enrolment	0.935***	0.945***	0.934***	0.963***
	(0.004)	(0.006)	(0.003)	(0.007)

Notes: As for Table 4.

Second, it is worth noting that although consideration of the control group adds power to the estimates, it is possible to dispense with the control group altogether and rely only on time variation in the receipt of treatment to identify performance effects. Once again, if the main pooled IV estimate is obtained with only the 94 treatment schools results are very similar—the average effect of academy conversion in this case is 0.09σ (0.03).

In terms of different performance measures two sets of additional tests were undertaken. The first looked at other measures of KS4 performance. These results are shown in Table A.4, where the following alternative measures of pupil performance were considered: GCSE Math, GCSE English, and 5 A*-C GCSEs including English and Math. Their use as dependent variable reveals a very similar pattern of results to those using the main KS4 points score.

Finally, a different measure of whether academisation under the labour programme resulted in improved school performance was considered. This looked at Ofsted inspections of schools before and after conversion, again relative to control schools.[27] The probability of transitioning between Ofsted grades was set up as a function of becoming an academy. Transitioning constitutes movements in inspection rankings (of outstanding, good, satisfactory or inadequate) before and after academy conversion for academies in the early and late 2000s and the same for comparison schools. Not all schools were inspected twice in this period, so the analysis is confined to the sub-set of schools that were inspected twice. Ordered probit estimates reported in Table 6 show a statistically significant improvement in inspection rankings of academies. This act as complementary and corroborative evidence in line with the KS4 performance gains already reported.

5. Mechanisms

The previous results uncovered evidence of significant performance improvements for pupils treated by academy conversion. They also showed these improvements to be more pronounced for those attending schools that gained the greatest autonomy. We now address the question—what use of academy freedoms can account for these findings? It has to be acknowledged that the analysis here is somewhat limited in what can be done with available data, but it is possible to offer three main sources of evidence on the question of mechanisms that may be at play. The first from comes from survey data on academies, the second on data on changes in head teachers and teaching staff before and after conversion, the third considering whether peer effects may have played a role.

27. Ofsted is the Office for Standards in Education, Children's Services and Skills which is a government department of Her Majesty's Chief Inspector of Schools in England which undertakes inspections of schools as part of the strongly enforced school accountability system that operates in England.

TABLE 6. Ordered probit estimates of change in Ofsted ranking, school level.

	Pr[Change in Ofsted Ranking]	
	(1)	(2)
Academy	0.865***	0.825***
	(0.241)	(0.241)
Control Variables	No	Yes
Number of Treatment and Control Schools	155	155
Marginal Effects:		
Pr[Change = 2‖Treatment = 1] − Pr[Change = 2|Treatment = 0]	0.329***	0.314**
	(0.091)	(0.102)
Pr[Change = 1‖Treatment = 1] − Pr[Change = 1|Treatment = 0]	−0.098**	−0.092*
	(0.048)	(0.047)
Pr[Change = 0‖Treatment = 1] − Pr[Change = 0|Treatment = 0]	−0.231***	−0.223***
	(0.049)	(0.050)

Notes: Ofsted is a nonministerial department of the government that inspects English schools and gives them an overall effectiveness rating (on a four point scale) based upon, amongst other things, teaching quality, leadership effectiveness, pupil outcomes, and personal development. A more thorough discussion of Ofsted is given in Section A.5. The dependent variable is coded as 0 for a reduction in Ofsted rating, 1 for no change, and 2 for an improvement. Robust standard errors in parentheses. The control variables included in specification (2) are proportion male, proportion White, and proportion of pupils eligible for free school meals—all of which are measured in the year of first inspection. The above is estimated on a subsample of treated schools for whom an Ofsted rating is observed before and after conversion; for control schools, all Ofsted inspections over the period 2000/2001 to 2009/2010 are used. Year of inspection dummies are included in all specifications. *Significant at 10%; **significant at 5%; ***significant at 1%.

5.1. Academies Survey

To begin this discussion of mechanisms, we first draw on the Department for Education's (2014) survey of academy schools "Do Academies Make Use of Their Autonomy?" This survey collected information on a wide array of changes that may have occurred following conversion.[28] These are summarised in Table 7 for 20 (and 3 comparable schools on which there is not full data) of the labour academies analysed in this paper, and for 148 academies (including the 23 labour academies) overall.

Table 7 ranks the responses in order of the percent making the particular change considered in the survey. The three most prominent changes, amongst the 23 converters in the sample, were "changed school leadership", "procured services that were previously provided by the local authority", and "changed the curriculum you offer". Over 75% of the schools said they made these changes pursuant to gaining the new academy freedoms. This ranking is broadly consistent with that of the 148 sponsored academies overall.

28. In May 2013 the Department for Education sent a questionnaire to all 2919 open academies. Of the 720 respondents, 148 were sponsored academies, with 74 of these being secondary schools. Of the 74, 23 converted pre-May 2010 and thus were academies at some point in the sample period.

TABLE 7. Department of education survey of changes after academy conversion, 23 labour academies, and 148 sponsored academies.

	23 Labour academies	148 Sponsored academies including the 23 labour academies	148 Sponsored academies including the 23 labour academies	148 Sponsored academies including the 23 labour academies
	% Making change	% Making change	% Say most important change	% Making change say linked to improved attainment
Changed school leadership	87	72	56	73
Procured services that were previously provided by the LA	78	83	5	17
Changed the curriculum you offer	74	61	26	77
Changed the performance management system for teachers	74	70	3	39
Collaborated with other schools in more formalised partnerships	70	68	8	45
Introduced savings in back-office functions	70	55	0	12
Added nonteaching positions	70	50	3	31
Reconstituted your governing body	65	76	0	26
Changed your pattern of capital expenditure	65	54	1	19
Increased the number of pupils on roll	61	41	0	12
Hired teachers without qualified teacher status (QTS)	48	24	0	14
Introduced or increased revenue-generating activities	48	34	0	8
Changed your admission criteria	43	20	0	7
Increased the length of the school day	39	18	0	63
Changed staff pay structures	30	24	0	9
Sought to attract pupils from a different geographical area	13	12	0	11
Changed the length of school terms	9	6	0	22
Reduced the number of pupils on roll	4	3	0	0

Notes: Taken from Department for Education (2014). The 23 labour academies comprise of 20 schools in the sample and 3 schools that are excluded due to incomplete data.

TABLE 8. Change in staff and pupils before and after academy conversion.

	Change in head teacher (1)	Log(Number of teachers) (2)	Log(Number of pupils) (3)	Log(Teachers per pupil) (4)
Academy	0.214***	0.067*	0.045	0.022
	(0.047)	(0.039)	(0.035)	(0.018)
School Fixed Effects	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes
Sample Size	1641	1641	1641	1641
Number of Treatment and Control Schools	208	208	208	208

Notes: Robust standard errors (clustered at the school level) are reported in parentheses. *Significant at 10%; ***significant at 1%. Columns (2)–(4) control for whether the schools, in each year, enrol pupils prior to grade 7 and post grade 11.

When asked what the most important change was, two answers dominate—"changed school leadership" (at 56%) and "changed the curriculum you offer" (at 26%). Furthermore, both of these were reported to be linked to improved outcomes (in 73% and 77% of cases respectively). Other changes that were notably linked to improved outcomes were "Increased the length of the school day" (63%) and "Collaborated with other schools in more formalised partnerships" (45%).

5.2. Changes in Head Teacher and Teaching Staff

It is also possible to look at statistical difference-in-differences estimates at school-level for three of the important factors identified in Table 8: whether a new head teacher is taken on upon conversion; whether more pupils are enrolled; and whether more teachers are taken on. This is facilitated by the availability of school level data over time on each of these.

Column (1) of Table 8 reports results for head teacher change. There is significant head teacher turnover before and after conversion to an academy. Over all post conversion years this averages out to 21% more head teacher turnover in academies. Further probing makes it clear that this substantial degree of turnover is very much concentrated in the conversion year. In treatment schools, 45% more head teacher turnover occurred in the year of conversion c as compared to the control schools. This is shown in event study estimates equivalent to Table 8 in Table A.5. It seems to be a one off change that occurred as the subsequent year treatment effects from ($c = 1$) to ($c = 3$) were all insignificantly different from zero.

Thus a strong feature of academy conversions is to replace the head teacher. There is a more modest turnaround of the rank and file teaching staff and, if there is an increase, it seems to be due to a need to take on more teachers if more pupils enrol in academies post conversion. This can be gleaned from looking at columns (2) to (4) in Table 8. The table shows D-i-D estimates of the effect of academy conversion on

the number of teachers, number of pupils and the teacher–pupil ratio. Looking first at column (2), there is evidence that the number of teachers rose once the school gained academy status. This is because, as shown in column (3), more pupils were enrolled once the academies were up and running, although this change itself is statistically insignificant. Finally, column (4) shows that the increased number of teachers went hand in hand with increases in pupil enrolments as the teacher–pupil ratio did not rise significantly post-conversion. Overall, the table shows far less clear evidence of post-conversion teacher turnover, certainly when compared to the very significant evidence of head teacher turnover that occurred upon academy conversion.[29]

5.3. Peer Effects

So far changes in staff, and changes enacted by management, have been considered as potential sources of performance improvements. A final potential mechanism is whether an increase in peer quality resulted from academy conversion. Earlier analysis has already shown that academy conversion resulted in an upward shift in the test scores of grade 7 entrants into academies. It also resulted in a downward shift in the proportion of free school meal eligible pupils entering the schools.

As well as entrants into grade 7, schools can also enrol students in other grades; therefore, new students do enrol in the same grades as legacy enrolled pupils in the years following conversion. There are not that many such students, but a comparison of the attributes of those joining the same grades as the legacy enrolled pupils reveals them to be similar in terms of gender and FSM eligibility. However, they are more likely to be non-White and, if anything, have lower prior test scores.[30]

Table 9 presents a more formal test of whether peer effects that could result from this post-conversion entry display a connection with the observed performance improvements. It does so by estimating a specification comparable to column (6) of Table 4, but with inclusion of an additional interaction between the treatment variable and the average standardised KS2 score of incoming pupils.[31] These averages can be separately calculated for two sets of peers, namely, for incoming pupils into the same grade as the legacy enrolled pupils (as one would expect peer effects to operate

29. Moving beyond this, we also re-estimated the main IV pupil performance specification with an interaction of treatment with the extent of teacher turnover. We found no evidence that schools who exhibited large changes in the numbers of pupils, teachers, or the ratio of the two generate larger gains for their pupils. The same exercise, but with an interaction for head teacher change, shows that those who do change head teacher generate slightly higher effects, but that the difference was not statistically significant. Of course, we cannot estimate the contribution of each mechanism separately without having separate "mechanism experiments".

30. It is also worth noting that inflows of pupils in grades other than grade 7 are small. In the sample, of all those with a KS4 record in an academy post conversion, fewer than 10% were not legacy enrolled.

31. To be precise, treatment status is interacted with measures of compositional changes in the treatment school and instrumented by legacy enrolment status interacted with the same measure of compositional change measured at the legacy enrolment school.

TABLE 9. Change in peer quality after academy conversion.

	Entrants into same grade (1)	Entrants into all grades (2)
Academy	0.126***	0.124*
	(0.035)	(0.064)
Academy × Average KS2 of New Entrants Post Conversion	0.030	0.027
	(0.072)	(0.134)
School Fixed Effects	Yes	Yes
Year Dummies	Yes	Yes
Sample Size	1,263,751	1,263,751
Number of Treatment and Control Schools	208	208

Notes: Robust standard errors (clustered at the school level) are reported in parentheses. The estimates here are equivalent to those in column (6) of Table 4. *Significant at 10%; ***significant at 1%.

primarily within grade), and for all pupils entering the schools post conversion (which would require cross-grade peer effects to operate).

Peer effects do not seem to account for the main pattern of results. As can be seen from columns (2) and (4) of Table 9, the interactions between treatment status and KS2 scores are small and insignificant. This is irrespective of whether or not KS2 scores are averaged over all post-conversion enrolees or only those that sit their grade 11 exams with the legacy enrolled pupils. These empirical tests imply that the estimated performance improvements do not seem to come about because performance effects and changes in peer composition are related.

6. Conclusions

Whether new school types can potentially alleviate poor education standards has become a question of significant interest to educators, policymakers and parents. This paper focuses on a school reform that has become a high profile in this regard—the introduction of academy schools to the English secondary school sector. The impact of academy school conversion on pupil performance is studied, using a legacy enrolment methodology free of bias from changing pupil composition. Academy conversion is seen to generate significant improvements in pupil performance for those who attended schools treated by academy conversion.

Transformation to an academy raised end of secondary school educational outcomes by 0.12σ on average, and by more for children receiving more years of treatment, rising to 0.28σ three years post-conversion. These findings complement existing work from different settings like that on US charter schools (both newly set up and more closely to takeovers of public schools) on whether different school types can affect pupil performance. They also add significantly to this literature as many of

the best identified studies of US charters are often focussed on a single city or state setting. The national scope of the effort studied in this paper makes the findings of the paper less likely to be driven by context-specific factors than some of that research. As well as finding larger gains for pupils who spend more time in an academy, the paper also reports suggestive evidence that schools in urban areas, and those that gain the most autonomy from conversion, are the most likely to benefit from the program. Finally, there is not any evidence that improvements in peer composition drive the results, suggesting that the programme effects can, at least to some extent, be scaled up.

Before finishing, it is appropriate to place these findings into their policy context, especially given the large and rapid education reforms that have occurred recently in England. This paper studies the sponsored academies set up under the labour government's programme, which had 203 academies up and running in May 2010 when a new coalition government was voted in. Since then, the academies programme has been massively expanded and taken on a new direction, with the number of conversions skyrocketing and with new convertors not only being in the secondary sector, but also covering primary schools, and even reaching outside the state sector to some private schools. Moreover, the new coalition academies need not have a sponsor when they are converted. In the 2010s, mass academisation has become the order of the day in English education.

A key feature distinguishing these new coalition academies is that, on average, they are not characterised by poor performance and disadvantage in their predecessor state like the sponsored academies introduced and approved under the previous labour government analysed in this paper.[32] The way some of them are run is also different with, for example, some of the post May 2010 academies being run as chains of schools by major sponsors. It will be an important future research challenge to determine whether or not these new converter and chain run academies are able to deliver the kinds of performance improvements for students enrolling in them that the labour programme analysed here seemed to deliver.

Appendix: Data Construction and Additional Results

A.1. Data on Academy Schools

We first identified all schools that became academies over the school years 2002/2003 to 2010/2011. Sources for this are Department for Education extracts that give information on all academies that have opened or are in the process of doing so. The extract gives the opening date of the academy, its URN (a unique identifier for the school allowing us to identify it in various governmental data sources such as the National Pupil Database

32. See Eyles, Machin, and Silva (2018) for an empirical analysis of the different nature of pre- and post-May 2010 academies.

TABLE A.1. The nature of academy conversions.

	All schools							
	Pre-academy school type							
	All	New	Independent	City technology college	Voluntary aided	Foundation	Voluntary controlled	Community
All academies	244	12	5	12	18	34	2	161
All schools with full data (pre- and post-academy conversion)								
All academies	220	0	0	12	15	33	2	158
Become academies, up to 2008/2009	106	0	0	12	10	15	1	68
Future academies, after 2008/2009	114	0	0	0	5	18	1	90

Notes: Source for upper panel, same as Table 2. Source for lower panel, own calculations from Edubase, School Performance Tables, and Annual Schools Census.

TABLE A.2. Sample composition in difference-in-differences analysis, conversion cohort S_t.

	Before	After
Treatment (conversion cohort S_t)	Grade 11 pupils in $t-1$ and prior in schools $j \in S_t$.	Pupils legacy enrolled in grades 7–10 in $t-1$ in schools $j \in S_t$.
Control (for conversion cohort S_t)	Grade 11 pupils in $t-1$ and prior in 2009/2010 and 2010/2011 converters.	Pupils enrolled in grades 7–10 in $t-1$ in 2009/2010 and 2010/2011 converters.

and the Pupil Level Annual Census data), DFE number (a second unique identifier combining school specific and local authority specific numbers), and the URN number of the predecessor school.

Using performance tables data from the Department for Education (DfE) we match in predecessor school types. This gives 244 schools that became academies between the first 3 academy openings in 2002/2003 and those that gained academy status by September 2010 (the beginning of the academic school year). Previously independent schools were omitted due to pupils in these schools not having exam information at KS4. Similarly, brand new academy schools were omitted as they have no predecessor school.

In order to have a balanced panel we focus on academies that have some form of predecessor school open from at least 1997 onwards. Any later and the school will not have KS4 results for 2002. In order for the sample to be balanced for intake we exclude academies who do not enrol pupils in grade 7. The final sample contains 106 (of which we use the 94 who were not CTCs in their predecessor state) treatment schools (those that opened as academies prior to, or in, September 2008) and 114 control schools with observations ranging over the years 2001/2002–2008/2009. None of the control schools become academies during these sample years, but convert by September 2010.

A.2. Pupil Level Data

We use data from PLASC (pupil level annual schools census) and the NPD (national pupil database). The NPD contains information on all key stage 2 (KS2) and key stage 4 (KS4) exams sat at the end of primary and secondary school respectively. Each pupil is identified by a unique reference number and the data gives the unique URN of the school in which they sat the exam. Although the NPD reports on pupils in examination years PLASC has a record for every pupil for each year that they are in the maintained school sector. PLASC data gives the pupil, grade and school as well as demographic variables such as ethnicity, gender, and free school meal eligibility. We can track pupils through secondary school using the unique pupil identifier. This identifier is common to the NPD enabling us to merge NPD and PLASC data. This

gives a panel of pupils with their demographic information, their KS2 and KS4 test results and the school(s) that they attended from grade 7 (first grade of compulsory secondary education) through to grade 11 (final grade of compulsory education). We then extract those pupils who attended the 208 treatment and control schools at some point over the sample period. We can now see which schools pupils attended in every secondary compulsory year of schooling,[33] their demographic information and their exams results at KS4 and KS2. The intake analysis focuses on those who enter as a grade 7 student in 2001/2002–2008/2009 whereas the results analysis focuses on those who sit exams, are legacy in one of the 94 treatment schools, or sit exams in one of the 114 control schools over the same period.

Finally it is worth noting that PLASC does not cover years prior to 2002. For the observations before then we do still have NPD data on KS2 and KS4 performance (we have these going back to 1997 for KS4 and 1996 for KS2). Therefore, in the cohort and school specific trend estimates we include school fixed effects and time effects, but no covariates.

A.3. Clustering

A final note relates to how we define "school". For each of the treatment and control schools we assign a unique number. It is possible that two pupils from different schools are given the same number should the two differing schools later become the same academy. We identify when schools merge by looking at linked schools in edubase (this is a Department for Education database of all open and closed maintained schools in England). In one case a single school becomes two separate academies (North Westminster Community School splits into Paddington Academy and Westminster Academy in 2006). Pupils attending the predecessor school are randomly assigned one of the two numbers given to the two academies that open later. In estimated specifications, standard errors are clustered on this unique school number resulting in 208 clusters.

A.4. Attainment Measures

The main variable in the analysis of intake is an average score across three subject specific key stage 2 tests: English, Maths, and Science. Test scores are reported in two ways: first, a level from 2–5 is awarded in each subject and second as a raw test score. The raw test score is graded out of 80 for science and is the sum of two separate science papers each marked out of 40 whereas the English test score is marked out of 100 and is composed of the sum of two separate test scores, each marked out of 50, in reading and writing. Finally Maths is composed of two marks out of 50 with one of the tests being in mental arithmetic. The levels are based upon these underlying test scores but

33. Strictly speaking this is not true. Some pupils enter the schooling system either from another country or from independent schools. We observe when the pupils enter but not precisely where they came from. These pupils are retained in the analysis.

TABLE A.3. Event study estimates, pupil performance, end of secondary school test scores, Year 11, 2001/2002 to 2008/2009.

	Standardised end of secondary KS4 test scores				
	All IV (1)	Pupils in community predecessor school IV (2)	Pupils in noncommunity predecessor school IV (3)	Pupils in urban schools IV (4)	Pupils in nonurban schools IV (5)
Takes KS4 in Academy × ($c = -4$)	−0.011 (0.020)	−0.023 (0.022)	0.024 (0.037)	−0.012 (0.023)	0.008 (0.027)
Takes KS4 in Academy × ($c = -3$)	0.011 (0.026)	0.034 (0.028)	−0.026 (0.052)	0.015 (0.029)	0.005 (0.049)
Takes KS4 in Academy × ($c = -2$)	0.040 (0.030)	0.062* (0.033)	0.011 (0.059)	0.049 (0.034)	−0.018 (0.063)
Takes KS4 in Academy × ($c = -1$)	0.037 (0.037)	0.066 (0.040)	0.002 (0.076)	0.044 (0.041)	−0.016 (0.090)
Takes KS4 in Academy × ($c = 0$)	0.061 (0.042)	0.099* (0.045)	0.003 (0.086)	0.060 (0.048)	0.028 (0.074)
Takes KS4 in Academy × ($c = 1$)	0.204**(0.052)	0.250** (0.060)	0.135 (0.098)	0.207** (0.058)	0.114 (0.102)
Takes KS4 in Academy × ($c = 2$)	0.260** (0.065)	0.281** (0.072)	0.264* (0.144)	0.255** (0.070)	−0.024 (0.091)
Takes KS4 in Academy × ($c = 3$)	0.283** (0.078)	0.317** (0.085)	0.235 (0.165)	0.266** (0.086)	0.239* (0.101)
Standardised End of Primary KS2 Test Score	0.599*** (0.005)	0.603*** (0.005)	0.582*** (0.012)	0.591*** (0.006)	0.625*** (0.009)
Control Variables	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes	Yes
Sample Size	1,263,751	981,249	282,502	974,434	289,317
Number of Treatment and Control Schools	208	158	50	170	38
First Stage Coefficient on Legacy Enrolment × ($c = 0$)	0.964*** (0.003)	0.962*** (0.003)	0.967*** (0.005)	0.961*** (0.003)	0.980*** (0.003)
First Stage Coefficient on Legacy Enrolment × ($c = 1$)	0.926*** (0.004)	0.926*** (0.004)	0.927*** (0.009)	0.924*** (0.004)	0.933*** (0.014)
First Stage Coefficient on Legacy Enrolment × ($c = 2$)	0.877*** (0.007)	0.879*** (0.007)	0.870*** (0.020)	0.876*** (0.007)	0.907*** (0.000)
First Stage Coefficient on Legacy Enrolment × ($c = 3$)	0.840*** (0.015)	0.837*** (0.017)	0.851*** (0.031)	0.836*** (0.015)	0.887*** (0.000)

Notes: Robust standard errors (clustered at the school level) are reported in parentheses. Control variables included are dummies for whether the pupil is male, the pupil's ethnicity group, and whether they are eligible for free school meals, entered together with end of primary school KS2 test scores and a dummy variable for pupils for whom KS2 data is unavailable. The dependent variable is the standardised best 8 examinations point score of the pupil—see the Data Appendix for precise definitions. *Significant at 10% level; **significant at 5% level; ***significant the 1% level.

TABLE A.4. Alternative end of secondary school outcomes.

	Alternative end of secondary KS4 outcomes					
	Five A*-C, with English and Math IV (1)	Five A*-C, with English and Math IV event study (2)	English GCSE IV (3)	English GCSE IV event study (4)	Math GCSE IV (5)	Math GCSE IV event study (6)
Takes KS4 in Academy	0.022* (0.011)		0.100** (0.031)		0.074* (0.033)	
Takes KS4 in Academy × ($c = -4$)		−0.003 (0.007)		0.018 (0.019)		0.012 (0.021)
Takes KS4 in Academy × ($c = -3$)		0.001 (0.008)		0.031 (0.021)		0.001 (0.023)
Takes KS4 in Academy × ($c = -2$)		0.007 (0.010)		0.029 (0.033)		0.032 (0.026)
Takes KS4 in Academy × ($c = -1$)		0.016 (0.011)		0.074* (0.034)		0.039 (0.031)
Takes KS4 in Academy × ($c = 0$)		0.014 (0.014)		0.111*** (0.039)		0.051 (0.043)
Takes KS4 in Academy × ($c = 1$)		0.039* (0.017)		0.165*** (0.053)		0.136*** (0.044)
Takes KS4 in Academy × ($c = 2$)		0.040 (0.021)		0.137* (0.064)		0.152*** (0.052)
Takes KS4 in Academy × ($c = 3$)		0.085** (0.023)		0.192* (0.082)		0.220*** (0.055)
Standardised End of Primary KS2 Test Score	0.212*** (0.005)	0.212*** (0.005)	0.581*** (0.006)	0.581*** (0.006)	0.656*** (0.006)	0.656*** (0.006)
Control Variables	Yes	Yes	Yes	Yes	Yes	Yes
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	1,263,751	1,263,751	1,263,751	1,263,751	1,263,751	1,263,751
Number of Treatment and Control Schools	208	208	208	208	208	208
First Stage Coefficient on Legacy Enrolment	0.938*** (0.003)		0.938*** (0.003)		0.938*** (0.003)	
First Stage Coefficient on Legacy Enrolment × ($c = 0$)		0.964*** (0.003)		0.964*** (0.003)		0.964*** (0.003)
First Stage Coefficient on Legacy Enrolment × ($c = 1$)		0.926*** (0.004)		0.926*** (0.004)		0.926*** (0.004)
First Stage Coefficient on Legacy Enrolment × ($c = 2$)		0.877*** (0.007)		0.877*** (0.007)		0.877*** (0.007)
First Stage Coefficient on Legacy Enrolment × ($c = 3$)		0.840*** (0.015)		0.840*** (0.015)		0.840*** (0.015)

Notes: Robust standard errors (clustered at the school level) are reported in parentheses. Control variables included are dummies for whether the pupil is male, the pupil's ethnicity group, and whether they are eligible for free school meals, entered together with end of primary school KS2 test scores and a dummy variable for pupils for whom KS2 data is unavailable. *Significant at 10% level; **significant at 5% level; ***significant at 1% level.

TABLE A.5. Event study estimates, change in staff, and pupils before and after academy conversion.

	Change in head teacher (1)		Log(Number of teachers) (2)		Log(Number of pupils) (3)		Log(Teachers per pupil) (4)	
Academy × ($c = -4$)	0.044	(0.043)	0.014	(0.012)	0.007	(0.011)	0.008	(0.009)
Academy × ($c = -3$)	0.064	(0.057)	0.012	(0.019)	−0.006	(0.016)	0.019	(0.016)
Academy × ($c = -2$)	0.032	(0.066)	0.004	(0.025)	−0.025	(0.024)	0.028	(0.019)
Academy × ($c = -1$)	0.076	(0.085)	0.002	(0.034)	−0.048	(0.033)	0.050**	(0.025)
Academy × ($c = 0$)	0.453***	(0.106)	0.036	(0.046)	−0.035	(0.047)	0.071**	(0.031)
Academy × ($c = 1$)	0.056	(0.113)	0.112*	(0.060)	0.040	(0.058)	0.073*	(0.038)
Academy × ($c = 2$)	0.102	(0.141)	0.147*	(0.078)	0.099	(0.073)	0.048	(0.043)
Academy × ($c = 3$)	0.143	(0.163)	0.277***	(0.077)	0.229***	(0.075)	0.048	(0.060)
School Fixed Effects	Yes		Yes		Yes		Yes	
Year Dummies	Yes		Yes		Yes		Yes	
Sample Size	1641		1641		1641		1641	
Number of Treatment and Control Schools	208		208		208		208	

Notes: Robust standard errors (clustered at the school level) are reported in parentheses. Columns (2)–(4) control for whether the schools, in each year, enrol pupils prior to grade 7 and post grade 11. *Significant at 10%; **significant at 5%; ***significant at 1%.

are not always consistent. For instance, after an initial level is assigned after grading the test there may be a review of the pupil's test score resulting in a higher or lower level being awarded even if the underlying raw test mark is not altered. Similarly the mark required for any one level varies both between subjects and within subjects across years. For these reasons we use standardised raw test scores as the main dependent variable in KS2 regressions.

When pupils are not awarded a test mark or are performing at a level below the level of the test we award pupils a mark of 0. Those who miss the tests are excluded from the sample for the purposes of the KS2 regressions but are included in the KS4 regressions where we include a dummy for those who do not have a KS2 record or who miss KS2 exams. The KS4 results are robust to rerunning the regressions omitting those without a KS2 record and those whose scores are below test levels.

The main KS4 qualification in the UK is the GCSE (General Certificate of Secondary Education). GCSEs are graded from A*-G. The current points score calculations give an A* a score of 58 and a G a score of 16 with grades in between going up in increments of 6 between adjacent grades as follows:

Grade	Points	Grade	Points
A*	58	D	34
A	52	E	28
B	46	F	22
C	40	G	16

New scale.

Prior to this an A* was given a score of 8 and G a score of 1 with scores rising in unit increments.

Grade	Points	Grade	Points
A*	8	D	4
A	7	E	3
B	6	F	2
C	5	G	1

Old scale.

As well as GCSEs there are a wide range of equivalent qualifications focusing on more vocational subjects. These include GNVQs and BTecs. Depending upon the type of equivalent these are often worth multiple GCSEs and are often graded as a combination of GCSE grades, that is, a distinction in an intermediate GNVQ is equivalent to gaining two GCSEs with one at grade A and the other at grade A*.[34]

34. Most equivalents are graded as pass, merit or distinction but the Department for Education equates these categories, combinations of, A*-G grades.

The points score given to the qualification reflects the underlying GCSE grades that it is based upon so that under the new scoring system the aforementioned qualification would be given a score of 110. The points system we use is as follows:

Grade	Points	Grade	Points
A*	10	D	4
A	8	E	3
B	7	F	2
C	6	G	1

Scale used in the paper.

The points system we use addresses some of the concerns expressed pertaining to the 16–58 and 1–8 scales used over the course of the sample.[35] The nonlinearity reflects the fact that it appears hardest to jump from grades D to C and from A to A*.

We cap points scores at best 8 qualifications. To do this we normalize raw point scores by their GCSE equivalent, that is, a qualification worth 4 GSCEs and 208 points (under the 16–58 scale) is normalized to be worth 52 points. We then convert these points to the new measure and rank them highest to lowest. We then add up the grade weightings (in terms of GCSEs), taking fractions of qualifications if need be, until we reach 8. All those in the top 8 are then multiplied through by their weight and summed to give the points score.

The decision to cap at 8 is motivated by two concerns. Total points scores have the problem that pupils can appear to do well by entering many exams and performing poorly in them. Similarly using, for instance, 5 best means that those who focus very narrowly on a small set of exams may appear better than those who perform well over a larger selection of subjects/qualifications. The decision to cap at 8 balances these two concerns.

Finally, it is worth noting that the point measures create some notable discrepancies with the official method. For instance, an equivalent qualification worth two GCSEs graded CD is worth 74 points under the 16–58 scale meaning that it is worth more than a A* at GCSE. Using the system such a qualification is worth 10 points (the sum of the points scores for grades of C and D)—the equivalent of a GCSE at grade A*. A further example is a BTEC that is worth 76 points on the old scale and equivalent to 4 GCSEs. This is the same as achieving grades of 2 Fs and 2 Gs. In the system this is equivalent to a point score of 6. Thus the points mean the qualification is the same as getting a C at GCSE whereas the old measure means that the qualification is again worth more than an A*. In general the system reduces the relative points scores of equivalent qualifications compared to the official method. Despite this the results remain unchanged when using the (standardized) old (1–8) and new (16–58) points systems and when using total

35. We are grateful to Tim Leunig and Mike Treadaway for very helpful correspondence on this.

rather than capped scores. We present results in Table A.4 using different dependent variables.

A.5. Ofsted Reports

Ofsted is the government department that carries out inspections of maintained schools in England and Wales and reports to Parliament.[36] Inspectors give schools minimal prior warning of inspection and proceed to inspect the school based upon a pre-set criteria before awarding the school and overall effectiveness rating.[37] Overall effectiveness is based upon many criteria like pupil achievement, the effectiveness of management and the level of well-being and personal development of the pupils.

Post 2005 there are 4 possible inspection ratings—outstanding, good, satisfactory, and inadequate. Prior to 2005 the possible ratings given were excellent, very good, good, satisfactory, unsatisfactory, poor, and very poor. To measure whether academies improve over time we equate the 7 ratings given prior to 2005 into the 4 categories given post 2005 in the following manner:

Prior to 2005	Post 2005
Excellent, very good	Outstanding
Good	Good
Satisfactory	Satisfactory
Unsatisfactory, poor, very poor	Inadequate

The main interest is whether schools converting to academies are more likely to improve their rating relative to the control schools. To study this question we use Ofsted ratings for the years 2000–2010. We limit the sample to the years 2000–2010 as post 2010 all the schools in the sample have converted to academies making any comparisons between converters and those yet to convert impossible.

For the estimates, we use all inspection outcomes available for control schools. For treatment schools, we use the latest preconversion inspection and the earliest post conversion one. These restrictions results in the sample of treatment schools falling to 45 with the first three cohorts not represented in the sample at all. For controls schools we omit those that only have a single inspection over the period thus reducing the sample of control schools to 110. For this sample we define a variable equal to 0 if the school's first inspection is worse than its last, 1 if the inspections are the same and 2 if the latter inspection is an improvement on the first. As a robustness check we replicate the results using the following two conversions for Ofsted scores:

36. Throughout this and Section 5 of the main text school refers to the variable school that we cluster on as described in Section A.3—all mechanism regressions are performed at this level.

37. Overall effectiveness ratings have been awarded since 2000.

Conversion 1

Prior to 2005	Post 2005
Excellent	Outstanding
Very good, good	Good
Satisfactory	Satisfactory
Unsatisfactory, poor, very poor	Inadequate

Conversion 2

Prior to 2005	Post 2005	New Scale
Excellent, very good	Outstanding	Good
Good	Good	Good
Satisfactory	Satisfactory	Good
Unsatisfactory, poor, very poor	Inadequate	Bad

The results prove robust to these changes.

A.6. Data on Mechanisms

As well as considering Ofsted reports we study mechanisms by looking at survey results from the Department for Education (2014), head teacher change and teacher turnover.

We collect data on head teachers using the School Workforce Census and match a head teacher to each of the schools for each year in the sample. For each year we define a binary variable equal to 1 if this year's head teacher is different from last years. When two schools merge we set this variable to 1 only if the head is not the head of either of the predecessors. When two separate schools are defined as being the same (with respect to the clustering variable) we set this variable to 1 if both schools change their head teacher in that year. We have also defined change as either one of the two schools changing their head—the results are unchanged by this.

For the teacher and pupil analysis we use data from the annual schools census. The data gives us the number of qualified and unqualified teachers in all maintained secondary schools for the years 2002–2009. The measure of teachers is the full time equivalent of both qualified and unqualified teachers, whereas the measure of pupils is full time equivalent pupils. Because schools often open sixth forms upon conversion to academy status, and a few schools merge with schools that enrol children pregrade 7, we include dummies in the regressions for schools that have an attached sixth form, and those that enrol children in grades lower than 7. These latter variables come from maximum and minimum age group variables in the school performance tables.

References

Abdulkadiroglu, Atila, Joshua Angrist, Susan Dynarski, Thomas Kane, and Parag Pathak (2011). "Accountability and Flexibility in Public Schools: Evidence from Boston's Charters and Pilots." *The Quarterly Journal of Economics*, 126, 699–748.

Abdulkadiroglu, Atila, Joshua Angrist, Yusuke Narita, and Parag Pathak (2017). "Research Design Meets Market Design: Using Centralized Assignment for Impact Evaluation." *Econometrica*, 85, 1373–1432.

Abdulkadiroglu, Atila, Joshua Angrist, Peter Hull, and Parag Pathak (2016). "Charters Without Lotteries: Testing Takeovers in New Orleans and Boston." *American Economic Review*, 106(7), 1878–1920.

Adonis, Andrew (2012). *Education, Education, Education: Reforming England's Schools*. Biteback Publishing, London.

Altonji, Joseph, Todd Elder, and Christopher Taber (2005). "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy*, 113, 151–184.

Angrist, Joshua, Sarah Cohodes, Susan Dynarski, Parag Pathak, and Christopher Walters (2016). "Stand and Deliver: Effects of Boston's Charter High Schools on College Preparation, Entry, and Choice." *Journal of Labor Economics*, 34, 275–318.

Angrist, Joshua, Susan Dynarski, Thomas Kane, Parag Pathak, and Christopher Walters (2010). "Inputs and Impacts in Charter Schools: KIPP Lynn." *American Economic Review*, 100(2), 239–243.

Angrist, Joshua, Parag Pathak, and Christopher Walters (2013). "Explaining Charter School Effectiveness." *American Economic Journal: Applied Economics*, 5, 1–27.

Betts, Julian, Lorien Rice, Andrew Zau, Emily Tang, and Cory Koedel (2006). *Does School Choice Work? Effects on Student Integration and Achievement*. Public Policy Institute of California.

Clark, Damon (2009). "The Performance and Competitive Effects of School Autonomy." *Journal of Political Economy*, 117, 745–783.

Center for Research on Education Outcomes (2009). *Multiple Choice: Charter Performance in Sixteen States, Stanford University*. CREDO, Stanford University.

Center for Research on Education Outcomes (2013). *National Charter School Study*. CREDO, Stanford University.

Department for Children, Schools and Families (2007). "Academies and Independent Schools: Prospectus." http://dera.ioe.ac.uk/6578/7/Academies_Prospectus_Redacted.pdf. Retrieved May 21, 2018.

Department for Children, Schools and Families (2008). "The Composition of Schools in England" http://webarchive.nationalarchives.gov.uk/20130321145625/https://media.education.gov.uk/assets/files/pdf/b022008pdf.pdf. Retrieved May 21, 2018.

Department for Education (2014). "Do Academies Make Use of Their Autonomy?" https://www.gov.uk/government/publications/do-academies-make-use-of-their-autonomy. Retrieved May 21, 2018.

Dobbie, Will and Roland Fryer (2011). "Are High Quality Schools Enough to Close the Achievement Gap? Evidence from a Social Experiment in Harlem." *American Economic Journal: Applied*, 3, 158–187.

Dobbie, Will and Roland Fryer (2013). "Getting Beneath the Veil of Effective Schools: Evidence from New York City." *American Economic Journal: Applied*, 5, 58–75.

Dobbie, Will and Roland Fryer (2014). "The Medium Term Impacts of High Achieving Charter Schools." *Journal of Political Economy*, 5, 985–1037.

Epple, Dennis, Richard Romano, and Miguel Urquiola (2017). "School Vouchers: A Survey of the Economics Literature." *Journal of Economic Literature*, 55, 441–492.

Epple, Dennis, Richard Romano, and Ron Zimmer (2016). "Charter Schools: A Survey of Research on their Characteristics and Effectiveness." In *Handbook of the Economics of Education*, Vol. 5, edited by E. Hanushek, S. Machin, and L. Woessmann. Elsevier, Amsterdam.

Eyles, Andrew, Claudia Hupkau, and Stephen Machin (2016). "Academies, Charter and Free Schools: Do New School Types Deliver Better Outcomes?" *Economic Policy*, 31, 453–501.

Eyles, Andrew and Stephen Machin (2015). "The Introduction of Academy Schools to England's Education." Discussion Paper 1368. Centre for Economic Performance, London School of Economics.
Eyles, Andrew, Stephen Machin, and Sandra McNally (2017). "Unexpected School Reform: Academisation of Primary Schools in England." *Journal of Public Economics*, 155, 108–121.
Eyles, Andrew, Stephen Machin, and Olmo Silva (2018). "Academies 2—The New Batch: The Changing Nature of Academy Schools in England." *Fiscal Studies*, 39, 121–158.
Evans, William and Robert Schwab (1995). "Finishing High School and Starting College: Do Catholic Schools Make a Difference?" *The Quarterly Journal of Economics*, 100, 941–974.
Fryer, Roland (2014). "Injecting Charter School Best Practices into Traditional Public Schools: Evidence from Field Experiments." *The Quarterly Journal of Economics*, 129, 1355–1407.
Gleason, Philip, Melissa Clark, Christina Clark Tuttle, Emily Dwoyer, and Marsha Silverberg (2010). "The Evaluation of Charter School Impacts: (NCES 2010-4029)." U.S. Department of Education, National Center for Education Statistics. U.S. Government Printing Office, Washington, DC.
Gorard, Stephen (2014). "The Link between Academies in England, Pupil Outcomes and Local Patterns of Socio-economic Segregation between Schools." *Research Papers in Education*, 29, 268–284.
Hanushek, Eric and Ludger Woessmann (2011). "The Economics of International Differences in Educational Achievement." In *Handbook of the Economics of Education*, Vol. 3, edited by Hanushek Eric, Stephen Machin, and Ludger Woessmann. Elsevier, Amsterdam.
Hanushek, Eric and Ludger Woessmann (2015). *The Knowledge Capital of Nations: Education and the Economics of Growth*. MIT Press.
Hoxby, Caroline, Sonali Murarka, and Jenny Kang (2009). *How New York City's Charter Schools Affect Achievement*. New York City Charter Schools Evaluation Project.
Machin, Stephen, Shqiponja Telhaj, and Joan Wilson (2006). "The Mobility of English School Children." *Fiscal Studies*, 27, 253–280.
Machin, Stephen and James Vernoit (2011). "Changing School Autonomy: Academy Schools and their Introduction to England's Education." Discussion Paper 123. Centre for the Economics of Education, London School of Economics.
Machin, Stephen and Joan Wilson (2008). "Public and Private Schooling Initiatives in England: The Case of City Academies." In *School Choice International*, edited by R. Chakrabarti and P. Peterson. MIT Press, Cambridge, MA.
National Audit Office (2010). "Department for Education: The Academies Programme." http://www.nao.org.uk/publications/1011/academies.aspx. Retrieved May 21, 2018.
Neal, Derek (1997). "The Effects of Catholic Secondary Schooling on Educational Achievement." *Journal of Labor Economics*, 15, 98–123.
OECD (2011). "School Autonomy and Accountability: Are They Related to Student Performance?" PISA Focus 2011/9.
Price Waterhouse Coopers (2008). *Academies Evaluation Fifth Annual Report*. DCSF Publications, Annesley.
West, Anne and Elizabeth Bailey (2013). "The Development of the Academies Programme: 'Privatising' School-Based Education in England 1986–2013." *British Journal of Educational Studies*, 61, 137–159.

SUPPLEMENTARY DATA

Supplementary data are available at *JEEA* online.