Economics of Education Review 94 (2023) 102366

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

Heterogeneous effects of school autonomy in England

Lorenzo Neri[a,c,*], Elisabetta Pasini[b]

[a] University of St Andrews, UK
[b] Alma Economics, London, UK
[c] IZA, Germany

ARTICLE INFO

JEL Classification:
I20
I21
I28

Keywords:
Autonomous schools
School governance
School performance

ABSTRACT

A 2010 education reform gave English schools the option to become academies, autonomous but state-funded schools. Academies can opt for two different models of governance by choosing to remain standalone schools or join an academy chain. We investigate the causal effect of the governance model on student achievement and school inputs. We find that students in academy chains have higher end-of-primary school test scores, with stronger effects for low achievers and early converter academies. School chains are more efficient than standalone academies, achieving better results while spending less overall. Survey data suggest that chains favor management changes, whereas standalone academies make changes related to educational practices.

1. Introduction

Over the past two decades, many countries have increasingly seen in school autonomy the way forward to raise student achievement. The main rationale for increasing autonomy is to transfer more power in the hands of those who are likely to have better information on how to run their school, such as principals or local governing bodies. Several US-based studies find positive effects of autonomous - or 'charter' - schools on student achievement, in particular for urban charters (Hoxby and Murarka, 2009; Angrist et al., 2010, 2013; Dobbie and Fryer, 2011, 2015; Abdulkadiroğlu et al., 2011; Abdulkadiroğlu, Angrist, Hull, and Pathak, 2016; Dynarski et al., 2018; Walters, 2018), as do Böhlmark and Lindahl (2015) for Swedish 'free schools' and Eyles and Machin (2019) and Eyles et al. (2016) for English secondary academy schools. Other studies find, however, no or negative achievement effects on average for students attending autonomous schools compared to traditional state-funded schools (Bettinger, 2005; Gleason et al., 2010; Eberts and Hollenbeck, 2001; Eyles et al., 2017; Regan-Stansfield, 2018).[1]

The specific channels through which school autonomy may work are still debated. Teachers' feedback, tutoring, longer school time, a culture of high expectations and *ad hoc* practices targeting disadvantaged pupils seem to be the most successful practices in charter schools (Dobbie and Fryer, 2013; Fryer, 2014). Angrist et al. (2013) link charter schools' success to the use of the 'No Excuses' approach. Other recent papers, however, suggest that management practices, such as changes in the management structure (Eyles and Machin, 2019) and stronger accountability and high quality school leadership (Bloom et al., 2015), can play an important role in explaining autonomous schools' success.

We study the impact of different governance models in primary autonomous schools on student achievement and school inputs using a difference in differences (DID) strategy. In England, a 2010 educational reform opened to schools - first to high-performing and then to all state-funded schools - the possibility to become 'academies', publicly funded but autonomous schools. Schools can decide to convert as standalone academies (SATs - single-academy trusts) or join a school chain (MATs - multi-academy trusts). This choice results in two distinct models of governance with different degrees of centralisation. Standalone academies are now responsible for all managerial functions, whereas MAT schools' managerial activities are coordinated by a centralised trust, with a clear separation of roles between members of the trust and headteachers, whose main role is to run their school. Our research design compares MATs and SATs with the closest school of similar pre-reform

* Corresponding author at: University of St Andrews, St Andrews KY16 9AJ, Fife, UK.
E-mail addresses: lorenzo.neri@st-andrews.ac.uk (L. Neri), elisabetta.pasini@almaeconomics.com (E. Pasini).

[1] Angrist et al. (2013), while finding positive effects for urban charters, find *negative* effects for nonurban charter schools. More recent works look at long-term outcomes, such as college attendance, and find positive effects on students who attended charter schools compared to those who did not (Dobbie and Fryer, 2015; Angrist et al., 2016). A related strand of literature focuses instead on spillover effects of autonomous schools on students attending traditional state-funded schools, finding negative (see e.g., Bettinger, 2005; Booker et al., 2008; Ni, 2009; Cordes, 2018) or limited effects (Sass, 2006; Zimmer and Buddin, 2009; Imberman, 2011; Winters, 2012; Ridley and Terrier, 2018).

https://doi.org/10.1016/j.econedurev.2023.102366
Received 1 October 2022; Accepted 24 January 2023
Available online 5 March 2023

quality located outside the academy's catchment area, but in the same school district, that did not become an academy. The control group therefore consists of students enrolled in untreated schools belonging to the same school market.

We focus on early Converter primary academies, defined as schools that became academies between 2010 and 2014.[2] Partly because at the beginning the government prioritised 'outstanding' schools for conversion, early Converters are largely represented by high-achieving schools, and the vast majority were rated either outstanding or good by Ofsted inspections.[3] In addition, as more schools become academies, the funds available to school districts are expected to decrease, as they are proportional to the number of students enrolled in state-funded schools maintained by them. This may affect the quality of the services provided, thus changing the incentives to become academies and leading more schools to convert over time.

The majority of academies are represented by pre-existing schools that become academies. In this respect, they are similar to US public schools that become charters following a school takeover. However, while US charter schools often serve a large fraction of disadvantaged students, academies include a large number of schools that serve more advantaged and high-achieving students. In addition - unlike charters and Swedish free schools - academies represent a unique case study, since England is progressively moving towards a fully decentralised system, with an increasing number of schools becoming academies every year. Although in the aftermath of the reform converter academies were more likely to convert as SATs, in recent years the proportion of schools belonging to MATs grew dramatically (Fig. 1). As of March 2022, MATs managed more than 5,800 primary schools (about 35%) enrolling more than 1,500,000 of children aged 5-11.

In this setting, both the school's governance model choice and individual students' enrolment decisions are endogenous. Although we do not have an instrument for the school's choice, we show that control group schools closely track student performance in MAT and SAT schools, providing a counterfactual for what would have happened absent the school's conversion. We further show that pre-existing trends in student composition and performance do not generally predict a school's conversion model. As for student self-selection into and out of academies, we adopt an Instrumental Variable (IV) approach that exploits the fact that students who were enrolled in a school before its conversion are guaranteed a school place (similar to, e.g. Abdulkadiroğlu et al., 2016). Given that it is hard for parents to anticipate the future conversion into academy, as well as the governance model's choice, early enrolment decisions can be considered orthogonal to the school's conversion decision.

We first find that schools belonging to chains improve student performance at the end of primary school relative to SAT academies. An event study analysis shows that MAT were not outperforming SAT schools before the conversion decision. On average, exposure to MATs with respect to SATs increases math and language test scores by about 0.066-0.048 standard deviation (σ), which corresponds to 1.32 and 0.37 points (about 1.9-1.2% of the average) respectively. Since in the most complete specifications we control for student baseline achievement, these estimates can be interpreted in terms of progress made by the children. Achievement gains are driven by early converter and, to a lesser extent, by low achievers. When we benchmark these effects against traditional state-funded schools, MAT schools outperform the latter in math by about 0.028σ (significant at 10%), while SAT performance in both math and language is lower. Descriptive evidence suggests that more effective MATs did not take over more schools over time.

Our second finding is that MATs are more efficient that SATs. We show that MAT schools seem to benefit from economies of scale and, while raising student achievement with respect to SAT schools, they do so spending less overall. SAT schools increase per-pupil expenditures (about £234 more) with respect to traditional state schools and spend substantially more for teaching staff, maintenance of premises and back-office items. MATs spend less than SAT overall (about £83 less, significant at 10%), and in particular in learning and ICT resources and back-office items.

We end our analysis by documenting how management practices differ across MAT schools and SATs to shed light on the potential mechanisms underpinning our findings. We exploit data from a unique survey conducted by the Department for Education (DfE) in 2017, the "Academy Trust Survey 2017" (Cirin, 2017), which focuses specifically on practices and changes introduced by MAT and SAT schools. We document that, after conversion, SAT schools are more likely to implement school level changes, such as changing the curriculum and increasing the number of pupils, while MATs favour organisational level changes, such as replacing school leaders, reconstituting the governing body and changing the performance management system for teachers. Even though these figures should be considered with caution given the limited coverage of the survey, they seem to suggest that managerial practices play a key role in boosting student achievement in MATs.

Why should the fact that different governance models affect student achievement be of interest to researchers and policy-makers? Following the introduction of academies, the public debate on school governance has emphasised the potential risks faced by standalone academies, mostly related to their lack of expertise in managerial functions. The main concern is that autonomy requires a body of expertise in managerial fields that traditional representatives of local schools' governing bodies may not be endowed with. Indeed, as the program developed, policy-makers have supported school chains as an efficient way to foster schools' collaboration and reduce the educational gap between advantaged and disadvantaged students (Hutchings et al., 2016). Such view has been reinforced by the March 2022 Schools White Paper for England (The House of Commons Library, 2022), which sets out government plans to have all schools join an academy chain by 2030.[4] In addition, although the number of academies has massively increased, by 2022 the vast majority of schools becoming academies were joining MATs (Fig. 1).

Over the recent years, however, the literature has mostly attempted to test the effectiveness of autonomous schools by focusing on the impact of autonomy *per se* on student achievement. Indeed, apart from few exceptions (Andrews and Perera, 2017; Eyles et al., 2017; Woodworth et al., 2017; Eyles et al., 2018) very little has been said on the role played by different governance models in autonomous schools.[5] We contribute to this growing, albeit relatively limited, literature by focusing on high-achieving autonomous schools characterised by different governance models. Our paper is closely related to Eyles et al.

[2] Similar to Eyles et al. (2017), we exclude from the sample 'sponsored' academies. These are represented by underperforming schools that often become academies following a government intervention, and are supported by a sponsor (a person or organisation).

[3] Ofsted - the Office for Standards in Education, Children's Services and Skills - regularly carries out inspections in schools. Following an inspection, a school is rated from outstanding to inadequate. Inspections are carried out on a rolling basis depending on the outcome of the previous inspection.

[4] This White Paper was predated by another White Paper, *Educational Excellence Everywhere* (Department for Education, 2016a). This states that MAT was the preferred model as collaboration allows schools to benefit from the most successful leaders and their expertise.

[5] The CREDO report (Woodworth et al., 2017) looks at the impact of different types of charter schools - independent, hybrid, Charter Management Organisations (CMOs), Vender Operated Schools (VOSs) - on student academic growth using matching methods. Andrews and Perera (2017) look at achievement effects of MATs in a report for the Education Policy Institute. However, they only perform a descriptive analysis in the 2015 cross-section using school-level data and focusing on the comparison between MATs and LA schools. Eyles et al. (2018) focus on the dissimilarities between the first (pre-2010 reform) and second (post-2010 reform) wave of academies, touching upon the MAT/SAT distinction but only for secondary schools, whose MATs disproportionately include sponsored academies.

(2017) and Regan-Stansfield (2018), which look at the effect of primary academy schools on student performance using future academies as control group, finding no achievement gains. While Eyles et al. (2017) touch upon the MAT/SAT distinction, Regan-Stansfield (2018) stresses how this aspect of the academy reform has not been fully analysed yet.

Our paper begins by diving more deeply into the impact that different school governance models have on student achievement, picking up where Eyles et al. (2017) left off. We propose an alternative empirical strategy that takes into account the change in schools' incentives to convert over time, and therefore the potential endogeneity in the timing of conversion. We also document whether MAT effects differ across subgroups of students and schools, and whether more effective MATs take over more schools over time. We then focus on school inputs and study how they have changed across MAT and SAT schools following a school conversion. To the best of our knowledge, our paper is the first that attempts at understanding whether MATs are able to obtain larger test score gains per pound spent with respect to SATs.

Finally, we contribute to the growing literature looking at the role management practices play in shaping student outcomes. Bloom et al. (2015) explore the determinants of autonomous schools' success in several countries and suggest that management practices can be crucial to foster school performance and explain disparities in the quality of education across schools. Eyles and Machin (2019), using survey data on academies (Cirin, 2014), find that changes in the management structure and curriculum are the main factors underpinning pupils' improvement in secondary academies. They study, however, a different set of academies - failing secondary schools that became academies before the 2010 reform - and do not focus on the SAT/MAT distinction, but rather on academies vis-a-vis traditional state schools. Exploiting the same survey, Eyles et al. (2017) document how primary academies used their freedoms compared to secondary academies. Using survey data on MATs and SATs newly collected by the Department for Education (Cirin, 2017), we investigate how academies characterised by different governance models differentially implement educational vis-a-vis managerial practices.

2. Institutional setting

2.1. The english school system and the academy reform

Primary education in England is organised in two phases, Key Stage 1 (KS1) and Key Stage 2 (KS2). Children enter primary school in Reception year, when they are aged 5. KS1 runs from Year 1 to Year 2, when students are aged 7. KS2 runs from Year 3, when students are aged 8, to Year 6, when students are aged 11. State-funded schools are the majority and enrol about 95% of all students (Department for Education, 2016b). The majority of students attend 'community' schools, which are managed by the school districts (Local Authorities, LAs). LAs recruit teachers and staff, provide schools with all the services they need and administer the school budget set by the central government. The other most common state-funded schools are faith schools, which enjoy some degree of autonomy from the LA (e.g. on admission criteria). We refer to these different schools as 'traditional' state schools in what follows, to distinguish them from fully autonomous state schools (academies).

The Labour government introduced secondary school academies in 2000 through the Learning and Skills Act 2000, with the aim of improving performance by providing head teachers with direct control over their schools. Similarly to US charter schools, for the first 10 years the reform targeted only low performing secondary schools classified as inadequate by Ofsted inspections. The reform was then expanded to all primary and secondary schools by the coalition government in July 2010. Beside sponsored academies, which are underperforming schools often forced to convert following government intervention, converter academies, for which the conversion is voluntary, appeared. At the beginning, only schools rated 'outstanding' by Ofsted could apply for conversion, but from April 2011 the possibility was expanded to all schools 'performing well' (see West and Wolfe, 2018). Since 2010 the academisation process grew dramatically, and as of March 2022 6,279 out of 16,766 primary schools have already become academies.[6]

Academies are independent from local and central government and are non-profit charitable trusts. The decision to become an academy is taken by the school's governing body.[7] Similarly to traditional state-funded schools, they are funded by the central government and funding is linked to the number of students on roll. However, academies benefit from a £25,000 grant to support the conversion process, receive funding directly from the central government and are autonomous in aspects such as staffing (recruiting and paying teachers and staff, staffing structures, career development, discipline and performance management), provision of services (e.g. maintenance, HR, audit, legal services), and setting the curriculum (with the exclusion of a few subjects they are free to diverge from the traditional curriculum). Academies are free to set their own admission criteria, but they are subject to the guidelines stated in the Admission Code and cannot select students based on ability.[8]

A significant number of studies have been conducted in the US, where, since the late 1990s, the government has targeted low performing schools in deprived areas and forced them to become autonomous with the aim of implementing *ad hoc* policies to boost pupils' results, and consequently reducing the achievement gap among students. Similarly to English academies, 'charter schools' are publicly funded and tuition-free, but enjoy substantial operational autonomy from local and central government on the decisions concerning school curriculum, staffing, and the educational approach (e.g. school day length, school philosophy). However, charter schools are often located in deprived areas and serve a large fraction of low performing or minority students, while English academies include a substantial number of high achieving schools. Additionally, while academies can only be nonprofit organisations, US charter schools can be run for profit.

2.2. Multi-Academy and single-academy trusts

Besides sponsored and converter academies, another important distinction arose after 2010. Together with the decision of converting, converter academies can choose between converting as a standalone academy or joining a chain of academies. Such distinction resulted in two different models of governance. Standalone schools become SATs and the governing body - that now no longer requires LA governors - takes on all the responsibilities. MATs have instead a single governing body that runs all the schools belonging to the chain - as of 2016, those

[6] Last update available at: https://www.gov.uk/government/publications/open-academies-and-academy-projects-in-development. Official data regarding March 2022 are available upon request.

[7] The Academies Act 2010 states that the school's governing body must 'consult such persons as they think appropriate'. There isn't an obligation to consult with any specific party during the process, although it is considered good practice to consult with school staff and parents.

[8] A recent paper by Machin and Sandi (2020) investigates the exclusion of poorly performing pupils in academies. They find that the exclusion rate is higher in schools converted before 2010 compared to those converted in the second phase of the program (post-2010). However, they argue that such exclusion does not aim at boosting schools' performance, but it is, instead, the result of enforcing rigorous discipline codes.

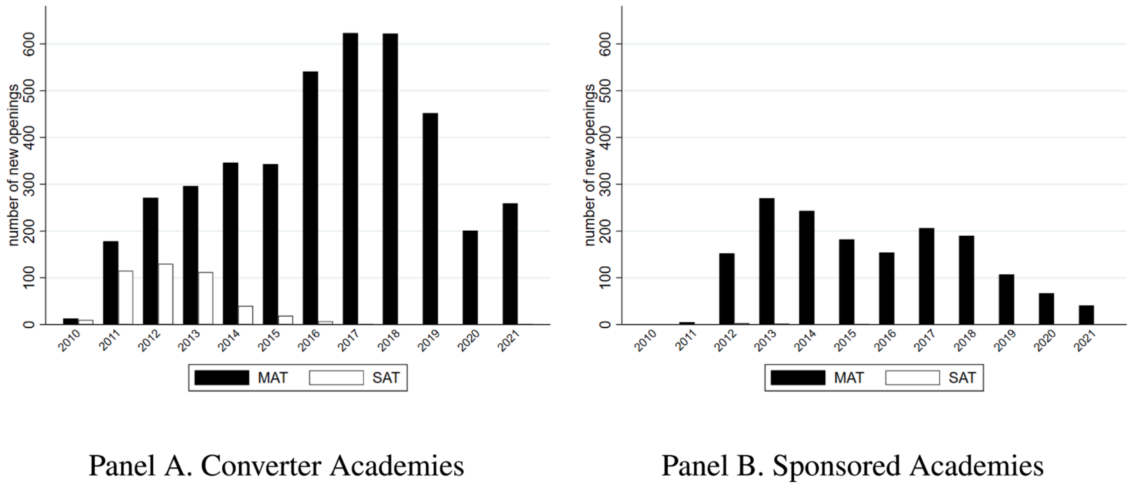


Fig. 1. Yearly openings of academies

Notes. The figure shows the number of openings of Converter (Panel A) and Sponsored (Panel B) academies by year of opening and academy status (MAT or SAT).

included in our sample managed about 7 schools on average.[9] Although right after the reform converter academies were more likely to convert into standalone academies, in recent years the proportion of schools belonging to MATs grew dramatically (Fig. 1).

The main distinction between MATs and SATs concerns the governance structure (Fig. 2). The trust running the MAT is responsible for all the academies in the chain. Schools belonging to the MAT share the same board of governors, which takes up most of the tasks previously performed by the local governing bodies of the single schools. The foundation members of the trust have ultimate control over the schools and appoint the board of governors (also called directors or trustees), which set the direction of the MAT, hold head-teachers accountable, and ensure financial probity. The presence of governors therefore creates an additional tier of governance between foundation members and the local governing bodies of the schools.[10] Trustees can delegate a number of functions to the local governing body of single schools, whose functions are now limited compared to local governing bodies of SATs. Indeed, the model introduced by MATs aims at removing pressures on local governing bodies and avoiding the recruitment of high skill governors for each single school (Grotberg and Robb, 2015).

The 2010 reform has therefore explicitly allowed for two different models. SATs stand for a decentralised system in which each single school is responsible for all the decisions and services, while MATs are based on a centralised system in which functions and operations are attributed to different actors along the 'governance chain'. In particular, managerial functions are carried out by the governors, whose knowledge of business practices can be expected to be better than that of school head-teachers. Indeed, as shown in Table 1, trust boards handle financial and legal compliance, senior appointments, and risk management, while schools mostly handle operational functions (e.g. school development plans, strategies, school staffing structures design), which in few cases are also carried out at regional level.[11]

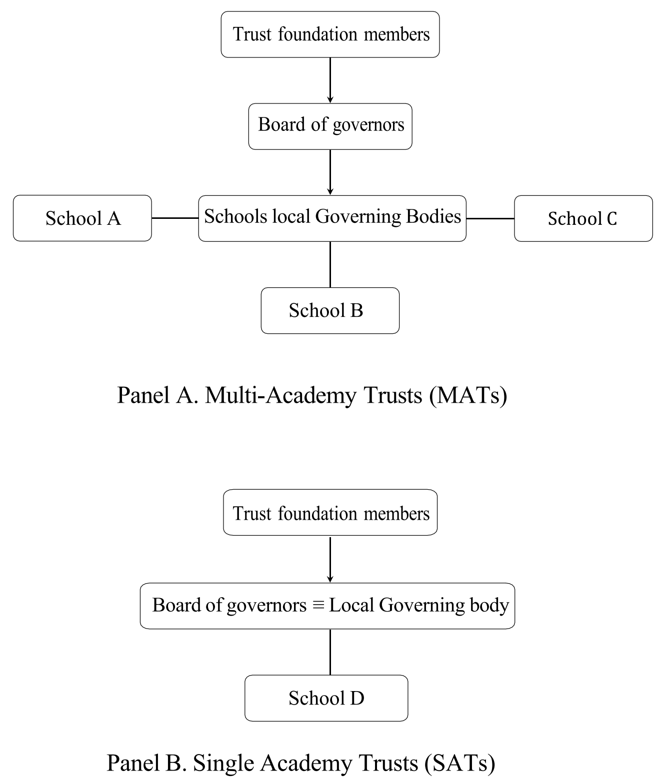


Fig. 2. Governance in SATs and MATs

A survey conducted by the Department for Education (Cirin, 2017) reveals different reasons for converting between MAT and SAT schools. The choice of the governance structure for converter academies

[9] MATs can include both primary and secondary schools and do not face any geographical constraint. Over time, the number of MATs including schools from different LAs has increased. As of 2022, and although in our sample the majority of schools belonging to a given MAT are located in the same LA, over time the number of schools joining MATs has increase and so has their geographical spread. In the early years after the reform (2010-2012), for instance, 61% of MATs had schools from only 1 Local Authority (LA), while about 29% of MATs managed schools from 2 or 3 LAs. As of March 2022, about 68% of MATs managed schools from only 1 LA, while about 26% of MATs managed schools from 2 or 3 different LAs.

[10] The board is made up of at least three signatory members, the CEO, and two elected parents. No more than 20% of trustees can be persons associated to a LA (e.g. head-teachers of community schools, LA officers). The average size of boards is 8 members, and more than half of MATs have between 7 and 10 members.

[11] Regional clusters represent a further tier of Governance between schools and the Trust. As they grow in size, trusts may choose to decentralise some functions to regional hubs whose proximity to schools makes the management more efficient.

Table 1
Location of responsibility within MATs

	Regional/ Cluster level (1)	Trust Board (2)	School level (3)
Financial compliance	5%	94%	1%
Legal compliance	5%	92%	3%
Appointing headteachers/principals	5%	90%	5%
Managing risks	5%	88%	7%
Holding individual headteachers/ principals to account	8%	82%	10%
Monitoring the performance of individual schools	14%	73%	14%
Human resources	10%	73%	17%
Allocating school budgets	6%	69%	25%
Directing school improvement support	18%	62%	20%
Setting academic targets	13%	44%	43%
Designing school staffing structures	15%	29%	57%
Setting individual school strategy/ objectives	8%	29%	62%
School development action plans	14%	8%	78%

Notes. The table presents the location of responsibility in multi-academy trusts (MATs) at Region/Cluster (column (1)), Trust Board (column (2)) and School (column (3)) level. Data source: Academy trust survey 2017. The sample of MAT respondents includes 237 MATs with at least 2 schools and 1 primary school.

ultimately rests on the school's governing body, which may opt for the preferred model. Schools applying to become part of a MAT need to obtain MAT's approval at the beginning of the process. 42% of MAT schools mention that the main reason to convert was to create opportunities for collaboration with other schools. Instead, about 26% and 20% of SAT schools mention obtaining more funds and using them as they see fit as the main reason to convert. SAT schools are also slightly more likely to convert to gain independence from the LA (about 14% vs 12%).

3. Empirical methodology and identification

3.1. Data

We use data from the National Pupil Database (NPD), a unique and rich dataset containing information at pupil and school level in England and covering all students attending publicly funded schools. The dataset contains detailed student demographics such as gender, ethnicity, language spoken at home, eligibility for free school meals (FSM) and special education needs (SEN) status, pupils' block of residence, and school attended.[12] The main source of information on academies is Edubase (now called 'Get information about schools'), a DfE-maintained dataset that contains school level data on single and multi-academy trusts (e.g. time of conversion, date in which the school has joined a chain, type of support, trust size).

The NPD includes information on student achievements at the end of KS1 and KS2. Although at the end of KS1 students are assessed by their own teachers, these assessments are low-stake evaluations. Hence, KS1 attainment can be considered a good proxy for pupils' performance at year 2. KS2 tests are instead national standardised tests in math and language taken at the end of primary school (year 6) and marked by external markers. At both stages students are also awarded a Level of attainment depending on the score they obtain - from Level 2 to Level 4 at KS1 and from Level 3 to Level 5 at KS2.[13]

We complement individual-level NPD data with school characteristics from several sources. The School Census and School Workforce Census (SWC) contain data on teacher characteristics, such as pupil-to-teacher ratio and teaching qualifications. The Consistent Financial Reporting (CFR) contains data on school expenditures by category (e.g., teaching staff, learning resources, back-office). We integrate the latter with publicly available data on academies expenditures. Finally, we link the NPD to Ofsted inspection reports from 2005 onwards, which are publicly available on the UK government website.[14]

We keep converter academies whose conversion is between July 2010 (the month in which the academy reform took effect) and December 2014. We retain only schools that remain continuously in the sample over the period considered.[15] The final sample includes 2,113 schools and 914,786 students enrolled in the last primary school grade over 2005-2016. Of the 2,113 schools, 260 became academies in 2010-2011, 626 in 2012-2013 and 298 in 2014. None of the schools in our final sample changes the governance model chosen initially over the years considered (i.e. from MAT to SAT or the opposite). Table 2 presents descriptive characteristics for traditional state schools as well as MAT and SAT schools in the final sample.

3.2. Empirical specifications

We investigate heterogeneous effects of conversion between stand-alone academies and academies in chains. In particular, we are interested in estimating the effects of exposure to MATs or SATs on KS2 test scores. We consider up to four years of exposure to an academy, which represent the length of KS2. This approach has the advantage that we can control for student baseline attainment at the end of KS1 measured before the conversion event. Our estimates of the impact of governance on student achievement can therefore be interpreted in terms of progress made by the student or value-added provided by the treatment.

We design a DID strategy that compares students exposed to a MAT vis-a-vis those enrolled in a SAT school before and after a conversion into academy, using students enrolled in traditional state schools as control group. As main control group, we consider students enrolled in traditional state-funded schools closely located to treated schools. Specifically, for each academy school we drop any state school that is located within the 80th percentile of the LA-specific student-school distance distribution. This is to avoid any spillover due to being in the immediate vicinity of a school becoming an academy. Amongst the remaining set of schools with the same Ofsted grade, we keep the closest state school. We use the last available Ofsted grade before the conversion took place. We conduct a series of robustness checks regarding the control group, which are outlined at the end of Section 4 below. In the early years of the academy reform the conversion decision was

[12] Blocks are Lower Layer Super Output Areas (LSOAs), a geographical unit created by the Office for National Statistics (ONS) for census reporting purposes containing 800 households on average.

[13] These levels are meant to capture the position of the student in the achievement distribution. Students awarded the lowest Level (2 and 3 at KS1 and KS2 respectively) are students performing below expectations; those awarded the middle Level (3 and 4 at KS1 and KS2 respectively) are students working at the expected level; those awarded the top Level (4 and 5 at KS1 and KS2 respectively) are students performing above the average.

[14] The School Census is available yearly from 2006 as part of the NPD. More recent data from the SWC and on academies funding are publicly available at https://www.gov.uk/government/collections/statistics-school-workforce and https://www.gov.uk/government/collections/statistics-local-authority-school-finance-data respectively. Ofsted reports can be accessed at https://www.gov.uk/government/statistical-data-sets/monthly-management-information-ofsteds-school-inspections-outcomes.

[15] A group of schools decided to boycott KS2 tests in 2010 by withdrawing their students from the text. Excluding from the sample the schools that have all missing scripts in 2010 (575 schools in our main sample) leaves the main estimates substantially unaffected. These results are presented in Appendix Table A.1, which replicates Table 4 for this restricted sample.

Table 2
Descriptives

	All schools Mean	S.D.	Academies Mean	S.D.	MAT Mean	S.D.	SAT Mean	S.D.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Students								
Male	0.50	0.50	0.51	0.50	0.50	0.50	0.51	0.50
Eligible for free school meals (FSM)	0.13	0.33	0.12	0.33	0.13	0.34	0.11	0.31
White	0.87	0.34	0.87	0.34	0.87	0.33	0.86	0.35
Black	0.03	0.16	0.03	0.16	0.02	0.16	0.03	0.17
Native	0.92	0.27	0.92	0.27	0.93	0.26	0.91	0.29
With special educational needs (SEN)	0.21	0.41	0.20	0.40	0.21	0.41	0.19	0.39
Panel B. Scores								
KS1 math Level 3	0.25	0.44	0.26	0.44	0.25	0.43	0.28	0.45
KS1 reading Level 3	0.30	0.46	0.30	0.46	0.29	0.45	0.32	0.47
KS2 math Level 5	0.40	0.49	0.41	0.49	0.39	0.49	0.44	0.50
KS2 reading Level 5	0.53	0.50	0.55	0.50	0.53	0.50	0.57	0.49
KS2 average score	49.04	12.82	49.47	12.70	48.83	12.88	50.39	12.38
KS2 math score	67.99	19.96	68.57	19.75	67.61	20.02	69.95	19.27
KS2 reading score	30.10	7.62	30.37	7.56	30.06	7.64	30.82	7.43
Panel C. Schools								
Community	0.58	0.49	0.56	0.50	0.56	0.50	0.55	0.50
Voluntary Controlled	0.13	0.34	0.11	0.31	0.10	0.30	0.13	0.33
Voluntary Aided	0.25	0.43	0.27	0.44	0.30	0.46	0.21	0.41
KS2 grade enrolment	35.13	21.18	37.55	22.41	35.16	21.89	41.62	22.70
Pupil-teacher ratio	21.67	3.34	22.04	3.26	21.80	3.35	22.44	3.06
Percent qualified teachers	0.96	0.08	0.96	0.08	0.96	0.08	0.95	0.08
Number of schools	2,113		1,184		746		438	
Number of students in final sample	914,786		494,777		293,027		201,750	

Notes. The table presents summary statistics for traditional state funded schools (columns (1) and (2)), converter academies (columns (3) and (4)), academies in MATs (columns (5) and (6)) and SATs (columns (7) and (8)). The academies considered are those included in the final sample (see Section 3). Accordingly, the table shows the number of schools and students in the final sample. Schools considered in columns (1) and (2) include all state-funded schools in the control group. Means and standard deviations are computed for 2009, the last pre-reform year.

essentially taken in conjunction with the decision concerning the governance model. None of the academies in our sample changed governance model or academy type (i.e. converter and sponsored) in the period considered. In contrast to previous works for England (e.g. Eyles et al., 2017), we do not use future converters as control group to account for potential endogeneity in the timing of conversion. Early converter academies, for instance, benefited from substantial financial incentives (West and Wolfe, 2018). In addition, as more schools become academies, the funds available to LAs will likely decrease since the latter are proportional to the number of students enrolled in state-funded schools maintained by them (Ladd and Fiske, 2016). This may affect the quality of services provided, thereby leading more schools to convert. In addition, Fig. 1 shows that the governance model chosen by academies was substantially different between early and late converters.

A recent econometric literature has highlighted several issues with TWFE estimators with variation in treatment timing and heterogeneous treatment effects.[16] In our context, schools enter treatment in different years, and we cannot rule out some degree of treatment effect heterogeneity. To deal with the pitfalls associated with TWFE estimation, we use a 'stacked-by-event' design, building 'placebo' events for control schools similar to Deshpande and Li (2019). We first create a separate dataset for each treated school, including all students in the treated school along with never-treated students enrolled in the control school. We define the relative time to event in each dataset with respect to the year when treatment starts. In this setting, it is possible to separately estimate year (cohort) and 'event time' fixed effects, which help control for additional (event) time trends in the run up to the conversion event. Then, we stack all datasets into one. In this procedure, one student in never-treated schools can, in principle, serve as control at different event times depending on the treatment wave considered. We follow Deshpande and Li (2019), Cengiz et al. (2019) and Fadlon and Nielsen (2019) and estimate the following model:

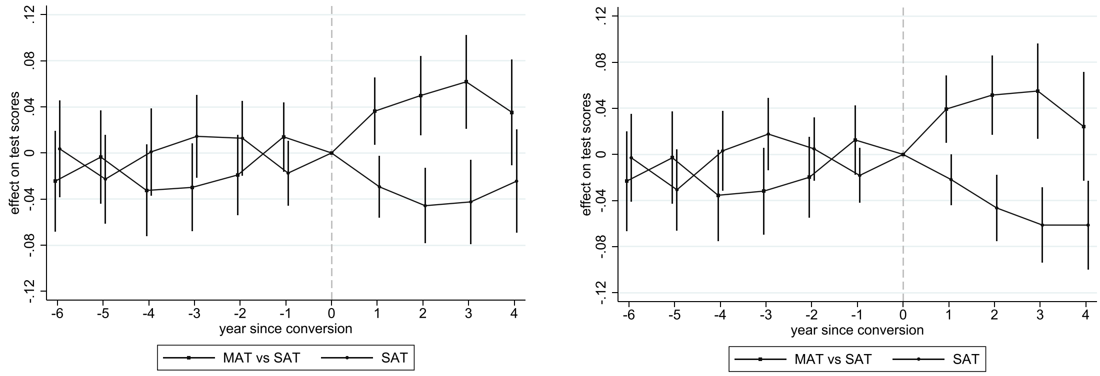
$$Y_{isct} = \beta_0 I(t \leq -7) + \sum_{k=-6}^{4} \beta_{1k} D_s d_k + \sum_{k=-6}^{4} \beta_{2k} D_s MAT_s d_k + \theta_1 X_{isct} + \gamma(s,c,t) + \varepsilon_{isct} \quad (1)$$

where Y_{isct} is the KS2 score of student i enrolled in school s and cohort c at time t. Test scores are standardised by subject and year. D_s is the treatment indicator, taking value 1 for all student enrolled in a school that becomes an academy. MAT_s is an indicator taking value one if the academy belongs to a chain and 0 otherwise. The coefficient of interest is β_2, representing the impact of being exposed to an academy that joined a MAT with respect to a SAT. β_{1k} and β_{2k} for $k = -6, ..., -1$ can be interpreted as placebo estimates of the effect of conversion into MAT or SAT before the actual time of conversion. We 'bin' together distant relative periods starting from -7 (see Sun and Abraham, 2021). X_{isct} is a vector of pupil characteristics including gender, ethnicity, language spoken at home, FSM eligibility, and KS1 achievement in writing, reading and math at KS1. Finally, $\gamma(s,c,t)$ is shorthand for sets of school, cohort and event time fixed effects. The time index c runs across cohorts of exam takers, with $c = 2005, ..., 2016$. Standard errors are clustered at the school level, with each school establishment counting for one cluster throughout the period of analysis.

The placebo estimates obtained lend validity to the suitability of the control group and the parallel trend assumption. Fig. 3, Panel A, plots event study estimates obtained by estimating Eq. (1) for KS2 average test scores, showing that before the actual conversion date the estimates of MAT and SAT effectiveness are all fairly close to zero (and never statistically significant) and, importantly, do not exhibit any trend before the conversion event. A further joint test for the significance of coefficients before conversion also rejects their statistical significance. To corroborate our 'stacked-by-event' design, Panel B of the same figure plots estimates of Eq. (1) obtained using the estimator proposed in Sun and Abraham (2021). The latter are always very close to those in Panel A.

Since both MAT_s and D_s are endogenous choice variables, one should still be cautious in interpreting estimates of β_{1k} and β_{2k} from regression (1). On the one hand, the choice of the school to join a MAT may be

[16] See De Chaisemartin and d'Haultfoeuille (2020); Baker et al. (2021); Borusyak and Jaravel (2021); Callaway and Sant'Anna (2021); Goodman-Bacon (2021); Sun and Abraham (2021).



Panel A. Stacked-by-event Panel B. Sun and Abraham (2021)

Fig. 3. Event study

Notes. The figure presents event study estimates using a 'stacked-by-event' design (Panel A) and the estimator proposed in Sun and Abraham (2021; Panel B). The school's conversion year is re-centred at zero. The round dots present estimates of the achievement gains in SAT schools with respect to traditional state schools; the square dots present estimates of MAT school gains with respect to SAT schools. The vertical bars show the 95% confidence interval. F-test and P-value for joint significance of pre-conversion coefficients in the stacked-by-event design are 1.46 and 0.13.

correlated to school observables or unobservables, such as trends in the strength and composition of student cohorts. On the other hand, D_s can be correlated with parents' decisions and timing of enrolment in an academy. We address these issues in the following two sections.

3.3. School decisions and intake

A school's decision to convert into an academy - and, particularly, whether converting as SAT or joining a MAT - could correlate with pre-existing characteristics in school performance and intake. School fixed effects in Eq. (1) control for any fixed unobserved school attributes that could possibly affect our estimates. In addition, our control group closely mimics trends in test scores in academy schools before the conversion year (see Fig. 3), suggesting that the former represent a sensible counterfactual to what would have happened to students in academy schools absent the treatment. Governance decisions, however, could still correlate with pre-existing trends in school performance and intake. For instance, one may worry that schools intentionally choose the year of conversion depending on the academic strength of the cohort taking KS2 tests around conversion year. This would then boost KS2 results of the school independently of the year of exposure.

We test for this by estimating the following multinomial logit regression, at school level:

$$ST_s = \eta_0 + \eta_1 \Delta M_s + \eta_2 \Delta W_s + \psi_s \tag{2}$$

where ST is a categorical variable individuating whether a school becomes a MAT academy, SAT academy or remains a traditional state school (baseline). ΔM_s is a vector of pre-conversion changes in school performance: it includes KS2 test scores and KS1 assessments by school teachers in math and language, as well as KS1 average point score. ΔW_s is a vector of changes in cohort composition before conversion and includes gender, FSM eligibility, ethnicity, language spoken at home, SEN status, and grade enrolment. The regression is estimated over the period 2009-2005, using differences between 2009 and 2005 (Table 3, columns (1) and (2)) as well as 2009 and 2007 (Table 3, columns (3) and (4)) in school performance and composition.[17] Neither changes in school composition nor in school performance consistently predict the governance model chosen upon conversion.

Table 3
Trends in characteristics and governance model

	4-year lag		2-year lag	
	MAT	SAT	MAT	SAT
	(1)	(2)	(3)	(4)
KS2 English scores	0.107	-0.001	0.015	-0.077
	(0.065)	(0.074)	(0.064)	(0.077)
KS2 math scores	-0.020	0.097	0.042	0.139
	(0.066)	(0.071)	(0.067)	(0.073)
KS1 English points	0.027	0.047	0.004	0.078
	(0.068)	(0.085)	(0.071)	(0.076)
KS1 math points	-0.089	0.017	-0.091	-0.068
	(0.068)	(0.078)	(0.068)	(0.075)
Male students	0.021	0.156	0.031	0.118
	(0.053)	(0.054)	(0.053)	(0.057)
Students eligible for free school meals	-0.044	-0.001	-0.023	-0.038
	(0.052)	(0.055)	(0.052)	(0.056)
Black students	-0.006	0.069	0.011	0.083
	(0.055)	(0.054)	(0.051)	(0.055)
Native students	0.013	-0.039	-0.030	-0.116
	(0.051)	(0.057)	(0.053)	(0.053)
Students with special educational needs	0.016	0.012	-0.037	-0.018
	(0.055)	(0.059)	(0.055)	(0.060)
KS2 grade enrolment	-0.077	0.088	0.045	0.061
	(0.049)	(0.062)	(0.049)	(0.060)
Number of schools	2,108		2,110	

Notes. The table shows regressions of a discrete variable individuating whether a school is not an academy (baseline), belongs to a MAT or is a SAT on changes in student and school characteristics. All independent variables are standardised to have zero mean and unit variance. The time period considered is 2009-2005. In column (1) changes are computed over 4 years (2009-2005), and in column (2) over 2 years (2009-2007). Results obtained for the period 2010-2005 are similar and are available upon request. Standard errors, shown in brackets, are clustered on schools.

3.4. Parental decisions and student selection

The second concern in this setting is the endogenous sorting of students across academies and traditional state schools (and MATs and SATs). On average, student mobility within the last phase of primary school (KS2) is low in England. Still, students can change school at any point in time, so that every year the fraction of students who spent all

[17] Since 2010 is the last year before the reform took effect, one could also estimate the same regressions over 2010-2005. However, in 2010 part of schools boycotted the KS2 tests, and therefore we would not be able to estimate this regression for our final sample. Results considering this time window are similar and are available upon request.

previous years in the school is likely to be mechanically lower. This implies that any estimate one would get with a simple OLS regression would not reflect the true impact of MATs or SATs. Indeed, one would not take into account that not all students taking KS2 tests have spent the same number of years in the school. Additionally, schools that decide to become academies may attract better students, so that the effects of being exposed to a MAT would be the result of self-selection of good students into a MAT or SAT school. If selection is correlated with pupils' unobservable characteristics, regression (1) does not estimate the causal effect of exposure to a MAT.

In order to deal with endogenous self-selection of pupils into academies we exploit the fact that enrolment decisions made by parents usually happen years before the conversion. Additionally, both the 2010 reform and subsequent single schools' conversion decisions could hardly be anticipated by parents.[18] Hence, we can safely assume that enrolment in a MAT or SAT school before the actual year of conversion is orthogonal to the school's decision of becoming an academy.

We exploit a grandfathering instrument as in Abdulkadiroğlu, Angrist, Hull, and Pathak, 2016. We condition on the sample of 'grandfathered' students (i.e. students who were already enrolled in their school before the conversion year), track the school where they seat KS2 tests, and instrument student enrolment in an academy at the time of KS2 tests with a variable indicating whether the student was enroled in an academy in the year before the conversion. As school conversion - as well as governance model's choice - could not be anticipated by parents, enrolment in a MAT or SAT school after conversion of students who took enrolment decisions for the pre-conversion school can essentially be considered passive.

4. Results

4.1. Governance, student achievement and school inputs

MAT schools outperform SAT schools after conversion in end-of-primary school tests in math and language. Table 4, Panel A reports OLS estimates obtained estimating, for the grandfathering sample, a 'pooled' version of regression (1) that aggregates the four post-conversion periods. The first row reports coefficients for the impact of being exposed to a SAT school with respect to a traditional state school. The second row shows the achievement gain for pupils exposed to a MAT school compared to those in SATs. For average test scores (column (2)), math (column (4)) and language (column (6)) there seems to be a negative effect of SATs on student achievement and a positive effect due to MAT exposure. These columns, presenting the most complete specification that also controls for student baseline achievement, suggest that being exposed to a MAT improves students' average test scores by 0.054σ (about 0.69 points, or 1.4% of the average). For math and language test scores the increase is about 1.8% and 1.2% respectively.

These estimates do not take into account that students can endogenously sort across academies as well as SAT and MAT schools. Hence, we use the Instrumental Variable (IV) strategy outlined in the previous Section and estimate regressions (1) using a 'grandfathering' instrument for the years of exposure. 2SLS estimates of the impact of different governance models on student achievement

(Panel B) are similar to OLS estimates. This is consistent with high first stage estimates (Panel C), which suggest that about 93% of students that take KS2 tests in an academy were enrolled there before conversion, implying relatively little student mobility.[19] Columns (4) and (6) suggest that MAT exposure raises test scores in math and language by 0.066-0.048σ, or 1.32 and 0.37 points respectively (about 1.9-1.2% of the average).

Overall, our results support the notion that MATs are more effective than SATs in raising student achievement. The improvement in KS2 test scores happens across the board, in both math and language. Since in the main specification (columns (2), (4) and (6)) we also control for student's baseline achievement (proxied with KS1 scores), the effects can be interpreted in terms of progress made by the children (or value-added). Such positive effects, however, seem to decrease after 4 years. Fig. 4 breaks down the effects by years of exposure, and re-estimates Eq. (1) using the grandfathering design and instrumenting D. For average score, math and language the positive effects persist, but they are increasing in the first years of exposure and then seem to decrease by the fourth year. This is partly due to an improvement in test scores of students attending SAT schools (see Appendix Fig. A.1).

Before moving to the second part of our analysis, which looks at school inputs, we discuss MAT effects relative to traditional state schools. Appendix Table A.3 presents 2SLS estimates of the impact of MAT and SAT relative to traditional state-funded schools. The results for KS2 average scores (Panel B, columns (1) and (2)) are similar to the findings in Eyles et al. (2017) and highlight that, while student achievement in SATs significantly worsens, MAT students do neither better nor worse than their peers in traditional state schools. Our results, however, also paint a more nuanced picture. Indeed, the table highlights that MAT schools are more successful than traditional state schools at improving math scores (about 0.028σ, significant at 10%), whereas the impact on language scores is 0.009σ and not statistically different from zero at any conventional level. The impact of SATs relative to traditional state schools is instead negative for both math and language test scores (-0.037 and -0.039σ respectively, consistently with Table 4). This implies that MATs improve student performance in one subject while leaving the other unaffected, whereas SAT schools perform worse than traditional state schools in both math and language.

We end our analysis by studying how different governance models affect school inputs. Table re-estimates Eq. (1) at school level, considering three categories: school composition, teaching staff and per-pupil expenditures. The composition of the last school grade - when the students considered in our analysis seat KS2 exams - is essentially unchanged, with no changes in the share of FSM eligible students and high achievers (columns (1) and (2)). Similarly, MATs do not differ from SATs in terms of the share of qualified teachers employed and pupil-to-teacher ratio (columns (3) and (4)). MATs, however, display substantial differences in school expenditures with respect to SATs, documented in columns (5) to (11). SATs increase per-pupil expenditures with respect to state schools (about £234) and spend substantially more for teaching staff, maintenance of school premises and back-office items. MATs, instead, spend less than SATs overall (about £83 less, significant at 10%). Although their teaching staff spending is similar, they spend less than SATs in learning and ICT resources (-£24) and back-office items (-£26). With respect to SATs they also spend more in development and training, although the difference is not significant. This pattern suggests that MATs benefit from economies of scale and can therefore be more efficient than SATs in the provision of certain services. MATs are successful at improving student achievement with respect to SATs by spending less overall and, in particular, by saving money in the provision of ICT and learning resources as well as back-office items.

4.2. Heterogeneous effects

The results presented mask some heterogeneity across different subgroups of children and schools. Understanding who benefited from MAT enrolment is particularly important in order to highlight whether the 2010 academy reform helped reducing the educational gap between students with different socio-economic backgrounds or baseline achievement. Fig. 5 plots 2SLS estimates of the impact of MATs on student achievement for different subsamples.

MAT schools seem to be slightly more effective at raising achievement for disadvantaged and native students. Students who are eligible

[18] As argued by Eyles et al. (2017), the reform proposal was first presented in April 2010 and implemented shortly afterwards.

[19] Appendix Table A.2 presents the full set of first stage estimates.

Table 4
Effect of MAT academies on student performance relative to SATs

	Average score		Math		Language	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. OLS						
Exposure to academy	0.018	-0.009	0.027	-0.002	0.009	-0.015
	(0.012)	(0.012)	(0.013)	(0.014)	(0.012)	(0.013)
Exposure to MAT academy	0.044	0.054	0.050	0.061	0.039	0.046
	(0.013)	(0.014)	(0.015)	(0.015)	(0.014)	(0.014)
Panel B. 2SLS						
Exposure to academy	-0.025	-0.038	-0.022	-0.037	-0.027	-0.039
	(0.013)	(0.014)	(0.014)	(0.015)	(0.013)	(0.014)
Exposure to MAT academy	0.049	0.057	0.055	0.066	0.042	0.048
	(0.014)	(0.015)	(0.016)	(0.016)	(0.014)	(0.015)
Panel C. First Stage						
Exposure to academy	0.934	0.934				
	(0.002)	(0.002)				
Exposure to MAT academy	0.936	0.936				
	(0.002)	(0.002)				
Observations	914,786	914,786	914,786	914,786	914,786	914,786
Student characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Baseline score	No	Yes	No	Yes	No	Yes
Event time FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes. The table shows OLS (Panel A) and 2SLS (Panel B) regressions of KS2 average scores (columns (1) and (2)), math scores (columns (3) and (4)) and language scores (columns (5) and (6)) on exposure to academy and MAT schools. Panel C (columns (1) and (2)) shows the corresponding first stage regressions. KS2 outcomes are standardised to have zero mean and unit variance. All columns control for student characteristics (gender, FSM eligibility, ethnicity, language spoken at home), event time, calendar year and school fixed effects. Columns (2), (4) and (6) add student baseline test scores, defined as the KS1 level obtained in math, writing and reading. Standard errors, shown in brackets, are clustered on schools.

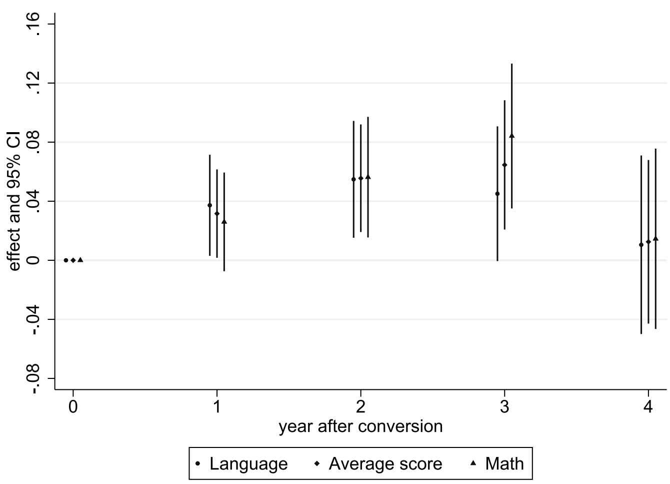


Fig. 4. Time-specific effects of MAT academies relative to SATs
Notes. The figure shows time-specific effects of attending a MAT with respect to a SAT school on student average, math and language test scores. The effects are obtained by estimating the 'pooled' version of Eq. (1) - see Table 4, Panel B - where the four post-treatment interactions are instrumented with the grandfathering instrument interacted with the four post-treatment time dummies. The vertical bars denote the 95% confidence interval. Appendix Fig. A.1 plots the corresponding time-specific estimates for SAT schools.

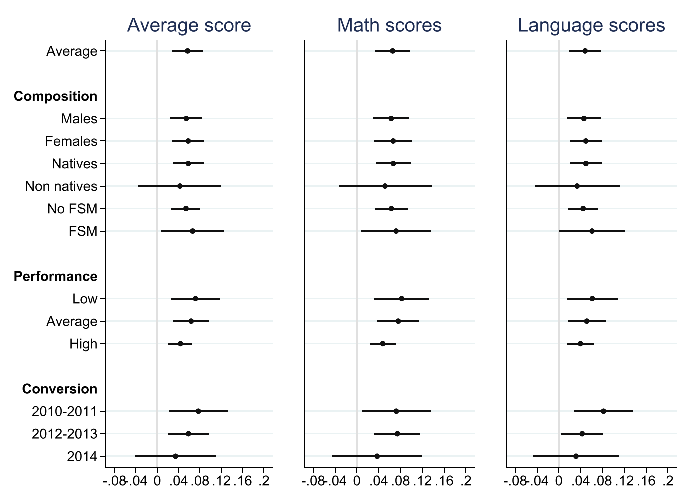


Fig. 5. Heterogeneous effects of MAT academies relative to SATs
Notes. The figure presents estimates of the impact of MAT school attendance on average, math and language test scores by different sub-groups. The top panel considers student characteristics and shows estimates for students by gender, language spoken at home (English or not) and for disadvantaged students (FSM eligible). The middle panel considers student baseline achievement. Students are defined as low, average and high achievers if they were awarded Level 1, 2 and 3 at KS1 respectively. The bottom panel considers the school conversion year. The horizontal lines show the 95% confidence interval.

for FSM gain about 0.066σ in terms of average test scores, while their non-eligible counterpart gain 0.054σ. Similarly, students who were low-achievers at baseline (KS1) experience a larger increase with respect to high-achievers (0.072σ vs 0.044σ). Although these differences are not statistically significant, they represent suggestive evidence that MATs seem to be slightly more effective for students who are more in need.

Academies converting very early seem also to be more effective than schools converting later. We split the sample in three waves - schools converted in 2010-2011, 2012-2013 and 2014. Appendix Table A.4 presents summary statistics for MAT and SAT schools belonging to the three waves.

Both early converter MAT and SAT schools are better in terms of student performance (e.g. share of top performer at KS2) with respect to late converter. Additionally, first wave's schools were more likely to be rated as 'outstanding' on a large number of characteristics by the Ofsted with respect to subsequent waves (see Appendix Table A.5), consistent with the initial stage of the reform that prioritised very good schools. Although the estimates for the three waves are not statistically different from each other, they suggest that early converters - and first wave in particular - were more effective at improving student achievement with respect to the last wave.

Table 6
School inputs by conversion wave: MAT academies relative to SATs

Dep. Var.:	School composition (shares)		Teachers		Per-pupil expenditure (£)						
	FSM	High achievers	Percent non-qualified	Pupil-to-teacher ratio	Total	Teaching staff	Educational support staff	Learning and ICT resources	Development and training	Premise maintenance	Back office
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Panel A. 2010-2011											
Exposure to academy	-0.003	0.003	0.002	0.053	199.200	36.984	10.949	13.055	-1.391	23.116	100.769
	(0.006)	(0.010)	(0.006)	(0.265)	(55.175)	(30.644)	(19.910)	(11.455)	(2.231)	(11.706)	(12.726)
Exposure to MAT academy	-0.006	0.007	0.004	-0.553	-269.673	-88.789	-47.244	-44.231	6.237	-20.906	-27.271
	(0.007)	(0.012)	(0.007)	(0.338)	(97.729)	(51.293)	(26.876)	(16.718)	(2.891)	(14.078)	(19.487)
Observations	5,297	5,297	4,739	4,743	4,741	4,713	4,532	4,725	4,644	4,741	4,741
Panel B. 2012-2013											
Exposure to academy	-0.011	0.004	0.002	-0.317	222.141	113.260	-10.082	-21.818	-1.102	6.522	90.187
	(0.004)	(0.007)	(0.004)	(0.190)	(47.263)	(22.972)	(17.484)	(8.372)	(1.518)	(10.176)	(13.695)
Exposure to MAT academy	0.005	-0.011	0.006	0.059	-41.167	1.369	7.561	-12.513	1.189	-2.333	-21.685
	(0.005)	(0.008)	(0.005)	(0.215)	(58.134)	(29.360)	(19.232)	(9.829)	(1.981)	(11.901)	(19.247)
Observations	13,732	13,732	12,455	12,465	12,321	12,294	12,043	12,265	12,317	12,321	12,321
Panel C. 2014											
Exposure to academy	-0.001	-0.007	0.006	0.060	551.516	204.210	71.413	24.430	0.887	93.725	85.944
	(0.010)	(0.013)	(0.010)	(0.396)	(92.301)	(54.702)	(35.212)	(21.084)	(3.706)	(37.094)	(28.231)
Exposure to MAT academy	0.007	0.005	-0.009	-0.289	-350.567	-61.596	-37.422	-51.701	-1.983	-87.861	-52.811
	(0.010)	(0.013)	(0.010)	(0.394)	(94.754)	(56.489)	(38.724)	(21.789)	(3.822)	(37.899)	(30.231)
Observations	6,682	6,682	6,081	6,095	6,043	6,043	6,001	6,011	6,043	6,043	6,043
Event time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. The table shows OLS regressions of school inputs on exposure to academy and MAT schools by conversion wave. Columns (1) and (2) consider school composition - FSM eligible and high achieving students, respectively. High achieving students are defined as students attaining Level 3 in at least one subject at the end of KS1. Columns (3) and (4) consider teachers' outcomes - percent of qualified teachers and pupil-to-teacher ratio, respectively. Columns (5) to (11) consider school expenditures for different items. School expenditure items' description in Table 5. The sample is at school/year/event time level. Only academies with at least 12 months of income and expenditure recorded in their accounts are included in the data published by the DfE. A further discrepancy between this sample and the main sample is that in the former not all the academies are consistently present in every year (e.g. because they miss the accounts deadline). All columns control for school, year and event time fixed effects. Standard errors, shown in brackets, are clustered on schools.

We then study, in Table 6, how school inputs changed in SATs and MAT schools across different conversion waves. Although the pattern documented in Table 5 generally holds, there are some notable differences across schools converting at different points in time. Early converter MATs spend significantly more than SATs on development and training (about £6.2) activities and substantially less in teaching and educational support staff. On the contrary, schools converting in 2012-2013 do not display substantial differences with respect to SATs. Late converter MAT schools spend less than SATs in learning and ICT resources (similarly to early converters) and premise maintenance, as well as back-office items (difference significant at 10%). Such findings suggest that schools converting across years differentially adjust their funding and cost patterns depending on their (different) needs.

Finally, we look at MAT-specific estimates of effectiveness (Fig. 6). The Figure is obtained as follows: i) we re-estimate the reduced-form of the model in Table 4 (Panel B) replacing the 'generic' treatment indicator with a full set of MAT school indicators; ii) we average school-specific estimate at MAT level and rank them in ascending order. From the Figure, it is apparent that effectiveness estimates are highly heterogeneous across MATs.

More effective MATs, however, did not necessarily take over more schools over time. In the same Figure, we plot the number of schools managed by MATs in 2022 (right-hand side axis). Although this Figure plots a simple correlation, it does suggest that high-performing MATs do not necessarily expand more (or less) than low-performing. We interpret this result as suggestive evidence that different MATs may exhibit preferences for a different 'optimal' size, and therefore may not necessarily choose to expand.[20]

4.3. Summary of main robustness checks

We end this Section by discussing two sets of robustness checks, one that relates to the control group and one to the set of controls, focusing on student performance. The results are presented in Table 7. Columns (2) includes the second closest school in the control group; column (3) drops any school within the 50*th* (instead of the 80*th*) percentile; column (4) keeps the two closest schools irrespective of their Ofsted grade. The main results for average scores presented in Table 7, which are also shown in column (1), are essentially unaffected and not sensitive to these alternative choices.

The second part of Table 7 documents that our results are not

[20] Fig. A.2 plots additional measures of MAT size, breaking down the number of schools by academy type (converter and sponsored) and plotting the number of pupils enrolled. These outcomes are measured using April 2022 Edubase.

Table 5
School inputs: MAT academies relative to SATs

Dep. Var.:	School composition (shares)		Teachers		Per-pupil expenditure (£)						
	FSM	High achievers	Percent non-qualified	Pupil-to-teacher ratio	Total	Teaching staff	Educational support staff	Learning and ICT resources	Development and training	Premise maintenance	Back office
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Exposure to academy	-0.007	0.003	0.002	-0.156	233.794	93.697	4.117	-7.149	-1.230	18.041	89.293
	(0.003)	(0.005)	(0.003)	(0.147)	(35.266)	(18.071)	(12.754)	(6.476)	(1.206)	(7.850)	(9.534)
Exposure to MAT academy	0.003	-0.005	0.003	-0.139	-83.401	2.609	-2.530	-24.121	2.092	-13.128	-26.009
	(0.003)	(0.006)	(0.004)	(0.162)	(44.901)	(23.155)	(14.429)	(7.669)	(1.461)	(8.944)	(13.010)
Observations	25,711	25,711	23,275	23,303	23,105	23,050	22,576	23,001	23,004	23,105	23,105
Event time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. The table shows OLS regressions of school inputs on exposure to academy and MAT schools. Columns (1) and (2) consider school composition - FSM eligible and high achieving students, respectively. High achieving students are defined as students attaining Level 3 in at least one subject at the end of KS1. Columns (3) and (4) consider teachers' outcomes - percent of qualified teachers and pupil-to-teacher ratio, respectively. Qualified teachers are those who have obtained Qualified Teacher Status (QTS). In academies, the QTS is not a legal requirement to teach since 2012. Columns (5) to (11) consider school expenditures for different items. Total expenditures (column (5)) includes any school expenditure. Educational support staff includes all staff who are not teachers (e.g. teaching assistants, laboratory technician); premises costs include cleaners, security staff, etc; back-office costs include administrative staff, supplies and legal and professional services. The number of observations in columns (6) to (11) may differ from column (5) when the expenditures for the specific item are reported as missing. The sample is at school/year/event time level. Only academies with at least 12 months of income and expenditure recorded in their accounts are included in the data published by the DfE. A further discrepancy between this sample and the main sample is that in the former not all the academies are consistently present in every year (e.g. because they miss the accounts deadline). All columns control for school, year and event time fixed effects. Standard errors, shown in brackets, are clustered on schools.

sensitive to the inclusion of alternative or additional sets of controls. Column (5) controls for standardised point scores in math, writing and language instead of a set of indicator variables for the level awarded (as in Table 4). Column (6) adds student residence (LSOA) fixed effects; finally, column (7) adds LA-specific time trends. Column (6) estimates are particularly interesting as they further control for student selection across neighbourhoods. Again, the main results are unaffected.

5. Governance and management practices

Our results suggest that MATs have a positive impact on student achievement compared to SATs. This positive impact was accompanied by different expenditure patterns, which suggest that MATs can be more efficient than SATs. These findings highlight that policies aimed at increasing school autonomy may require particular organisational arrangements to be effective and a one-size-fits-all approach is not likely to work. In this section we explore possible mechanisms that may explain the success of MATs, focusing on the role played by managerial practices. We exploit highly detailed data from a unique survey recently conducted by the DfE, the "Academy Trust Survey 2017" (Cirin, 2017).

The survey focused on practices of MAT and SAT schools converted into academies before February 2016. As the previous DfE survey ("Do academies make use of their autonomy?", Cirin,2014) focused on academies without distinguishing MAT and SAT academies, this is the first available information on how the latter differ in terms of practices since their introduction.[21] The survey was conducted on 326 MATs and 542 SATs, both primary and secondary schools. We consider the subset of respondents that includes all MATs with at least two schools and one primary schools (237) and primary SATs (167). We also consider an additional subset of respondents represented by the 129 schools that became SAT academies between 2010 and 2014. Although we cannot link the respondent schools to those in our sample, the former sample can be considered an approximation of the latter. We cannot distinguish MATs by conversion date in the survey data.

For the purpose of our study, we focus on the changes implemented by trusts after conversion. Table 8 shows that school chains and standalone academies differ quite significantly in terms of the changes they put in place. In particular, while SAT schools are more likely to make changes at school level, MATs are more likely to make organizational-level changes, mostly related to the reconstitution of the governing body.

When asked to rank the 5 most important changes made after conversion (Table 8, columns (1) to (3)), 61% and 62% of MATs mentioned changes in school leadership and reconstituting the governing body respectively, compared to 25% and 34% of SATs (26% and 34% in the 2010-2014 subsample). 27% of MATs also mentioned changes in the performance management system for teachers, compared to 14% of SAT schools. More than 60% of SAT schools, instead, mentioned changing the curriculum as one of the most important changes available, compared to 36% of MATs.[22] SAT schools also mention to have introduced or increased revenue-generating activities more than MATs (44% and 27% respectively).

Columns (4) to (6) show the most important change implemented after conversion. Once again MATs are more likely to mention changing school leadership (24%) and reconstituting the governing body (14%) as the most important change. SAT schools, instead, mention changing the curriculum (29%) and the procurement of services previously provided by the LA (26%). Overall, these figures show that academy chains prioritised changes at the managerial level rather than focusing on traditional school level changes, such as changing the curriculum offered or school day length. This suggests that the implementation of different managerial practices between MATs and SATs may explain the

[21] Only two questions in "Do academies make use of their autonomy?" are further split by MAT and SAT status, the first covering the reasons and 'main reason' for conversion, and the second covering the type of support provided by MAT schools to other schools.

[22] We note, however, that survey responses for MAT schools were completed at MAT level rather than by each school belonging to the chain, potentially leading to some under-reporting of school-specific curriculum changes within MATs.

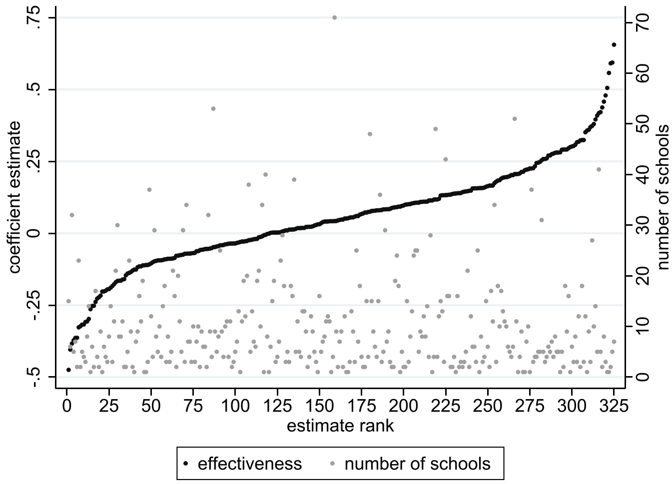


Fig. 6. MAT performance and school takeover
Notes. The figure plots MAT-specific estimates of the impact of MAT schools on student performance (black dots) against the number of schools belonging to the MAT in 2022 (grey dots). MAT estimates are obtained by i) estimating the reduced-form of the model in Table 4 (Panel B) and replacing the treatment indicator with a full set of MAT school indicators and ii) averaging the school-specific effects by MAT. Coefficient estimates are ranked in ascending order. The number of schools includes converter and sponsored academies covering the primary and secondary school phase. Appendix Fig. A.2 replicates this figure separating converter and sponsored academies (Panel A) and for the number of students enrolled (Panel B).

differences in performance and efficiency after conversion.

6. Conclusions

While most of the previous literature has mainly focused on the effectiveness of autonomous schools, this paper sheds light on potential mechanisms underpinning their success. We exploit a recent reform introduced in England that gave primary schools the opportunity to become autonomous and choose the preferred model of governance. Following the reform, the rapid expansion of chains of autonomous schools, the so-called MATs, brought in a new model of governance characterised by the separation of roles and responsibilities along the governance chain. Focusing on primary schools, we explore whether the performance and efficiency of academies belonging to chains differ compared to standalone academies.

We show that pupils exposed to schools belonging to chains perform better in both math and language. Exposure to a MAT compared to a SAT school increases test scores by 0.066-0.048σ in math and language, respectively (about 1.32 and 0.37 points). Effects are stronger for low achievers and early converters. Although schools still face a non-trivial choice to become academies, such positive effects indicate that MATs have the potential to improve student achievement. They do so while spending less overall and, in particular, for the provision of ICT and learning resources and back-office items. We also present suggestive evidence that more successful MATs do not necessarily take over more schools. Such findings highlights that studying the dynamics of large scale educational reforms is important to evaluate their overall effectiveness (see Baude et al., 2020).

Less obvious is the mechanism underpinning our findings. Using recent survey data collected by the DfE, we show that while SATs are more likely to make changes at school level (e.g. changing the curriculum offered, introducing revenue generating activities, adding non-teaching positions), MATs are more likely to make changes related to managerial practices (e.g. reconstituting the governing body, changing the school leadership, creating formal networks between schools). In its 2022 Schools White Paper, the English government set out its aim of expanding the academy program, expecting that MATs will "run at least 10 schools". Overall, our results suggest that policy-makers should be cautious in pushing for an increase in school chains' size and may want to focus instead on interventions at the managerial level, which may improve school effectiveness and thereby student outcomes. In addition, our evidence points to substantial heterogeneity across MATs in affecting student achievement. Hence, understanding 'what works' within MATs is important to design effective schools. We hope to address this in future works.

Table 7
Effect of MAT academies on student performance relative to SATs: robustness checks

Dep. Var.: average score	Main sample	Control groups Use 2 closest schools	Drop within 50th pctile	Closest irrespective of Ofsted grade	Control variables Standardised scores	Residence FE	LA trends
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Exposure to academy	-0.038	-0.027	-0.024	-0.058	-0.040	-0.038	-0.036
	(0.014)	(0.013)	(0.014)	(0.014)	(0.014)	(0.014)	(0.013)
Exposure to MAT academy	0.057	0.057	0.055	0.062	0.057	0.059	0.058
	(0.015)	(0.014)	(0.015)	(0.014)	(0.015)	(0.014)	(0.014)
Observations	914,786	1,330,375	940,784	949,431	914,786	908,859	908,859
LA-specific time trends	No	No	No	No	No	No	Yes
Student residence (LSOA) FE	No	No	No	No	No	Yes	Yes
Student characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline score	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Event time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. The table shows 2SLS regressions of KS2 average scores on exposure to academy and MAT schools. The outcome is standardised to have zero mean and unit variance. All columns control for student characteristics (gender, FSM eligibility, ethnicity, language spoken at home), event time, calendar year and school fixed effects. Column (1) shows the main estimate from Table 4, Panel B (Column (2)). Column (2) considers an alternative control group including the two closest schools to a treated academy. Column (3) drops nearby schools located within the 50th percentile of the student-school LA-specific distance distribution. Column (4) keeps the closest school to a treated school irrespective of the Ofsted grade. Column (5) controls for pseudo-continuous KS1 scores in math, writing and reading instead of KS1 level indicators. Column (6) add student residence (LSOA) fixed effects. Column (7) adds LA-specific time trends. Standard errors, shown in brackets, are clustered on schools.

Table 8
Most important changes introduced by MATs and SATs

	Top 5 most important changes			Most important change		
	Multi-Academy Trusts	Single-Academy Trusts		Multi-Academy Trusts	Single-Academy Trusts	
		2010-2015	2010-2014		2010-2015	2010-2014
	(1)	(2)	(3)	(4)	(5)	(6)
Procuring services that were previously provided by the LA	68%	77%	78%	11%	26%	26%
Changing the curriculum you offer	36%	62%	63%	10%	29%	29%
Introducing back-office savings (e.g. human resources, ICT, payroll)	78%	60%	60%	27%	9%	9%
Introducing or increasing revenue-generating activities	27%	44%	44%	1%	6%	6%
Changing the pattern of capital expenditure	31%	43%	43%	3%	9%	9%
Reconstituting the governing body	62%	34%	34%	14%	4%	4%
Changing school leadership	61%	25%	26%	24%	4%	4%
Adding non-teaching positions	12%	19%	19%	0%	0%	0%
Changing staff pay structures	19%	17%	18%	2%	1%	2%
Changing the admission criteria	11%	17%	16%	1%	2%	2%
Increasing the number of pupils on roll	18%	16%	16%	10%	29%	29%
Changing the length of school terms	7%	15%	15%	0%	1%	1%
Changing the performance management system for teachers	27%	14%	14%	3%	1%	1%
Hiring teachers without qualified teacher status (QTS)	8%	8%	8%	1%	1%	2%
Seeking to attract pupils from a different geographical area	3%	6%	5%	0%	1%	1%
Changing the length of the school day	5%	4%	3%	n/a	n/a	n/a
Number of respondents	237	140	129	237	140	129

Note. The table presents the proportion of MATs and SATs that endorse a change as being one of the five most important (columns (1) to (3)) and the most important (columns (4) to (6)) available to them after conversion. The sample of respondents includes 237 MATs with at least two schools and one primary school and 140 primary SAT schools. Other changes includes pay staff structure, admission criteria, hiring teachers without qualified teacher status (QTS), and seeking pupils from a different geographical area. Data source: Academy trust survey 2017.

Data availability

The authors do not have permission to share data.

Acknowledgements

Special thanks to Erich Battistin and Barbara Petrongolo for their invaluable support and guidance. We would like to thank Robert Cirin for his help in accessing survey data on school practices. We also thank Sebastian Axbard, Sarah Cattan, Gianni De Fraja, Francesco Fasani, Andrea Ichino, Steve Machin, Marco Manacorda, Sandra McNally, Marco Ovidi, Parag Pathak, Anna Raute, Olmo Silva, Andrea Tesei, David Wilkinson, and seminar participants at MIT Sloan, LSE, Queen Mary University, the EEA-ESEM 2018 Congress, the LSE-Oxford Graduate Conference, and the RES 2018 Annual Conference for many helpful comments. This work was supported by the Economic and Social Research Council. The views expressed here are those of the authors alone.

Appendix A. Additional Tables and Figures

/

Table A.1
Effect of MAT academies on student performance relative to SATs (no 2010 boycotters)

	Average score		Math		Language	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. OLS						
Exposure to academy	0.011	-0.013	0.020	-0.006	0.002	-0.020
	(0.014)	(0.014)	(0.015)	(0.016)	(0.014)	(0.015)
Exposure to MAT academy	0.036	0.044	0.038	0.047	0.035	0.040
	(0.015)	(0.016)	(0.017)	(0.018)	(0.015)	(0.016)
Panel B. 2SLS						
Exposure to academy	-0.031	-0.044	-0.027	-0.042	-0.035	-0.046
	(0.015)	(0.016)	(0.017)	(0.018)	(0.015)	(0.016)
Exposure to MAT academy	0.039	0.045	0.041	0.048	0.037	0.041
	(0.016)	(0.017)	(0.018)	(0.018)	(0.016)	(0.017)
Panel C. First Stage						
Exposure to academy	0.934	0.934				
	(0.002)	(0.002)				
Exposure to MAT academy	0.937	0.937				
	(0.002)	(0.002)				
Observations	680,169	680,169	680,169	680,169	680,169	680,169
Student characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Baseline score	No	Yes	No	Yes	No	Yes
Event time FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes. The table shows OLS (Panel A) and 2SLS (Panel B) regressions of KS2 average scores (columns (1) and (2)), math scores (columns (3) and (4)) and language scores (columns (5) and (6)) on exposure to academy and MAT schools. Panel C (columns (1) and (2)) shows the corresponding first stage regressions. The sample excludes schools that boycotted 2010 KS2 tests, proxied with schools having all missing scripts in 2010. KS2 outcomes are standardised to have zero mean and unit variance. All columns control for student characteristics (gender, FSM eligibility, ethnicity, language spoken at home), event time, calendar year and school fixed effects. Columns (2), (4) and (6) add student baseline test scores, defined as the KS1 level obtained in math, writing and reading. Standard errors, shown in brackets, are clustered on schools.

Table A.2
First stage estimates

	Exposure to: Academy	MAT academy	Exposure to MAT academy after: One year	Two years	Three years	Four years
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Average exposure						
Grandfathering	0.934	-0.003				
	(0.002)	(0.000)				
Grandfathering x MAT	-0.004	0.936				
	(0.003)	(0.002)				
Panel B. Exposure to MAT by year						
Grandfathering	0.934		-0.001	-0.001	-0.001	-0.001
	(0.002)		(0.000)	(0.000)	(0.000)	(0.000)
Grandfathering x MAT after:						
One year	0.021		0.968	0.001	0.001	0.000
	(0.002)		(0.001)	(0.000)	(0.000)	(0.000)
Two years	-0.001		0.000	0.938	0.001	0.001
	(0.003)		(0.000)	(0.002)	(0.000)	(0.000)
Three years	-0.021		0.000	0.001	0.912	0.001
	(0.004)		(0.000)	(0.000)	(0.003)	(0.000)
Four years	-0.045		0.000	0.001	0.001	0.879
	(0.006)		(0.000)	(0.000)	(0.000)	(0.005)
Observations	914,786	914,786	914,786	914,786	914,786	914,786
Student characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Baseline score	Yes	Yes	Yes	Yes	Yes	Yes
Event time FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes. The table shows the full set of first stage estimates. Panel A presents the average estimates and coincide with those shown in Table 4, Panel C. For MAT schools, Panel B breaks down the effect by the year in which the student seats KS2 tests after a school conversion. All columns control for student characteristics (gender, FSM eligibility, ethnicity, language spoken at home), student baseline test scores (KS1 level obtained in math, writing and reading and schools), and school calendar year and event time fixed effects. Standard errors, shown in brackets, are clustered on schools

Table A.3
Effect of SAT and MAT academies relative to traditional state schools

	Average score		Math		Language	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. OLS						
Exposure to SAT academy	0.018	-0.009	0.027	-0.002	0.009	-0.015
	(0.012)	(0.012)	(0.013)	(0.014)	(0.012)	(0.013)
Exposure to MAT academy	0.063	0.045	0.077	0.059	0.048	0.031
	(0.012)	(0.012)	(0.013)	(0.013)	(0.012)	(0.012)
Panel B. 2SLS						
Exposure to SAT academy	-0.025	-0.038	-0.022	-0.037	-0.027	-0.039
	(0.013)	(0.014)	(0.014)	(0.015)	(0.013)	(0.014)
Exposure to MAT academy	0.024	0.019	0.034	0.028	0.015	0.009
	(0.013)	(0.013)	(0.015)	(0.015)	(0.013)	(0.013)
Panel B. First Stage						
Exposure to SAT academy	0.937	0.937				
	(0.002)	(0.002)				
Exposure to MAT academy	0.933	0.933				
	(0.002)	(0.002)				
Observations	914,786	914,786	914,786	914,786	914,786	914,786
Student characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Baseline score	No	Yes	No	Yes	No	Yes
Event time FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes. The table shows OLS (Panel A) and 2SLS (Panel B) regressions of KS2 average scores (columns (1) and (2)), math scores (columns (3) and (4)) and language scores (columns (5) and (6)) on exposure to a SAT and MAT academy, relative to traditional state schools. Panel C (columns (1) and (2)) shows the corresponding first stage regressions. KS2 outcomes are standardised to have zero mean and unit variance. All columns control for student characteristics (gender, FSM eligibility, ethnicity, language spoken at home), event time, calendar year and school fixed effects. Columns (2), (4) and (6) add student baseline test scores, defined as the KS1 level obtained in math, writing and reading. Standard errors, shown in brackets, are clustered on schools.

Table A.4
Academy characteristics by year of conversion

	MATs First wave	Second wave	Third wave	SATs First wave	Second wave	Third wave
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Students						
Male	0.51	0.50	0.51	0.51	0.51	0.50
Eligible for free school meals (FSM)	0.12	0.15	0.11	0.08	0.12	0.15
White	0.88	0.86	0.90	0.88	0.83	0.87
Black	0.05	0.02	0.01	0.02	0.03	0.03
Native	0.95	0.91	0.94	0.93	0.89	0.94
With special educational needs (SEN)	0.21	0.21	0.22	0.17	0.20	0.20
Panel B. Scores						
KS1 math Level 3	0.26	0.25	0.25	0.31	0.27	0.26
KS1 reading Level 3	0.30	0.28	0.30	0.35	0.31	0.30
KS2 math Level 5	0.42	0.38	0.39	0.47	0.43	0.39
KS2 reading Level 5	0.57	0.51	0.54	0.60	0.57	0.52
KS2 average score	50.01	48.43	48.80	51.41	50.08	48.31
KS2 math scores	69.23	67.18	67.34	71.52	69.50	66.71
KS2 language scores	30.79	29.67	30.26	31.30	30.67	29.92
Panel C. Schools						
Community	0.72	0.59	0.44	0.49	0.56	0.70
Voluntary Controlled	0.11	0.08	0.11	0.10	0.16	0.05
Voluntary Aided	0.08	0.29	0.42	0.21	0.20	0.23
KS2 grade enrolment	41.64	35.43	31.82	45.13	39.65	41.33
Pupil-teacher ratio	22.06	21.94	21.48	22.53	22.37	22.53
Percent qualified teachers	0.95	0.96	0.95	0.95	0.95	0.96
Number of schools	116	375	255	144	251	43
Number of students	48,798	151,405	92,824	66,487	114,220	21,043

Notes. The table presents characteristics for MAT (columns (1) to (3)) and SAT schools (columns (4) to (6)) converted in 2010/2011 (first wave), 2012/2013 (second wave) and 2014 (third wave)columns (1) and (2)). Means are computed for 2009, the last pre-reform year.

Table A.5
Fraction of schools by Ofsted judgement and conversion wave

	Outstanding First wave	Second wave	Third wave	Good First wave	Second wave	Third wave	Satisfactory First wave	Second wave	Third wave
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. Multi-academy trusts (MAT)									
Overall grade	0.32	0.17	0.15	0.45	0.42	0.47	0.22	0.38	0.38
Behaviour and safety of pupils	0.53	0.33	0.37	0.44	0.60	0.58	0.04	0.07	0.05
Quality of teaching	0.26	0.15	0.13	0.52	0.46	0.50	0.21	0.38	0.36
Quality of pupils' learning	0.29	0.17	0.14	0.49	0.43	0.48	0.21	0.38	0.38
Quality of SEN pupils' learning	0.30	0.18	0.15	0.50	0.51	0.55	0.19	0.30	0.30
Pupils' attendance	0.11	0.12	0.12	0.52	0.46	0.50	0.34	0.39	0.37
Leadership and management	0.34	0.20	0.16	0.46	0.45	0.48	0.19	0.33	0.35
Effectiveness of Governing Body	0.31	0.16	0.14	0.48	0.48	0.49	0.20	0.33	0.35
Panel B. Single-academy trusts (SAT)									
Overall grade	0.49	0.19	0.14	0.37	0.54	0.62	0.13	0.27	0.21
Behaviour and safety of pupils	0.69	0.45	0.48	0.30	0.52	0.48	0.01	0.02	0.05
Quality of teaching	0.44	0.15	0.12	0.42	0.58	0.64	0.13	0.27	0.24
Quality of pupils' learning	0.45	0.16	0.12	0.41	0.56	0.64	0.13	0.28	0.21
Quality of SEN pupils' learning	0.47	0.18	0.14	0.42	0.60	0.71	0.11	0.22	0.12
Pupils' attendance	0.25	0.16	0.17	0.57	0.49	0.50	0.18	0.34	0.33
Leadership and management	0.52	0.19	0.14	0.36	0.56	0.62	0.11	0.25	0.24
Effectiveness of Governing Body	0.44	0.18	0.24	0.42	0.53	0.45	0.14	0.28	0.31

Notes. The table shows the fraction of schools by Ofsted judgement and wave of conversion. The sample includes schools that became academies between July 2010 and December 2014. Only Converter academies are considered. Panel A presents grades for multi-academy trusts (MATs) and Panel B for single-academy trusts (SATs). Columns (1) to (3) show the fraction of schools judged outstanding; columns (4) to (6) show the fraction of schools judged good; columns (7) to (9) show the fraction of schools judged satisfactory or inadequate. The table uses the last available Ofsted ranking before the conversion took place.

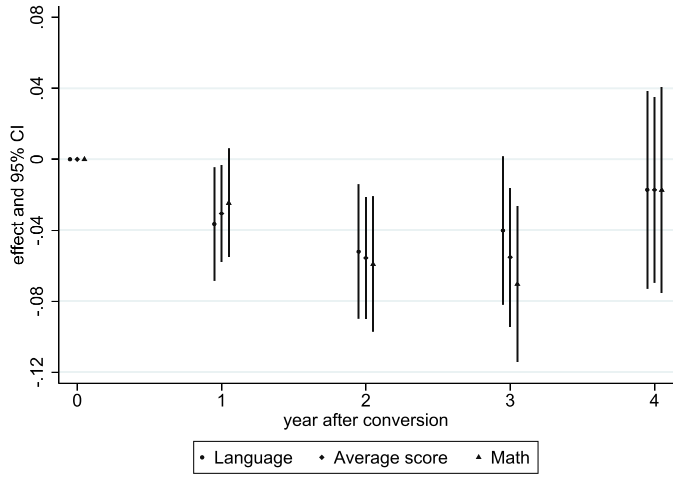


Fig. A.1. Time-specific effects for SATs relative to MAT academies
Notes. The figure shows time-specific effects of attending a SAT with respect to a state-funded non autonomous school on student average, math and language test scores. The effects are obtained by estimating the 'pooled' version of Eq. (1) - see Table 4, Panel B - where the four post-treatment interactions are instrumented with the grandfathering instrument interacted with the four post-treatment time dummies. The vertical bars denote the 95% confidence interval.

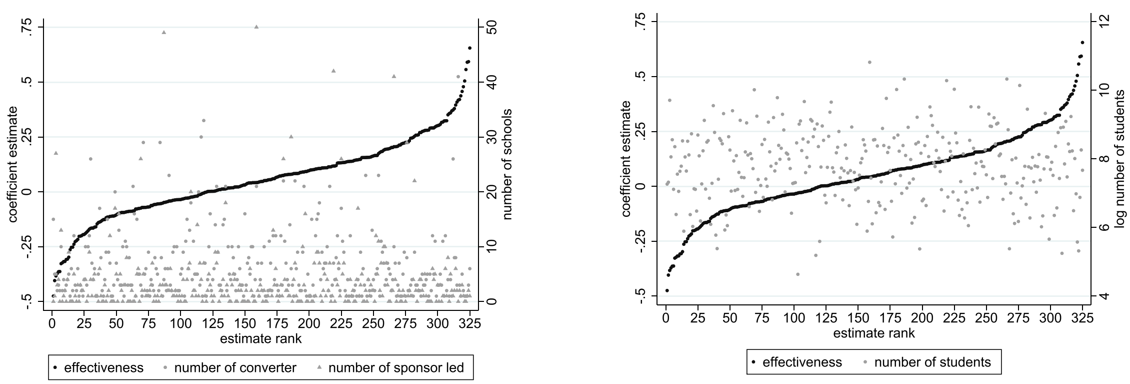


Fig. A.2. MAT performance, school takeover and number of students
Notes. Panel A of the figure plots MAT-specific estimates of the impact of MAT schools on student performance (black dots) against the number of converter (circle) and sponsored (triangle) schools belonging to the MAT in 2022. Panel B replicates Panel A but plotting the number of students enrolled in the MAT instead. Coefficient estimates are ranked in ascending order. MAT estimates are obtained by i) estimating the reduced-form version of regression (1) and replacing the treatment indicator with a full set of MAT school indicators and ii) averaging the school-specific effects by MAT.

References

Abdulkadiroğlu, A., Angrist, J. D., Dynarski, S. M., Kane, T. J., & Pathak, P. A. (2011). Accountability and flexibility in public schools: Evidence from Boston's charters and pilots. *The Quarterly Journal of Economics, 126*(2), 699–748.

Abdulkadiroğlu, A., Angrist, J. D., Hull, P. D., & Pathak, P. A. (2016). Charters without lotteries: Testing takeovers in New Orleans and Boston. *The American Economic Review, 106*(7), 1878–1920.

Andrews, J., & Perera, N. (2017). *Edu- cation Policy Institute Consolidated Report*.

Angrist, J. D., Cohodes, S. R., Dynarski, S. M., Pathak, P. A., & Walters, C. R. (2016). Stand and deliver: effects of Boston's charter high schools on college preparation, entry, and choice. *The Journal of Labor Economics*.

Angrist, J. D., Dynarski, S. M., Kane, T. J., Pathak, P. A., & Walters, C. R. (2010). Inputs and impacts in charter schools: KIPP Lynn. *American Economic Review, 100*(2), 239–243.

Angrist, J. D., Pathak, P. A., & Walters, C. R. (2013). Explaining charter school effectiveness. *American Economic Journal: Applied Economics, 5*(4), 1–27.

Baker, A., Larcker, D. F., and Wang, C. C. (2021). How much should we trust staggered difference-in-differences estimates? *Available at SSRN 3794018*.

Baude, P. L., Casey, M., Hanushek, E. A., Phelan, G. R., & Rivkin, S. G. (2020). The evolution of charter school quality. *Economica, 87*(345), 158–189.

Bettinger, E. P. (2005). The effect of charter schools on charter students and public schools. *Economics of Education Review*, 24(2):133–147.

Bloom, N., Lemos, R., Sadun, R., & Reenen, J. V. (2015). Does management matter in schools? *The Economic Journal, 125*(584), 647–674.

Böhlmark, A., & Lindahl, M. (2015). Independent schools and long-run educational outcomes: evidence from Sweden's large-scale voucher reform. *Economica, 82*(327), 508–551.

Booker, K., Gilpatric, S. M., Gronberg, T., & Jansen, D (2008). The effect of charter schools on traditional public school students in Texas: Are children who stay behind left behind? *Journal of Urban Economics, 64*(1), 123–145.

Borusyak, K., & Jaravel, X. (2021). Revisiting event study designs: Robust and efficient estimation. *Working paper*.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics, 225*(2), 200–230.

Cengiz, D., Dube, A., Lindner, A., & Zipperer, B. (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics, 134*(3), 1405–1454.

Cirin, R. (2014). Do academies make use of their autonomy? *Research Report*.

Cirin, R. (2017). Academy trust survey 2017. *Research Report*.

Cordes, S. A. (2018). In pursuit of the common good: The spillover effects of charter schools on public school students in New York City. *Education Finance and Policy, 13* (4), 484–512.

De Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review, 110*(9), 2964–2996.

Department for Education (2016a). Educational excellence everywhere – The education white paper.

Department for Education. (2016b). Schools, pupils and their characteristics. *Statistical First Release*, 20.

Deshpande, M., & Li, Y. (2019). Who is screened out? Application costs and the targeting of disability programs. *American Economic Journal: Economic Policy, 11*(4), 213–248.

Dobbie, W., & Fryer, R. G. (2011). Are high-quality schools enough to increase achievement among the poor? Evidence from the Harlem children's zone. *American Economic Journal: Applied Economics, 3*(3), 158–187.

Dobbie, W., & Fryer, R. G. (2013). Getting beneath the veil of effective schools: Evidence from New York City. *American Economic Journal: Applied Economics, 5*(4), 28–60.

Dobbie, W., & Fryer, R. G. (2015). The medium-term impacts of high-achieving charter schools. *Journal of Political Economy, 123*(5), 985–1037.

Dynarski, S., Hubbard, D., Jacob, B., and Robles, S. (2018). Estimating the effects of a large for-profit charter school operator. *NBER Working Paper No. 24428*.

Eberts, R. W. and Hollenbeck, K. (2001). An examination of student achievement in Michigan charter schools.

Eyles, A., Hupkau, C., & Machin, S. (2016). Academies, charter and free schools: Do new school types deliver better outcomes? *Economic Policy, 31*(87), 453–501.

Eyles, A., & Machin, S. (2019). The introduction of academy schools to England's education. *Journal of the European Economic Association, 17*, 1107–1146.

Eyles, A., Machin, S., & McNally, S. (2017). Unexpected school reform: academisation of primary schools in England. *Journal of Public Economics, 155*, 108–121.

Eyles, A., Machin, S., & Silva, O. (2018). Academies 2–the new batch: The changing nature of academy schools in england. *Fiscal Studies, 39*(1), 121–158.

Fadlon, I., & Nielsen, T. H. (2019). Family health behaviors. *American Economic Review, 109*(9), 3162–3191.

Fryer, R. G. (2014). Injecting charter school best practices into traditional public schools: Evidence from field experiments. *Quarterly Journal of Economics, 129*(3), 1355–1407.

Gleason, P. M., Clark, M. A., Tuttle, C. C., & Dwoyer, E. (2010). The evaluation of charter school impacts final report. *Education*.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. Technical Report 2. *National Bureau of Economic Research*.

Grotberg, A. and Robb, M. (2015). Education in chains.

Hoxby, C. M. and Murarka, S. (2009). Charter schools in New York City: who enrolls and how they affect their students' achievement. *NBER Working Paper No. 14852*.

Hutchings, M., Francis, B., & Kirby, P. (2016). The impact of academy chains on low income students. *The Sutton Trust*.

Imberman, S. A. (2011). The effect of charter schools on achievement and behavior of public school students. *Journal of Public Economics, 95*(7-8), 850–863.

Ladd, H. F., & Fiske, E. B. (2016). England confronts the limits of school autonomy. *National Center for the Study of Privatization in Education. Working Paper 232*.

Machin, S. J., & Sandi, M. (2020). Autonomous schools and strategic pupil exclusion. *The Economic Journal, 130*, 125–159.

Ni, Y. (2009). The impact of charter schools on the efficiency of traditional public schools: Evidence from Michigan. *Economics of Education Review, 28*(5), 571–584.

Regan-Stansfield, J. (2018). Does greater primary school autonomy improve pupil attainment? evidence from primary school converter academies in England. *Economics of Education Review, 63*, 167–179.

Ridley, M., & Terrier, C. (2018). Fiscal and education spillovers from charter school expansion. *Working Paper 25070*. National Bureau of Economic Research.

Sass, T. R. (2006). Charter schools and student achievement in Florida. *Education Finance and Policy, 1*(1), 91–122.

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics, 225*(2), 175–199.

The House of Commons Library. (2022). March 2022 schools white paper (England). *Com- Mons Library Research Briefing*.

Walters, C. R. (2018). The demand for effective charter schools. *Journal of Political Economy*. Forthcoming.

West, A., & Wolfe, D (2018). Academies, the school system in England and a vision for the future. *Clare Market Papers. No. 23*.

Winters, M. A. (2012). Measuring the effect of charter schools on public school student achievement in an urban environment: Evidence from New York City. *Economics of Edu- cation review, 31*(2), 293–301.

Woodworth, J. L., Raymond, M. E., Han, C., Negassi, Y., Richardson, W., & Snow, W. (2017). Charter management organizations. *Center for Research on Education Outcomes (CREDO)*.

Zimmer, R., & Buddin, R. (2009). Is charter school competition in California improving the performance of traditional public schools? *Public Administration Review, 69*(5), 831–845.