ENGLISH ACADEMIES' IMPACT ON PUPILS: COGNITIVE AND NON-COGNITIVE SKILLS, SUBJECT CHOICE, AND SCHOOL MANAGEMENT PRACTICES

Nuno Manuel da Costa Braz

Thesis submitted to the Social Research Institute

in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

UNIVERSITY COLLEGE LONDON

August 2024

I, Nuno Manuel da Costa Braz, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Date: 12 August 2024

ABSTRACT

Academies are independent state-funded schools that are managed outside of Local Authority control and enjoy greater autonomy. While most of the empirical literature on academies focuses on exam performance, this thesis contains three papers that evaluate the impact of academy conversion on a wider range of outcomes.

The first paper examines the effect of academy attendance on decision-making skills, mental health, and social behaviour. Using the Millennium Cohort Study (MCS) in a difference in difference framework, my analysis focuses on pupils who are already attending the school prior to academy conversion. The overall impact of academy attendance is non-significant, but there are marked differences in outcomes between converter and sponsor-led academies. My analysis suggests that transfer to converter academy status is associated with significantly raised decision-making skills for their pupils, whereas conversion to sponsor-led academy status is associated with an increase in pupils' self-esteem. These results could suggest that sponsor-led academies are using ranking maximising strategies and converter academies are not targeting non-cognitive outcomes.

The second paper, also using the MCS and legacy enrolment, explores how academy conversion shapes the subjects that pupils choose at age 14. The overall impact of academy attendance is driven by converter academies. Pupils at converter academies are significantly more likely to study science subjects and facilitating subjects than their peers at maintained schools. Since

converter academies have more advantaged intakes, these results may raise concerns for social mobility.

The third paper explores how academy conversion shapes school policies, management practices, and the learning environment. Using the Programme for International Student Assessment (PISA) data, a focus is placed on maintained schools that convert to academy status. Academies are compared to schools that become academies after the sample period. Academies develop distinctive managerial structures, learning environments, and school policies. Converter and sponsor-led academies have distinct admission policies, use diverse teaching methods, focus on different subjects, face dissimilar problems, but generally share a positive school climate. Multi-academy trusts have fostered strong leadership and management practices.

Keywords: academies, school autonomy, cognitive and non-cognitive skills, GCSE subject choice, management.

IMPACT STATEMENT

The research presented in this thesis may have both an academic and policy impact. It helps to improve understanding of education quasi-markets, focusing on the impact of giving schools autonomy. This has been a policy priority in England and follows an international trend of market-based reforms in education. Given the politically charged nature of this policy and the scale of its implementation, this work provides an important contribution to understanding what drives an autonomous school system and its wider outcomes beyond test scores. Having empirical evidence on the impact of autonomy is important if the school system is to deliver the results policymakers seek.

In chapter 3, I show that pupils at converter academies significantly improve their decision-making skills in relation to their peers at maintained schools. Chapter 4 also shows that pupils at these schools are more likely to take science subjects and facilitating subjects at age 14. On the other hand, despite improvements in exam performance after conversion, pupils at sponsor-led academies have similar cognitive skills and make similar subject choices to peers at maintained schools. Given increased social stratification among academies (converter academies have more advantaged intakes), these results have implications for social mobility and need to be addressed by policymakers, potentially improving social cohesion and welfare.

Chapter 3 also shows that sponsor-led academies improve their pupils' self-esteem. This is an interesting result, but that is not observed at converter academies. In effect, there are no meaningful differences in mental health, social behaviour or risk-taking between pupils at academies and those at maintained schools. The similarity is consistent with a lack of quasi-market incentives for schools to target non-cognitive outcomes. This result is important for policymakers when designing schools' accountability framework, but also for parents whose choice of schools is supposed to drive the system. Given the importance of non-cognitive skills for later outcomes, this work might stimulate general awareness of this issue and influence public policy.

Chapter 5 explores distinctive features of academies that could explain previous findings. It shows that converter and sponsor-led academies have distinct admission policies, use different teaching methods, and face different problems. Interestingly, pupils at sponsor-led academies spend more time on mathematics, which has recently become a key policy priority. Considering the divergent performance of sponsor-led and converter academies, these features are of interest to policymakers, to those responsible for school intervention, and to school leaders and teachers. Chapter 5 also reveals that schools that joined multi-academy trusts have strong leadership and more structured management practices. This is a relevant result for policymakers currently encouraging schools to join a multi-academy trust but also for school leaders working within school networks.

I have presented the findings of this work at international conferences and academic seminars and will seek publication in leading peer-reviewed journals. Dissemination of results online and in the media plus public engagement with policymakers and education stakeholders will maximize research impact. As result of this thesis, future generations will benefit from evidence-based policies and an improved school system.

ACKNOWLEDGEMENTS

First, I would like to express my deepest gratitude to Alice Sullivan and Nikki Shure for their insightful guidance and unwavering support throughout the PhD. This thesis would not have been possible without their help and collaboration. Much of what I have learnt results from their suggestions and example. Their patience in dealing with issues and their constructive feedback was very much appreciated.

I would also like to thank Lorraine Dearden and Hedvig Horvath for their comments and direction at upgrade. This thesis benefited from discussions at the 2022 International Workshop on Applied Economics of Education, the Children of the noughties: a conference to celebrate 21 years of the Millennium Cohort Study, the 36th Annual Conference of the European Society for Population Economics, the 2023 European Association of Labour Economists, and the education work in progress seminar at the Centre for Economic Performance. Generous financial support from the Portuguese Foundation for Science and Technology is gratefully acknowledged.

I would be remiss in not mentioning the colleagues at UCL, who shared with me this challenging journey, for their friendship and encouragement.

I am deeply indebted to the support of my friends and family in this endeavour. They have been amazingly understanding and generous in accommodating the impact of this project in my life. None of this would be possible without the love and support of my mother and father over the years.

Lastly, I would like to recognise the role of the Christian faith in setting the purpose and horizon of my accomplishments.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

APU	Applied Psychology Unit
BA	Bachelor of Arts
BAS	British Ability Scales
CANTAB	Cambridge Neuropsychological Test Automated Battery
CEO	Chief Executive Officer
CGT	Cambridge Gambling Task
CLOSER	Cohort and Longitudinal Studies Enhancement Resources
CLS	Centre for Longitudinal Studies
CSD	Consistent School Database
DfE	Department for Education
DID	Difference in Differences
EAL	English as an Additional Language
EBacc	English Baccalaureate
FSM	Free School Meals
GCSE	General Certificate of Secondary Education
IMD	Index of Multiple Deprivation
ITT	Intention to Treat
IV	Instrumental Variable
KS	Key Stage
LA	Local Authority
LATE	Local Average Treatment Effect
LSOA	Lower Super Output Area
MAT	Multi-Academy Trust
MCS	Millennium Cohort Study
NCDS	National Child Development Study
NPD	National Pupil Database
NS-SEC	National Statistics Socio-Economic Classification
NVQ	National Vocational Qualifications
Ofsted	Office for Standards in Education, Children's Services and Skills
OLS	Ordinary Least Squares
ONS	Office for National Statistics
PISA	Programme for International Student Assessment
PSM	Propensity Score Matching
OECD	Organisation for Economic Cooperation and Development
SAT	Single-Academy Trust
SD	Standard Deviation
SDQ	Strengths and Difficulties Questionnaire
SEN	Special Educational Needs

STEM	Science, Technology, Engineering, and Maths
SWM	Spatial Working Memory
URN	Unique Reference Number
WMS	World Management Survey

1. INTRODUCTION

English academies are state-funded schools that are independent of Local Authorities (LA) and enjoy greater autonomy (Bolton, 2015). They were introduced in 2002 by the Labour government to raise standards in underperforming disadvantaged schools (A. West & Bailey, 2013). The programme was greatly expanded after 2010 by the Conservative and Liberal Democrat Coalition government, allowing every school to apply for academy conversion (Bolton, 2015; National Audit Office, 2018). Aiming at a system wide change, the subsequent Conservative government has envisioned a system where all state schools become academies (National Audit Office, 2018; A. West & Bailey, 2013). The academy policy follows the international neoliberal trend towards school autonomy and education quasi-markets (Chapman & Salokangas, 2012; A. West & Bailey, 2013). Similar reforms can be found in the United States and Sweden, namely American charter schools and Swedish free schools (Fenwick-Sehl, 2013). The lessons learnt from the expansion of academies in England, unique in scale and speed (Eyles et al., 2018), are paramount for policymakers considering a more autonomous school system. It is not only about assessing if this school model is able to raise standards across schools but also identifying the mechanisms and incentives leading to this improvement.

The academy programme seeks the improvement of educational attainment through an increase in school diversity, parental choice and a market-like system (A. West & Nikolai, 2017). The basic assumption is that autonomous schools, funded on a per pupil basis, are compelled to improve so as to attract more pupils (Machin & Silva, 2013). Autonomous schools are supposed to better understand pupils' needs and make decisions that meet their preferences and improve performance (OECD, 2016c). In effect, the introduction of academies reinforces education quasi-market mechanisms in England and seeks to foster an innovative school system driven by parental choice (Greany & Waterhouse, 2016; Woods & Simkins, 2014).

This thesis consists of three empirical papers that explore the variation in pupil outcomes and school characteristics associated with the timing of academy conversion. Selection issues common in the literature complicate efforts to establish causality. I attempt to approximate it using rich longitudinal data coupled with staggered treatment and legacy enrolment. I study outcomes not available in administrative data to shed light on how schools respond to academy reform. The thesis contributes to a literature on English academies that is mostly based on administrative data. It is also relevant to the wider literature on autonomous schools, especially US charter schools, and on education quasi-markets.

Chapter 2 sets the context for the academies programme and comparable schools in Sweden and in the United States. This review of the literature highlights concepts and features that are useful in understanding empirical results across the three empirical papers.

Chapter 3 examines the impact of academy conversion on pupils' cognitive and non-cognitive skills. Previous studies use administrative data and present mixed results on the effects of academy conversion on exam performance (Andrews et al., 2017; Eyles et al., 2016b, 2017, 2018; Eyles & Machin, 2019). I use the Millennium Cohort Study to study the impact of academies on pupils' problem-solving skills, social behaviour, mental health, and self-esteem. I attempt to establish causality through a difference in difference approach. To avoid selection issues, I focus on pupils that (at the start of the secondary phase) were enrolled in LA maintained schools that converted to academy status in the following years.

The overall impact of academy attendance (using difference in differences and legacy enrolment) on pupils' cognitive and non-cognitive skills is non-significant. Results show, however, that converter academies considerably raise the decision-making skills of their pupils and that sponsor-led academies significantly increase their pupils' self-esteem in comparison with similar pupils at LA maintained schools. Chapter 3 shows wider effects of school autonomy. This is also relevant to wider research considering how secondary schools foster cognitive and non-cognitive skills.

Chapter 4 explores how academies shape pupils' subject choice at secondary phase. I investigate the odds of pupils at academies choosing facilitating subjects, the English Baccalaureate, science subjects, and vocational subjects at age 14. Using the Millennium Cohort Study, I identify differences associated with academy attendance through a comprehensive regression model on subject choice. To avoid selection issues, I focus on pupils that at start of secondary phase enrol in LA maintained schools that convert to academy status in the following years.

The overall impact of academy attendance on subject choice is driven by converter academies. Pupils at converter academies are significantly more likely to study science subjects and facilitating subjects at age 14 than their peers at LA maintained schools. This suggests that converter academies offer a more academically demanding curriculum. Chapter 4 adds to literature on subject choice at secondary school, highlighting how schools shape those choices.

Chapter 5 investigates management practices, school policies, and learning environment at academies. I use school level variables from the Programme for International Student Assessment (PISA) to compare schools that become academies and LA maintained schools. To ensure schools are comparable, I use as control group schools that become academies after the sample period. School management quality is measured by an adaptation of the World Management Survey management index, following Bloom et al. (2015) and Leaver et al. (2022), Other variables revealing management practices include responsibility of school staff over resources and over the curriculum as well as school leadership on curricular development, instruction, professional development, and teacher participation. Outcome variables also include data on curriculum, admissions, and assessment policies. The learning environment is described by the disciplinary climate, inquiry-based learning practices, teacher-directed instruction, adaptive instruction as well as pupil and teacher behaviour hindering learning.

The results suggest that academies develop distinct managerial structures and policies. Multi-academy trusts develop strong leadership and adopt more structured management practices. Moreover, converter and sponsor-led academies have distinct admissions criteria and face different problems. Both converter and sponsor-led academies enjoy, albeit differently, a better school climate in terms of discipline and pupil behaviour. Pupils at sponsor-led academies devote more time to mathematics and problem-based learning. Converter academies use instead more teacher-directed and adaptive instruction. Chapter 5 reveals features of schools that are unobserved in most analyses. This adds to an emerging literature on school management.

The thesis concludes with a discussion of policy implications and equity concerns raised by findings, setting out future research. A recurring theme across this work is that sponsor-led and converter academies are different and offer dissimilar educational experiences, possibly with lasting effects for pupils. This raises concerns for social mobility since academy conversion has increased socioeconomic stratification between sponsor-led academies and converter academies, with the latter having more advantaged intakes (Braz, 2018; Eyles et al., 2018). Another area of concern is the lack of significant effects of academy conversion on non-cognitive skills, considering their longstanding relevance. This shows that quasi-market incentives are not directing schools towards these critical outcomes. On a more positive note, considering current policy driving the expansion of multi-academy trusts (Department for Education, 2022), the higher management quality observed there is encouraging.

2. ACADEMY PROGRAMME

2.1 INTRODUCTION

The purpose of this chapter is to set the context for the academies programme, highlighting features and concepts that are useful in understanding empirical results and the underlying mechanisms. Section 2.2 presents academies and free schools. Acknowledging the shifts in government policy, especially the Academies Act 2010 and the following heterogeneity in academies, section 2.3 outlines policy phases with an emphasis on the distinctive characteristics of sponsor-led academies, converter academies, and on the development of academy chains. Section 2.4 reviews how schools are using increased autonomy. Section 2.5 describes main features of the English education system where academies are introduced, presenting different types of schools and other educational actors in the institutional framework. Some information on school accountability is provided, revealing constraints to school autonomy. Section 2.6 gives a brief overview of the theoretical arguments motivating the academy policy, underlining the conditions for education quasi-markets to work efficiently and their inherent dynamics. The chapter concludes by presenting international autonomous schools that are in some respects similar to academies and that are used as models in the political debate.

2.2 ACADEMIES

Academies are state-funded schools that operate outside the control of LAs and are managed by non-profit trusts (Bolton, 2015; National Audit Office, 2018; A. West, 2015). Academies are funded directly by the Department for Education (DfE) according to the number of pupils and operate with considerably more autonomy than LA maintained schools (Eyles & Machin, 2019; A. West & Bailey, 2013). Academies have greater discretion over their budget, employment, pay and working conditions, staffing, career development, discipline and performance management, admissions and taught curriculum (Bolton, 2015; Eyles & Machin, 2019; National Audit Office, 2018). They are subject to statutory assessment, Ofsted (Office for Standards in Education, Children's Services and Skills) inspections and the school admissions code (A. West & Nikolai, 2017). Academies are not required to follow the national curriculum[1] and the School Teachers Pay and Conditions statutory guidance (A. West & Wolfe, 2019). Academies are responsible for procuring services previously provided by LAs such as administration, finance, and human resources (National Audit Office, 2018).

Free schools also have academy status (Harris, 2012; Higham, 2014). These are newly established state-funded schools, which open in response to local demand (Cirin, 2014; Esper, 2023; Green et al., 2015; Harris, 2012; Higham, 2014; R. Morris & Perry, 2019). Free schools are proposed, set up and managed by groups of parents and teachers, educational institutions,

[1] Academies must offer a broad and balanced curriculum, including English, mathematics, science, and religious education (A. West & Wolfe, 2019).

religious groups, businesses, community groups or charities (Higham, 2014; R. Morris & Perry, 2019; Wiborg et al., 2018). Free schools receive the same funding and autonomy as academies and operate within the same framework (Allen & Higham, 2018; Wiborg et al., 2018).

The focus of this thesis is on academies, since free schools are new schools, not conversions of existing schools. But evidence from free schools, given their similarity and legal status, is used to provide supporting examples.

2.3 Policy Development

I next describe the phases of development of the academy policy and the features of different types of academies. The emergence of academy chains is explained.

Academies were introduced by the Labour government in 2002 as a targeted remedial intervention to address persistent academic underperformance and educational inequality (Eyles et al., 2018; Woods & Simkins, 2014). Conversion to academy status is forced on failing LA maintained secondary schools, mainly in highly deprived urban neighbourhoods (Bolton, 2012; Cirin, 2014; National Audit Office, 2018). Academies replace schools with a disadvantaged pupil intake (high proportion of pupils underperforming at Key Stage (KS) 2, pupils eligible for free school meals (FSM), pupils from a minority ethnic background and pupils with English as an additional language (EAL)) and a record of low KS4 performance (Bolton, 2012; Eyles & Machin, 2019; Hatton et al., 2019). Conversion to academy status requires a government-approved independent sponsor that can be a philanthropist, a business, a charity or an educational institution (National Audit Office, 2018; A. West & Bailey, 2013). The sponsor should press for organisational restructuring and introduce innovative management and best practices (Bertoni et al., 2021; Regan-Stansfield, 2018). These academies are known as “sponsor-led academies”.

The Academies Act of 2010, enacted by the Conservative and Liberal Democrat Coalition government, introduces free schools, streamlines the process of academy conversion, and allows LA maintained schools (including primary schools and special schools) to apply for conversion without an external sponsor — known as “converter academies” (Bolton, 2015; Higham, 2014; National Audit Office, 2018). This voluntary process has become the preferred route for conversion (Bolton, 2012). DfE approval can take between two and five months after application, and academies open on average five months after approval[2] (Bertoni et al., 2021). Schools rated “Outstanding” by Ofsted are pre-approved for academy conversion (Bolton, 2015). Other maintained schools are allowed to convert, provided their exam performance is improving, have sound finances and a “Good” Ofsted report (Bolton, 2015; Eyles et al., 2017). Most grammar schools and foundation schools have converted (Bolton, 2015). In 2016, the Conservative government set the target, later modified, of all schools becoming academies by 2022 (National Audit Office, 2018). Eyles et al. (2018) show that their predecessor schools are high performing schools that have more advantaged intakes. Hicks (2014) also reports that schools with fewer FSM pupils are more likely to apply for

[2] DfE aims for schools converting to open as academies within 12 months of their application being approved (National Audit Office, 2018).

conversion. According to Hicks (2014), academy conversion has occurred disproportionately in Conservative constituencies, especially if schools are under a Labour LA. According to DfE surveys, schools convert to get more funding and greater control over resources or to create opportunities for collaboration (Cirin, 2014, 2017).

Building on Labour's academy programme, underperforming maintained schools continue to be forced to convert under a sponsor, as a result of government intervention (National Audit Office, 2018). The median time for academy conversion is 17 months (Duchini et al., 2023). Since 2016, the DfE aims to convert schools rated "Inadequate" within nine months of Ofsted inspection (National Audit Office, 2018). Unsurprisingly, sponsor-led academies converting before and after 2010 have fairly similar predecessor schools (Eyles et al., 2018).

The number of academies has grown steadily (National Audit Office, 2018). Initially academies are introduced gradually (four percent of secondary schools in 2009) but after 2010 the programme is massively expanded (Eyles & Machin, 2019; National Audit Office, 2018). In 2018, 35% of English state schools (72% of secondary schools and 27% of primary schools) have academy status[3] (National Audit Office, 2018). Most academies are converters (Bolton, 2015; National Audit Office, 2018). In most geographic areas the majority of secondary schools are academies, reaching 100% in some areas (National Audit Office, 2018).

[3] According to the 2022 School Census, 40% of all schools are academies (80% of secondary schools and 39% of primary schools), accounting for 53% of pupils.

Schools have been encouraged to join academy chains in order to foster collaboration among schools, enhance their management capabilities, and prevent the atomisation of the school system (Department for Education, 2016a; Eyles et al., 2017; Greany et al., 2023; Higham & Earley, 2013; Male, 2022; Neri & Pasini, 2018). An increasing proportion of academies have joined a multi-academy trust (MAT) (Cirin, 2014, 2017; National Audit Office, 2018). Some underperforming schools are required to join a high-performing MAT and most outstanding converter academies support other schools through one (Andrews et al., 2017; Cirin, 2014). Schools join a MAT seeking to build alliances with other schools and share practice, or to create economies of scale and an effective back office (Salokangas & Chapman, 2014). In 2017, 73% of academies belong to a MAT (A. West & Wolfe, 2019).[4] In 2022, the Conservative government set the target for all schools to join a MAT by 2030 (Department for Education, 2022). MATs are expected to improve professional development and educational standards through economies of scale, reinvigorated school leaders, and best practice dissemination (Greany & Waterhouse, 2016; National Audit Office, 2018; Rayner et al., 2018; Regan-Stansfield, 2018).

Schools within MATs have no legal identity (A. West & Wolfe, 2019). They are governed under a shared vision and brand by a centralised governance structure, having varying levels of autonomy (Boyask, 2018; Chapman & Salokangas, 2012; Neri & Pasini, 2018; Salokangas & Chapman, 2014; Wilkins, 2017; Woods & Simkins, 2014). MATs increasingly acquire the

[4] In 2021, 60% of secondary schools are part of a MAT (Neri et al., 2022).

role previously held by LAs (Boyask, 2018; Chapman & Salokangas, 2012; Wilkins, 2017). They are responsible for finance, management, and support services, but often delegate operational functions and daily management to schools that follow their strategic direction and operating procedures (Chapman & Salokangas, 2012; Cirin, 2017; Male, 2022; Neri & Pasini, 2018; A. West & Wolfe, 2019; Wilkins, 2017). MATs vary considerably in the number of schools and pupils,[5] governance model, and geographical spread (Andrews et al., 2017; Simkins, 2015; Wilkins, 2017). Large MATs tend to develop a regional management structure, whereas small MATs typically allow schools full discretion over teaching and learning (Cirin, 2017). Most MATs allow a degree of flexibility in teaching and curriculum delivery (Cirin, 2017).[6] School autonomy, though, is constrained by MAT's policies and practices (Salokangas & Chapman, 2014).

Overall, there are considerable differences between sponsor-led academies and converter academies, reflecting the 2010 policy shift. The expansion of academies and the development of MATs denote a major reorganisation of schools. Next, I review available evidence on how schools have changed due to this policy.

---

[5] Most MATs are responsible for between three to 20 schools (Male, 2022). MATs have on average seven schools that generally operate in the same region (Greany et al., 2023).

[6] For example, Neri et al. (2022) show that MAT procurement decisions over assessment and curriculum tend to be more decentralised.

2.4 AUTONOMY IN PRACTICE

A survey of school leaders, conducted by Higham and Earley (2013), reveals most believe autonomy improves teaching and learning and the use of resources. School leaders anticipate greater control over school management (Higham & Earley, 2013). However, several schools have not changed teaching and learning or pay and conditions after academy conversion (Higham & Earley, 2013).

According to a DfE survey to academies in 2013 (Cirin, 2014), sponsor-led academies and early converter academies are more innovative. Following conversion, academies change their curriculum, change how they monitor pupils' attainment and teacher performance (Cirin, 2014). Some academies alter term dates or the length of the school day (Cirin, 2014). School leaders claim that those changes improve attainment (Cirin, 2014).

In a posterior DfE survey to academies, conducted in 2016, Cirin (2017) reports that most academies procure services previously provided by LAs and introduce savings in administrative functions. Academies in MATs are more likely to make organisational changes (Cirin, 2017). Interestingly, Eyles et al. (2017) note that, after conversion, school expenditure at primary schools increases chiefly on administrative and routine operations. Davies et al. (2021) also show that academies increase spending on back office and spend less on teacher salaries. Similarly, Duchini et al. (2023) find that after conversion sponsor-led academies increase expenditure on non-teaching personnel, including headteachers. In fact, there is a high headteacher turnover in these academies[7] (Duchini et al., 2023; Eyles et al., 2016a) and headteachers at academies have on average a higher salary (Duchini et al., 2023; Telhaj et al., 2022). According to Telhaj et al. (2022), headteachers at pre-2010 sponsor-led academies receive a wage premium, whereas post-2010 academies are more likely to employ headteachers that were already at the top of the wage distribution.

Conversion to academy status under a sponsor is also related to teacher turnover (Duchini et al., 2023). Older and underperforming teachers leave sponsor-led academies around the timing of conversion and are substituted by newly hired teachers, mostly from outstanding schools (Duchini et al., 2023). Moreover, Duchini et al. (2023) show that changes of the teaching body at sponsor-led academies are associated with restructuring teachers' pay and higher pay dispersion.

How these changes are reflected in teaching and learning at academies remains unclear. Some studies seem to suggest that academies are adopting a conservative curricular stance. For example, Greany and Waterhouse (2016) claim that few academies devise innovative curricula. In addition, qualitative studies like Wiborg et al. (2018) and Esper (2023) reveal a perception of limited curriculum and pedagogical innovations at free schools (that also enjoy academy status).[8] In fact, according to Cirin (2014), academies tend to follow the national curriculum for English, mathematics and science. Wiborg et al.

[7] Sponsor-led academies and those that join a MAT are likely to change chool leadership (Cirin, 2014, 2017).

[8] According to headteachers interviewed by Wiborg et al. (2018), free schools innovate predominantly in management practices. Similarly, Esper (2023) highlights innovative marketing strategies in three London free schools.

(2018) and Esper (2023) also suggest that schools in academy chains adhere more closely to the national curriculum and operate in a more standardized teaching environment. However, studies focusing on free schools or based on the perceptions of a small sample of teachers and headteachers may not reflect curricular development at academies more broadly. The role of school leaders in explaining curricular innovation is actually emphasised by Greany and Waterhouse (2016). Besides, descriptive analyses do not reveal changes due to academisation. Academies have greater autonomy over taught curriculum and school competition may incentivise diversification and curricular development to attract pupils and improve overall performance. Yet autonomous schools might not innovate. Chapters 3 and 4 investigate the impact of academisation on pupils' outcomes and subjects studied at secondary phase, echoing an eventual transformation in teaching and learning after conversion. Basically, any effect of academy status on pupil outcomes requires schools to have changed and implies a degree of innovation.

Abovementioned surveys do not portray academies' climate and practices clearly. There is indeed little school level data on this, making it difficult to draw conclusions on school changes from academy conversion. The school questionnaire in PISA provides insight into schools' learning environment. For instance, OECD (2016c) shows that autonomous schools generally use adaptive instruction more often. Using PISA 2015, Jerrim and Shure (2016) show that the staff of converter academies is deemed more resistant to change. They also reveal that a lack of teacher preparation and staff not meeting individual pupils' needs remain key concerns for headteachers of sponsor-led academies. Chapter 5 elaborates on these findings using additional variables and a sharper identification strategy.

2.5 Institutional Context

This section outlines basic features of the English education system where academies are introduced. It mentions the different types of state schools and the role of LAs. It also describes the accountability framework limiting school autonomy.

2.5.1 Schooling

State schools in England vary in autonomy and governance, comprising LA maintained schools, city technology colleges, free schools, and academies (Eyles & Machin, 2019). LA maintained schools are part of a local education system administered by democratically elected LAs that redistribute state funds, coordinate pupil admissions, employ staff, and provide administrative and management services to schools (Eyles et al., 2017; Harris, 2012; A. West & Bailey, 2013; A. West & Wolfe, 2019). City technology colleges, free schools, and academies operate independently of LAs and are funded directly by central government (Fenwick-Sehl, 2013; A. West & Bailey, 2013).

LA maintained schools include community schools, voluntary-controlled schools, voluntary-aided schools, and foundation schools. Community schools and voluntary-controlled schools are managed by LAs that are responsible for admissions and staffing (Machin & Silva, 2013). Voluntary-aided schools and foundation schools are partnerships between the state and the voluntary sector and are responsible for admissions and employing the school staff

(Machin & Silva, 2013). Most voluntary-aided schools and voluntary-controlled schools have a religious basis (Allen & West, 2009; Machin & Silva, 2013).

City technology colleges and grant-maintained schools (later replaced by foundation schools) are forerunners of academies (Fenwick-Sehl, 2013; Walford, 2014; A. West & Bailey, 2013). Most foundation schools and city technology colleges have actually become academies (Bolton, 2015; A. West & Bailey, 2013). Academies are state partnerships with the private and voluntary sectors, where the central government is responsible for performance evaluation, funding, and policy direction (Granoulhac, 2017; Harris, 2012; A. West & Bailey, 2013). Basically, the expansion of academies has greatly reduced the role of LAs in school-based education (A. West & Bailey, 2013). Services previously provided by LAs have been replaced with market supply or MAT's central services (Higham & Earley, 2013).

2.5.2 Accountability Framework

League tables and school inspections form a dual accountability system (Allen & Burgess, 2012; Greany & Waterhouse, 2016). The School Inspectorate, Ofsted,[9] visits schools every three to five years and publishes a school report with a broad assessment on the quality of teaching and effectiveness (Eyles et al., 2018). League or performance tables include information regarding average school Key Stage[10] attainment, school value-

[9] Ofsted is an independent national board responsible for monitoring and reporting on educational standards within schools (Allen & Burgess, 2012).

[10] Schooling in England is organised around KS. KS1 and KS2 are undertaken in primary school (up to age 10) and KS3 and KS4 in secondary school (up to age 16) (Eyles et al., 2018). Pupils are assessed through externally assessed standardised national exams at the end of KS2 and KS4 (Eyles et al., 2018).

added, and school composition (Bertoni et al., 2021). League tables and school inspections inform parental choice and government interventions, monitor school progress, and uphold minimum standards (Greany & Waterhouse, 2016; Machin & Silva, 2013).[11] Graduation rates and average test scores are indeed useful predictors of school effectiveness, but partly reflect school composition (Allen & Burgess, 2013).

National exams and school inspections inject a degree of standardisation into schools (Greany & Waterhouse, 2016; Wiborg et al., 2018; Woods & Simkins, 2014). Competitive pressures further reinforce the significance of performance data for academies (Higham & Earley, 2013). Since parents use league tables and Ofsted reports to choose schools, academies must perform against national standards (Greany & Waterhouse, 2016; Woods & Simkins, 2014). Some argue such a performance-driven environment hinders schools' incentives to innovate and promotes curricula homogenisation (Esper, 2023; Greany & Waterhouse, 2016; Shah, 2018; Wiborg et al., 2018). Contrariwise, pupil performance at PISA 2015 improves with school autonomy where achievement data is tracked and published (OECD, 2016c). This is confirmed in Hanushek et al. (2013), suggesting external accountability discourages opportunistic behaviour.

Independence from LAs is a key distinctive feature of the academies programme. However, Ofsted and league tables, coupled with quasi-market mechanisms, constrain school autonomy.

---

[11] Schools that are considered “Good” or “Outstanding” by Ofsted have less pressure and more resources (Greany & Waterhouse, 2016). On the other hand, a judgement of “Inadequate” triggers a set of policy actions (Allen & Burgess, 2012).

2.6 Policy Rationale

Next, I look to the theoretical arguments motivating the academy policy. The conditions for education quasi-markets are mentioned.

The introduction of new autonomous school types intends to raise standards and efficiency through parental choice and school competition (Elwick, 2018; Fenwick-Sehl, 2013; Machin & McNally, 2012; R. Morris & Perry, 2019; OECD, 2019). Reformers envision an innovative school-led system of state-funded autonomous schools (Greany & Waterhouse, 2016; Woods & Simkins, 2014), emulating high-performing private schools (Green et al., 2011). An autonomous state sector is supposed to be cost-effective (Granoulhac, 2017).

The operation of education quasi-markets requires parental choice, high-quality performance information,[12] supply flexibility, and school autonomy (Braconier, 2012). Parents ought to have adequate information and choose schools based on their quality (Allen & Burgess, 2010). Schools should have incentives to be popular and to expand through higher standards (Allen & Burgess, 2010). Underperforming schools must be allowed to close, new schools to open, and popular schools to expand (Braconier, 2012). Because of quasi-market mechanisms, autonomous schools are expected to be more responsive to parental preferences, pupils' needs, and local circumstances

[12] Market-oriented reforms are often linked to accountability systems based on school inspections and pupil performance at standardised and externally developed exams (OECD, 2017a).

(Greany & Waterhouse, 2016; Harris, 2012; OECD, 2016c; Woods & Simkins, 2014). If high performing schools attract more pupils and funding follows pupils, autonomous schools will be incentivised to monitor and improve their practices (Machin & Silva, 2013).

In essence, the expansion of academies, in conjunction with parental choice and the accountability framework, creates the foundation for education quasi-markets in England.

2.7 International Examples of School Autonomy

This chapter concludes acknowledging the international political trends and policy-borrowing process behind the academy policy. Attention is given to international autonomous schools that are comparable to academies and that have been mentioned in the political debate. Key differences in policy are highlighted.

Neoliberal policies underpin the international movement towards autonomous schools and education quasi-markets (Chapman & Salokangas, 2012; A. West & Bailey, 2013). Neoliberalism, under neoclassical assumptions, pursues market rationality and consumer choice as organising principles, and has been associated with private provision of welfare services (including education) and the introduction of market competition in the public sector (A. West, 2014; A. West & Bailey, 2013; Wilkins, 2017). This means reshaping welfare institutions through market relationships, such as when introducing choice, diversity, and competition in state education (Alexiadou et al., 2016; Higham & Earley, 2013). Still, market-based reforms have not been followed solely by right-wing governments (Hicks, 2015; A. West, 2014).

English policymakers that have shaped the policy on academies have cited similar reforms in the United States (US) and Sweden to support or oppose reform (Fenwick-Sehl, 2013; Williams, 2023). The policy debate in England has indeed discussed international evidence and specific features of the US and Swedish education systems that could affect the implementation of similar reforms (Fenwick-Sehl, 2013). Fenwick-Sehl (2013) notes that, despite differences in policy across countries, Swedish and US schools have been used to legitimise and serve as models for English academies, so as to generate public support.

Autonomous state-funded schools in the US and Sweden, like academies and free schools in England, are introduced to raise educational standards through parental choice and school competition (Chapman & Salokangas, 2012; Elwick, 2018). These are known respectively as “charter schools” and “free schools”. The majority of US charter schools and Swedish free schools are new schools, not conversions of existing schools as most academies (Eyles et al., 2016a). Charter schools and Swedish free schools enjoy operational autonomy and are managed out of local government control (Chapman & Salokangas, 2012; Elwick, 2018). These schools are managed by independent entities that administer admissions, curriculum, budget, and staffing (Chapman & Salokangas, 2012; Fenwick-Sehl, 2013; A. West, 2014).

US charter schools are state-funded schools that operate outside school district control and are funded on a per pupil basis (Abdulkadiroğlu et al., 2011; Walters, 2018). These schools are managed by charities, universities, or groups of teachers and parents (Abdulkadiroğlu et al., 2011). Charter schools enjoy greater operational autonomy, having control over the curriculum, budget, staffing, and the length of school day and year (Abdulkadiroğlu et al., 2011; Walters, 2018). Most charter schools serve impoverished and ethnic minority urban communities (Abdulkadiroğlu et al., 2016; Krowka et al., 2017). Charter schools that follow the "No Excuses" approach feature strict discipline, high academic expectations, intensive monitoring and tutoring, emphasis on literacy and mathematics, extended school day and year, teacher training and parental involvement (Abdulkadiroğlu et al., 2011, 2016; Krowka et al., 2017; Walters, 2018).

Swedish free schools are funded by local government on a per pupil basis and are managed by parents, teachers, corporations, religious or community groups (Alexiadou et al., 2016; Böhlmark & Lindahl, 2015; A. West, 2014; A. West & Nikolai, 2017). Free schools enjoy autonomy over budget and curriculum, but comply with general regulatory framework and inspections (Böhlmark & Lindahl, 2015; A. West, 2014; A. West & Nikolai, 2017). Most free schools are located in affluent urban areas (A. West, 2014).

Despite many common features and a similar political discourse, the academy policy has distinctive features (Fenwick-Sehl, 2013). Focusing on the sponsor-led academies introduced by the Labour government, Fenwick-Sehl (2013) points out that academies are non-profit and that independent providers are not allowed to establish new schools. The introduction by the Coalition government of free schools with academy status, proposed and established by independent providers, attenuate those differences (Fenwick-Sehl, 2013; A. West & Bailey, 2013).

As shown above, US charter schools and Swedish free schools are in many ways like academies. Basically, academies are out of LA control and their autonomy is limited by the accountability regime. However, sponsor-led academies are significantly different from converter academies. Academies are grouped in MATs and standalone academies.

3. Academies' Impact on Pupils: Looking Beyond Exams

3.1 Introduction

This chapter examines the impact of academy conversion on pupils' cognitive and non-cognitive skills. Previous studies focus on the effects of academy attendance on exam performance (Andrews et al., 2017; Eyles et al., 2016b, 2017, 2018; Eyles & Machin, 2019). Many would argue, however, that education aims go beyond exam performance, covering wider cognitive and non-cognitive skills, mental health as well as social and moral development. According to these views, the purpose of education is the formation of the whole person or to help pupils to become independent and self-regulated lifelong learners that are able to fully participate in society (Arends, 2006; Delors, 1996; OECD, 2017a). Indeed, the school curriculum is supposed to promote the "spiritual, moral, cultural, mental and physical development of pupils" (Department for Education, 2014). Besides, mental health provision in schools is a UK policy priority (Ford et al., 2021).

So, even if education quasi-market reforms improve exam performance, how are pupils faring in those other aspects? Given the pivotal role of league tables for school choice (Allen & Burgess, 2013), academies may strategically use their autonomy to target pupils at the threshold of performance measures, drawing on traditional teaching focused on exams preparation, instead of on developing skills and other important educational dimensions. Are high performing pupils at national exams also developing key social and problem-solving skills? How well are they coping with stress, and pressure to get higher marks? Pupils' anxiety and work-leisure imbalance may become a problem at schools that stress rankings and pupil competition. In fact, academic pressure is often associated with high anxiety (Marcus et al., 2020). On the other hand, an increased focus on national exams may lead underperforming pupils to lower self-esteem (OECD, 2017a). Is there a trade-off between skill development and wellbeing? Evaluating the full impact of academy conversion entails understanding academies accomplishments in relation to pupils' social and cooperative skills, problem-solving skills, mental health, and self-esteem.

This chapter sets out to answer the following question: Do pupils at academies have better cognitive skills not easily measured in national exams (problem solving) and non-cognitive skills (social behaviour, mental health, self-esteem) than their peers at Local Authority (LA) maintained schools?

The outcomes of interest are problem solving skills, social behaviour, mental health, and self-esteem. Using the Millennium Cohort Study, I attempt to establish causality through difference in differences. In order to avoid selection issues, the treatment group is pupils that, at the start of secondary schooling, were enrolled in LA maintained schools that converted to academy status in the following years. The treatment group includes pupils attending academies, and the control group includes pupils attending LA maintained schools. A subgroup analysis for sponsor-led and converter academies reveals heterogeneous effects. Results suggest that converter academies considerably raise the decision-making skills of their pupils and that sponsor-led academies significantly increase their pupils' self-esteem, in comparison with similar pupils at LA maintained schools.

The structure of this chapter is as follows. Section 3.2 reviews previous studies on academies with a focus on pupil performance regarding cognitive skills and non-cognitive skills and highlights existing gaps in the literature. Section 3.3 explores possible school effects and underlying mechanisms on problem-solving skills, mental health, and social skills. Sections 3.4 and 3.5 present data and the econometric model used to answer the research question. Section 3.6 discusses empirical findings. Finally, section 3.7 offers some concluding comments.

3.2 Performance Effects of Academy Conversion

This section reviews evidence on the impact of academies on pupils' outcomes. Acknowledging the difficulty in establishing causality, sub-section 3.2.1 reviews the literature on the effects of academy conversion on pupil intakes. Evidence of change in school composition following academy conversion raises selection issues that require appropriate evaluation methods. Information on parental choice and school selection is provided. Sub-sections 3.2.2 and 3.2.3 consider studies assessing the effects of academy attendance on cognitive and non-cognitive skills. Some methodological approaches addressing selection issues are mentioned. Most literature analyses high stakes exam performance using administrative data. Highlighting the gaps in the literature, this section concludes by focusing on cognitive skills not easily captured in national exams and on non-cognitive skills.

Since academies were introduced in England in 2002, this literature review focuses on studies from 2002 onwards. Evidence from Swedish free schools and US charter schools is also included, as these schools are somewhat comparable to academies and are referred to as their models. English free schools, that followed academies and enjoy academy status, are not the focus of this work, but relevant findings on free schools are included due to their similarity.

3.2.1 Effects on Pupil Intake

Reform advocates predict that school autonomy, coupled with parental choice and accountability, would result in a more efficient school allocation of pupils where individual preferences and needs are matched to differentiated schools (Machin & Silva, 2013). However, critics claim that the expansion of academies leads to school segregation (Machin & Silva, 2013). If affluent parents are better informed or more effective at exploiting school choice, underperforming schools will attract disadvantaged intakes (Machin & Silva, 2013; OECD, 2019). Moreover, since schools are rated on pupil attainment, academies have incentives to compete for local pupils and screen for high achieving pupils, so as to maximise average performance, minimise teaching costs, and become more attractive to parents (Allen & Burgess, 2013; Gibbons et al., 2008; Machin & Silva, 2013; OECD, 2019).

This sub-section reviews empirical evidence on the effects of academy conversion on pupil intakes, ascertaining concerns on potential selection bias affecting research. It includes a discussion of underlying factors, namely parental choice and school selection.

3.2.1.1 School Composition

Early sponsor-led academies improve their pupil intake. Eyles and Machin (2019) and Eyles et al. (2016a), through a pupil-level difference in differences comparison of early academies (between 2002/03 and 2008/09) relative to predecessor schools and maintained schools that become academies after the sample period ends, show that pre-2010 sponsor-led academies significantly reduce the share of FSM eligible pupils. Despite the declining share of FSM pupils, sponsor-led academies still have more disadvantaged pupils than local schools (Bolton, 2012). Eyles and Machin (2019) and Eyles et al. (2016a) also find that after conversion sponsor-led academies attract and admit higher ability pupils. Similarly, Andrews et al. (2017) report a significant change in pupil ability intake following academy conversion for sponsor-led academies both before and after 2010. This confirms findings in Eyles et al. (2018).

According to Hicks (2017), converter academies have not changed their pupil intakes. Eyles et al. (2017) and Regan-Stansfield (2018) also find no evidence that primary converter academies alter their intakes. Further, Eyles et al. (2018) reveal that secondary converter academies do not experience a significant change in the ability of their intakes. However, Eyles et al. (2018) show that, at secondary phase converter academies reduce their share of FSM eligible pupils.[13] In contrast, they reveal that sponsor-led academies converting after 2010, unlike early sponsor-led academies (Eyles & Machin,

---

[13] A more advantaged pupil intake is likely to lower the number of pupils with special educational needs (SEN) (Liu et al., 2020). Accordingly, converter academies have lower proportions of pupils with SEN statement (Norwich & Black, 2015).

2019), attract more disadvantaged pupils. Hence, secondary schools becoming academies after 2010 are increasingly socially stratified, as sponsor-led academies attract more disadvantaged pupils and converter academies attract fewer disadvantaged pupils (Eyles et al., 2018). Braz (2018) confirms this using the Millennium Cohort Study (MCS).

The abovementioned studies show that before 2010 sponsor-led academies improve their pupil intakes. After 2010, secondary sponsor-led academies enrol more disadvantaged pupils whereas converter academies admit more advantaged intakes. This support claims of selection at academies either owing to parental preferences or to school admission policies. Hence, appropriate identification strategies in the impact evaluation design are required.

3.2.1.2 School Choice and Selection

Disentangling parent choice from school selection in explaining pupil intakes is not an easy exercise. In England, school admission is based on parental preferences, unless applications exceed school capacity (Allen et al., 2013; Bertoni et al., 2021; Burgess et al., 2015; Gibbons et al., 2013; Gibbons & Silva, 2008; Leroux, 2015). Generally, parental preferences depend on schools' academic performance, socio-economic composition and distance to home (Bertoni et al., 2021; Burgess et al., 2015; R. Morris & Perry, 2019). Oversubscription criteria include sibling attendance, proximity from home, catchment areas, feeder schools, banding, lotteries, religious attendance (faith schools), admission tests (grammar schools), and aptitude for the school specialism (Allen et al., 2013; Allen & West, 2009; Bertoni et al., 2021, 2023; R. Morris, 2014).

Academy conversion is perceived by parents as a school rebranding (Bertoni et al., 2021). Parents interpret academy conversion as a signal of quality when combined with a record of high attainment and popularity (Bertoni et al., 2021). Public perception associating academies with preferred school features leads to an increase in applications. In effect, Cirin (2014) reports an increase in first-choice applications after academy conversion. Based on MCS, Leroux (2015) shows that parents who prioritise academic performance are more likely to apply to an academy. These parents are on average wealthier and more educated (Leroux, 2015). Similarly, Bertoni et al. (2021) find that following conversion affluent parents are more likely to prefer academies. These parents regard academic performance and popularity as important choice factors (Bertoni et al., 2021).

Bertoni et al. (2021) reveal that pupils who prefer academies are more likely to enrol in one. So, they conclude that preferences are not offset by school admission practices and largely explain school segregation. Eyles and Machin (2019) also attribute composition effect to a post-conversion change in local preferences. Moreover, West (2014) shows that converter academies tend to keep predecessor schools' admissions criteria. However, more research is needed to establish the role that parental preferences and admissions practices have in explaining the changes in school composition after academy conversion. Regardless, there are selection issues associated with academy conversion that must be addressed to identify academy attendance causal effects.

3.2.2 Effects on Cognitive Skills

3.2.2.1 High Stakes Exam Performance

When we compare the average exam performance of academies with LA maintained schools, we generally observe that sponsor-led academies underperform and that converter academies have higher attainment (Bolton, 2012, 2015; Department for Education, 2018). Early sponsor-led academies have generally improved their performance after conversion (Hatton et al., 2019). Yet, as shown above, these comparisons reflect pupil intakes and features of predecessor schools.

Andrews et al. (2017), Eyles et al. (2016a), Eyles et al. (2016b), Eyles et al. (2017), Eyles et al. (2018), Eyles and Machin (2019), and Machin and Sandi (2020) ascertain the causal impact of academy conversion using difference in differences. These authors use as control group LA maintained schools that become academies in the following period, assuming that the timing of conversion is as good as random and excluding schools that do not convert. Since school composition may change after conversion, they focus on pupils who are already enrolled in the predecessor school, as conversion is exogenous to them. They estimate a local average treatment effect (LATE) on pupils that would not have attended an academy had they not been pre-enrolled, irrespective of whether they take their exams in that school. Using the National Pupil Database (NPD), Eyles and Machin (2019), Eyles et al. (2016b), Eyles et al. (2017), Eyles et al. (2018) and Andrews et al. (2017) verify that schools in the treatment and control groups have similar pre-conversion trends.

Sponsor-led academies that opened before 2010 experience a significant improvement in pupil performance[14] at GCSEs (General Certificate of Secondary Education) and in the likelihood of university attendance and degree completion (Andrews et al., 2017; Eyles et al., 2016a, 2016b; Eyles & Machin, 2019; Machin & Sandi, 2020). Sponsor-led academies are also more likely to improve their Ofsted rating (Duchini et al., 2023; Eyles & Machin, 2019). According to Machin and Silva (2013), those positive effects come from pupils at the top of the ability distribution. While Eyles et al. (2016b) identify larger effects for disadvantaged pupils. Higher performance could be explained by better management[15] and curriculum flexibility, but Andrews et al. (2017) point out that early academies receive more resources. Machin and Sandi (2020) refute suggestions that performance improvements at academies result from year 11 exclusions shaping the pool of GCSE takers.[16]

Sponsor-led academies that opened after 2010 experience a performance improvement in the year immediately before conversion and in the conversion year, but it tails off over the following years (Andrews et al., 2017). Andrews et al. (2017) argue that this pre-treatment effect may be related to early intervention from sponsors or to school action to avoid forced conversion. It is also possible that schools in the control group improve in

[14] The improvement is greater the longer an academy has been in operation and increases with years of exposure to treatment (Andrews et al., 2017; Eyles et al., 2016a; Eyles & Machin, 2019). Eyles and Machin (2019) find larger effects for schools that gained relatively more autonomy and for those located in urban areas. Eyles et al. (2016b) also find larger effects for London.

[15] Eyles and Machin (2019) and Duchini et al. (2023) identify a significant headteacher turnover after academy conversion.

[16] Simulations reinstating excluded pupils in the excluding school show that negative peer effects would have to be implausibly large to cancel out performance gains (Machin & Sandi, 2020). Besides, performance gains are not associated with changes in the exclusion rate after academy conversion (Machin & Sandi, 2020).

anticipation of their own conversion. Hatton et al. (2019), using propensity score matching, also show some performance improvements at these schools.

According to Andrews et al. (2017), converter academies have a smaller effect on GCSE performance and there is considerable variation. Only outstanding converter academies experience a significant improvement (Andrews et al., 2017). This result is confirmed in Bertoni et al. (2021). On the other hand, Hicks (2017) use a weighting scheme to ensure covariate balance prior to conversion and finds improvements only for previously underperforming schools. Interestingly, Bertoni et al. (2023), using admission lotteries and a distance-based regression discontinuity design, show that two free schools, in many ways similar to converter academies, significantly improve pupil GCSE scores.

Eyles et al. (2017) find no evidence of improvement at KS2 performance resulting from primary schools' academy conversion. Overall findings in Neri and Pasini (2018) are consistent with this. Regan-Stansfield (2018) also finds no effect on average pupil performance following conversion, but identifies a small positive effect for disadvantaged pupils. Interestingly, Neri and Pasini (2018) reveal that pupils in primary converter academies that joined a MAT have higher performance than those in standalone academies. This is confirmed in Neri et al. (2022).

Andrews et al. (2017) compare school value-added in LAs and MATs and find little difference in performance. The variation within MATs and LA maintained schools is greater than the variation between the two groups (Andrews et al., 2017). Andrews et al. (2017) argue that establishment type is less meaningful than within-school practices. Similarly, Bloom et al. (2015) link pupil achievement not to school autonomy but to how autonomy is used by schools, stressing management quality. Labour productivity, quality of provision, and financial performance are indeed correlated with pupil attainment (Bryson et al., 2023). On the other hand, Bertoni et al. (2023) attribute free schools' impact to different educational approaches. The "No Excuses" approach, for example, provides a homogeneous teaching model followed by many charter schools and that resemble the approach adapted by several academies (Bertoni et al., 2023; Eyles et al., 2016a). In fact, a systematic review reveals that pupils attending "No Excuses" charter schools have better outcomes[17] (Krowka et al., 2017). Despite institutional differences, evidence from US charter schools provides an important benchmark for academies.[18]

In brief, the abovementioned studies show that early academy conversions have a significantly positive impact on pupils' exam performance, whereas there are mixed results for conversions occurring after the academy programme was widely expanded in 2010. However, it should be noted that academy conversions are recent, and studies do not capture long term effects.

---

[17] Positive effects of "No Excuses" charter schools remain significant for three years and then stabilise or decrease (Krowka et al., 2017). A case study on a free school suggests that this approach mostly benefits boys and underachieving pupils (Bertoni et al., 2023).

[18] Using admission lotteries as instruments for charter attendance, Abdulkadiroğlu et al. (2011) find large test score gains, particularly for low ability pupils. This is consistent with other studies exploring admissions lotteries, such as Angrist et al. (2016), Dobbie and Fryer (2015), and West et al. (2016). However, observational analysis suggests that charter schools using admission lotteries have larger impacts than other charter schools (Abdulkadiroğlu et al., 2011). Abdulkadiroğlu et al. (2016), using matching and a grandfathering instrument, find that pupils enrolled at state schools that become charter schools experience achievement gains as large or larger than those admitted in charter schools through a lottery. They argue that pupils with the largest potential gains are the least likely to apply to a charter school. Similarly, Walters (2018) uses instrumental variables based on admission lottery offers and school distance and finds that low ability disadvantaged pupils experience larger test score gains but are less likely to apply to a charter school. Walters (2018) maintains that the expansion of charter schooling would increase average treatment effects.

3.2.2.2 Low Stakes Test Performance

Given the predominance in the literature of administrative data, there are very few empirical studies assessing the impact of academy attendance on low stakes tests. When we compare the performance of academies with LA maintained schools at the Programme for International Student Assessment (PISA), that assesses pupils' ability to apply knowledge to solve real world problems, we observe that sponsor-led academies and converter academies achieve the lowest and the highest average scores respectively (Jerrim & Shure, 2016). However, as mentioned before, such a comparison may simply reflect pupils' prior attainment and features of predecessor schools.

Academies are sometimes associated with a traditional academic curriculum and more demanding standards (Barker & Hoskins, 2017; R. Morris & Perry, 2019). To maximise their performance, perhaps academies rely on teaching focused on exam preparation. As learning and teaching become attached to high stakes tests, exam contents become the curriculum (Pinto & Santos, 2006; Shah, 2018). Hence, it is possible that important skills are overlooked by academies. On the other hand, some schools may be better across educational dimensions and there may be a high correlation between different outcomes.[19] This research proposes to clarify how well pupils at academies do on wider cognitive skills and to establish a causal link to academy conversion, filling a gap in the literature as there is no published evidence on this.

---

[19] For example, there is a strong correlation between GCSE grades and performance at PISA (Jerrim & Shure, 2016).

3.2.3 Effects on Non-cognitive Skills

As in the case of low stakes tests, there is little empirical evidence regarding the impact of academy attendance on non-cognitive skills. Given the pivotal role of league tables, schools perhaps neglect pupils' social development. The focus on exam performance may also create mental health issues for pupils under stress and affect their wellbeing.

Frostick et al. (2018) use the Well London Adolescent Survey to compare the mental health, school connectedness, and aspirations of secondary school pupils attending pre-2010 sponsor-led academies and LA maintained schools in deprived neighbourhoods. This study shows that pupils attending sponsor-led academies report significantly more feelings of school connectedness,[20] but there is not a significant difference on mental health, self-esteem, and aspirations. Frostick et al. (2018) test if academies have an indirect effect on aspirations and mental health through school connectedness and sustain that it has a positive mediating effect on pro-social behaviour. Frostick et al. (2018) point out efforts of early sponsor-led academies to develop new identities through rebranding, new buildings and disciplinary policies, which foster feelings of belonging and safety within school. Differences in teacher relationships, school structure, and ethos may also be a factor (Frostick et al., 2018). However, there are sample imbalances and results do not have a causal interpretation (Frostick et al., 2018).

[20] School connectedness refers to pupils feeling accepted, respected and supported at school (Frostick et al., 2018). It affects pupils' attitudes, motivation and behaviour, comprising teacher supportiveness, peer relationships, engagement in education, safety, and feelings of belonging (Frostick et al., 2018).

This chapter elaborates on those findings using a sharper identification strategy and focusing on post-2010 academies. Actually, the impact of converter academies on non-cognitive skills has not yet been studied. However, a case study of two free schools, that also have academy status, reveals a reduction in pupil absences and school mobility, possibly reflecting an improvement in pupil behaviour (Bertoni et al., 2023). Therefore, attending an academy may have a positive impact on pupils' non-cognitive skills. There is, however, mixed evidence from the US regarding the impact of charter schools on non-cognitive skills,[21] suggesting a more nuanced reality.

3.2.4 Summary

Most literature on academies is based on administrative data. Pupil intakes reveal a change in school composition after academy conversion that raises selection issues. There are positive performance effects in pre-2010 sponsor-led academies. The evidence is less clear for schools that convert after 2010. Post-2010 sponsor-led academies experience weaker effects and there are mixed results for converter academies. There is very little evidence regarding skills not measured by national exams. Despite some evidence suggesting positive effects on social behaviour, the impact of academies remains unclear. This chapter aims to shed light on that.

---

[21] Sass et al. (2016) find significant effects of charter school attendance on secondary school graduation, university attendance, and earnings, but not on test scores. Hence, Sass et al. (2016) conclude that charter schools foster non-cognitive skills which affect pupils' later outcomes. Dobbie and Fryer (2015) show that pupils attending charter schools alter their risk aversion, are less likely to get pregnant, or to be incarcerated. Nevertheless, charter schools report significantly lower non-cognitive skills such as grit or persistence (Dobbie & Fryer, 2015). Similarly, West et al. (2016) find that charter schools have positive effects on school attendance, but negative effects on conscientiousness, self-control, and grit.

3.3 Beyond Exams

I now explore cognitive and non-cognitive skills that national exams do not assess effectively. I start highlighting those skills and mention educational methods appropriate to develop them. I then discuss the relevance of problem-solving skills, mental health, and social skills.

3.3.1 Scope of High Stakes Exams

National exams are important tools in social selection and certification, validating professional qualifications and determining progression to higher education (OECD, 2016c; Pinto & Santos, 2006). Exams are also frequently used for monitoring school performance and informing government interventions (OECD, 2016c).

Exams are designed to measure, in an accurate and standardised way, to what extent pupils attain educational objectives, in relation to their cohort (Arends, 2006; Pinto & Santos, 2006). However, there are skills that exams do not assess effectively (Arends, 2006; Pinto & Santos, 2006). These may include increasingly important transferable skills, including higher order thinking, problem-solving skills, creativity, judgement, and autonomy (Arends, 2006; Delors, 1996). Of course, these skills partly rely on subject knowledge. Frequently, the best pupils in mathematics have excellent problem-solving skills (OECD, 2014). Nevertheless, the association between mathematics and problem-solving skills in PISA 2012 is weak (OECD, 2014). This suggests that problem-solving skills might not be fully captured in subject tests (OECD, 2014). The same can be said of collaborative problem-solving skills in PISA 2015 (Department for Education, 2017). Non-cognitive skills such as self-control, grit, social awareness, and empathy might also not be effectively captured by exams (Arends, 2006; Braconier, 2012; Kautz et al., 2014; M. R. West et al., 2016).

Some educational methods target those skills. "Problem-based learning" fosters transferable skills such as problem-solving and decision making skills, higher order thinking, critical thinking, metacognition and self-regulation (Arends, 2006; Davis & Harden, 1999; Hmelo-Silver, 2004; OECD, 2013, 2014, 2016c). It is designed to produce creative independent self-directed lifelong learners, capable of solving problems and acquire new skills (Arends, 2006; Dahlgren & Dahlgren, 2002; Davis & Harden, 1999; Pawson et al., 2006). Pupils are expected to develop those skills while working through a complex, ill-structured, and open-ended problem[22] that is challenging and relevant to them (Barrows, 2002; Chulkov & Nizovtsev, 2015; Hmelo-Silver, 2004). "Problem-based learning" and "cooperative learning"[23] also foster mental health, self-esteem, and social skills such as interpersonal, communication and cooperation skills (Arends, 2006; Dahlgren & Dahlgren, 2002; Davis & Harden, 1999; Gillies & Boyle, 2010; Graaff & Kolmos, 2003; Hmelo-Silver, 2004; Johnson & Johnson, 2014; Schul, 2011; Sharan, 2014). It has been claimed that pupils working cooperatively are also more creative,

[22] The problem analysis is at the basis of the problem based learning process and individual work supplements group work (Graaff & Kolmos, 2003). Pupils identify learning issues, search the knowledge they need to resolve the problem, analyse possible solutions and reflect on their effectiveness, while teachers guide and facilitate learning, creating a productive learning environment, helping pupils learn necessary skills, questioning and challenging them, supporting positive group interactions and providing feedback (Barrows, 2002; Chng et al., 2011; Davis & Harden, 1999; Hmelo-Silver, 2004; Wilkerson, 1995).

[23] Cooperative learning is characterised by positive interdependence, individual accountability, and small group interaction (Arends, 2006).

and attain higher cognitive and moral development (Johnson & Johnson, 2014).

"Problem-based learning" and "cooperative learning" are not designed for exam preparation. These methods are less structured, may lack a broad and organized knowledge framework, and are not adequate to convey a great amount of information or to cover a long curriculum (Arends, 2006; Chulkov & Nizovtsev, 2015; Davis & Harden, 1999; Graaff & Kolmos, 2003; OECD, 2016c; Pawson et al., 2006). Besides, these methods are more time-consuming and harder to implement, requiring more resources, smaller classes and training (Arends, 2006; Davis & Harden, 1999; Gillies & Boyle, 2010; OECD, 2016c; Pawson et al., 2006).

Moreover, assessment methods must be consistent with learning objectives (Graaff & Kolmos, 2003; Pawson et al., 2006). Multiple choice tests may fail to test deep understanding, and do not capture the full range of academic skills (Hmelo-Silver, 2004; Sullivan et al., 2014). It has been claimed that essay questions, tests in two phases, and portfolios can be more appropriate to assess higher order thinking and application of knowledge (Davis & Harden, 1999; Pinto & Santos, 2006). Project reports and research work are suitable for the assessment of skills (Pinto & Santos, 2006). Cooperative work is assessed through criterion-references, authentic assessments, case studies, portfolios, and exhibitions (Gillies & Boyle, 2010).

In brief, there is a risk that teaching focused on exam preparation may narrow education goals and overlook key transferable skills and educational aspects. This section will next discuss the importance of problem-solving skills, social skills, and mental health.

3.3.2 Problem-solving Skills

A successful adaptation to a changing world requires continuously acquiring new knowledge and skills and applying them to novel problems and settings, evaluating choices, and making decisions (OECD, 2013). "Learning to learn" and problem-solving skills, including critical thinking and complex reasoning, are critical for dynamic careers (Pinto & Santos, 2006). As economic and technological development enhances interpersonal and non-routine analytic tasks in the workplace, "learning to do" becomes more intellectual and focused on transferable skills (Carneiro, 2003; Delors, 1996; OECD, 2014). These skills are increasingly demanded for managerial, professional, and technical occupations (OECD, 2014). Workers are expected to be lifelong learners who can solve non-routine problems and think critically and creatively (OECD, 2014, 2017c). Furthermore, concerned citizens must be skilled problem solvers (OECD, 2013).

Problem-solving skills imply a capacity to engage in creative and critical thinking to understand and resolve problems where a solution is not obvious (OECD, 2013). It involves cognitive processes such as exploring, understanding, and representing the problem, formulating hypotheses, devising and executing a strategy while monitoring its progress, and critically reflecting on alternatives (OECD, 2013, 2014). Problem-solving is affected by motivational and affective factors, including personal beliefs on ability and interest (OECD, 2013).

3.3.3 Social Skills and Mental Health

Social skills such as communication, conflict management, teamwork, and cooperation are increasingly demanded in the labour market (Carneiro, 2003; Delors, 1996; Deming, 2017; Kautz et al., 2014; OECD, 2017a, 2017c). Professional careers often require interpersonal skills and the ability to work effectively in groups (OECD, 2017a). Individuals with social skills are more likely to work in high-paying occupations (Deming, 2017). In effect, non-cognitive skills at youth affect a range of economic, health, and social outcomes later in life (Kautz et al., 2014; Patalay & Fitzsimons, 2016; Siddiqui & Ventista, 2018). Moreover, democracy and social cohesion require communication and collaboration skills (Carneiro, 2003; Schul, 2011).

Non-cognitive skills significantly affect pupils' educational choices and school achievements (Braconier, 2012; Kautz et al., 2014). Conscientiousness, self-control, grit, and growth mindset are positively correlated with school attendance, behaviour, and test score gains (M. R. West et al., 2016). Adamecz-Volgyi et al. (2023), using Next Steps, show that inner locus of control, academic self-concept, work ethic, and self-esteem are positively correlated to educational attainment at age 16 and university attendance. Having low self-esteem significantly reduces the likelihood of going to university for high achieving pupils (Adamecz-Volgyi et al., 2023). Non-cognitive skills may be especially important for first generation university students (Adamecz-Volgyi et al., 2023).

Schools have an important role in the development of non-cognitive skills (Kautz et al., 2014). Ford et al. (2021) show that school features explain a small but significant variation in pupils' mental health. Advantaged pupil intakes, a more positive school climate, school sports, and good teacher-pupil relationships are related to better mental health (Ford et al., 2021; Marcus et al., 2020). On the other hand, exams and more instruction time increase pupils' stress (Marcus et al., 2020). In fact, a systematic review of literature reveals that non-cognitive skills can be improved by school-based interventions (Siddiqui & Ventista, 2018). Kautz et al. (2014) show that during adolescence non-cognitive skills are actually more malleable than cognitive skills, and that pupil mentoring, guidance and information are effective interventions.

Holmlund and Silva (2014) use matched difference in differences and double differences (comparing progression before and after) to study an intervention targeting non-cognitive skills of pupils at risk of truancy, exclusion, and underperformance with the aim of improving their GCSE grades. They report positive effects on school attendance, partly captured by an improved attitude toward education, increased motivation, self-esteem, and confidence. Wellbeing and mental health are known predictors of truancy (Children's Society, 2018). Holmlund and Silva (2014) only find, however, positive effects on the test scores of non-treated pupils, possibly owing to a better learning environment at school (Holmlund & Silva, 2014).

Mental health, social skills, and problem-solving skills are not fully captured by national exams. Their relevance in evaluating the impact of schooling has an important theoretical support.

3.4 Data

The Millennium Cohort Study (MCS) draws its participants from children born in the United Kingdom (UK) over a 16 month period around the year 2000 (Centre for Longitudinal Studies, 2020; Fitzsimons, Agalioti-Sgompou, et al., 2020). In England, where the sample of interest is drawn, this corresponds to a single school year, from 1 September 2000 to 31 August 2001, comprising 11,695 children. The MCS follows cohort members through their lives, currently including seven sweeps taken at nine months, age three, five, seven, 11, 14 and 17. The variables of interest are mainly taken from surveys at age 11 and 14, but information from surveys at age five and seven is also considered. At age 14, 7,814 cohort members participated in the survey in England. The MCS sample is stratified to ensure economically deprived regions and areas with high proportions of ethnic minorities are overrepresented (Centre for Longitudinal Studies, 2020; Fitzsimons, Agalioti-Sgompou, et al., 2020). Because of MCS sampling design and attrition, the data is weighted to provide a representative random sample of the population.

Despite providing extensive information on inequality and education, the MCS has very few school level data (Platt & Centre for Longitudinal Studies, 2014). The Centre for Longitudinal Studies (CLS) has provided me with the list of schools from Edubase that is used at the MCS sixth survey (MCS6)[24] interview, at age 14. I use it to create a school variable for

[24] University of London, Institute of Education, Centre for Longitudinal Studies (2020c).

establishment type group in 2014/15.[25] In order to identify school type in 2012/13 (when cohort members start secondary phase), I use the Consistent School Database (CSD)[26] that allows matching of school records between years when school identifiers (unique reference number, URN) change, as in the case of academy conversion.[27] It allows me to identify secondary school type at starting year (2012/13) and the set of secondary schools that convert to academy status in the following years (2013/14 and 2014/15).[28] The resulting school list is matched to MCS cohort members by CLS and has been used in Braz (2018). This identifies both the school attended in 2015 and the school where cohort members started secondary phase. It allows the identification of the treatment group: pupils attending LA maintained schools that become academies in 2013/14 and 2014/15. Cohort members that are not in a secondary school or moving to a secondary school in 2012/13 have been excluded from the estimation sample[29] to ensure the integrity of the treatment group (exclude pupils joining a school after academy conversion). Only cohort members whose secondary school is known are included in the sample.[30] The sample is limited to England, as there are no academies in Scotland, Wales, and Northern Ireland.

---

[25] Establishment type groups: converter academies, sponsor-led academies, LA maintained schools, free schools, independent schools.

[26] Cohort and Longitudinal Studies Enhancement Resources, CLOSER (2018).

[27] The CSD contains school level data linked over time between 1998/99 and 2016/17, identifying establishment type and academy conversion date.

[28] For the few cases where CLOSER (2018) do not identify establishment type in 2012/13, academies' predecessor school and conversion date are identified using the list provided in the Academies Management Information Data of the DfE (May 2018).

[29] At age 11, 95% cohort members are in school year six, one percent are in school year five, and four percent have already started secondary school (Platt & Centre for Longitudinal Studies, 2014). Pupils attending middle schools are excluded.

[30] Information on school where started secondary phase is missing for 207 cohort members. Plus, 15 cases are excluded for missing data on school attended at age 14. Missingness is typically from lower income families.

The estimation sample is not randomly recruited from MCS and intends to represent only the population of interest (pupils attending LA maintained schools in 2012/13). As such, it may have distinctive features without hindering external validity. In effect, when I compare this sample with the full MCS cohort for England, through a probit model (see appendix A Table A1), I find significant socioeconomic selection. This suggests pupil intakes at LA maintained schools have a more disadvantaged background, with slightly less educated parents in non-managerial/professional occupations, living in more deprived urban areas (see sub-section 3.2.1). National statistics are consistent with this. In 2012/13, 13.7% of pupils at secondary academies were eligible for FSM compared with 16.3% across all secondary state schools (Department for Education, 2013b).

Table 3.1 presents basic information on the distribution of pupils across schools in the estimation sample. The 3,125 cohort members in the sample of interest are distributed across 827 schools. The sample focuses on pupils that enrol in LA maintained schools at the start of secondary phase. In 2012/13, 3,443 (44%) cohort members enrol in LA maintained schools and 3,557 (47%) in academies. The remainder 7% are in free schools and independent schools.[31] The 179 cohort members enrolled in middle schools are excluded. This is consistent with the national distribution. At secondary phase in 2012/13 there were 1,266 academies (963 converter and 303 sponsor-led academies) (Eyles et al., 2018), representing 50% of secondary schools (National Audit Office, 2018).

[31] School type at start of secondary phase is unknown for 248 cohort members.

Table 3.1 Sample distribution per school type in 2012/13

	Number of pupils
LA maintained schools (CONTROL GROUP)	2,563
Schools that convert in 2013/14 and 2014/15 (TREATMENT GROUP)	
Prospective sponsor-led academies	199
Prospective converter academies	363
TOTAL	**3,125**
	Number of schools
LA maintained schools	696
Prospective sponsor-led academies	46
Prospective converter academies	85
TOTAL	**827**

Note: Statistics do not reflect MCS survey settings (unweighted measures). The sample is restricted to pupils reported as attending a secondary school or moving to secondary school in 2012/13 in England (University of London, Institute of Education, Centre for Longitudinal Studies, 2020c).

In the period of interest (2013/14 and 2014/15), 228 LA maintained schools convert – 151 converter and 77 sponsor-led academies (Eyles et al., 2018). The treatment group, drawn from this set, comprises 562 pupils enrolled in 131 LA maintained schools that later become academies. More than half of schools that convert during this period are represented in the sample of interest. The fact that 185 cohort members have moved secondary school raises issues of non-compliance with treatment assignment that will be considered in the methodology.

Focusing on pupils that remain in the same school,[32] MCS6 reveals that 7% of the treatment group attend faith schools. 7% of those attending prospective converter academies are in grammar schools and 15% in single-sex schools. In comparison, no one in grammar schools and only 3% in single-sex schools attend prospective sponsor-led academies. These figures roughly reflect contemporaneous national trends of academy conversion at the time.

[32] MCS6 has information on secondary schools attended at age 14, but not for schools at start of secondary phase (except for grammar schools).

Comparing KS4 performance tables in 2013 and 2015, 14% of the increase in the number of academies is observed in faith schools. Moreover, the number of grammar schools and single-sex schools with converter academy status has increased (6% and 10% respectively) between 2013 and 2015. In contrast, the number of single-sex sponsor-led academies marginally decreased nationwide in that period. Besides, sponsor-led academies are not allowed to be selective (Long et al., 2023). The treatment group thus seem a fitting representation of schools becoming academies in 2013/14 and 2014/15.

3.4.1 Outcomes of Interest

While most of the literature focuses on national exams, this chapter studies a wider range of outcomes such as problem-solving skills, social skills, and mental health. Variables ought to be available at two points in time (pre-treatment and post-treatment). See

Table 3.2 for a list of outcome variables, their basic meaning and range. To allow for comparison and interpretation, outcome variables are standardised to mean zero and standard deviation one.

The outcomes of interest cover different aspects of pupil performance. Word Activity, CGT Quality of Decision Making, and CGT Risk Adjustment reveal a set of cognitive skills, whereas CGT Risk-taking, CGT Deliberation Time, SDQ Total Difficulties, SDQ Pro-social, and the Rosenberg Self-esteem Scale focus on mental health and non-cognitive skills. A more detailed consideration of each outcome of interest follows.

Table 3.2 Outcome variables list

	Original Scale	Sample Range
COGNITIVE SKILLS		
Word Activity	number of correct responses out of 20	0 – 18
CGT Quality of Decision Making	mean proportion of trials most likely outcome chosen	0 – 1
CGT Risk Adjustment	degree betting behaviour varies with boxes ratio	-2 – 5
NON-COGNITIVE SKILLS		
CGT Risk Taking	mean proportion gambled on trials most likely outcome chosen	0 – 1
CGT Deliberation Time	mean time to choose box colour (milliseconds)	362 – 23,691
SDQ Total Difficulties	3-point Likert scale (20 items)	0 – 38
SDQ Pro-social	3-point Likert scale (5 items)	0 – 10
Shortened Rosenberg Self-esteem Scale	4-point Likert scale (5 items)	5 – 20

Note: Outcome variables are taken from MCS6, age 14. (Cambridge Cognition, 2019; Closs, 1976; R. Goodman, 1997; Rogers et al., 1999; Rosenberg, 1965).

3.4.1.1 Cognitive Skills

The MCS has administered several cognitive tests to cohort members. These include the British Ability Scales (BAS) Picture Similarity at age five, BAS Pattern Construction at age five and seven, BAS Word Reading at age seven, BAS Verbal Similarities at age 11, Cambridge Neuropsychological Test Automated Battery (CANTAB) Spatial Working Memory Task at age 11, CANTAB Cambridge Gambling Task at age 11 and 14, and Word Activity at age 14 (Centre for Longitudinal Studies, 2020; Fitzsimons, Agalioti-Sgompou, et al., 2020).

The Word Activity, at age 14, is a shortened version of the Applied Psychology Unit (APU) Vocabulary Test (Closs, 1976) used in the 1970 British Cohort Study. It measures understanding of the meaning of words. At this activity, cohort members select in a multiple-choice list the word with the same meaning as the one presented to them. Word Activity, at age 14, can be compared to the BAS Verbal Similarities, at age 11 (fifth survey, MCS5).[33] The BAS Verbal Similarities (Elliott, 1996) measures verbal reasoning and acquired knowledge. At BAS Verbal Similarities, cohort members are given three stimulus words and asked to name the class to which they belong.

The CANTAB Cambridge Gambling Task (CGT) (Cambridge Cognition, 2019; Rogers et al., 1999), at age 11 and 14, assesses decision making skills under uncertainty, outside a learning context. Cohort members must place a bet on whether a token is hidden under a red or blue box, and if the bet is correct (incorrect) the number of points is added (subtracted) to their score. The likelihood of each choice is indicated by the ratio of red to blue boxes. CGT Quality of Decision Making shows the mean proportion of trials a cohort member chooses the most likely option (more probable colour), denoting better decision making. CGT Risk Adjustment measures the extent to which cohort members bet higher proportions on larger ratio trials. It shows how risk taking varies in response to setting and available information. Adapting reward seeking behaviour to statistical risk implies a greater propensity to seek reward when the probability of obtaining it is high. CGT Quality of Decision Making and CGT Risk Adjustment reveal wider cognitive skills, including problem solving and decision making (Platt & Centre for Longitudinal Studies, 2014). They are also positively associated with intelligence (Deakin et al., 2004; Flouri et al., 2019).

A few pre-treatment variables can be used as a proxy for these cognitive skills. For example, BAS Picture Similarity measures problem solving

[33] University of London, Institute of Education, Centre for Longitudinal Studies (2022).

ability and non-verbal reasoning at the third survey (MCS3),[34] age five. BAS Pattern Construction measures spatial problem solving at age five and at the fourth survey (MCS4),[35] age seven. CANTAB Spatial Working Memory (SWM) measures ability to retain spatial information in working memory and to employ a problem-solving strategy at age 11. These variables are included in the model as explanatory variables.

3.4.1.2 Non-cognitive Skills

The MCS includes the Big Five, a taxonomy of non-cognitive skills comprising openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism (Fitzsimons, Haselden, et al., 2020). Unfortunately, parents and cohort members are only asked the Big Five at age 14 and 17, not allowing for a comparison before secondary phase. Those five descriptors of personality are, however, related with measures of non-cognitive skills available both at age 11 and 14. Basically, conscientiousness is related to deliberation, extraversion to thrill-seeking, agreeableness to altruism and cooperation, and neuroticism to self-esteem and mental disorders (Kautz et al., 2014).

The MCS administers the Strengths and Difficulties Questionnaire (SDQ) (R. Goodman, 1997, 2001), a behaviour screening questionnaire filled out by the parents at age three, five, seven, 11 and 14 (Centre for Longitudinal Studies, 2020; Fitzsimons, Agalioti-Sgompou, et al., 2020). The SDQ Total

[34] University of London, Institute of Education, Centre for Longitudinal Studies (2020d).

[35] University of London, Institute of Education, Centre for Longitudinal Studies (2020a).

Difficulties subscale is used to reveal behaviour issues regarding emotional problems, conduct problems, hyperactivity/inattention, and peer problems. Higher scores indicate greater symptoms. It is a measure of overall mental health issues, correlated with child mental disorder (A. Goodman et al., 2010). The SDQ Pro-social subscale identifies pro-social behaviour, revealing social skills and adaptive behaviour. Teacher reported SDQ variables at age 11 are used for robustness checks.

CGT also provides behaviour indicators at age 11 and 14 (Platt & Centre for Longitudinal Studies, 2014). CGT Deliberation Time measures the mean time taken to make a response, indicating decision making latency. CGT Risk Taking measures the mean proportion gambled on trials where the most likely option is chosen, implying higher (lower) sensitivity to reward (punishment). CGT Risk Taking, typically taken to approximate thrill-seeking, is significantly associated with socially inappropriate behaviour such as misbehaving in class and bullying (Brandt et al., 2019; Flouri & Papachristou, 2019).

Additionally, MCS includes the Shortened Rosenberg Self-esteem Scale (Rosenberg, 1965) that measures self-satisfaction at age 11 and 14, based on self-reported agreement or disagreement to five positively worded statements (Fitzsimons, Agalioti-Sgompou, et al., 2020; Platt & Centre for Longitudinal Studies, 2014). Higher scores indicate greater self-esteem. Low self-esteem is associated with anxiety and depression (Bannink et al., 2016).

3.4.2 Explanatory Variables

Administrative data includes information on pupils' gender, ethnic group, language, SEN statement, and FSM eligibility. The MCS contains data that allows to construct comparable variables. Aggregate ethnic groups (six categories) are based on the Census categories (Office for National Statistics, 2014). In general, the construction of variables follows previous literature using the MCS (Dearden et al., 2011; Sullivan et al., 2013).[36]

MCS also includes the Organisation for Economic Cooperation and Development (OECD) equivalised weekly family income (Centre for Longitudinal Studies, 2020; Hagenaars et al., 1994). Whereas FSM eligibility reveals how the most disadvantaged pupils are sorted (families living below a poverty threshold), income also describes the pattern followed by more advantaged families. Furthermore, equivalisation[37] allows a comparison of income between households of different sizes (Department for Work and Pensions, 2012). I use the log of OECD equivalised weekly family income.[38]

---

[36] In the construction of variables, I mainly use the MCS5 and MCS6. MCS5 is collected in 2012 at age 11, the last year of primary phase and before transition to secondary school (Centre for Longitudinal Studies, 2020). MCS6 is taken in 2015 at age 14, when cohort members are already attending a secondary school (Fitzsimons, Agalioti-Sgompou, et al., 2020). A few lagged variables of interest are included from MCS3 and MCS4, at age five and seven respectively. Since cohort members attending schools that become academies after 2015 are not identified in available data, variables from the seventh survey (MCS7) at age 17 (University of London, Institute of Education, Centre for Longitudinal Studies, 2020b), taken in 2018, are not included in order to ensure that schools in the control group have not converted to academy status before pupil assessment.

[37] The OECD modified equalisation scale assumes that the needs of a household grow with each additional member but not in a proportional way. The OECD modified equalisation scale is re-scaled to take a couple without children as the reference group, assigning a value of one to the main carer, of 0.33 to each additional adult member and of 0.2 to each child (aged under 14).

[38] Income is a strictly positive ratio scale variable that is skewed to the right. The logarithmic transformation reduces data heteroscedasticity or skewedness and makes the estimates less sensitive to outliers, narrowing the variable range (Wooldridge, 2014). Moreover, when using logs the coefficient does not depend on the unit of measurement and approximates a percentage change (Wooldridge, 2014).

Alternative measures include the OECD below 60% median indicator and the OECD equivalised income quintiles that are used in robustness checks.

MCS data on parents' education and occupation allows to further differentiate pupils' background. I consider both the mother and the father (where present in the sample)[39] to derive parents' socio-economic status. Parental occupation (current job) is derived from the National Statistics Socio-economic Classification (NS-SEC) (Rose & Pevalin, 2003). The NS-SEC has 13 operational categories and is used in a five-category version. Since it excludes parents that are not in work nor on leave, a category identifying those cases is added. Similarly, I use the highest educational qualification of parents (across all sweeps), looking at both academic and vocational qualifications, grouped in the five National Vocational Qualifications (NVQ) levels (Agalioti-Sgompou et al., 2017).[40] An alternative variable focused on academic qualifications is used in robustness checks.

Related variables include parents' age, labour market status, hours worked (part-time or full-time),[41] family composition (number of parents, grandparents, siblings), housing tenure, main carer's mental health and longstanding illness. Housing tenure types follow those used in the English Housing Survey (Ministry of Housing, Communities and Local Government, 2020). The main respondent's mental health is screened using a standardised

[39] The cohort members' main carer is typically the mother. Item nonresponse is common in data from partners.

[40] NVQ level one corresponds to approximately GSCE D – G, NVQ level two to GSCE A* – C, NVQ level three to A levels, NVQ level four to a Bachelor of Arts (BA) degree and NVQ level five to a postgraduate degree.

[41] The OECD defines a part-time job as working less than 30 hours a week (Bastelaer et al., 1997).

Kessler Six Scale (Kessler et al., 2003) that quantifies psychological distress based on self-reported symptoms of depression and anxiety.

Geographical data[42] available in the MCS identifies urban context and local level of deprivation. Postcodes have been classified into rural and urban Lower Super Output Areas (LSOA), based on the Office for National Statistics (ONS) Rural/Urban Classification (Bibby & Shepherd, 2004).[43] Six domain deciles of the Index of Multiple Deprivation (IMD) 2004 (Office of the Deputy Prime Minister & Neighbourhood Renewal Unit, 2004) measure relative levels of different types of deprivation at LSOAs.

A more precise matching exercise is possible due to MCS rich data on cohort members. Available information includes cohort members' age, general level of health, bedtime rules, how often someone helps doing homework, hours per weekday spent watching television or videos, number of books at home, visits to library, and church attendance. Most variables used are available both at age 11 and 14, but data from earlier sweeps is also considered such as whether received childcare, looked after by grandparents, and how often someone reads to cohort member at age five. There is not much school level information available on the MCS, but it is possible to identify cohort members that have attended more than two primary schools and those

---

[42] Geo-referenced data at MCS is derived using mean unit postcode centroids based in the ONS Postcode Directory and point data generated from postcodes collected during interviews (Church, 2017).

[43] LSOAs forming settlements with populations of over 10,000 are urban, while the remainder are defined as one of the three rural types: town and fringe, village or hamlet and isolated dwelling.

that have attended independent schools, single sex schools or faith schools in primary phase.[44]

To minimise missing data due to item nonresponse, an additional missing category for categorical variables and a dummy flag for continuous variables (combined with mean imputation) is created. Item nonresponse is assumed to be missing at random.

3.4.3 Descriptive Statistics

The outcomes of interest at the end of primary school and towards the end of secondary school are presented in Table 3.3. The performance of pupils at LA maintained schools that convert to academy status in 2013/14 and 2014/15 (treatment group) is compared to that of pupils at LA maintained schools that have not converted (control group). The treatment group is separated into sponsor-led academies and converter academies. Figure 3.1 represents graphically the distribution of each outcome variable before and after treatment.

Before secondary school there are no significant differences in cognitive skills between the treatment and the control groups, even when considering sponsor-led and converter academies separately. Likewise, non-cognitive skills at the end of primary school are similar in both groups. However, the measure of self-esteem is lower for pupils attending prospective sponsor-led academies and higher for those attending prospective converter

[44] Similar information is gathered for secondary schools attended at age 14, but not for schools where cohort members start secondary phase (grammar school variables are an exception). Using these variables creates missing data for pupils that change secondary school.

academies. Incidentally, this difference disappears at secondary school. Pupils at prospective sponsor-led academies also have more mental health issues at age 11.

Pupils' performance at secondary school differs somewhat between the treatment and the control groups, particularly when we look to sponsor-led and converter academies separately. Sponsor-led academies significantly underperform at the more conventional cognitive test (Word Activity). On wider cognitive skills such as decision making and risk adjustment, pupils at converter academies perform better than their peers at LA maintained schools. On the other hand, pupils at sponsor-led academies adjust less their behaviour to available information. These differences are consistent with school performance at high stakes exams, where converter academies generally outperform LA maintained schools and sponsor-led academies (Bolton, 2012, 2015; Department for Education, 2018).

Regarding non-cognitive skills, the pattern emerging at secondary school requires further investigation. Overall, pupils at academies score significantly less in risk taking and deliberation time. Basically, pupils at sponsor-led academies are more risk averse, while those at converter academies make decisions faster. This may denote a distinct behavioural profile. Pupils at converter academies also show fewer mental health issues. On the other hand, the evolution of self-esteem since primary school indicates that pupils' self-worth at converter academies has worsen and at sponsor-led academies has improved. This is interesting since satisfaction with life generally decreases during adolescence (OECD, 2017b). However, at age 14 pupils' self-esteem is similar across groups, suggesting mean reversion.

Table 3.3 Outcome variables descriptive statistics

	CONTROL		TREATMENT			SPONSOR-LED			CONVERTER		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
VARIABLES	N	Mean	N	Mean	Diff	N	Mean	diff	N	Mean	diff
POST TREATMENT (AGE 14) COGNITIVE SKILLS											
Word Activity	2,364	-0.102	523	-0.192	-0.0898	193	-0.333	-0.231***	330	-0.093	0.00931
		[0.972]		[0.946]	(0.0648)		[0.832]	(0.0476)		[1.013]	(0.0950)
CGT Quality of Decision Making	2,379	-0.022	521	0.008	0.0303	188	-0.138	-0.116	333	0.107	0.129*
		[0.978]		[0.936]	(0.0655)		[0.961]	(0.100)		[0.895]	(0.0710)
CGT Risk Adjustment	2,378	-0.031	521	-0.007	0.0235	188	-0.201	-0.171**	333	0.125	0.155**
		[0.986]		[1.031]	(0.0646)		[0.878]	(0.0866)		[1.113]	(0.0787)
NON-COGNITIVE SKILLS											
CGT Risk Taking	2,378	0.084	521	-0.026	-0.111**	188	-0.051	-0.135	333	-0.010	-0.0944
		[1.012]		[0.972]	(0.0561)		[0.894]	(0.0820)		[1.025]	(0.0656)
CGT Deliberation Time	2,379	0.081	521	-0.060	-0.141**	188	-0.033	-0.114	333	-0.078	-0.159**
		[1.219]		[0.824]	(0.0579)		[0.757]	(0.0742)		[0.870]	(0.0743)
SDQ Total Difficulties	2,508	0.136	553	0.083	-0.0533	198	0.244	0.108	355	-0.026	-0.162*
		[1.092]		[1.012]	(0.0756)		[0.961]	(0.0959)		[1.029]	(0.0972)
SDQ Pro-social	2,515	-0.085	553	-0.018	0.0671	198	-0.017	0.0685	355	-0.019	0.0661
		[1.055]		[0.996]	(0.0552)		[0.892]	(0.0825)		[1.066]	(0.0604)
Rosenberg Self-esteem Scale	2,407	-0.045	533	-0.051	-0.00569	193	-0.137	-0.0921	340	0.008	0.0530
		[1.014]		[1.043]	(0.0725)		[0.959]	(0.116)		[1.095]	(0.0776)
PRE-TREATMENT (AGE 11) COGNITIVE SKILLS											
Verbal Similarities Ability Score	2,513	-0.090	548	-0.213	-0.123	196	-0.226	-0.136	352	-0.205	-0.115
		[1.008]		[0.990]	(0.0810)		[0.924]	(0.135)		[1.033]	(0.0942)
CGT Quality of Decision Making	2,411	0.048	539	-0.012	-0.0601	194	-0.028	-0.0760	345	-0.002	-0.0494
		[1.005]		[1.046]	(0.0607)		[0.942]	(0.0853)		[1.117]	(0.0778)
CGT Risk Adjustment	2,411	-0.023	539	-0.017	0.00546	194	-0.014	0.00860	345	-0.019	0.00333
		[0.970]		[0.927]	(0.0521)		[0.900]	(0.0826)		[0.941]	(0.0600)

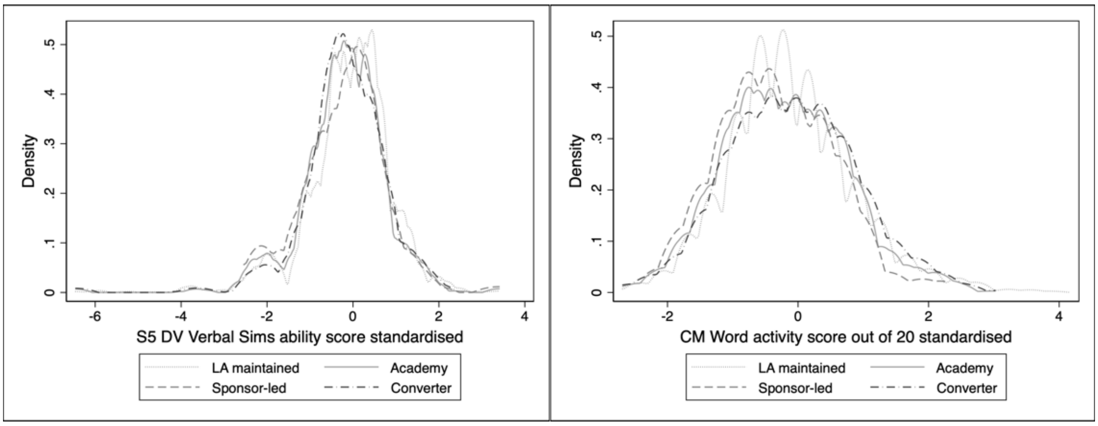
	CONTROL		TREATMENT			SPONSOR-LED			CONVERTER		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
VARIABLES	**N**	**Mean**	**N**	**Mean**	**Diff**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**
NON-COGNITIVE SKILLS											
CGT Risk Taking	2,411	0.075	539	0.043	-0.0321	194	-0.016	-0.0911	345	0.083	0.00794
		[0.984]		[0.913]	(0.0497)		[0.875]	(0.0872)		[0.934]	(0.0581)
CGT Deliberation Time	2,411	-0.010	539	-0.053	-0.0433	194	-0.035	-0.0253	345	-0.065	-0.0555
		[0.886]		[0.863]	(0.0508)		[0.757]	(0.0837)		[0.936]	(0.0516)
SDQ Total Difficulties	2,476	0.172	541	0.151	-0.0213	190	0.335	0.162*	351	0.034	-0.139
		[1.110]		[1.065]	(0.0776)		[1.042]	(0.0939)		[1.058]	(0.0943)
SDQ Pro-social	2,485	-0.093	545	-0.066	0.0271	193	-0.087	0.00520	352	-0.051	0.0412
		[1.111]		[0.993]	(0.0687)		[0.970]	(0.123)		[1.005]	(0.0661)
Rosenberg Self-esteem Scale	2,329	-0.080	497	-0.106	-0.0261	165	-0.401	-0.321***	332	0.076	0.156***
		[1.067]		[0.983]	(0.0800)		[0.925]	(0.117)		[0.969]	(0.0563)

Note: The treatment group includes pupils enrolled in 2012/13 at LA maintained secondary schools that become academies in 2013/14 and 2014/15, whereas the control group includes pupils enrolled at LA maintained secondary schools that do not convert to academy status. The treatment group is composed by pupils attending prospective sponsor-led and converter academies. All variables are standardised to mean zero and standard deviation one. All statistics, except the number of observations (N) that is unweighted, reflect MCS survey settings. The quantity in parenthesis below the mean is the standard deviation. The difference (diff) refers to the difference in means between the treatment group and the control group and the quantity in parenthesis below is the (robust) standard error clustered at the school level and reflecting MCS survey settings. (University of London, Institute of Education, Centre for Longitudinal Studies, 2020c, 2022). * significant at 10% level; ** significant at 5% level; *** significant at 1% level or below.

Figure 3.1 Density of outcome variables before and after treatment

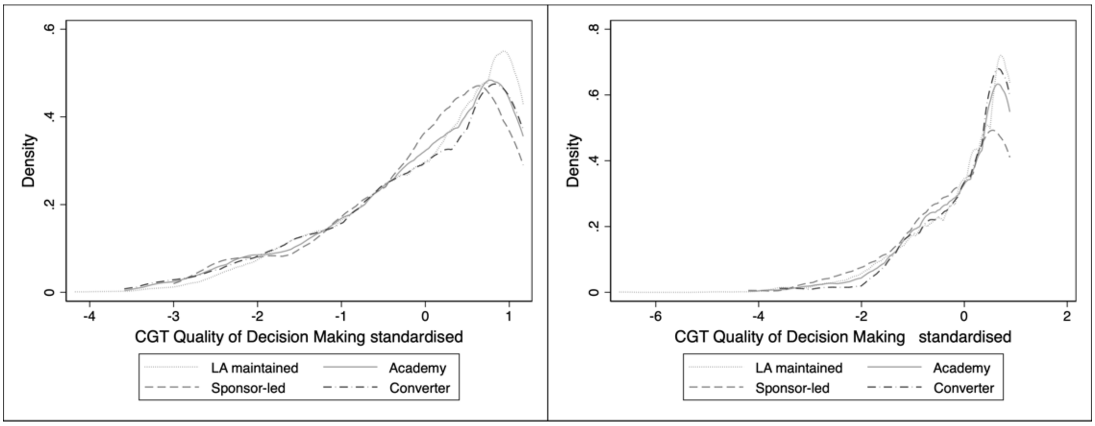
Word Activity

Age 11 **Age 14**



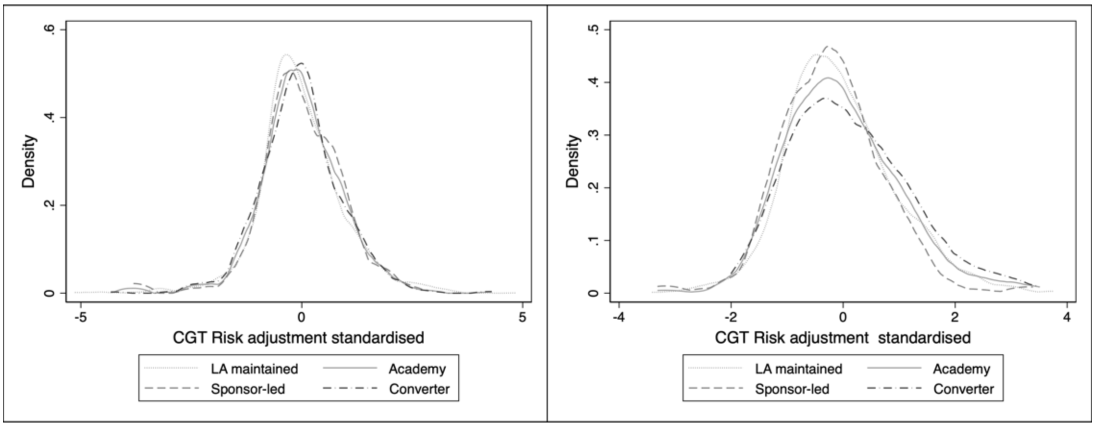
CGT Quality of Decision Making

Age 11 **Age 14**



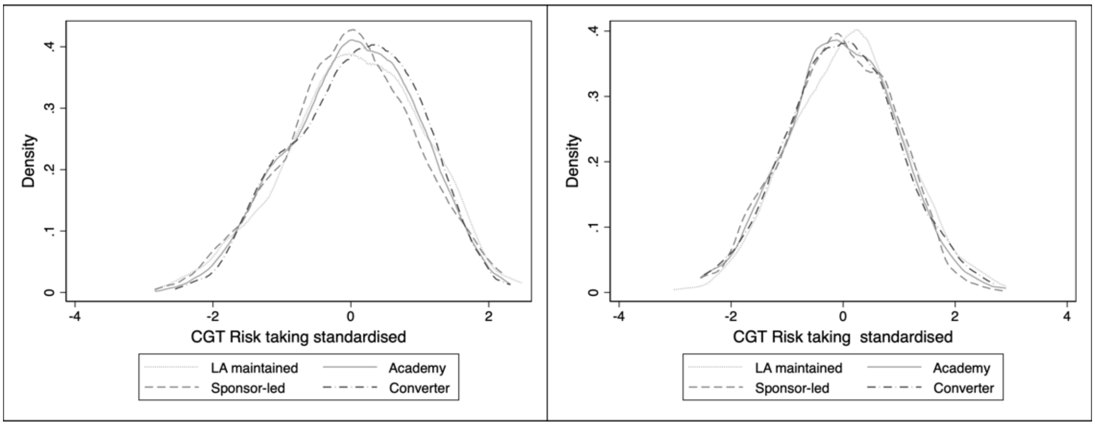
CGT Risk Adjustment

Age 11 **Age 14**



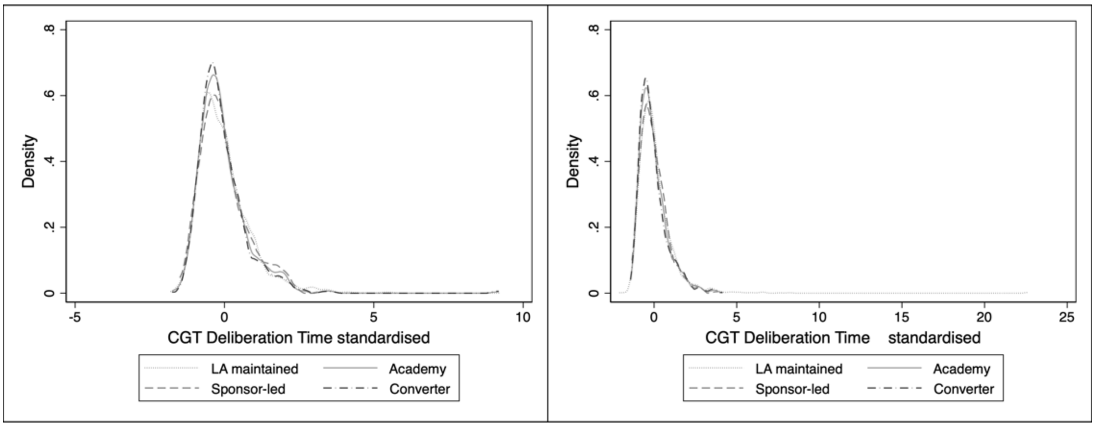
CGT Risk Taking

Age 11 **Age 14**



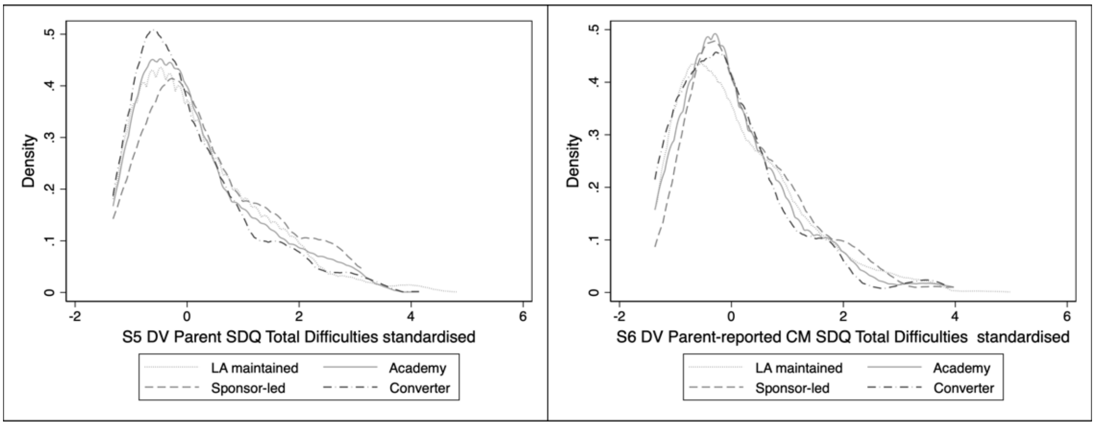
CGT Deliberation Time

Age 11 **Age 14**



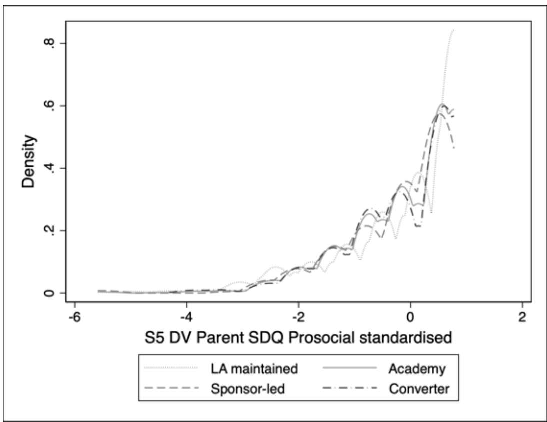
SDQ Total Difficulties

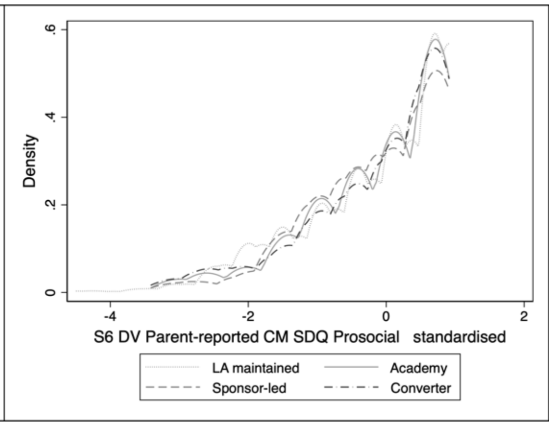
Age 11 **Age 14**



SDQ Prosocial

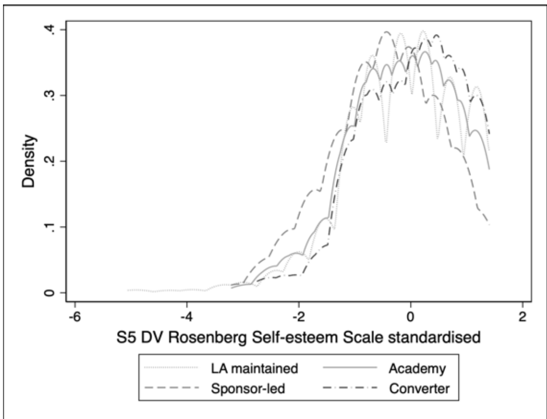
Age 11 | **Age 14**

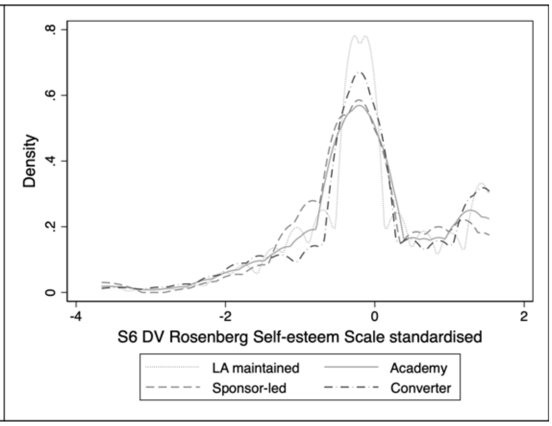




Rosenberg Self-esteem Scale

Age 11 | **Age 14**





Note: The treatment group includes pupils enrolled in 2012/13 at LA maintained secondary schools that become academies in 2013/14 and 2014/15, whereas the control group includes pupils enrolled at LA maintained secondary schools that do not convert to academy status. The treatment group is composed by pupils attending prospective sponsor-led and converter academies. The kernel density estimate for each outcome variable is displayed before (age 11) and after treatment (age 14). The variables reflect MCS survey weights. (University of London, Institute of Education, Centre for Longitudinal Studies, 2020c, 2022).

Possible selection issues undermine any conclusion we might draw from raw comparisons of pupil performance between academies and LA maintained schools. Table 3.4 presents descriptive statistics[45] on the characteristics of pupils in the treatment group and in the control group (see appendix A Table A2 for descriptive statistics on the full list of explanatory variables and Table A3 for Chi-square tests[46] on categorical variables).

A raw comparison suggests that pupils in LA maintained schools that convert to academy status in 2013/14 and 2014/15 are not dissimilar. Most variables are not significantly different and there is no headline difference between the treatment and the control groups. There are, though, a few gaps. Pupils in the treatment group are less likely to go to church, or to visit a library, and have a more regular bedtime. Their fathers are more likely to have stayed in education after age 16. However, when the treatment group is divided according to the academy route taken, the differences are more evident.

Pupils attending LA maintained schools that later become sponsor-led academies come from larger families, with younger parents who have a lower income and are less likely to have a degree or a managerial/professional occupation. These families live in somewhat more deprived urban communities. Pupils in prospective sponsor-led academies are less likely to have attended an independent primary school or a faith primary school. Looking into these pupils' prior attainment is clear they have significantly

---

[45] The sample is weighted according to MCS survey design and attrition.

[46] Chi-square tests are used to verify independence between categorical variables and treatment assignment. Since this does not clearly identify the source of a statistically significant difference, a further investigation of individual categories is required when that happens (Sharpe, 2015).

underperformed at reading (age seven) and problem-solving (age five). They also show more mental health issues at age five. Nevertheless, fewer pupils in these schools have a statement of SEN. This profile is consistent with national statistics. Schools becoming sponsor-led academies in 2013/14 and 2014/15 have 23% FSM eligible pupils (Hatton et al., 2019).

Pupils attending LA maintained schools that later become converter academies have more advantaged characteristics. This is also consistent with national trends (Bolton, 2015). There are fewer pupils eligible for FSM at these schools and they usually live with two carers in their own house, at slightly less deprived areas. Their fathers are slightly more educated, and their mothers have better mental health. Pupils attending prospective converter academies have more regular bedtimes and spend less time watching television or videos. These pupils show fewer mental health issues at age five.

The abovementioned differences between the groups are based on many independent tests. However, the probability of finding at least one statistically significant difference across a set of tests, even when there is no actual difference, increases with each additional test (Armstrong, 2014; Gelman et al., 2012). This implies that, given the number of tests performed, a few statistically significant differences would always be expected, regardless of the true associations. As suggested by Armstrong (2014), a Bonferroni correction[47] is appropriate when testing the global null hypothesis that multiple

[47] The Bonferroni correction adjusts the p-value (the significance level is divided by the total number of tests performed) so as to maintain the overall significance level (Armstrong, 2014; Gelman et al., 2012; Sharpe, 2015; VanderWeele & Mathur, 2019; Wright, 1992). It ensures that the global null hypothesis is rejected no more than the significance level when it in fact holds.

tests are not significant, so as to ascertain if a particular treatment have the estimated effect. In this case, the Bonferroni correction shows that no variable remains significant (at 10% significance level) when comparing the treatment and the control groups before secondary phase.

This result reinforces the initial impression that there is no headline difference between pupils attending LA maintained schools and those attending schools that later become academies. Nevertheless, the apparent different profile of pupils attending sponsor-led and converter academies, that is consistent with the way schools are selected for conversion, renders this distinction very relevant and requires a study of possible heterogenous effects. Next section will consider how to address possible selection issues and the identification strategy.

Table 3.4 Descriptive statistics pre-treatment

	CONTROL		TREATMENT			SPONSOR-LED			CONVERTER		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
VARIABLES	**N**	**Mean**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**
AGE 11											
CM CHARACTERISTICS											
CM Sex (Male)	2,563	0.531	562	0.525	-0.00525	199	0.480	-0.0509	363	0.555	0.0247
		[0.502]		[0.498]	(0.0368)		[0.471]	(0.0582)		[0.513]	(0.0386)
SEN statement	2,558	0.056	560	0.035	-0.0209	199	0.014	-0.042***	361	0.049	-0.00704
		[0.232]		[0.184]	(0.0131)		[0.111]	(0.00881)		[0.223]	(0.0195)
FAMILY CONTEXT											
One parent/carer in Household	2,563	0.306	562	0.269	-0.0368	199	0.325	0.0197	363	0.232	-0.074***
		[0.464]		[0.443]	(0.0280)		[0.442]	(0.0467)		[0.436]	(0.0283)
Number of siblings in Household	2,563	1.604	562	1.737	0.133	199	1.875	0.270**	363	1.647	0.0428
		[1.174]		[1.203]	(0.0832)		[1.138]	(0.125)		[1.236]	(0.103)
Age (Main)	2,563	39.007	562	38.543	-0.464	199	37.232	-1.775***	363	39.403	0.396
		[6.448]		[6.046]	(0.436)		[5.285]	(0.494)		[6.390]	(0.586)
SOCIOECONOMIC STATUS											
OECD equiv weekly family income (log)	2,563	5.829	562	5.813	-0.0159	199	5.698	-0.131*	363	5.889	0.0598
		[0.478]		[0.474]	(0.0487)		[0.432]	(0.0683)		[0.485]	(0.0659)
Free School Meal Eligible	2,555	0.248	560	0.220	-0.0278	198	0.284	0.0362	362	0.178	-0.0698**
		[0.435]		[0.413]	(0.0265)		[0.425]	(0.0547)		[0.395]	(0.0291)
Housing Tenure	2,537		557			196			361		
Rent privately		0.186		0.145	-0.0404*		0.158	-0.0273		0.137	-0.0489**
		[0.393]		[0.353]	(0.0224)		[0.348]	(0.0417)		[0.355]	(0.0212)
Owner occupier		0.511		0.560	0.0496		0.437	-0.0739*		0.639	0.129***
		[0.505]		[0.498]	(0.0351)		[0.472]	(0.0442)		[0.496]	(0.0426)
LA or HA rent		0.304		0.295	-0.00918		0.405	0.101**		0.224	-0.0797*
		[0.464]		[0.457]	(0.0354)		[0.467]	(0.0445)		[0.431]	(0.0440)
Occupation (Main)	2,519		553			197			356		
Managerial and Professional		0.194		0.158	-0.0357*		0.096	-0.098***		0.199	0.00545
		[0.399]		[0.364]	(0.0201)		[0.278]	(0.0202)		[0.413]	(0.0229)

	CONTROL		TREATMENT			SPONSOR-LED			CONVERTER		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**
VARIABLES	**N**	**Mean**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**
Intermediate		0.150		0.154	0.00390		0.158	0.00828		0.151	0.000993
		[0.360]		[0.360]	(0.0206)		[0.344]	(0.0352)		[0.370]	(0.0205)
Small employers		0.060		0.066	0.00658		0.021	-0.038***		0.096	0.0364
and self-employed		[0.239]		[0.248]	(0.0188)		[0.136]	(0.00959)		[0.304]	(0.0274)
Lower supervisory		0.017		0.028	0.0107		0.042	0.0245		0.019	0.00150
and technical		[0.131]		[0.164]	(0.00961)		[0.188]	(0.0195)		[0.140]	(0.00786)
Semi-routine and		0.219		0.232	0.0124		0.247	0.0277		0.222	0.00222
Routine		[0.417]		[0.421]	(0.0256)		[0.406]	(0.0414)		[0.429]	(0.0303)
Not in work nor on		0.360		0.362	0.00214		0.435	0.0752		0.313	-0.0465
Leave		[0.484]		[0.479]	(0.0396)		[0.467]	(0.0641)		[0.479]	(0.0484)
Education (Main)	2,144		461			147			314		
NVQ equiv. level 1		0.124		0.116	-0.00826		0.136	0.0120		0.105	-0.0193
(incl. GSCE D – G)		[0.334]		[0.323]	(0.0183)		[0.330]	(0.0311)		[0.318]	(0.0215)
NVQ equiv. level 2		0.334		0.388	0.0536*		0.452	0.118**		0.353	0.0188
(incl. GSCE A* – C)		[0.479]		[0.492]	(0.0316)		[0.478]	(0.0510)		[0.495]	(0.0368)
NVQ equiv. level 3		0.175		0.193	0.0185		0.203	0.0286		0.188	0.0131
(incl. A levels)		[0.385]		[0.399]	(0.0241)		[0.387]	(0.0475)		[0.405]	(0.0237)
NVQ equiv. level 4		0.296		0.242	-0.0535*		0.181	-0.115***		0.275	-0.0202
(incl. BA degree)		[0.463]		[0.433]	(0.0290)		[0.370]	(0.0332)		[0.463]	(0.0347)
NVQ equiv. level 5		0.071		0.061	-0.0103		0.028	-0.043***		0.079	0.00764
(incl. postgraduate)		[0.261]		[0.242]	(0.0133)		[0.158]	(0.0136)		[0.279]	(0.0171)
HOME ENVIRONMENT											
Visited library at	2,563	0.279	562	0.222	-0.0575*	199	0.253	-0.0257	363	0.201	-0.0784**
least monthly		[0.452]		[0.414]	(0.0298)		[0.410]	(0.0433)		[0.413]	(0.0367)
Attended religious	2,561	0.275	562	0.198	-0.077**	199	0.170	-0.104**	363	0.216	-0.0589
service monthly		[0.450]		[0.398]	(0.0360)		[0.355]	(0.0488)		[0.425]	(0.0459)
Regular bedtime	2,563	0.869	562	0.904	0.0354*	199	0.887	0.0182	363	0.915	0.0467**
(usually, always)		[0.340]		[0.294]	(0.0186)		[0.299]	(0.0275)		[0.288]	(0.0212)

	CONTROL		TREATMENT			SPONSOR-LED			CONVERTER		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
VARIABLES	**N**	**Mean**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**
AGE 7											
COGNITIVE SKILLS											
Word Reading	2,312	111.316	515	110.365	-0.951	187	108.681	-2.635**	328	111.529	0.213
Standardised		[17.611]		[17.698]	(1.020)		[16.708]	(1.282)		[18.256]	(1.550)
NON-COGNITIVE SKILLS											
SDQ Total	2,240	0.136	501	0.049	-0.0871	175	0.115	-0.0207	326	0.004	-0.132
Difficulties		[1.088]		[0.953]	(0.0587)		[0.795]	(0.0688)		[1.056]	(0.0874)
AGE 5											
COGNITIVE SKILLS											
Picture Similarities	2,408	54.254	528	53.027	-1.226	188	51.177	-3.077**	340	54.273	0.0190
T-scores		[11.532]		[10.859]	(0.962)		[9.349]	(1.520)		[11.696]	(0.860)
NON-COGNITIVE SKILLS											
SDQ Total	2,262	0.146	496	0.103	-0.0424	174	0.441	0.295***	322	-0.130	-0.276***
Difficulties		[1.087]		[0.985]	(0.0935)		[0.909]	(0.0957)		[0.960]	(0.100)

Note: The treatment group includes pupils enrolled in 2012/13 at LA maintained secondary schools that become academies in 2013/14 and 2014/15, whereas the control group includes pupils enrolled at LA maintained secondary schools that do not convert to academy status. The treatment group is composed by pupils attending prospective sponsor-led and converter academies. All statistics, except the number of observations (N) that is unweighted, reflect MCS survey settings. The quantity in parenthesis below the mean (or proportion in the case of categorical variables) is the standard deviation. The difference (diff) refers to the difference in means between the treatment group and the control group and the quantity in parenthesis below is the (robust) standard error clustered at the school level and reflecting MCS survey settings. (University of London, Institute of Education, Centre for Longitudinal Studies, 2020a, 2020d, 2022). * significant at 10% level; ** significant at 5% level; *** significant at 1% level or below.

3.5 Methodology

The purpose of this chapter is to investigate the impact of academies on pupils' problem-solving skills, social behaviour, and mental health. I next outline the empirical strategy used to accomplish this, highlighting identification issues for a causal interpretation.

Pupils are not randomly assigned to schools, and schools do not randomly become academies. Academy conversion is either voluntary for high performing schools (converter academies) or forced onto underperforming schools (sponsor-led academies). Pupils are distributed across schools as a result of a selection process and parental choice. If high performing pupils are more likely to attend an academy, they will have higher outcomes, regardless of the school they attend. So, comparing school outcomes from pupils at academies to those of pupils at maintained schools would exaggerate the impact of academy attendance (Angrist & Pischke, 2009).

Since the MCS has little school-level information, it is not possible to directly account for selection of schools into academisation. Schools that become academies may be different from LA maintained schools that do not convert. Available data do not allow to further differentiate schools, preventing school-level matching or a quasi-experimental design of the control group. MCS has, nonetheless, very rich pupil-level longitudinal data that could help account for school selection. Effectively, if pupils are as good as randomly distributed across treatment and control groups, they are unlikely to select into different schools. The methods that follow attempt to account for pupil selection, indirectly also accounting for school selection.

Because of self-selection into and out of academies,[48] pupils are only included in the treatment group if they are enrolled in the secondary school prior to academy conversion (2012/13) and are affected by academy conversion in subsequent years of their secondary schooling (2013/14 to 2016/17)[49] – known as "legacy" enrolment or "grandfathered" pupils (Abdulkadiroğlu et al., 2016; Andrews et al., 2017; Eyles et al., 2016b, 2017; Eyles & Machin, 2019). Since I am using MCS6 (age 14, taken in 2015), I focus on pupils attending LA maintained schools that become academies in 2013/14 and 2014/15. This ensures academy conversion is exogenous to secondary school enrolment decision.

Most academies open between July and September (Bertoni et al., 2021). If an academy opens in 2013 (considered as an academy in operation from 2013/14 and included in the treatment group), its conversion application process would have started after cohort members have chosen schools by October 2011 (starting secondary phase in September 2012). In effect, DfE approval takes on average 2.3 months after application (standard deviation 3.7) and converter academies open on average 4.5 months after approval (standard deviation 3.4) (Bertoni et al., 2021). 85% of academies approved in 2015/16 open within 12 months of approval (National Audit Office, 2018). Schools that convert in 2016/17 take on average 10.7 months between applying and opening as a converter academy (National Audit Office, 2018). Moreover, schools rated as "inadequate", that are directed to convert, open as

[48] See sub-section 3.2.1 on pupil composition change after academy conversion.

[49] Pupils spend five years in secondary school. So, pupils enrolled at the predecessor school in year seven have up to four years of potential treatment exposure. Of course, the length of treatment exposure varies with the time of conversion.

sponsor-led academies on average 18 months after Ofsted inspection (National Audit Office, 2018). The median time for conversion under a sponsor is 17 months (Duchini et al., 2023).

Following a bad Ofsted inspection report there might be a period of troubled discussion and sometimes resistance to the prospect of academisation (Keddie, 2019). The decision to apply for academy conversion could also be predicated by debate and controversy. However, there is a lot of uncertainty in that period involving various players and many schools end up not converting, as shown in the case study presented in Rayner et al. (2018). So, even if some parents were aware of any plans to convert at the time they apply to secondary schools it would be hard to know if and when a school would become an academy. Therefore, given time taken to process applications, academisation is not likely to have been anticipated by parents at the time of secondary school application. The similar profile of pupil intakes across maintained schools that become academies and those that do not, described in section 3.4.3, corroborates this argument.

Although pupils tend to remain in the same school from school year seven to school year 11, some cohort members change secondary schools and are surveyed in another school. Hence, I estimate the intention to treat (ITT) effect of academy conversion on pre-enrolled pupils, irrespective of whether they actually stay in the school until the end of secondary phase (Andrews et al., 2017; Eyles et al., 2016b, 2017; Eyles & Machin, 2019).[50]

---

[50] This avoids issues related with self-selection or schools strategically sorting out pupils. Schools may, for instance, transfer pupils because of low academic achievement, behavioural problems, or SEN (OECD, 2016c).

Taking advantage of the rich MCS database, I use the explanatory variables presented in the previous section to attempt an identification strategy based on selection on observables (Dearden et al., 2009). Hence, the treatment impact is at first estimated using linear regression (OLS) and propensity score matching (PSM).

OLS valued-added specification, including a comprehensive set of pre-treatment variables plus a lagged outcome measure, should capture earlier factors and estimate the treatment effect under the conditional independence assumption (Angrist & Pischke, 2009). That is, conditional on observed characteristics treatment status (attending a school that later becomes an academy) is independent of potential outcomes. However, despite the rich longitudinal data I am able to control for, some relevant variables are possibly not observed. Todd and Wolpin (2003) show that value-added specifications rely on strong assumptions and are prone to endogeneity bias when relevant variables are not observed. Basically, for baseline performance (age 11, at end of primary phase) to capture all relevant past unobserved variables, the impact of those variables on the outcomes of interest would have to vary at a constant rate with time (Todd & Wolpin, 2003). Moreover, consistent estimates require all relevant contemporaneous variables to be included (Todd & Wolpin, 2003). According to Angrist and Pischke (2009), value-added OLS sets the lower bound for the causal effect of interest.

The OLS model is based on the following value-added relationship for pupil *i* in legacy enrolment school *s*:

$$Y(14)_{is} = \alpha_{01} + \alpha_1 A_s + \sum_{j=1}^{J} \alpha_{2j} X_{jis} + \alpha_3 Y(11)_{is} + \varepsilon_{1is} \qquad (3.1)$$

where Y_{is} are pupil outcomes measured at different points in time (age 11 and 14), X_{jis} is a set of *j* pre-treatment pupil and school characteristics, α_{01} is an intercept, and ε_{1is} is an error term. X_{jis} includes pupil characteristics (sex, age, ethnic group, language, SEN, health, SWM scores), family context (number of parents, grandparents and siblings, parents' age, mother's illness and mental health), socioeconomic status (income, FSM, housing tenure, parents' labour market status, full-time work, occupation and education), home environment (number of books, library and church attendance, hours watching TV, bedtime, help with homework), and geographical variables (IMD domains, urban) at age 11, prior performance (Word Reading, Maths, Pattern Construction, Picture Similarities, SDQ total difficulties and pro-social scores) at age five and seven, childcare (providers, grandparents, reading), and primary school characteristics (independent, single-sex or faith school) (see appendix A Table A4 for a detailed list of variables). A_s is a dummy variable equal to one for secondary schools that become academies in the sample period (treatment group). The coefficient of interest is α_1. The error term is assumed to be uncorrelated with treatment, but a degree of autocorrelation between pupils within schools is allowed, denoting clustered standard errors at school level (Abadie et al., 2017).[51]

PSM selects from LA maintained schools that do not convert those pupils that have observed characteristics as similar as possible to those

[51] Due to MCS geographical clustering, most cohort members are in the same school as at least one other cohort member. In the estimation sample, 382 pupils are the only cohort members at their school. Pupils are distributed across 827 schools, with an average of four pupils per school and a median of two pupils per school.

enrolled in LA maintained schools that become academies.[52] The probability of treatment (attending a school that later becomes an academy) is estimated using a probit regression with a set of pre-treatment explanatory variables that influence outcomes and school enrolment but are unaffected by that decision. If conditional on these variables academy status is independent of potential outcomes, controlling for the probability of treatment identifies the causal effect (Angrist & Pischke, 2009). PSM weights define the control group. The matching procedure should balance the distribution of variables in both groups.[53] Conditional on matching variables, it is assumed that the assignment to treatment is random and the comparison between the performance of the treatment group and of the control group is an unbiased estimate of the treatment effect.[54] Since PSM also relies on selection on observables, this model allows to check the robustness of OLS estimates but is also prone to the possible omission of relevant variables.

The PSM estimate is based on the following specification, weighted by the propensity score:

$$Y(14)_{is} = \alpha_{02} + \alpha_4 A_s + \varepsilon_{2is} \qquad (3.2)$$

where the coefficient of interest is α_4. The treatment effect is the comparison of means between the treatment and control group weighted by the propensity score which reflects the probability of being in the treatment

[52] Available data do not allow for school-level matching.

[53] Kernel matching includes all cases. I use an Epanechinikov kernel with a bandwidth of 0.06. A few different matching methods are tested (see appendix A Table A5).

[54] Following DuGoff et al. (2014), survey weights and strata are included among matching variables used in the calculation of PSM weights, but survey settings are not used in the matching process. PSM weights and survey weights are multiplied, and the resulting composite weight is applied in conjunction with survey design elements to estimate the treatment effect (DuGoff et al., 2014; Ridgeway et al., 2015).

group, according to a set of *j* pre-treatment variables (same as in OLS model and listed in appendix A Table A4), resulting from the next equation:

$$A_s = \alpha_{03} + \sum_{j=1}^{J} \alpha_{5j} X_{jis} + \alpha_6 Y(11)_{is} + \varepsilon_{3is} \qquad (3.3)$$

Previous methods based on selection on observables require the inclusion of all relevant variables for a causal interpretation (conditional independence assumption) and are subject to endogeneity. In effect, including lagged dependent variables controls for fixed unobserved variables only under stringent assumptions on the impact of variables over time (Todd & Wolpin, 2003). Acknowledging these limitations and conditions, difference in differences (DID) is a preferred alternative. DID controls for fixed unobservable variables that may be correlated with treatment assignment and the outcomes of interest (Angrist & Pischke, 2009). In the absence of academy conversion, pupil performance is assumed to be determined by a time-invariant school effect and a time effect that is common across schools. This implies a similar trend in both groups in the absence of treatment (common trends assumption). Under parallel trends, the change in average outcomes in the control group provides a reliable counterfactual. For obvious reasons that is not a testable assumption. I show however in section 3.4.3 that pupils attending schools that become academies and pupils in maintained schools were very similar prior to secondary phase across many variables and over time. It would be interesting to make this same comparison using pre-treatment school-level variables, but available data do not allow for this. Having similar baseline characteristics is not enough to say these groups follow parallel trends in outcomes but is reassuring that observed variables between age five and 11 are generally not significantly different between the treatment and control groups.

Having one single cohort at MCS, the first difference refers to the difference in individual performance measured at different ages, and the second difference to the difference between the performance of pupils in the treatment group and in the control group. The treatment effect is estimated using DID and matched DID. The latter estimate is presented to show the robustness of findings.

The DID estimate, my preferred specification, uses staked data (two records per cohort member, one for each year) and is based on the following relationship for pupil *i* in legacy enrolment school *s* and age *t*:

$$Y_{ist} = \omega_0 + \beta_1 A_s + \omega_1 T_t + \omega_2 (A_s * T_t) + \sum_{j=1}^{J} \delta_{1j} X_{jist} + \varepsilon_{4ist} \qquad (3.4)$$

that includes main effects for the treatment group A_s (pupils attending schools that become academies) and age T_t (dummy variable equal to one for age 14) plus an interaction term. The coefficient of interest is ω_2. The model also controls for a set of *j* pupil and school characteristics X_{jist} (similar to those used in OLS model but stacked in two periods (age 11 and 14), listed in appendix A Table A4) and includes an intercept ω_0. In the absence of time-varying effects, DID estimates remove effects from time-invariant variables, including fixed variables between age 11 and 14, such as ethnicity, and variables observed at one point in time, namely those observed at age five and seven. For the common trends assumption to hold the effect of included variables on the outcome must be the same across time or those variables must have the same average across groups (Zeldow & Hatfield, 2021). To ensure adding these covariates do not confound treatment effect estimates, an unconditional DID model is presented for comparison. When estimating the matched DID, the equation is weighted by the propensity score estimated in (3.3).

ITT estimates calculate the effect of starting secondary phase in a maintained school that later becomes an academy, ignoring non-compliance with treatment assignment and providing a lower bound on the (assumed positive) treatment effect. In order to estimate the actual impact of attending an academy, I use legacy enrolment as an instrument for academy attendance to correct for the fact that not all legacy enrolled pupils remain in the school and receive treatment. This assumes the exclusion restriction that here means legacy enrolment is as good as randomly assigned and it has no effect on outcomes except through academisation (Angrist & Pischke, 2009). The observed similarity between treatment and control groups, discussed earlier, lends plausibility to this assumption.

The first stage equations, underlying the reduced form above (3.4), are as follows:

$$I_s = \omega_{01} + \beta_2 A_s + \omega_3 T_t + \omega_4 (A_s * T_t) + \sum_{j=1}^{J} \delta_{2j} X_{jist} + \varepsilon_{5ist} \quad (3.5)$$

$$(I_s * T_t) = \omega_{02} + \beta_3 A_s + \omega_5 T_t + \omega_6 (A_s * T_t) + \sum_{j=1}^{J} \delta_{3j} X_{jist} + \varepsilon_{6ist} \quad (3.6)$$

where I_s is a dummy variable that indicates if a pupil is enrolled in an academy at age 14. This reveals if pupils enrolled in the predecessor school at the start of secondary phase receive treatment by academy conversion. β_2 is the proportion of pupils in the ITT group that stay in the school and are surveyed while there,[55] indicating the strength of the instrument. First stage equations also control for the set of variables used in the reduced form (3.4).

Second stage estimates, obtained replacing A_s by the predicted I_s in (3.4), provide the local average treatment effect (LATE) on pupils that would not have attended an academy had they not been pre-enrolled. It measures the impact on those pupils where treatment status changes with conversion, excluding those pupils that would always move to or from an academy (Angrist & Pischke, 2009). Given the high compliance with treatment assignment,[56] contamination adjusted estimates are expected to be similar to ITT estimates. Therefore, in the remainder of the text both will referred to as the effect of attending an academy.

I use alternative regression methods and several specifications as robustness checks to the preferred model and its findings. Separate subgroup analyses for sponsor-led and converter academies identify heterogeneous effects.

3.6 Empirical Results

Table 3.5 shows the estimated impact of attending an academy on pupils' cognitive and non-cognitive skills. The table compares the results obtained when using linear regression (OLS), propensity score matching (PSM), difference in differences (DID), matched DID, and DID using

[55] Pupils in the treatment group that move to another academy and pupils in the control group that move to another LA maintained school are also included among compliers.

[56] Focusing on the school type attended in 2015, there are 85 pupils that do not comply with treatment assignment. At age 14, 70 pupils in the control group attend an academy and 15 pupils in the treatment group do not attend an academy.

instrumental variables (IV)[57] (see appendix A Table A4 for a full list of explanatory variables used in each method). The first two methods follow an identification strategy based on observables, using the rich MCS database.[58] DID controls for any fixed individual unobservable and common time effects not related with treatment. In addition to overall estimates, a separate analysis for sponsor-led and converter academies is provided in Table 3.6 and Table 3.7.

I consider the impact of academy attendance on cognitive skills and non-cognitive skills separately. To conclude, I summarize results from robustness checks undertaken to test the sensitivity of the main results to different model specifications.

---

[57] DID using IV corrects ITT estimates for non-compliance with treatment assignment.

[58] See appendix A Figure A1 for a comparison of propensity scores in treatment and control groups before and after matching and Table A5 for alternative matching methods.

Table 3.5 Impact of academy attendance on outcome variables

		(1) OLS		(2) PSM		(3) DID		(4) MATCHED DID		(5) DID/IV	
VARIABLES	**N**	**Treatment**	**N**	**Treatment**	**N**	**Treatment**	**N**	**Treatment**	**N**	**Treatment**	**First Stage**
COGNITIVE SKILLS											
Word Activity	2,842	-0.0290	2,820	-0.0455	2,842	0.0574	2,820	-0.0303	2,842	0.0610	0.934***
		(0.0480)		(0.0730)		(0.0898)		(0.0855)		(0.0953)	(0.0119)
CGT Quality of Decision Making	2,750	0.0531	2,733	0.0476	2,750	0.0703	2,733	0.0994	2,750	0.0744	0.939***
		(0.0495)		(0.0813)		(0.0661)		(0.0847)		(0.0699)	(0.0113)
CGT Risk Adjustment	2,749	0.0352	2,733	0.0429	2,749	0.0389	2,733	0.0731	2,749	0.0412	0.939***
		(0.0458)		(0.0723)		(0.0733)		(0.0720)		(0.0776)	(0.0113)
NON-COGNITIVE SKILLS											
CGT Risk Taking	2,749	-0.0733	2,734	-0.0903	2,749	-0.0859	2,734	-0.0681	2,749	-0.0909	0.939***
		(0.0539)		(0.0803)		(0.0609)		(0.0774)		(0.0646)	(0.0113)
CGT Deliberation Time	2,750	-0.125**	2,735	-0.181**	2,750	-0.0797	2,735	-0.163**	2,750	-0.0845	0.939***
		(0.0535)		(0.0910)		(0.0614)		(0.0786)		(0.0647)	(0.0113)
SDQ Total Difficulties	2,961	-0.0193	2,946	-0.0389	2,961	-0.0279	2,946	-0.0711	2,961	-0.0295	0.941***
		(0.0393)		(0.0740)		(0.0482)		(0.0512)		(0.0509)	(0.0104)
SDQ Pro-social	2,979	0.0584	2,964	0.0289	2,979	0.0511	2,964	0.0720	2,979	0.0539	0.941***
		(0.0496)		(0.0669)		(0.0636)		(0.0647)		(0.0672)	(0.0103)
Rosenberg Self-esteem Scale	2,696	0.0168	2,676	0.0249	2,696	-0.00832	2,676	-0.00399	2,696	-0.00865	0.943***
		(0.0572)		(0.0936)		(0.0645)		(0.0785)		(0.0676)	(0.0104)

Note: The treatment group includes pupils enrolled in 2012/13 at LA maintained secondary schools that become academies in 2013/14 and 2014/15, whereas the control group includes pupils enrolled at LA maintained secondary schools that do not convert to academy status. The treatment group is composed by pupils attending prospective sponsor-led and converter academies. All outcome variables are standardised to mean zero and standard deviation one. The quantity in parenthesis below the treatment coefficient is the (robust) standard error clustered at the school level and reflecting MCS survey settings. Propensity score matching (PSM) uses an Epanechinikov kernel with a bandwidth of 0.06 (Leuven & Sianesi, 2003). MCS survey weights and strata are included in matching variables. PSM weights and survey weights are multiplied and form the weight applied to matched regressions with other survey design elements (DuGoff et al., 2014). For PSM and matched DID, standard errors are bootstrapped using 1,000 and 10,000 replications respectively (Kolenikov, 2010). Bootstrap samples respect school clustering. (University of London, Institute of Education, Centre for Longitudinal Studies, 2020a, 2020c, 2020d, 2022). * significant at 10% level; ** significant at 5% level; *** significant at 1% level or below.

Table 3.6 Impact of academy attendance: Sponsor-led academies

	(1) OLS		(2) PSM		(3) DID		(4) MATCHED DID		(5) DID/IV		
VARIABLES	**N**	**Treatment**	**N**	**Treatment**	**N**	**Treatment**	**N**	**Treatment**	**N**	**Treatment**	**First Stage**
COGNITIVE SKILLS											
Word Activity	2,519	-0.0991 (0.0646)	2,495	-0.0356 (0.152)	2,519	-0.0478 (0.142)	2,495	-0.0326 (0.138)	2,519	-0.0502 (0.149)	0.943*** (0.0152)
CGT Quality of Decision Making	2,428	-0.0703 (0.0785)	2,406	-0.0502 (0.172)	2,428	-0.0987 (0.120)	2,406	-0.0452 (0.145)	2,428	-0.104 (0.127)	0.943*** (0.0155)
CGT Risk Adjustment	2,427	-0.0915 (0.0684)	2,406	-0.0285 (0.139)	2,427	-0.174* (0.0965)	2,406	-0.0275 (0.138)	2,427	-0.183* (0.101)	0.943*** (0.0155)
NON-COGNITIVE SKILLS											
CGT Risk Taking	2,427	-0.0293 (0.0831)	2,406	-0.0497 (0.224)	2,427	-0.0891 (0.0981)	2,406	-0.0196 (0.162)	2,427	-0.0938 (0.103)	0.943*** (0.0155)
CGT Deliberation Time	2,428	-0.109 (0.0864)	2,408	-0.125 (0.185)	2,428	-0.0459 (0.0923)	2,408	-0.174 (0.136)	2,428	-0.0483 (0.0968)	0.943*** (0.0155)
SDQ Total Difficulties	2,617	0.0122 (0.0672)	2,590	-0.00114 (0.151)	2,617	0.0125 (0.0792)	2,590	-0.117 (0.102)	2,617	0.0128 (0.0830)	0.945*** (0.0147)
SDQ Pro-social	2,634	0.0981 (0.0829)	2,606	0.0678 (0.179)	2,634	0.0513 (0.129)	2,606	0.171 (0.139)	2,634	0.0537 (0.135)	0.944*** (0.0144)
Rosenberg Self-esteem Scale	2,377	0.129 (0.0874)	2,352	0.109 (0.186)	2,377	0.196** (0.0797)	2,352	0.00477 (0.131)	2,377	0.204** (0.0816)	0.953*** (0.0115)

Note: The treatment group includes pupils enrolled in 2012/13 at LA maintained secondary schools that become academies in 2013/14 and 2014/15, whereas the control group includes pupils enrolled at LA maintained secondary schools that do not convert to academy status. The treatment group is composed by pupils attending prospective sponsor-led academies. All outcome variables are standardised to mean zero and standard deviation one. The quantity in parenthesis below the treatment coefficient is the (robust) standard error clustered at the school level and reflecting MCS survey settings. Propensity score matching (PSM) uses an Epanechinikov kernel with a bandwidth of 0.06 (Leuven & Sianesi, 2003). MCS survey weights and strata are included in matching variables. PSM weights and survey weights are multiplied and form the weight applied to matched regressions with other survey design elements (DuGoff et al., 2014). For PSM and matched DID, standard errors are bootstrapped using 1,000 replications and 10,000 replications respectively (Kolenikov, 2010). Bootstrap samples respect school clustering. (University of London, Institute of Education, Centre for Longitudinal Studies, 2020a, 2020c, 2020d, 2022). * significant at 10% level; ** significant at 5% level; *** significant at 1% level or below.

Table 3.7 Impact of academy attendance: Converter academies

		(1) OLS		(2) PSM		(3) DID		(4) MATCHED DID		(5) DID/IV	
VARIABLES	**N**	**Treatment**	**N**	**Treatment**	**N**	**Treatment**	**N**	**Treatment**	**N**	**Treatment**	**First Stage**
COGNITIVE SKILLS											
Word Activity	2,652	0.0251 (0.0649)	2,634	-0.00208 (0.0970)	2,652	0.133 (0.0900)	2,634	0.0231 (0.0980)	2,652	0.143 (0.0959)	0.928*** (0.0149)
CGT Quality of Decision Making	2,566	0.136*** (0.0496)	2,551	0.108 (0.0883)	2,566	0.181*** (0.0560)	2,551	0.193** (0.0895)	2,566	0.192*** (0.0589)	0.935*** (0.0135)
CGT Risk Adjustment	2,565	0.130** (0.0603)	2,550	0.112 (0.106)	2,565	0.183** (0.0807)	2,550	0.162 (0.101)	2,565	0.194** (0.0858)	0.935*** (0.0135)
NON-COGNITIVE SKILLS											
CGT Risk Taking	2,565	-0.0977 (0.0676)	2,551	-0.0899 (0.0961)	2,565	-0.0893 (0.0685)	2,551	-0.0618 (0.0884)	2,565	-0.0949 (0.0731)	0.935*** (0.0135)
CGT Deliberation Time	2,566	-0.137** (0.0624)	2,552	-0.137 (0.102)	2,566	-0.107 (0.0828)	2,552	-0.165* (0.0924)	2,566	-0.114 (0.0875)	0.935*** (0.0135)
SDQ Total Difficulties	2,772	-0.0344 (0.0514)	2,754	-0.0749 (0.0953)	2,772	-0.0492 (0.0601)	2,754	-0.0666 (0.0616)	2,772	-0.0522 (0.0639)	0.938*** (0.0123)
SDQ Pro-social	2,787	0.0257 (0.0499)	2,768	0.0312 (0.0796)	2,787	0.0553 (0.0528)	2,768	0.0642 (0.0742)	2,787	0.0587 (0.0561)	0.939*** (0.0123)
Rosenberg Self-esteem Scale	2,537	-0.0324 (0.0673)	2,515	-0.0177 (0.125)	2,537	-0.130 (0.0803)	2,515	-0.0525 (0.105)	2,537	-0.138 (0.0844)	0.936*** (0.0134)

Note: The treatment group includes pupils enrolled in 2012/13 at LA maintained secondary schools that become academies in 2013/14 and 2014/15, whereas the control group includes pupils enrolled at LA maintained secondary schools that do not convert to academy status. The treatment group is composed by pupils attending prospective converter academies. All outcome variables are standardised to mean zero and standard deviation one. The quantity in parenthesis below the treatment coefficient is the (robust) standard error clustered at the school level and reflecting MCS survey settings. Propensity score matching (PSM) uses an Epanechinikov kernel with a bandwidth of 0.06 (Leuven & Sianesi, 2003). MCS survey weights and strata are included in matching variables. PSM weights and survey weights are multiplied and form the weight applied to matched regressions with other survey design elements (DuGoff et al., 2014). For PSM and matched DID, standard errors are bootstrapped using 1,000 replications and 10,000 replications respectively (Kolenikov, 2010). Bootstrap samples respect school clustering. (University of London, Institute of Education, Centre for Longitudinal Studies, 2020a, 2020c, 2020d, 2022). * significant at 10% level; ** significant at 5% level; *** significant at 1% level or below.

3.6.1 Cognitive Skills

Table 3.5 shows there are no significant differences in cognitive performance between the treatment group and the control group, regardless of the method used. However, when we look to sponsor-led academies and converter academies separately, the cognitive differences become clear.

The cognitive underperformance of pupils at sponsor-led academies, anticipated in the descriptive analysis, is not confirmed after controlling for pupil characteristics, background, and prior performance, as seen in Table 3.6 columns (1) and (2). However, Todd and Wolpin (2003) show that in the presence of omitted variables, a value-added OLS specification does not provide consistent estimates. In fact, when controlling for fixed unobservable variables, the DID estimate shows that pupils at sponsor-led academies adjust -0.17 of a standard deviation (SD) less their behaviour to available information than peers at LA maintained schools. Yet, the estimate from matched DID on risk adjustment is also much lower and not significant, suggesting sensitivity of the DID estimate to functional form.[59]

Table 3.7 shows that pupils at converter academies perform significantly better at decision making (DID estimate 0.18 SD) than their peers at LA maintained schools. Pupils at converter academies are more likely to choose the most probable outcome. This result, projected in descriptive statistics and linear regression,[60] is confirmed by the matched DID estimate.

[59] I must also acknowledge the low statistical power of the estimation sample, which makes it difficult to identify small effects.

[60] The estimates from PSM are generally not very different from OLS, but their bootstrapped standard errors are systematically higher, making it difficult to reach conventional levels of significance.

DID also reveals that pupils at converter academies adjust significantly more their behaviour to risk (0.18 SD), suggesting these pupils are rational and flexible decision makers that make higher bets when odds are favourable. Again, confirming indication from raw comparisons and OLS. The estimate from matched DID on risk adjustment is slightly lower and marginally not significant,[61] but roughly corroborates the DID estimate. These results are consistent with school performance at high stakes exams, where converter academies generally outperform LA maintained schools and sponsor-led academies (Bolton, 2012, 2015; Department for Education, 2018).

The IV estimates in column (5) broadly align with DID estimates in column (3), because of the high rate of compliance for legacy enrolled pupils (most pupils stay in same school). The IV estimates change only slightly compared with DID estimates, decreasing to -0.18 SD in risk adjustment for sponsor-led academies, and increasing to 0.19 SD in decision making and risk adjustment for converter academies. These are the preferred average causal estimates of academy conversion.

Estimated effects are large in relation to most education interventions (Kraft, 2020).[62] To contextualise, it is helpful to compare with results found in the literature. Eyles et al. (2016a), Eyles et al. (2016b), Andrews et al. (2017), and Eyles and Machin (2019) find gains between 0.08 SD and 0.15 SD in GCSE exam scores for pre-2010 sponsor-led academies. Regarding post-

[61] Coefficient level of significance 11%.

[62] Kraft (2020) proposes the following benchmarks for causal studies of education interventions on standardised pupil achievement: less than 0.05 is small, between 0.05 and 0.20 is medium, more than 0.20 is large. He also suggests that for large studies using broad achievement measures an effect of 0.15 should be considered large.

2010 conversions, Andrews et al. (2017) report gains of 0.06 SD and 0.03 SD in GCSE exam scores for sponsor-led academies and outstanding converter academies respectively. Although I use different outcomes, these studies report lower estimates and show that effect sizes are indeed large. In order to make better sense of results, as an illustration based in Harrison et al. (2022),[63] the odds ratio of a pupil at a converter academy showing symptoms of eating pathology is 0.01 because of the improvement in decision making.

3.6.2 Non-cognitive Skills

In line with descriptive statistics, linear regression and matching show that pupils attending academies make decisions faster. However, when controlling for fixed unobservable variables this difference is no longer significant, except when applying PSM weights to DID. Table 3.5 shows that there is no other significant difference in non-cognitive skills between the treatment and the control groups. As before, subpopulation analysis reveals important variation.

The OLS estimate suggests that pupils at converter academies make decisions faster (see Table 3.7), driving the overall academy effect mentioned earlier. Nevertheless, this is again not verified by DID. It is thus possible that some unobservable explains that result. If a relevant input is not fully captured in included variables, OLS value-added estimates will be biased (Todd & Wolpin, 2003), possibly explaining larger and significant OLS estimates on deliberation time. No other significant difference in non-cognitive skills is found

---

[63] Harrison et al. (2022), using the MCS, estimate a odds ratio of 0.46 for prodromal eating pathology at age 14 and 17 of CGT quality of decision making scores at age 14.

between pupils attending converter academies and their peers at LA maintained schools.[64]

The DID estimate in Table 3.6 reveals that self-esteem is 0.20 SD higher for pupils attending sponsor-led academies.[65] The IV estimate is similar, due to the high rate of compliance. This result confirms that pupils' self-worth improved at sponsor-led academies, as hinted by descriptive statistics. The OLS estimate, however, is smaller and not significant. Todd and Wolpin (2003) explain that in a value-added specification, optimising behaviour creates a correlation between any omitted variable and the baseline outcome. As seen in descriptive statistics, pupils at sponsor-led academies report lower self-esteem at age 11. This is typical of disadvantaged intakes (Bannink et al., 2016). If those parents respond to low self-esteem by decreasing family inputs (less parental time or worse family environment), omitted variables will be correlated with the outcome at baseline and affect OLS estimates. Regarding the other non-cognitive skills, pupils at sponsor-led academies are not dissimilar from their peers attending maintained schools.

As mentioned previously, estimated effects are large in comparation with most education interventions (Kraft, 2020). It is worth, though, to look specifically to non-cognitive outcomes. A systematic review of school-based interventions for the improvement of non-cognitive skills reports effect sizes between 0.01 SD and 0.62 SD (Siddiqui & Ventista, 2018). However, many of these studies have a small sample size or use outcomes directly related to the intervention and are likely to have bigger effect sizes (Kraft, 2020). From the

[64] Given the sample size, only large effects are able to reach statistical significance.
[65] This is not confirmed by matched DID, suggesting sensitivity to functional form.

literature on academies, Frostick et al. (2018) identify an effect of 0.02 SD on pro-social scores for pupils attending pre-2010 sponsor-led academies. This shows that the estimated effect on self-esteem is considerable. As an illustration of its consequence, using the results of Flouri (2006),[66] pupils attending sponsor-led academies are expected to raise their NVQ level by 0.02 due to their self-esteem.

3.6.3 Sensitivity Analysis

Robustness checks are motivated by the concern that results might be sensitive to the variables used to measure parental education, income, and pupils' mental health and behaviour, or to the inclusion of school level variables. An unconditional DID model (without covariates) is also used for comparison, checking if adding variables affects identification. The results from alternative models are presented in Table A6, Table A7 and Table A8 in appendix A.

The sensitivity analysis reveals a pattern of results very similar to the main specification. The different models slightly shift the estimate of interest by a small amount, as expected, but do not alter the effects meaningfully. A revealing case is the inclusion of a dummy identifying grammar schools. The grammar school variable is not included in the main specification because it is not available pre-treatment. In effect, there are no grammar schools at primary phase. Additionally, there are no grammar schools among sponsor-led academies. But, as shown in appendix A, the results of including this variable

---

[66] Based on the 1970 British Cohort Study, Flouri (2006) estimates that one SD in self-esteem at age 10 raises NVQ by age 26 in 0.10 and 0.09 for males and females respectively.

are not qualitatively different from those of the main specification. Similarly, the estimates from the unconditional DID model largely confirm main results.

In brief, I have not found statistically significant differences in outcomes between pupils attending academies and those at LA maintained schools. However, when considering sponsor-led and converter academies separately, a significant effect of academy attendance emerges. Pupils at sponsor-led academies report significantly higher self-esteem. Whereas attending a converter academy significantly improves pupils' decision-making skills.

3.7 Discussion

The introduction and expansion of academies in England facilitate the study of the implications of school autonomy and education quasi-markets for pupils' outcomes. Since academies' impact evaluation literature dwells essentially on performance at high stakes exams, this chapter focuses on a wider range of outcomes including problem-solving skills, mental health, and social behaviour.

Using the rich MCS database and a difference in differences framework combined with legacy enrolment, I find that the cognitive and non-cognitive skills of comparable pupils attending academies and LA maintained schools are identical. Differences in certain outcomes arise when we distinguish sponsor-led academies from converter academies. Academy conversion raises satisfaction for pupils at sponsor-led academies and raises decision making skills for those at converter academies. So, conversion affects pupil outcomes differently depending on the academy route taken.

Despite mixed results in the literature regarding GCSE performance (Andrews et al., 2017), pupils attending converter academies significantly improve their problem-solving and decision making skills in comparison with peers at LA maintained schools. Pupils at converter academies make optimal choices more often. These typically high performing schools foster wide cognitive skills that go beyond exam preparation. However, I find no evidence of improvement on non-cognitive skills at converter academies. This suggests that the incentives in place do not encourage academies to target non-cognitive outcomes. However, at secondary phase, when non-cognitive skills are more malleable than cognitive skills, focusing on developing non-cognitive skills is more effective (Kautz et al., 2014). Considering the longstanding relevance of these skills, it would be important to review those incentives.

Pupils enrolled in sponsor-led academies have greater self-esteem than their peers at LA maintained schools. Since the literature shows that sponsor-led academies somewhat improve pupil performance at GCSEs (Andrews et al., 2017), it is interesting to note that this is not at the cost of pupils' satisfaction. Perhaps high performance at GCSEs boosts self-esteem. Alternatively, as suggested in Frostick et al. (2018), this positive effect might originate from efforts of sponsor-led academies to develop new identities, disciplinary structures, teacher relationships, and ethos. Pupils' wellbeing is indeed associated with the disciplinary climate, sense of belonging to school (OECD, 2017b), bullying, and school connectedness (Patalay & Fitzsimons, 2016). Yet, these self-reported measures may instead reflect reference bias and school context (Kautz et al., 2014).[67] On the other hand, the academic gains described in the literature for sponsor-led academies (Andrews et al., 2017) are not reflected in the cognitive effects reported here. The only significant cognitive effect I find for sponsor-led academies (risk adjustment) is negative. This may imply that gains at national exams are achieved through a teaching focused on test preparation or other ranking maximising strategies.[68] Careful attention should be given to this matter to ensure that schools provide a well-rounded education.

An important caveat in this work is the lack of detailed school level data, preventing a comparison of schools in the treatment and the control groups. Comparing pupils at academies with their peers at LA maintained schools is not ideal, particularly when analysing sponsor-led and converter academies that followed very different routes for conversion. A methodological refinement using additional data could address this issue by redefining the control group as pupils attending LA maintained schools that become academies in the following period, as implemented in chapter 5. This is an important future development to consolidate the causal interpretation of results.

[67] Pupils in more demanding schools (higher work ethic, teacher strictness, clarity of rules) tend to rate their non-cognitive skills more critically (M. R. West et al., 2016).

[68] Exam preparation and frequent tests can also help minimise pupils' anxiety that affects wellbeing (OECD, 2017b).

4. ACADEMIES' IMPACT ON PUPILS: SUBJECT CHOICE AT AGE 14

4.1 INTRODUCTION

This chapter explores how academies shape pupils' subject choice at age 14. Subject choice is a pivotal element of secondary phase in England. However, given the key role of league tables, schools might offer different pathways to pupils from different socioeconomic backgrounds, entering them for qualifications likely to improve school ranking scores. School competition may also lead to diversification and specialization, allowing schools to develop expertise in a particular field. A specialism creates a niche in the local education market in which the school may have no direct competitors. If parents are seeking a school that offers that specific curriculum, they may have no other local option. Besides, if a school attracts most local pupils who are adept or motivated towards a subject, other local schools will find it difficult to work on that subject. This issue is particularly acute in the case of academies because of school autonomy and increased stratification (Braz, 2018; Eyles et al., 2018).

If pupils from different backgrounds are sorted into differentiated schools, the choice set of subjects available to them, the expectations, and the guidance received will differ (Dilnot, 2016). School leadership, teachers, and ethos influence pupils' choices and aspirations. Besides, pupils may find themselves in a school where their preferred subjects are not available. Perhaps school counselling and the curriculum offer at academies limit subject choices, requiring pupils to study certain subjects and preventing them from taking others. Some pupils might be pointed to university with time-tested strategies of success whereas others lack this guidance, subsequently directing them to divergent professional careers. Vocational courses may be appropriate for some pupils, preventing unemployment and unskilled work, but divert others from high status occupations. Nevertheless, some pupils might have no real choice or are led not to consider alternatives.

This chapter will thus address the following question: Are pupils at academies more likely to choose academically oriented subjects at age 14 (Key Stage 4, KS4) than their peers at LA maintained schools?

The outcomes of interest are facilitating subjects, the English Baccalaureate (EBacc), science technology engineering and maths (STEM) subjects, and vocational subjects. Previous studies have shown that taking these subjects at age 14 influences post-16 outcomes (Anders et al., 2018a; Henderson et al., 2018). Facilitating subjects and the EBacc, for instance, are supposed to facilitate university entry. Pupils are examined in chosen subjects at age 16, and school results are published in league tables.

Using the Millennium Cohort Study, I attempt to approximate causality through a comprehensive regression model on subject choice at age 14, identifying differences associated with academy attendance. To avoid selection issues, I focus on pupils that at start of secondary phase enrol in LA maintained schools that become academies in the following years (treatment group). The control group includes pupils attending LA maintained schools. A subgroup analysis for sponsor-led and converter academies reveals heterogeneity. Results suggest that pupils at converter academies are more likely to take science subjects and facilitating subjects.

The structure of this chapter is as follows. Section 4.2 reviews previous studies on subject choice with a focus on the influence of school types, and highlights existing gaps in the literature. Sections 4.3 and 4.4 present data and the econometric model used to answer the research question. Section 4.5 discusses empirical findings. Section 4.6 offers some concluding comments.

4.2 Subject Choice at Age 14

In this section, I explore possible factors influencing subject choice at English secondary schools. I start with a brief description of policy affecting curriculum offer and present some evidence on the importance of chosen subjects for pupils' later outcomes. I then discuss factors that have been identified as influencing subject choice. I am especially interested in how schools condition pupils' choices.

4.2.1 Policy Context

In England, pupils generally choose the subjects they study in KS4 at age 14 and are examined in the General Certificate of Secondary Education (GCSE) and equivalents at age 16 (Jin et al., 2011). That choice affects their future academic and professional path (Anders et al., 2018a; Henderson et al., 2018; Jin et al., 2011; Sullivan, Zimdars, et al., 2010). In effect, some schools require a threshold GCSE score or having taken a particular GCSE subject before allowing an A level subject to be chosen at Key Stage 5 (KS5) (Alldritt & Taylor, 2015; Dilnot, 2016).

In 2002, the Labour government introduced numerous new qualifications, including applied GCSEs and vocational courses, to promote curricular diversity (Henderson et al., 2018; Moulton et al., 2018) and address pupils' diverse interests and abilities (Pring, 2018). GCSEs have remained popular, but the number of pupils taking vocational qualifications has increased (Jin et al., 2011; Wolf, 2011).

In 2010, the Coalition government reverted to a more traditional curriculum and introduced the EBacc (Anders et al., 2018a; Henderson et al., 2018; Long, 2019a; Moulton et al., 2018). The EBacc measures the percentage of pupils in a school that achieve a good grade in English, mathematics, history or geography, two sciences, and a modern or ancient foreign language (Anders et al., 2018a; Jin et al., 2011; Long, 2019a). In line with government policy, 90% of pupils are expected to enter the EBacc by 2025 (Long, 2019a). In the same vein, the Wolf (2011) review calls for a common core curriculum, centred around English and mathematics, and more restrictive GCSE equivalences. In 2014, the number of counted GCSEs in league tables were capped (Henderson et al., 2018). Reformed GCSEs[69] in English language, English literature, and mathematics were introduced in 2015, with other reformed subjects phased in for teaching between 2016 and 2018 (Department for Education, 2018; Long, 2017).[70]

[69] Reformed GCSEs are linear (assessment at the end of the course), have reformed content and a new grading scale (Long, 2017). Tiering and resits are limited (Long, 2017).

[70] In 2016, the Conservative government also introduced the school value-added measure "Progress 8", based on pupil performance across eight subjects (Attainment 8), including English, mathematics and three further EBacc subjects (Department for Education, 2016b, 2018).

4.2.2 Impact on Later Outcomes

Pupils who have studied vocational subjects are less likely to stay in full time education post-16 (Jin et al., 2011; Vidal Rodeiro & Vitello, 2021). Moreover, Vidal Rodeiro and Vitello, using the National Pupil Database (NPD), show that pupils who only have taken academic GCSEs are more likely to take A levels[71] whereas those that have taken vocational subjects are more likely to progress to vocational courses at age 16. In addition, Moulton et al. (2018), using Next Steps linked to the NPD, find that pupils pursuing an EBacc curriculum have a greater likelihood of continuing in full time education, taking A levels, and taking facilitating A level subjects, while the reverse is true for those taking applied GCSE subjects.

Using propensity score matching with Next Steps, Anders et al. (2018a) find that pupils taking the EBacc are slightly more likely to attend university, while those taking applied GCSE subjects are less likely to go to one. According to De Philippis (2021), taking triple science at GCSE increases the likelihood of attending a Russell group university and of graduating with a STEM degree. Dilnot (2018), using a random effect model based on the NPD linked to Higher Education Statistics Agency data and The Times/ Sunday Times Good University Guide, also shows that holding facilitating A levels, especially mathematics, is associated with attending a high ranking university. This is confirmed by Henderson et al. (2020).

---

[71] Dilnot (2016) shows that the probability of taking facilitating A levels increases with having facilitating GCSEs. Similarly, taking GCSE sciences is highly predictive of studying sciences at age 18 (De Philippis, 2021; Sullivan, Joshi, et al., 2010).

The impact of subject choice at age 14 is not restricted to further education options. Iannelli (2013), using the National Child Development Study (NCDS), shows that studying English, mathematics, science, and languages at age 16 increases the likelihood of being in high status occupations. Similarly, Vidal Rodeiro and Williamson (2019) use the NPD linked to Destination Measures to show that pupils following academic pathways have higher progression to education, employment, and training. However, for underperforming pupils the likelihood of progression is greater if they take vocational courses at GCSE (Vidal Rodeiro & Williamson, 2019).

The abovementioned evidence emphasises the importance of subject choice at GCSE for pupils' academic and professional careers and serves as a key motivation for the research question addressed in this chapter.

4.2.3 Factors Influencing Subject Choice

Rational choice models[72] explain pupil choices at secondary school (Breen et al., 2014; Breen & Goldthorpe, 1997; Gabay-Egozi et al., 2010; Jæger, 2007). Pupils' choices are influenced by expected returns (Belfield et al., 2020; Bleemer & Zafar, 2018; Delavande & Fumagalli, 2019; Flannery & O'Donoghue, 2013; Wilson et al., 2005). More prestigious academic routes have higher expected returns and higher risk of failure (Breen et al., 2014; Gabay-Egozi et al., 2010; Jæger, 2007). Breen and Goldthorpe (1997) predict that advantaged pupils have a preference for academic subjects because of

[72] Forward-looking rational pupils evaluate the options available to them and maximise utility based on the expected costs and benefits, net of uncertainties, in line with risk aversion and time discounting preferences and subject to socioeconomic constraints (Breen et al., 2014; Breen & Goldthorpe, 1997; Gabay-Egozi et al., 2010; Jæger, 2007; Wilson et al., 2005).

differences in ability, available resources, and relative risk aversion. Academic options allow them to minimise the risk of downward social mobility and preserve their social position. Additionally, given vocational courses' secure labour market prospects, Breen et al. (2014) expect risk aversion to prevent disadvantaged pupils from choosing more challenging academic subjects.[73]

Nevertheless, decision making is not perfectly rational. Individuals use rules of thumb and do not have perfect information (Barrance & Elwood, 2018; Jin et al., 2011). Incomplete information leads families to underestimate the benefits and overestimate the costs of academic routes (Bleemer & Zafar, 2018; Delavande & Fumagalli, 2019). Pupils overweight immediate factors, overestimate the likelihood of performing well and not fully appreciate the way their preferences change over time (Jin et al., 2011). Default options, recent advice, and vivid outcomes also influence their decisions (Bleemer & Zafar, 2018; Jin et al., 2011). On the other hand, school practices such as using option blocks and tiering restrict available routes (Barrance & Elwood, 2018; Smyth, 2018).

According to Jæger (2007), secondary pupils respond both to economic and peer incentives. He shows that pupils opt for academic subjects because of expected returns and when they are commonly chosen in their peer group. In fact, Hedges and Speckesser (2017) find that having high ability peers reduces the likelihood of taking a vocational course. Moreover, Belfield et al. (2020) reveal that expected non-monetary benefits and costs of education are significant predictors of pupils' choices.

[73] Yet, some disadvantaged pupils combine more demanding secondary subjects with those with a low risk of failure (Gabay-Egozi et al., 2010).

The economic and cultural capital available within families also influences pupils' attitudes, aspirations, and occupational choices (Barker & Hoskins, 2017; Dilnot, 2016). Disadvantaged pupils have fewer opportunities and resources to support their aspirations (Archer et al., 2012).[74] Woods (1976) observes that many follow a pre-ordained route, as if there was not much choice and certain subjects were not appropriate for them. Science, for instance, remains an unthinkable aspiration for some (Archer et al., 2012). Advantaged pupils who have graduated parents are more likely to expect to work in science (Jerrim & Shure, 2016; OECD, 2016b). This partly reflects the opportunities and support available to them (OECD, 2016b).[75] Access to science-related resources and activities may foster a strong science orientation (Archer et al., 2012). But attitudinal factors also deter disadvantaged pupils from science subjects (OECD, 2016b).

The most common reasons pupils give for choosing subjects are their usefulness for future careers, interest in the subject, their ability, and subjects' difficulty (Jin et al., 2011; Opposs, 2016; Tripney et al., 2010). Pupils' perception of their ability is thus central. Academic self-concept is related to prior attainment and parents' education (Sullivan, 2009). Interestingly, subject choices of pupils with high academic self-concept are greatly influenced by parents and teachers (Sullivan, Zimdars, et al., 2010).

[74] Delavande and Fumagalli (2019) show that disadvantaged pupils' expectations regarding university admission are less closely linked to prior attainment.

[75] Parents in professional occupations and with higher levels of cultural capital have a more positive attitude to science, but fathers with a science background are particularly influential (Archer et al., 2012). Knowing the future implications of subject choices, they help their children navigate the curricular structure and secure the best options (Lucas, 2001).

Socioeconomic differences in cognitive development emerge at an early age (Barker & Hoskins, 2017). Advantaged pupils generally perform well in mathematics and have a comparative advantage in science (Werfhorst et al., 2003). Werfhorst et al. (2003), using data from the NCDS, show that pupils choose subjects where they have a comparative advantage. Codiroli (2017) reveals that pupils who do well in English at age 11 are less likely to choose STEM A level subjects. Pupils with comparative advantage in reading at age 11 or in humanities at age 16 are more likely to take social studies and arts at university (Werfhorst et al., 2003). On the other hand, girls with a comparative advantage in science at age 11 and that take triple science at age 16 are more likely to enrol in STEM degrees (De Philippis, 2021).

A systematic review of UK literature has shown that gender and prior ability are important determinants of subject choice at age 14 (Tripney et al., 2010). Boys are more likely to take triple science and girls to take modern foreign languages, whilst high ability pupils are more likely to take both (Jin et al., 2011; Tripney et al., 2010). Sullivan (2009) and Sullivan, Joshi, et al. (2010), using the NCDS, similarly reveal that girls are less likely to study mathematics and science. More recently, Anders et al. (2018b) also find that gender is associated with secondary subject choice. Girls study a more academically selective curriculum, are less likely to study three separate sciences, and are more likely to take applied subjects (Anders et al., 2018b). While qualitative data from secondary schools suggests most pupils believe ability is unrelated to gender and that pupils' favourite subjects are not clearly gendered (Francis, 2000), Henderson et al. (2018) confirm that girls are less likely to study STEM subjects and are more likely to study applied subjects.

Anders et al. (2018b) also find that socioeconomic background and prior attainment are associated with subject choice at secondary school. Advantaged and high performing pupils are more likely to take academically selective subjects, facilitating subjects, and EBacc subjects, while those with low prior attainment are more likely to take applied subjects (Anders et al., 2018b). Jin et al. (2011) confirm that advantaged pupils and those with high prior attainment are more likely to take the EBacc. Similarly, Henderson et al. (2018) show that pupils from wealthier backgrounds and with more educated parents study a more academically selective curriculum or the EBacc and take less applied subjects.[76] Henderson et al. (2018) reveal, though, that socio-economic differences in choosing STEM subjects are largely driven by prior attainment. Triple science, for example, is predominantly taken by high ability pupils (De Philippis, 2021; Department for Education, 2018).

4.2.4 School Effect

Anders et al. (2018b) show that GCSE subject choice is shaped by the schools pupils attend. They use a random effects model based on the NPD and focus on academically selective subjects,[77] facilitating subjects, the EBacc, triple science subjects, and applied subjects. Anders et al. (2018b) reveal a significant variation in subject choice between schools. They find that pupils attending schools with high achieving advantaged intakes are more

---

[76] This is consistent with Sullivan, Zimdars, et al. (2010) that, through a multiple correspondence analysis on the same data, show pupils with graduated parents clustering around science subjects, modern languages, and history.

[77] Academic selectivity is measured by the average KS3 score of pupils taking the subject. The academic selectivity of a pupil set of subjects is measured by the standardised sum of the academic selectivity score of the eight most selective subjects one takes (Anders et al., 2018a; Henderson et al., 2018).

likely to take academically selective subjects, facilitating subjects, and EBacc subjects and are less likely to take applied subjects. Similarly, Henderson et al. (2018), using Next Steps linked with the NPD, show that pupils in schools with a high proportion of free school meals (FSM) eligible pupils are less likely to study academically selective subjects and the EBacc. This is consistent with Davies et al. (2008).

Anders et al. (2018b) also reveal that pupils attending single sex schools and grammar schools take more academically selective subjects, facilitating subjects, EBacc subjects, and triple science (all-male schools) and take less applied subjects (Anders et al., 2018b). Besides, they note a difference in subjects studied at non-selective schools in local areas where grammar schools have a meaningful share (Anders et al., 2018b). Similarly, Henderson et al. (2018) find that pupils attending grammar schools or single-sex schools study a more academically selective curriculum and the EBacc and take less applied subjects.[78] Academies are understandably absent from these studies which refer to an earlier period.

Discussing their findings, Henderson et al. (2018) draw attention to school policies and curriculum offer that affect pupils' choices. Jin et al. (2011), using Next Steps, show that some pupils are actually unable to choose subjects that they would like to study because of limited school offer. Many pupils are in schools that, for instance, do not offer triple science (Jin et al., 2011). Since school offer is related to school composition, Anders et al.

[78] In the NCDS (an earlier cohort), pupils at grammar schools are more likely to study science subjects and pupils at single sex schools are more likely to study gender atypical subjects (Davies et al., 2008; Sullivan, 2009; Sullivan, Joshi, et al., 2010).

(2018b) suggest that it reflects what is deemed appropriate for pupil intakes, considering their average ability and socioeconomic background.[79] Changes in the pupil intake might, therefore, alter teachers' expectations and what is deemed to meet pupils' needs (Davies et al., 2003). On the other hand, school offer depends on the expertise of teachers (Davies et al., 2003). Anders et al. (2018b) highlight the struggle of disadvantaged schools in recruiting and retaining qualified teachers for subjects like languages and sciences. This is corroborated by Martindale (2019). Hence, schools make decisions about the subjects they offer to meet the needs and interests of their pupils, within the constraints of their resources and the local education market (Anders et al., 2018b; Davies et al., 2003; Dilnot, 2016; Jin et al., 2011; Woods, 1976).

Accordingly, qualitative data collected by Davies et al. (2003) suggests that access to academic and vocational options varies across schools. Some schools focus on academic subjects while others offer a large range of academic and vocational options (Jin et al., 2011). Jin et al. (2011) find that larger schools offer a wider range of subjects. It is difficult for small schools to offer many options (De Philippis, 2021).[80] Davies et al. (2003) argue that popular high-performing schools set the trend and the other local schools seize niches pre-empting options.

---

[79] An earlier qualitative study shows teachers monitoring subjects' selection and encouraging choices according to ability and family background (Woods, 1976). More recently, Opposs (2016) reveals that some schools do not offer subjects considered difficult and that teachers discourage some pupils from taking those. Inter-subject comparisons present art, fine art, physical education, and home economics consistently among the easiest GCSE subjects, and statistics, modern foreign languages, and Latin among the most difficult (He et al., 2018).

[80] Given their size, small schools do not have capacity to employ a range of specialists and have teachers working in subjects other than their degree subject (Davies et al., 2003). Small teaching groups are subsidised by larger groups elsewhere (Davies et al., 2003).

Given league tables' benchmarks and in order to maximise their ranking position, some schools might pressure high performing pupils to take certain subjects and encourage pupils at the margin to take easier options and avoid more challenging ones (Barrance & Elwood, 2018; Henderson et al., 2018; Jin et al., 2011). In effect, as suggested in Allen and Burgess (2013), school exam entry policies respond to changes in performance metrics. Schools have the incentive to offer subjects that can improve their performance and encourage pupils taking these (Jin et al., 2011; Long, 2019a). For example, headline measures focused on facilitating subjects, such as the EBacc, are likely to cause a shift towards them (Department for Education, 2018; Dilnot, 2016; Long, 2019a). Restrictions on the inclusion of vocational qualifications[81] have actually led sponsor-led academies to adapt a more academic curriculum (Hatton et al., 2019; Hutchings & Francis, 2018).

Considering how school policies shape subject choice and the underlying incentives and constraints schools face, one cannot but wonder how autonomous schools act. Hence, this chapter looks at how secondary academies use their autonomy on school provision. Previous studies refer to the period before the massive academisation of English secondary schooling and thus not address this issue.

[81] Wolf (2011) advises that pupils are entered for some vocational courses, that they are deemed to pass easily, to improve school performance at league tables. In fact, Jin et al. (2011) reveal that the increase in pupils meeting the performance benchmark has been driven by the rise of vocational qualifications and that schools that have been swiftest to adopt these have made larger gains in performance.

4.3 DATA

The Millennium Cohort Study (MCS) draws its participants from children born in the UK over a 16 month period around the year 2000 (Centre for Longitudinal Studies, 2020; Fitzsimons, Agalioti-Sgompou, et al., 2020). In England, where the sample of interest is drawn,[82] this corresponds to a single school year, from 1 September 2000 to 31 August 2001, comprising 11,695 children. The MCS follows cohort members through their lives, currently including seven sweeps taken at nine months, age three, five, seven, 11, 14, and 17. The variables of interest are mainly taken from surveys at age 11 and 17, but information from surveys at age five, seven, and 14 is also considered. At age 17, 7,198 cohort members have participated in the survey in England. The MCS sample is stratified to ensure economically deprived regions and areas with high proportions of ethnic minorities are overrepresented (Centre for Longitudinal Studies, 2020; Fitzsimons, Agalioti-Sgompou, et al., 2020). Because of MCS sampling design and attrition, the data is weighted to provide a representative random sample of the population.

Despite providing extensive information on education, the MCS has very little school level data (Platt & Centre for Longitudinal Studies, 2014). Establishment type[83] is derived from the list of schools used at the MCS sixth survey (MCS6)[84] in 2015. I use the Consistent School Database (CSD)[85] to

[82] The sample is limited to England where academies were introduced.

[83] Establishment type groups: converter academies, sponsor-led academies, free schools, LA maintained schools, independent schools.

[84] University of London, Institute of Education, Centre for Longitudinal Studies (2020c).

[85] Cohort and Longitudinal Studies Enhancement Resources (2018).

identify secondary school type at starting year (2012/13) and the set of secondary schools that convert to academy status in 2013/14 and 2014/15.[86] The resulting school list is matched to MCS, identifying both the school attended in 2015 and the school where cohort members started secondary phase. This allows the identification of the treatment group – pupils attending LA maintained schools that become academies in 2013/14 and 2014/15. Cohort members that are not in a secondary school or moving to a secondary school in 2012/13 are excluded to ensure the integrity of the treatment group.[87] Only cohort members whose subject choice and secondary school attended are known are included.[88]

The estimation sample is not randomly recruited from MCS and intends to represent only the population of interest (pupils attending LA maintained schools). As such, it may have distinctive features without hindering external validity. When I compare this sample with the full MCS cohort for England, through a logit model, I find little evidence of selection (see appendix B Table B2). Cohort members in the sample have a different ethnic background, live in

[86] The CSD allows matching of school records between years when school identifiers (URN) change, as in the case of academy conversion. It contains school level data linked over time between 1998/99 and 2016/17, identifying establishment type and academy conversion date. For the few cases where the CSD does not identify establishment type in 2012/13, academies' predecessor school and conversion date are identified using the list provided in the DfE Academies Management Information Data (May, 2018).

[87] At age 11, 95% cohort members are in school year six, one percent are in school year five, and four percent have already started secondary school (Platt & Centre for Longitudinal Studies, 2014). Pupils attending middle schools are excluded.

[88] Information on GCSEs (or equivalent qualifications) pupils take at age 16 is missing for 602 cohort members. Information on moving to secondary phase is missing for 119 cohort members. There is also no information on school started secondary phase for 151 cohort members and school type attended at age 14 for 11 cohort members. A logit analysis reveals that missing cases are more likely to be male, white, and from low income families.

more deprived areas, and their parents have slightly lower occupations. This to some extent reflects pupil intakes at maintained schools nationwide.[89]

Table 4.1 presents basic information on the distribution of pupils across schools in the estimation sample. The 2,524 cohort members in the sample are distributed across 728 schools. The sample focuses on pupils that enrol in LA maintained schools at the start of secondary phase. In 2012/13, 3,443 (44%) cohort members enrol in LA maintained schools and 3,557 (47%) in academies. The remainder 7% are in free schools and independent schools.[90] The 179 cohort members enrolled in middle schools are excluded. This is consistent with the national distribution. At secondary phase in 2012/13 there were 1,266 academies – 963 converter and 303 sponsor-led academies (Eyles et al., 2018), representing 50% of secondary schools (National Audit Office, 2018).

In the period of interest (2013/14 and 2014/15), 228 LA maintained schools convert – 151 converter and 77 sponsor-led academies (Eyles et al., 2018). The treatment group comprises 454 pupils enrolled in 118 LA maintained schools that later become academies. Half of schools that convert during this period are represented in the sample of interest. The fact that 146 cohort members have moved secondary school raises issues of non-compliance with treatment assignment that are addressed in the next section.[91]

---

[89] In 2012/13, 13.7% of pupils at secondary academies are eligible for FSM compared with 16.3% across all secondary state schools (Department for Education, 2013b).

[90] School type at start of secondary phase is unknown for 248 cohort members, and age 17 outcomes are missing for 602 cohort members.

[91] In 2015, 10 pupils in the treatment group do not attend an academy and 53 pupils in the control group attend an academy (missing information for 11 cohort members).

Focusing on pupils that remain in the same school,[92] MCS6 reveals that 8% of the treatment group attend faith schools. 8% of those attending prospective converter academies are in grammar schools and 16% in single-sex schools. In comparison, no one in grammar schools and only 2% in single-sex schools attend prospective sponsor-led academies. These figures roughly reflect contemporaneous national trends of academy conversion at the time. Comparing KS4 performance tables in 2013 and 2015, 14% of the increase in the number of academies is observed in faith schools. The number of converter academies that are selective and single-sex has also increased (6% and 10% respectively) in this period. In contrast, the number of single-sex sponsor-led academies marginally decreased nationwide. Besides, sponsor-led academies are not allowed to be selective (Long et al., 2023). The treatment group thus seem a fitting representation of schools becoming academies in 2013/14 and 2014/15.

Table 4.1 Sample distribution per school type in 2012/13

	Number of pupils
LA maintained schools (CONTROL GROUP)	2,070
Schools that convert in 2013/14 and 2014/15 (TREATMENT GROUP)	
Prospective sponsor-led academies	161
Prospective converter academies	293
TOTAL	**2,524**
	Number of schools
LA maintained schools	610
Prospective sponsor-led academies	40
Prospective converter academies	78
TOTAL	**728**

Note: Statistics do not reflect MCS survey settings (unweighted measures). The sample is restricted to pupils reported as attending a secondary school or moving to secondary school in 2012/13 in England (University of London, Institute of Education, Centre for Longitudinal Studies, 2020c).

---

[92] MCS6 has information on secondary schools attended at age 14, but not for schools at start of secondary phase (except for grammar schools).

4.3.1 Outcomes of Interest

The outcomes of interest are binary variables identifying whether cohort members choose EBacc subjects, facilitating subjects, STEM subjects, and vocational subjects at age 14. See appendix B Table B1 for a list of subjects entered in each outcome variable. Information is drawn from the seventh survey (MCS7),[93] at age 17, on the GCSEs (or equivalent qualifications) pupils take at age 16.

Facilitating A levels are considered a good general preparation for university study (Dilnot, 2018). Facilitating A level subjects (recommended by the Russell Group) include mathematics, English, English literature, biology, chemistry, physics, history, geography, modern and classical languages (Anders et al., 2018b; Dilnot, 2016; Moulton et al., 2018). A level subjects depend on those taken at GCSE and are the standard qualifications for university entry (Anders et al., 2018a; Moulton et al., 2018). GCSE precursors to facilitating A level subjects include English, mathematics, history, geography, languages, physics, chemistry, and biology (Anders et al., 2018b; Dilnot, 2016). Pupils are identified as taking facilitating subjects (binary variable) if they study at least three facilitating subjects at GCSEs (or equivalent qualifications), excluding English and mathematics.[94]

---

[93] University of London, Institute of Education, Centre for Longitudinal Studies (2020b).

[94] English and mathematics are taken by 90% and 89% of sample respectively. GCSEs in English and mathematics have typically take-up rates of over 90% (Carroll & Gill, 2018; Department for Education, 2018; Jin et al., 2011). In effect, if pupils fail to achieve a good pass in English or mathematics, they are required to retake the subject (Vidal Rodeiro & Vitello, 2021). This is the first cohort to take the reformed GCSEs in English language, English literature, and mathematics (Carroll & Gill, 2018; Department for Education, 2018).

The EBacc curriculum at GCSE comprise English language and English literature, mathematics, history or geography, sciences (core and additional science, double science award, or three single sciences), and a modern or ancient foreign language (Anders et al., 2018a; Henderson et al., 2018; Long, 2019a; Moulton et al., 2018). Since subjects available to pupils at age 16 depend on those taken earlier, undertaking the EBacc allows pupils to choose facilitating A level subjects (Anders et al., 2018a; Long, 2019a; Moulton et al., 2018). Pupils are identified as taking the EBacc (binary variable) if they have studied the full set of eligible subjects (listed above).

The EBacc science component identifies students taking two or more sciences subjects, including core and additional science GCSEs, double science awards across all three major science subjects, or three single sciences in biology, chemistry, physics, or computer science (Anders et al., 2018a; Long, 2019a; Moulton et al., 2018). STEM GCSE subjects also include information and communications technology, and statistics (Henderson et al., 2018). These subjects are of interest given the need for a workforce trained for STEM-based occupations which have high labour market returns (Dilnot, 2016; Sullivan, 2009). Pupils are identified as taking STEM subjects (binary variable) if they study at least three STEM subjects, taking into account double awards in science.[95]

Vocational qualifications offered at age 14 are broad in scope (Vidal Rodeiro & Vitello, 2021). GCSE applied subjects include applied art and

[95] Triple science implies longer instruction time and more complex topics (De Philippis, 2021). Some pupils take instead core and additional science with further additional science (Department for Education, 2018).

design, applied business, engineering, health and social care, applied ICT, leisure and tourism, manufacturing, and applied science (Moulton et al., 2018). Pupils taking these courses may not be able to take facilitating A level subjects (Moulton et al., 2018). Vocational options at age 16 (KS5) include National Vocational Qualifications, Business and Technology Education Courses, City and Guilds, and vocational courses at Further Education Colleges (Moulton et al., 2018). Pupils are identified as studying a vocational course (binary variable) if they have taken a vocational course or at least one applied GCSE subject.

4.3.2 Explanatory Variables

Administrative data includes information on pupils' gender, ethnic group, language, and FSM eligibility. Comparable variables are derived from MCS data.[96] In general, the construction of variables follows previous literature using the MCS (Dearden et al., 2011; Sullivan et al., 2013). Most variables used are available at the fifth survey (MCS5),[97] age 11, but data from earlier sweeps is also considered.[98]

MCS includes the Organisation for Economic Cooperation and Development (OECD) equivalised weekly family income (Centre for Longitudinal Studies, 2020; Hagenaars et al., 1994). Whereas FSM eligibility reveals how the most disadvantaged pupils are sorted, income also describes

---

[96] Aggregate ethnic groups (five categories) are based on the Census categories (Office for National Statistics, 2014).

[97] University of London, Institute of Education, Centre for Longitudinal Studies (2022).

[98] In the construction of variables, I mainly use the MCS5. MCS5 was collected in 2012 at age 11, the last year of primary schooling and before transition to secondary school (Centre for Longitudinal Studies, 2020). A few lagged variables of interest are included from MCS3 and MCS4, at age five and seven respectively.

more advantaged families. Furthermore, equivalisation[99] allows a comparison between households of different sizes (Department for Work and Pensions, 2012). I use the log of OECD average equivalised weekly family income.[100]

MCS data on parents' education and occupation allows to further differentiate pupils' background. I consider both the mother and the father (where present) to derive parents' socio-economic status.[101] Parental occupation (current job) is derived from a five-category version of the National Statistics Socio-economic Classification (NS-SEC) (Rose & Pevalin, 2003), based on the Standard Occupational Classification. Since NS-SEC excludes parents that are not in work nor on leave, a category identifying those cases is added. Similarly, I use the highest educational qualification of parents, including both academic and vocational qualifications, grouped in the National Vocational Qualifications (NVQ) levels (Agalioti-Sgompou et al., 2017).[102] Related variables include parents' average hours worked, family composition

---

[99] The OECD modified equalisation scale assumes that the needs of a household grow with each additional member but not in a proportional way. The OECD modified equalisation scale is re-scaled to take a couple without children as the reference group, assigning a value of one to the main carer, of 0.33 to each additional adult member, and of 0.2 to each child (aged under 14).

[100] Income is a strictly positive ratio scale variable that is skewed to the right. The logarithmic transformation reduces data heteroscedasticity or skewedness and makes the estimates less sensitive to outliers, narrowing the variable range (Wooldridge, 2014). Moreover, when using logs the coefficient does not depend on the unit of measurement and approximates a percentage change (Wooldridge, 2014).

[101] When there is information on both parents, the higher status occupation is chosen.

[102] NVQ level one corresponds to approximately GSCE D – G, NVQ level two to GSCE A* – C, NVQ level three to A levels, NVQ level four to a BA degree, and NVQ level five to a postgraduate degree.

(number of parents, grandparents, siblings), housing tenure,[103] main carer's age, longstanding illness and mental health.[104]

Geographical data available in the MCS identifies the local level of deprivation where cohort members reside.[105] The domain deciles of the Index of Multiple Deprivation (IMD) 2004 (Office of the Deputy Prime Minister & Neighbourhood Renewal Unit, 2004) measure relative levels of deprivation at LSOA and allow a comparison of deprivation types.

Available information at MCS includes cohort members' age, general level of health, bedtime rules, how often someone helps doing homework, hours per weekday spent watching television or videos, number of books at home, visits to library, and church attendance. Early child variables such as whether received childcare, looked after by grandparents, and how often someone read to cohort member at the third survey (MCS3),[106] age five, are also included.

The MCS has administered several cognitive tests to cohort members. These include the British Ability Scales (BAS) Picture Similarity[107] at age five, BAS Pattern Construction[108] at age five and at the fourth survey (MCS4),[109] age seven, BAS Word Reading at age seven, Number Skills at age seven,

---

[103] Housing tenure types follow those used in the English Housing Survey (Ministry of Housing, Communities and Local Government, 2020).

[104] The cohort members' main carer is typically the mother. The main respondent's mental health is screened using a standardised Kessler Six Scale (Kessler et al., 2003) that quantifies psychological distress based on self-reported symptoms of depression and anxiety.

[105] Geo-referenced data at MCS is derived using mean unit postcode centroids based in the ONS Postcode Directory and postcodes collected during interviews (Church, 2017).

[106] University of London, Institute of Education, Centre for Longitudinal Studies (2020d).

[107] BAS Picture Similarity measures problem solving ability and non-verbal reasoning.

[108] BAS Pattern Construction measures spatial problem solving.

[109] University of London, Institute of Education, Centre for Longitudinal Studies (2020a).

BAS Verbal Similarities[110] at age 11, and Cambridge Neuropsychological Test Automated Battery (CANTAB) Cambridge Gambling Task (CGT)[111] at age 11 (Centre for Longitudinal Studies, 2020; Fitzsimons, Agalioti-Sgompou, et al., 2020). The MCS also administers several behaviour assessments. It includes the Strengths and Difficulties Questionnaire (SDQ)[112] filled out by the parents, at age three, five, seven, and 11 (Centre for Longitudinal Studies, 2020; Fitzsimons, Agalioti-Sgompou, et al., 2020).

There is not much school level information available on the MCS, but it is possible to identify cohort members that have attended more than two primary schools and those that have attended single sex schools or faith schools in primary phase. Similar information is gathered for secondary schools attended at age 14 but not for schools where cohort members have started secondary phase.

Multiple imputation is used to handle missing data due to item nonresponse, when is not possible to impute missing values from previous

[110] BAS Verbal Similarities measures verbal reasoning and acquired knowledge.

[111] The CGT assesses decision making skills under uncertainty. CGT Quality of Decision Making denotes better decision making. CGT Risk Taking measures sensitivity to reward and punishment.

[112] The SDQ Total Difficulties subscale reveals emotional problems, conduct problems, hyperactivity/inattention, and peer problems. Parent reported SDQ variables were removed from the fifth edition of MCS5 (age 11), used in this chapter.

sweeps.[113] Item nonresponse is assumed to be missing at random.[114] A complete case analysis is provided for comparison.[115]

4.3.3 Descriptive Statistics

Table 4.2 presents the characteristics of pupils in the treatment and control groups.[116] A raw comparison suggests that pupils in LA maintained schools that convert to academy status in 2013/14 and 2014/15 are not dissimilar. Most variables are not significantly different and there is no headline difference between the treatment and the control groups. There are, though, a few differences. Pupils in the treatment group report worse general health and go to church less often. However, when the treatment group is divided according to the academy route taken, differences are clearer.

[113] The imputation model includes the outcome variables and explanatory variables used in analysis, MCS sampling variables and survey weights, plus earlier measures of imputed variables. 70 imputations are estimated using chained equations. Imputed values of outcome variables and of treatment assignment are not used in analysis.

[114] Covariates have less than 5% missing cases. Only one variable has more than 100 missing cases.

[115] There are 1,728 complete cases, representing 68% of sample of interest. They take more EBacc subjects, facilitating subjects, and STEM subjects. Complete cases are higher performing pupils from a white English-speaking advantaged background (higher income, more educated parents in managerial positions).

[116] See appendix B Table B3 for descriptive statistics on the full list of explanatory variables and Table B4 for Chi-square tests on categorical variables. Chi-square tests are used to identify independence between categorical variables and treatment assignment. Since this does not clearly identify the source of a statistically significant difference, a further investigation of individual categories might be required (Sharpe, 2015).

Table 4.2 Descriptive statistics pre-treatment

	CONTROL		TREATMENT			SPONSOR-LED			CONVERTER		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
VARIABLES	**N**	**Mean**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**
AGE 11											
CM CHARACTERISTICS											
CM Sex (Male)	2,070	0.508	454	0.494	-0.0135	161	0.411	-0.097**	293	0.538	0.0305
		(0.539)		(0.527)	(0.0361)		(0.525)	(0.0452)		(0.522)	(0.0430)
Verbal Similarities Ability	2,041	-0.0149	447	-0.178	-0.164*	159	-0.259	-0.244	288	-0.136	-0.121
Score		(1.049)		(1.020)	(0.0893)		(1.033)	(0.166)		(1.011)	(0.0994)
SOCIOECONOMIC STATUS											
OECD equiv weekly family	2,070	5.972	454	5.934	-0.0378	161	5.813	-0.159*	293	5.998	0.0264
income (log)		(0.492)		(0.487)	(0.0528)		(0.510)	(0.0857)		(0.461)	(0.0660)
Free School Meal Eligible	2,065	0.175	452	0.191	0.0160	160	0.283	0.108*	292	0.143	-0.0323
		(0.410)		(0.414)	(0.0307)		(0.480)	(0.0650)		(0.366)	(0.0302)
Parents Occupation	2,048		453			160			293		
Managerial and		0.420		0.366	-0.0543		0.207	-0.21***		0.450	0.0295
professional		(0.534)		(0.508)	(0.0405)		(0.432)	(0.0329)		(0.521)	(0.0464)
Parents Education	1,854		403			131			272		
NVQ equiv. level 4		0.404		0.361	-0.0429		0.253	-0.15***		0.412	0.00752
(incl. BA degree)		(0.522)		(0.502)	(0.0360)		(0.460)	(0.0363)		(0.511)	(0.0415)
NVQ equiv. level 5		0.142		0.157	0.0143		0.0833	-0.0590		0.191	0.0483
(incl. postgraduate degree)		(0.372)		(0.380)	(0.0266)		(0.292)	(0.0439)		(0.408)	(0.0296)

Note: The treatment group includes pupils enrolled in 2012/13 at LA maintained secondary schools that become academies in 2013/14 and 2014/15, whereas the control group includes pupils enrolled in LA maintained secondary schools that do not convert to academy status. The treatment group is composed by pupils attending prospective sponsor-led and converter academies. The quantity in parenthesis below the mean (or proportion in the case of categorical variables) is the standard deviation. All statistics, except the number of observations (N) that is unweighted, reflect MCS survey settings. The difference (diff) refers to the difference in means between treatment group and control group and the quantity in parenthesis below is the (robust) standard error clustered at the school level and reflecting MCS survey setting (University of London, Institute of Education, Centre for Longitudinal Studies, 2022). * significant at 10% level; ** significant at 5% level; *** significant at 1% level or below.

Pupils attending LA maintained schools that later become sponsor-led academies have rather disadvantaged characteristics. They are more likely to be eligible for FSM. Their parents have a lower income and are less likely to have a degree or a managerial/professional occupation. This is consistent with national statistics. Schools becoming sponsor-led academies in 2013/14 and 2014/15 have 23% FSM pupils (Hatton et al., 2019). Moreover, pupils at prospective sponsor-led academies have lower prior attainment in reading and maths. They also report worse health and more mental health issues at age five. Their mothers are younger, have worse mental health, have fewer books at home, and live in more deprived communities.

In contrast, pupils attending LA maintained schools that later become converter academies live in somewhat less deprived areas. This is also consistent with national trends (Bolton, 2015). They have more regular bedtimes and spend less time watching television. These pupils report better health and show fewer mental health issues at age five.

The abovementioned baseline differences between the groups are based on many independent tests. As suggested by Armstrong (2014), a Bonferroni correction is appropriate when testing the global null hypothesis that multiple tests are not significant, so as to ascertain if a particular treatment has the estimated effect.[117] In this case, the Bonferroni correction shows that

---

[117] The probability of finding at least one statistically significant difference across a set of tests, even when there is no actual difference, increases with each additional test (Armstrong, 2014; Gelman et al., 2012). This implies that, given the number of tests performed, a few statistically significant differences would always be expected, regardless of the true associations. The Bonferroni correction adjusts the p-value (divides the significance level by the total number of tests being performed) so as to maintain the overall significance level (Armstrong, 2014; Gelman et al., 2012; Sharpe, 2015; VanderWeele & Mathur, 2019; Wright, 1992). It ensures that the global null hypothesis is rejected no more than the significance level when it in fact holds.

no difference remains significant (at 10 percent significance level) when comparing the treatment and the control groups. This reinforces the impression that there is no headline difference between pupils attending LA maintained schools and those attending schools that later become academies. Nevertheless, the apparent difference in pupil intakes at sponsor-led and converter academies, which is consistent with academy routes, requires a study of possible heterogenous effects. The next section will consider the identification strategy and how to address possible selection issues.

4.4 Methodology

The purpose of this chapter is to investigate the impact of academy conversion on subject choice at age 14. I next outline the empirical strategy used, highlighting identification issues for a causal interpretation.

Pupils are not randomly assigned to schools and schools do not randomly become academies. Academy conversion is either voluntary for high performing schools (converter academies) or forced onto underperforming schools (sponsor-led academies). Pupils are distributed across schools as a result of selection and parental choice. If pupils who prefer an academic curriculum are more likely to attend an academy, comparing subject choices of pupils at academies to those of pupils at maintained schools will likely exaggerate the impact of academy attendance (Angrist & Pischke, 2009).

Since the MCS has little school-level information, it is not possible to directly account for selection of schools into academisation. Schools that become academies may be different from LA maintained schools that do not convert. Available data do not allow to further differentiate schools. MCS has though very rich pupil-level longitudinal data that could help account for school selection. If pupils are as good as randomly distributed across groups, they are likely to attend similar schools. The methods that follow attempt to account for pupil selection, effectively also accounting for school selection.

Because of selection into and out of academies, pupils are only included in the treatment group if they are enrolled in the secondary school prior to conversion and are affected by academy conversion in subsequent years of their secondary schooling – known as "legacy" enrolments or "grandfathered" pupils (Abdulkadiroğlu et al., 2016; Andrews et al., 2017; Eyles et al., 2016b, 2017; Eyles & Machin, 2019). Although pupils tend to remain in the same school from school year seven to school year 11, some change secondary school. Hence, I estimate the intention to treat (ITT) effect of academy conversion on pre-enrolled pupils, irrespective of whether they actually take their GCSEs in that school (Andrews et al., 2017; Eyles et al., 2016b, 2017; Eyles & Machin, 2019).[118]

Since I am looking at the GCSE (or equivalent) entries of the MCS cohort,[119] I focus on pupils that at start of secondary phase (2012/13) enrol in LA maintained schools that become academies in 2013/14 and 2014/15. This ensures academy conversion is exogenous to secondary school enrolment decision. Most academies open between July and September (Bertoni et al., 2021). If an academy has opened in 2013/14 or 2014/15 (treatment group), its conversion application process will have started after pupils apply to

[118] This avoids issues related with self-selection or schools sorting out pupils.
[119] GCSE subjects are chosen in 2014/15 (year 9) and examined in 2016/17.

schools by October 2011 (starting secondary phase in September 2012). DfE approval takes on average two months after application and converter academies open on average five months after approval (Bertoni et al., 2021). In 2017, the conversion process takes on average 11 months (National Audit Office, 2018). Schools that are directed to convert open as sponsor-led academies on average 18 months after being rated “inadequate” at inspection (National Audit Office, 2018). The median time for conversion under a sponsor is 17 months (Duchini et al., 2023). These schools may go through a period of discussion and resistance before conversion (Keddie, 2019). Decision of converter academies to apply for conversion may also be preceded by controversy. This is a period of uncertainty involving several entities and many schools at the end do not convert (Rayner et al., 2018). So, even if some parents were aware of these discussions at the time they choose secondary school it would be difficult to say if a specific school will become an academy. Therefore, given time taken to process applications, academisation is not likely to have been anticipated by parents at the time of secondary school choice. This is consistent with similar pupil intakes at maintained schools that become academies and at those that do not, as described in section 4.3.3.

Taking advantage of the rich MCS database, I use the explanatory variables presented in the previous section to attempt an identification strategy based on selection on observables (Dearden et al., 2009). The treatment impact is estimated using a logistic model. The logistic regression estimates the odds ratio of a specific subject choice associated with academy attendance, controlling for a comprehensive set of pre-treatment variables. Under the conditional independence assumption, treatment estimates have a causal interpretation (Angrist & Pischke, 2009). This requires including all relevant variables so that potential outcomes are independent of treatment assignment (enrol in a school that becomes an academy). MCS gives some credibility to this claim, given the rich longitudinal data we can draw variables from. It is, nonetheless, a strong assumption.

The model of interest is based on the following relationship for pupil *i* in legacy enrolment school *s*:

$$Y_{is} = \delta_{01} + \beta_1 A_s + \sum_{j=1}^{J} \delta_j X_{jis} + \varepsilon_{1is} \qquad (4.1)$$

where Y_{is} are binary variables denoting whether a pupil has chosen a particular subject group at age 14 (EBacc, STEM subjects, facilitating subjects, applied subjects), X_{jis} is a set of *j* pre-treatment pupil and school characteristics, δ_{01} is an intercept, and ε_{1is} is an error term. X_{jis} includes pupil characteristics (sex, age, ethnic group, language, health, BAS and CGT scores), family context (number of parents, grandparents and siblings, mother's age, illness and mental health), socioeconomic status (income, FSM, housing tenure, parents' hours of work, occupation and education), home environment (number of books, library and church attendance, hours watching TV, bedtime, help with homework), and geographical variables (IMD domains) at age 11, prior performance (Word Reading, Maths, Pattern Construction, Picture Similarities, SDQ total difficulties scores) at age five and seven, childcare (providers, grandparents, reading), and primary school characteristics (single-sex or faith school) (see appendix B Table B3 for a detailed list of variables). A_s is a dummy variable for legacy enrolment at secondary schools that become academies in the sample period (treatment group). The coefficient of interest is β_1. The error term is assumed to be uncorrelated with treatment, but a degree of autocorrelation within schools is allowed (Abadie et al., 2017).[120]

ITT estimates calculate the effect of starting secondary phase in a school that later converts to academy status. I also estimate the actual effect of attending an academy using legacy enrolment (ITT) as an instrument for academy attendance, within a linear probability model (OLS). The effect estimated through instrumental variables (IV) corrects for the fact that not all legacy enrolled pupils remain in the school and receive treatment. Given the high rate of compliance with treatment assignment,[121] ITT and IV estimates are expected to be similar. Hence, in what follows both will be referred to as the effect of attending an academy.

IV measure the local average treatment effect (LATE) on pupils that would not have attended an academy had they not been pre-enrolled, excluding those pupils that would always move to or from an academy (Angrist & Pischke, 2009). It assumes that legacy enrolment is as good as randomly assigned and it has no effect on outcomes except through academisation (Angrist & Pischke, 2009). The observed similarity between treatment and control groups, discussed earlier, lends plausibility to this assumption.

The first stage equation, underlying the reduced form above (4.1), is as follows:

$$I_s = \delta_{02} + \beta_2 A_s + \sum_{j=1}^{J} \delta_{2j} X_{jis} + \varepsilon_{2is} \qquad (4.2)$$

[120] Due to MCS geographical clustering, most are in the same school as at least one other cohort member. 343 are the only cohort members at their school. Pupils are distributed across 728 schools with an average of three pupils per school.

[121] Focusing on the school type attended in 2015, there are 63 pupils that do not comply with treatment assignment. At age 14, 53 pupils in the control group attend an academy and 10 pupils in the treatment group do not attend an academy.

where I_s is a dummy variable that indicates if a pupil is enrolled in an academy at age 14. This reveals if pupils enrolled in the predecessor school at the start of secondary phase receive treatment by academy conversion. $ß_2$ is the proportion of pupils in the ITT group that stay in the school and are surveyed while there.[122] The first stage also controls for the set of variables used in the reduced form. Second stage estimates are obtained replacing A_s by the predicted I_s in (4.1).

Propensity score matching (PSM) at pupil level is used to check robustness of results.[123] PSM selects from maintained schools that do not convert those pupils that have observed characteristics as similar as possible to those attending schools that become academies. The probability of treatment (attending a school that becomes an academy) is estimated using a probit regression with a set of pre-treatment variables that influence outcomes and school enrolment but are unaffected by that decision.[124] PSM weights define the control group.[125] Conditional on matching variables, it is assumed that the assignment to treatment is random (independent of outcomes) and the comparison between treatment and control groups is an unbiased estimate of the treatment effect (Angrist & Pischke, 2009). Since PSM relies on selection on observables is also prone to the omission of relevant variables.

[122] Pupils in the treatment group that move to another academy or pupils in the control group that move to another LA maintained school are also included among compliers.

[123] Available data do not allow for school-level matching.

[124] Following DuGoff et al. (2014), survey weights and strata are included among matching variables used in the calculation of PSM weights, but survey settings are not used in the matching process. PSM weights and survey weights are multiplied, and the resulting composite weight is applied in conjunction with survey design elements to estimate the treatment effect (DuGoff et al., 2014; Ridgeway et al., 2015).

[125] Kernel matching includes all cases. I use an Epanechinikov kernel with a bandwidth of 0.06. Alternative matching methods are tested (see appendix B Table B7).

The PSM estimate is based on the following univariate specification, weighted by the propensity score:

$$Y_{is} = \delta_{03} + \beta_3 A_s + \varepsilon_{3is} \qquad (4.3)$$

where the coefficient of interest is β_3. The treatment effect is the odds ratio between the treatment and control group weighted by the propensity score which reflects the probability of being in the treatment group. The propensity score is based on a set of *j* pre-treatment variables (same as in main model and listed in appendix B Table B3) as in the next equation:

$$A_s = \delta_{04} + \sum_{j=1}^{J} \delta_{3j} X_{jis} + \varepsilon_{4is} \qquad (4.4)$$

I use alternative methods and model specifications as robustness checks to the main analysis. Heterogeneity among sponsor-led and converter academies is investigated.

4.5 Empirical Results

4.5.1 Descriptive Analysis

The outcomes of interest are presented in Table 4.3. In the estimation sample 26% of cohort members study the EBacc, 40% take three or more STEM subjects, 66% take three or more facilitating subjects, and 40% take at least one applied subject. The subjects studied by pupils at LA maintained schools that convert to academy status in 2013/14 and 2014/15 (treatment group) is compared to those of pupils at LA maintained schools that have not converted (control group). The treatment group is separated into sponsor-led and converter academies.

Table 4.3 Outcome variables descriptive statistics

	CONTROL		TREATMENT			SPONSOR-LED			CONVERTER		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**
VARIABLES	**N**	**Mean**	**N**	**Mean**	**Diff**	**N**	**Mean**	**Diff**	**N**	**Mean**	**diff**
AGE 17											
KS4 SUBJECT CHOICE											
English Baccalaureate	2,070	0.263	454	0.257	-0.00530	161	0.187	-0.0757**	293	0.295	0.0320
		(0.475)		(0.461)	(0.0299)		(0.416)	(0.0368)		(0.477)	(0.0385)
STEM subjects	2,070	0.405	454	0.443	0.0382	161	0.317	-0.0883*	293	0.510	0.105**
		(0.530)		(0.524)	(0.0349)		(0.496)	(0.0480)		(0.523)	(0.0436)
Facilitating Subjects	2,070	0.656	454	0.658	0.00205	161	0.567	-0.0890	293	0.706	0.0502
		(0.513)		(0.500)	(0.0348)		(0.529)	(0.0704)		(0.477)	(0.0340)
Vocational subjects	2,070	0.403	454	0.379	-0.0244	161	0.453	0.0500	293	0.339	-0.0637
		(0.529)		(0.511)	(0.0394)		(0.531)	(0.0703)		(0.496)	(0.0426)

Note: The treatment group includes pupils enrolled in 2012/13 at LA maintained secondary schools that become academies in 2013/14 and 2014/15, whereas the control group includes pupils enrolled in LA maintained secondary schools that do not convert to academy status. The treatment group is composed by pupils attending prospective sponsor-led and converter academies. All statistics, except the number of observations (N) that is unweighted, reflect MCS survey settings. The quantity in parenthesis below the mean (proportion) is the standard deviation. The difference (diff) refers to the difference in means between the treatment group and the control group and the quantity in parenthesis below is the (robust) standard error clustered at the school level and reflecting MCS survey setting (University of London, Institute of Education, Centre for Longitudinal Studies, 2020b). * significant at 10% level; ** significant at 5% level; *** significant at 1% level or below.

Subject choices at secondary school are not significantly different in the treatment and the control groups. However, this changes when we look to sponsor-led and converter academies separately. Pupils at sponsor-led academies take significantly less STEM subjects and the EBacc, while pupils at converter academies take significantly more STEM subjects than their peers at LA maintained schools. Given the increased school stratification between sponsor-led and converter academies,[126] these differences support the hypothesis that pupils from different socioeconomic backgrounds are being directed to different career paths. However, possible selection issues undermine any conclusion we might draw from these comparisons. Hence, I next control for observables in estimating the academy attendance impact.

4.5.2 Logistic Regression Model

Table 4.4 estimates the effect of attending an academy on subject choice at age 14, using a logistic regression model based on the rich MCS database. In addition to the estimates of average effects for academies, a separate analysis for sponsor-led and converter academies is provided.

Table 4.4 column (1) shows that pupils in the treatment group have higher odds of taking STEM subjects, facilitating subjects, and the EBacc and lower odds of taking vocational subjects. However, these differences are only significant for science. Looking to sponsor-led and converter academies separately, curricular differences are clearer. As shown in Table 4.4 column (2), there are no significant differences between the options taken by pupils at

[126] See sub-section 3.2.1 on pupil composition change after academy conversion.

sponsor-led academies and their peers at LA maintained schools. The differences found in the descriptive analysis (fewer STEM subjects and the EBacc) disappear when controlling for background variables. On the other hand, Table 4.4 column (3) shows that pupils at converter academies are more likely to take three STEM subjects (as projected in descriptive statistics) and three facilitating subjects than their peers at LA maintained schools. These curricular differences, coupled with findings from previous studies that show pupils at converter academies generally outperforming those at LA maintained schools and at sponsor-led academies on cognitive skills (see chapter 3), suggest that pupils at converter academies have an advantage in pursuing prestigious professional careers.

Table 4.4 Impact of academy attendance on subject choice

LOGIT OUTCOME	(1) ACADEMIES Treatment odds	(2) SPONSOR-LED Treatment odds	(3) CONVERTER Treatment odds
AGE 17			
English Baccalaureate	1.075 (0.233)	1.033 (0.340)	1.097 (0.299)
STEM subjects	1.271* (0.166)	0.766 (0.157)	1.582*** (0.259)
Facilitating Subjects	1.148 (0.171)	0.738 (0.194)	1.482** (0.294)
Vocational subjects	0.881 (0.133)	0.953 (0.263)	0.832 (0.143)
Observations	**2,524**	**2,231**	**2,363**

Note: The treatment group includes pupils enrolled in 2012/13 at LA maintained secondary schools that become academies in 2013/14 and 2014/15, whereas the control group includes pupils enrolled in LA maintained secondary schools that do not convert to academy status. The treatment group is composed by pupils attending prospective sponsor-led and converter academies. Columns (2) and (3) restrict treatment group to pupils attending prospective sponsor-led academies and prospective converter academies respectively. The quantity in parenthesis below the treatment coefficient is the (robust) standard error clustered at the school level and reflecting MCS survey setting. Estimates are based on 70 imputations. The multiple imputation model uses chained equations. (University of London, Institute of Education, Centre for Longitudinal Studies, 2020a, 2020b, 2020c, 2022). * significant at 10% level; ** significant at 5% level; *** significant at 1% level or below.

4.5.3 Alternative Models

Propensity score matching at pupil level is used as an alternative approach. Table B6 in appendix B shows that estimated odds ratios using PSM are generally similar to main results.[127]

Instrumental variables are used to correct ITT estimates for non-compliance with treatment assignment. Estimates from a linear probability model[128] and a model using instrumental variables are shown in Table B6 in appendix B. The IV estimates are roughly similar to the linear probability model estimates, because of the high rate of compliance for legacy enrolled pupils.

Since main analysis is conducted using multiple imputation, a complete case analysis is provided in appendix B Table B5.[129] The odds ratios for science subjects are slightly higher than in the main model. But the odds of taking facilitating subjects at converter academies are no longer significantly different than at maintained schools. This could imply that the observed academy effect on facilitating subjects is driven by science subjects. Overall results in complete case analysis are not qualitatively different.

I also test the sensitivity of the main results to the inclusion of secondary school level variables in the model.[130] The results are presented in Table B6 column (4) in appendix B. The inclusion of these variables slightly shifts the estimate of interest by a small amount. This is expected but including these variables do not substantially change results.

---

[127] See appendix B Table B7 for alternative matching methods and Figure B1 for a comparison of propensity scores in the treatment and control groups before and after matching.

[128] Coefficients cannot be compared with odds ratios estimated in main regression.

[129] Complete cases are 68% of sample of interest.

[130] Dummy variables for faith schools, single sex schools, and grammar schools.

In brief, I have found that differences in subject choice at age 14 between pupils attending academies and those at LA maintained schools are driven by converter academies. Pupils attending converter academies are more likely to study science subjects and facilitating subjects at GCSE.

4.6 Discussion

The academies programme in England helps understanding of how school autonomy affects pupils' subject choice at secondary phase. Using the rich MCS database and legacy enrolment, I find that subject choices of comparable pupils attending academies and LA maintained schools are generally identical. The only exception is on science subjects and is driven by converter academies. Pupils at converter academies are more likely to take science subjects than their peers at LA maintained schools. They also have higher odds of studying facilitating subjects. Pupils at these schools may have a different choice set of options, expectations, and guidance (Dilnot, 2016). This may reflect school competition, school leadership, teachers, ethos, and school constraints. It might also echo what is deemed appropriate for their pupil intakes.

Interestingly, chapter 3 shows that converter academies significantly raise the decision-making skills of their pupils. This effect could be explained by a more science-based curriculum, assuming that science teachers are more likely to use learning methods that promote problem solving skills. Research on school level practices and features of academies is required to expand this supposition further (see chapter 5).

These results raise important equity concerns. In effect, academy conversion has been associated with increased stratification (Braz, 2018; Eyles et al., 2018). If converter academies have more advantaged intakes, a more academically demanding curriculum at these schools creates obvious issues for social mobility. Pupils at converter academies are more likely to study facilitating GCSE subjects, particularly STEM subjects, possibly affecting their educational transitions. These subjects might facilitate future access to university and prestigious STEM-based occupations (Dilnot, 2016) which confer a wage premium related to the shortage of quantitative skills in the labour market (Wolf, 2011). Given technology's increasing social role (OECD, 2016b), science skills are fundamental and equality of opportunity should be assured.

An important caveat in this chapter is the lack of school level data, preventing a comparison of schools in the treatment and control groups. Comparing pupils at academies with their peers at LA maintained schools is not ideal, particularly when considering sponsor-led and converter academies that followed different routes for academy conversion. Using as the control group pupils attending LA maintained schools that become academies in the following period, as implemented in chapter 5, provides a more robust identification. This methodological refinement using additional data is an important future development.

5. ACADEMIES' SCHOOL MANAGEMENT PRACTICES, POLICIES AND LEARNING ENVIRONMENT

5.1 INTRODUCTION

This chapter investigates academies' management practices, school policies, and learning environment. Previous studies show a change in pupil intakes and mixed results on exam performance after academy conversion (Andrews et al., 2017; Eyles et al., 2016b, 2017, 2018; Eyles & Machin, 2019). There is also some evidence of academisation having an impact on pupils' decision-making skills, self-esteem, and subject choice (see chapters 3 and 4). If academy status has an impact on pupils' outcomes, there must have been a change in school policies, management, and climate after conversion. That is precisely the point of giving autonomy to schools. Schools with good management practices have higher pupil achievement (Bloom et al., 2015; Leaver et al., 2022). However, an international comparison of education systems suggests that the effects of school autonomy depend on how prepared and how accountable schools are (OECD, 2016c). It is possible that academies and LA maintained schools are not that different and that academisation is mainly a school rebranding. Autonomy allows academies to experiment with managerial practices and school policies, but it is not clear if they do so (Bloom et al., 2015; Bryson et al., 2023).

This chapter intends to answer the following question: Do academies have distinctive management practices, school policies, and learning environment?

School management quality is measured by an adaptation of the World Management Survey index, following Bloom et al. (2015) and Leaver et al. (2022), This index focuses on management practices known to be important across economic sectors (Bloom et al., 2015). Other outcome variables describing school management practices include the responsibility and leadership indices that indicate the responsibility of school staff over resources and the curriculum, and school leadership on curricular development, instruction, professional development, and teacher participation. Outcomes of interest also include policies schools follow on curriculum, admissions, and assessment, namely the time pupils spend in various subjects, pupil admissions criteria, and the assessment mode frequency. Good management is expected to promote a motivating learning environment (Leaver et al., 2022). Therefore, outcome variables include indices describing the learning environment and the school climate. Learning environment indices portray the disciplinary climate, inquiry-based learning, teacher-directed instruction, and adaptive instruction. Pupil and teacher behaviour that hinders learning at school is reflected in school climate indices. The abovementioned variables are the outcomes of interest, revealing school management practices, policies, and learning environment.

I use school level variables from the Programme for International Student Assessment (PISA) to compare academies and LA maintained schools. To ensure groups are comparable, I use schools that become academies after the sample period as the control group. Subgroup analyses for sponsor-led and converter academies and for multi-academy trusts (MATs) reveal heterogeneous effects.

Results suggest academies are organised differently. Multi-academy trusts have fostered strong leadership and more structured management practices. Converter and sponsor-led academies have distinct admission criteria and face different problems. Both converter and sponsor-led academies seem to enjoy, albeit differently, a better school climate in terms of discipline and pupil behaviour. They have a diverse learning environment. Pupils at sponsor-led academies devote more time to mathematics and problem-based learning. Converter academies use instead more teacher-directed and adaptive instruction.

The structure of this chapter is as follows. Section 5.2 reviews literature on school management. Sections 5.3 and 5.4 present data and the econometric model used to answer the research question. Section 5.5 discusses empirical findings. Section 5.6 offers some closing comments.

5.2 School Management Practices

This section reviews literature on school management, highlighting its importance for school policies, learning environment, and educational outcomes. A particular attention is given to evidence regarding management and leadership at autonomous schools.

5.2.1 Overview

Bloom et al. (2015), using the World Management Survey (WMS) index for schools based on 20 basic management practice measures derived from interviews with secondary school headteachers, show that managerial practices vary significantly across and within countries and that schools' adoption of more structured managerial practices is limited, especially in people management. Large urban schools and those with a lower pupil-teacher ratio have higher management scores (Bloom et al., 2015).[131] Economies of scale and more resources could thus give some schools a managerial advantage (Bloom et al., 2015). Furthermore, Bloom et al. (2015) estimate school level regressions and find that school management quality is positively correlated with pupil achievement.

Similarly, Leaver et al. (2022) use a PISA adaptation of the WMS management index and find a positive relationship between management quality and test scores at PISA 2012. Good management is associated with a decrease in teacher shortages, an increase in teacher motivation and effort, and an increase in household effort (Leaver et al., 2022). Leaver et al. (2022) argue that good management practices increase the effort of parents and pupils by promoting a stimulating learning environment.

Focusing on people management, Bryson et al. (2023) use Workplace Employment Relations Surveys and show that intensive use of human resource management practices is correlated with higher labour productivity, quality of provision, and financial performance in schools. It is not, however, associated with higher pupil attainment (Bryson et al., 2023). Interestingly, headteachers' evaluation of schools, that is actually associated with pupil attainment, increases with the use of human resource management practices (Bryson et al., 2023). Bryson et al. (2023) believe that the benefits of human resource management might take time to materialise.

[131] Female headteachers are also more likely to implement more structured managerial processes (Bloom et al., 2015).

5.2.2 Autonomous Schools

Bloom et al. (2015) show that autonomous state schools have higher management quality, largely due to leadership and governance. The headteacher having a clear and coherent long-term school strategy and being accountable for pupil performance to institutional stakeholders accounts for most of the difference in management scores between autonomous and other state schools (Bloom et al., 2015). Bloom et al. (2015) estimate this relationship for the UK and find a positive but not significant difference in management scores for autonomous state schools.[132] However, results for England are of limited interest because of the low response rate (Bloom et al., 2015). Besides, Bloom et al. (2015) group academies together with foundation schools and voluntary-aided schools, making it difficult to draw conclusions on academisation.

Bryson et al. (2023) reveal that human resource management practices at academies are similar to those observed at LA maintained schools. Nevertheless, improvement plans at sponsor-led academies have often overhauled school workforce (Martindale, 2019). ponsor-led academies employ a higher percentage of teachers without Qualified Teacher Status (Martindale, 2019). The prospect of academisation induces older and underachieving teachers to leave sponsor-led academies but there is an increase in newly hired teachers after conversion (Duchini et al., 2023). Moreover, Duchini et al. (2023) show that sponsor-led academies restructure

[132] The same happens in other OECD countries, with the exception of Sweden where difference is significant (Bloom et al., 2015).

teachers' pay leading to higher pay dispersion among equally experienced teachers. A case study reveals that at a sponsor-led academy teachers' performance management is centred on pupil attainment (J. P. Morris, 2020). On the other hand, Hicks (2017) reveals a reduction in average teacher qualification at converter academies.

Sponsor-led academies experience substantial management restructuring (Bertoni et al., 2021). Most sponsor-led academies change their senior leadership and introduce fundamental operating changes (Cirin, 2014).[133] This is observed by Morris (2020). Eyles et al. (2016a) and Duchini et al. (2023) also note a high headteacher turnover at conversion year in sponsor-led academies. According to Telhaj et al. (2022), these academies offer headteachers a wage premium.[134] New headteachers tend to have a higher salary and come from an outstanding school (Duchini et al., 2023).

Academies that joined a MAT are also likely to change school leadership (Cirin, 2017).[135] Leadership is really a key variable in successful school networks (Greany et al., 2023).[136] Professional school governance focused on results is an emerging feature of MATs (Healey, 2022). The proportion of teachers in leadership roles is indeed higher at these schools

---

[133] The appointment of a new headteacher is significant in the realization of the proposed ethos and vision for the school (Gibson, 2015).

[134] Sponsor-led academies have generally less experienced headteachers (Telhaj et al., 2022).

[135] School leaders' alignment with the chain's vision is important for MATs (Hetherington & Forrester, 2023; Keddie, 2019).

[136] Culpin and Male (2022) consider MATs "loosely coupled organisations" that require leadership competencies beyond what is outlined in the National Professional Qualification for Executive Leadership, introduced to support professional development of school leaders.

(Davies et al., 2021).[137] Despite little evidence of economies of scale (Davies et al., 2021), school leaders believe MAT's structure facilitates collaboration and efficiency through improved procurement, back-office, and monitoring, professional development, and pooling resources, teachers, and support staff (Cirin, 2017; Simon et al., 2021).

A few qualitative studies suggest that MATs have developed new management practices. A case study shows schools within MATs have common corporate policies and procedures (Gibson, 2015). Another case study reveals that MAT schools are expected to adhere to standardised programmes and monitoring (Keddie, 2019). In effect, a mixed-methods study shows that schools that joined a MAT operate in a more standardised environment (Wiborg et al., 2018). A chain-wide standardised approach is indeed recommended by the DfE as good practice (Culpin & Male, 2022).[138] However, a comparative case study conducted by Salokangas and Chapman (2014) finds that while some MATs focus on standardised procedures and have centralised and prescriptive performance management, accountability, and support mechanisms, in other MATs prevails a collaborative culture.[139] Effectively, Neri et al. (2022), using data from British Educational Suppliers Association linked to NPD, show that MATs where school leaders are aligned

---

[137] Academies have reduced the number of teachers and teaching assistants per pupil but have higher percentage of teachers in leadership positions (Martindale, 2022). In effect, after conversion expenditure on non-teaching personnel increases (Duchini et al., 2023).

[138] Effective MATs share successful strategies to support school improvement (Hutchings & Francis, 2018). Improvement plans often include tightening processes and looking for administrative efficiencies (Simon et al., 2021).

[139] For example, in the MAT studied by Constantinides (2022) accountability is centred on internal school evaluation and peer reviews. The use of systematically collected data is encouraged to inform decision-making at school level, allowing headteachers to adapt MAT policies locally, but underperforming schools have less autonomy (Constantinides, 2022).

with the chain board and those that are closer to the value-added frontier are more decentralised.

In brief, academies might have slightly more structured managerial practices. However, human resource management at academies seems to be like maintained schools, despite evidence of a school workforce overhaul. Sponsor-led academies have undergone extensive reorganisation. The emergence of professional leadership and standardised practices within MATs imply changes in school management.

Academy conversion might alter management practices, which in turn shape school policies. Disciplinary policies are a good example. Exclusion rates are higher at academies (Machin & Sandi, 2020). Using the NPD, Machin and Sandi (2020) show that following academy conversion schools exclude more pupils, but the impact is much larger for pre-2010 academies. Higher exclusion rates reflect the rigorously enforced disciplinary policy of early academies (Machin & Sandi, 2020). Some of these schools have adopted a "No Excuses" culture (Machin & Sandi, 2020). Another example is SEN provision. Liu et al. (2020), also using NPD, show that sponsor-led academies are more likely to change the SEN status of pupils. Basically, pupils at sponsor-led academies are less likely to receive SEN support (Liu et al., 2020). On other areas the change in school policy after academy conversion is not clear. For instance, West (2014) shows that converter academies tend to keep predecessor school's admissions criteria. Further, looking into headteacher pay, Telhaj et al. (2022) show that differentials between academies and LA maintained schools are largely attributable to teacher and school characteristics.[140] Academy conversion seems therefore to have a second order role in some areas.

This chapter investigates school polices at academies, regarding curriculum, admissions, and assessment, in comparison with LA maintained schools. Differences possibly reflect dissimilar management practices. The chapter also examines how academy conversion is affecting school management and learning environment.

5.3 DATA

The Programme for International Student Assessment (PISA) measures pupils' achievement in science, reading, and mathematics at age 15 (OECD, 2017d, 2017a). PISA surveys take place every three years (OECD, 2017d). This chapter uses the sixth survey[141] that is conducted in 2015 across 72 countries. The sample of interest is restricted to England, comprising 5,194 pupils and 206 schools.[142] Information on pupils and schools is gathered from pupils and headteachers (OECD, 2017a).[143] This includes information on school composition, resources, staffing, decision-making, curriculum, and teaching practices. Data is weighted to reflect PISA sampling design and provide a representative sample of the population (Jerrim et al., 2017).

---

140 Headteachers at academies have on average a higher salary (Telhaj et al., 2022).
141 OECD (2016a).
142 Schools in England are randomly selected, stratified by school type, location, gender, and GCSE performance (Jerrim & Shure, 2016). Within each school a random sample of 30 pupils is selected (Jerrim & Shure, 2016).
143 170 headteachers have answered the school questionnaire in England.

To identify academy conversion year and the characteristics of predecessor schools, I use a 2014/15 list of schools from Edubase and construct a comprehensive school level database with publicly available information.[144] [145] It identifies schools that become academies, academy route, academy trusts, and timing of conversion.[146] This school level database has been matched to PISA 2015 by the DfE.[147]

The school database allows me to follow the identification strategy of Eyles and Machin (2019), Eyles et al. (2016b), Eyles et al. (2017), Eyles et al. (2018) and Andrews et al. (2017) who define the treatment group as pupils attending academies and the control group as pupils attending LA maintained schools that become academies in the following period. It also allows a comparison of schools in 2011/12 (pre-conversion) and school level multivariate regressions.

---

[144] I use the Consistent School Database (CSD) to link information from several sources. The CSD enables matching of school records between years when the school identifier (URN) changes. I have checked online information on individual schools at Get Information About Schools from DfE to minimise missing data and matching inconsistencies. The Academies Management Information Data (May 2018) from DfE identifies conversion year, academy route and academy trusts in 2018.

[145] School mergers and splits between 2011 and 2017 are identified. These schools must be treated with care, since they have undergone changes that go beyond academy conversion. Five schools in estimation sample have undergone a merger or a split in the period. I impute hypothetical values (weighted average based on the number of pupils) for one sample school that was merged between 2012 and 2015.

[146] Since some academies have an academy as predecessor school, conversion year refers to the first academy conversion or the first time a school acquired academy status. Following Andrews et al. (2017), academies that are open in year X are considered to be in operation from the following academic year (X+1).

[147] This chapter was produced using statistical data accessed via the Secure Research Service, part of the Office for National Statistics (ONS). The use of this data, which is Crown Copyright, in this chapter does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This chapter uses research datasets which may not exactly reproduce National Statistics aggregates.

5.3.1 Outcomes of Interest

The outcomes of interest are variables describing (1) school management practices, (2) school policies, and (3) learning environment.[148] Information is drawn from PISA 2015 student and school questionnaires.

The variables used to study (1) school management practices include responsibility indices, leadership indices, and a variation of the World Management Survey (WMS) management index.

Responsibility indices indicate the relative level of responsibility of school staff in allocating resources and over the curriculum. Responsibilities comprise selecting and firing teachers, setting teachers' salary and the school budget, disciplinary and assessment policies, admissions, and choosing textbooks, course content, and school offer (OECD, 2017d). Higher values on the indices denote greater school responsibility.[149] These measures are standardised to mean zero and standard deviation one.

School leadership indices regarding curricular development, instruction, professional development, and teacher participation are based on headteachers' self-reported frequency of relevant activities (OECD, 2017d). Curricular development refers to using pupil performance and teachers' professional development to attain school goals. Instructional leadership comprises promoting evidence-based teaching practices and commending teachers for pupils' learning. Professional development includes discussing

---

[148] A detailed list of outcome variables is presented in Table C1 in appendix C.

[149] The indices compare the level of responsibility of the school governing board, the headteacher and teachers to that of the LA and the government. Headteachers may identify several entities as taking a role in each task.

and solving classroom problems with teachers. Teacher participation entails involvement in school decision-making and a culture of continuous improvement. These indices are standardised to mean zero and standard deviation one.[150]

The school management index is an adaptation of the WMS management index for PISA 2015, following Bloom et al. (2015) and the work of Leaver et al. (2022) for PISA 2012. The WMS is used as a benchmark, focusing on management practices known to be important in other sectors (Bloom et al., 2015; Leaver et al., 2022). 46 questions in the PISA 2015 school questionnaire are identified in relation with 14 WMS topics.[151] More structured practices are assigned higher scores, and each topic score is standardised. The overall management index is the average of topic scores.[152] Separate indices for operations and people management are also presented.

The variables used to examine (2) school policies include pupil learning time, admissions criteria, and assessment mode. Pupils report how many lessons in science, mathematics, and English they are required to attend per week and how much time they spend per lesson.[153] Total learning time is derived and presented in hours. Information on admissions is provided by headteachers. Binary variables identify admissions criteria, including prior ability, feeder schools, faith, aptitude for school specialism, siblings' attendance, and catchment area. Headteachers also state school assessment

---

[150] Positive scores mean headteachers report it more often than the English average.

[151] See Table C2 in appendix C for a detailed list of the variables and scores applied in the construction of the management index.

[152] Schools with missing data in more than one question are dropped.

[153] Learning time is calculated by multiplying the reported number of minutes on average in class by the number of class periods per week (OECD, 2017d).

policy. They report the frequency of using mandatory and non-mandatory standardised tests, teacher developed tests, and teacher ratings.[154]

The variables used to observe (3) learning environment include learning environment indices and school climate indices.

Learning environment indices are drawn from pupils' reports regarding science lessons (OECD, 2016c)[155] and reveal disciplinary climate, inquiry-based learning, teacher-directed instruction, and adaptive instruction (OECD, 2017d). Disciplinary climate refers the efficacy of classroom management, allowing pupils to concentrate on tasks without noise or distractions. Inquiry-based learning includes investigations where pupils explain and test their ideas, debate, and draw conclusions. Teacher-directed instruction involves explanation, demonstration, and class discussion. Adaptive instruction refers to teachers providing individual help and changing lessons to meet pupils' needs. These indices are standardised to mean zero and standard deviation one.[156]

School climate is affected by the behaviour of teachers and pupils. Behaviour that hinders learning at school is flagged by headteachers (OECD, 2017d). Pupils' behaviour hindering learning includes truancy, skipping classes, lack of respect for teachers, alcohol or drug use, and bullying. Teachers' behaviour hindering learning includes not meeting pupils' needs,

[154] Binary variables identify schools where methods are more frequently used.

[155] Pupils in the same school report a great variety of teaching methods in their science classes, which likely extends to other subjects (OECD, 2017c). The representation of the school-wide ethos is not straightforward. Besides, pupils' reports on teaching may not be always technically accurate (OECD, 2016c).

[156] Positive scores denote pupils respond more positively than the English average.

absenteeism, resisting change, being too strict, and not being well prepared. Indices are standardised to mean zero and standard deviation one.

5.3.2 Pre-treatment Variables

Comparing academies with LA maintained schools assumes they are similar prior to conversion. The school level data matched to PISA 2015 includes information from 2012 Performance Tables, 2012 School Workforce and Ofsted, allowing a comparison of LA maintained schools that become academies in the sample period (between 2012/13 and 2014/15) and those that become academies afterwards (2015/16 and 2016/17). I also use this data in multivariate regressions, controlling for pre-treatment school characteristics, when comparing academies with LA maintained schools.

Performance Tables include school level information on GCSE performance (percentage achieving five good grades including English and mathematics, capped average score, capped value added) and pupil characteristics (number of pupils, percentage of boys, percentage of pupils eligible for FSM, percentage of pupils with English as first language, percentage of pupils with SEN statement, KS2 average score). The School Workforce indicates the pupil-teacher ratio at each school. A list of Ofsted inspections between September 2010 and March 2014 is used to derive the last Ofsted report in 2012.

5.3.3 Descriptive Statistics

Looking into school type in 2015 (see appendix C Table C3), academies have higher management quality, more leadership in curricular development, and increased responsibility over resources, particularly those that have joined a MAT. In effect, better leadership seems to be a feature of MATs. Sponsor-led academies have strong leadership and take more responsibility over the curriculum. Converter academies take more responsibility over resources.

Focusing on school policies, we find no meaningful difference between academies and LA maintained schools in 2015, with the only exception being mandatory standardised tests are used less often at academies. Converter academies are less likely to select pupils based on school specialism. Sponsor-led academies devote more time to mathematics and English.

On school climate, we do not find significant differences between academies and LA maintained schools. However, converter academies have a better classroom disciplinary climate and fewer issues with pupil behaviour.

Differences might reflect the characteristics of predecessor schools and the academy route. Since I have information on schools in 2012 and the survey is taken in 2015, I compare schools that become academies between 2012/13 and 2014/15 (treatment group) and schools that become academies in the following period.

Table 5.1 presents basic information on the distribution of pupils. The 2,969 pupils in the sample are distributed across 116 schools.[157] The sample of interest focuses on LA maintained schools and schools that become academies after 2011/12, corresponding to 57% of PISA cohort. There are 1,719 (33%) cohort members in pre-2012 academies. The remainder 9% are in free schools and independent schools.[158] This is consistent with the national distribution. At secondary phase in 2011/12 there are 878 academies,

[157] Average of 26 pupils per school.
[158] 24 pupils are in a school for which there is no information regarding school type.

including 629 converter and 249 sponsor-led academies (Eyles et al., 2018), representing 36% of secondary schools (National Audit Office, 2018). Nationwide LA maintained schools in 2011/12 have on average 18% of pupils eligible for FSM, 14% with English as additional language and 60% achieving five good GCSEs (Department for Education, 2012, 2013a). This compares well with our sample schools that in 2012 have 23% of FSM pupils, 83% of pupils with English as first language, and 64% achieving five good GCSEs.

Table 5.1 Sample distribution per school type

	Number of pupils	Number of schools
Schools that convert between 2012/13 and 2014/15 (TREATMENT GROUP)		
Converter academies	495	20
Sponsor-led academies	380	15
SAT	287	12
MAT	588	23
TOTAL	**875**	**35**
Schools that convert in 2015/16 and 2016/17 (CONTROL GROUP)		
Prospective converter academies	180	7
Prospective sponsor-led academies	101	4
Prospective SAT	0	0
Prospective MAT	281	11
TOTAL	**281**	**11**
LA maintained schools in 2014/15	**2,094**	**81**

Note: Statistics do not reflect PISA 2015 survey settings (unweighted measures). The sample is restricted to England (OECD, 2016a).

In the period of interest (between 2012/13 and 2014/15), 616 LA maintained schools convert – 485 converter and 131 sponsor-led academies (Eyles et al., 2018). The treatment group, drawn from this set, comprises 875 pupils enrolled in 35 academies. In 2014/15, 2,094 pupils are enrolled in 81 maintained schools, 11 of which become academies in 2015/16 and 2016/17 (control group).

The treatment group includes 20 converter academies and 15 sponsor-led academies. In 2012, 18% of sample pupils at prospective converter academies are eligible for FSM and 68% achieve five good GCSEs. 79% of these schools are rated “Good” or “Outstanding” by Ofsted. In contrast, 31% of pupils attending prospective sponsor-led academies are eligible for FSM and only 47% achieve good GCSEs. 38% of these schools are deemed “Inadequate” by Ofsted. Comparing these unweighted figures with national statistics gives a sense of the population. Schools that become converter academies are more advantaged (below average FSM rates) and high performing (Bolton, 2015). 86% of these schools are rated “Good” or “Outstanding” by Ofsted (National Audit Office, 2018). Moreover, 45% of schools that become sponsor-led academies nationwide in this period are deemed “Inadequate” by Ofsted, on average have disadvantaged intakes (24% FSM pupils) and underperform at GCSE (Hatton et al., 2019). Considering the sample size, the treatment group seem to roughly represent schools becoming academies between 2012/13 and 2014/15.

Table 5.2 presents descriptive statistics on the outcomes of interest for schools that become academies between 2012/13 and 2014/15, for schools that become academies in 2015/16 and 2016/17, and for all maintained schools in 2014/15.

Table 5.2 Descriptive statistics by school type

OUTCOME VARIABLES	TREATMENT (1) N	(2) Mean	CONTROL (3) N	(4) Mean	LA MAINTAINED (5) N	(6) Mean
MANAGEMENT						
Responsibility index						
Curriculum	656	0.013	281	0.612	1,826	0.003
		(1.038)		(0.000)		(0.994)
Resources	656	-0.058	281	0.082	1,826	-0.222
		(1.039)		(1.003)		(1.071)
Leadership index						
Curricular Development	608	0.115	281	-0.174	1,798	-0.124
		(0.941)		(1.175)		(1.037)
Instructional	608	0.202	281	-0.007	1,748	-0.053
		(1.015)		(1.280)		(0.962)
Professional Development	581	-0.043	281	0.073	1,798	0.034
		(1.285)		(0.851)		(0.880)
Teacher Participation	608	0.085	281	-0.324	1,798	-0.095
		(1.218)		(0.765)		(0.901)
WMS Management index						
Overall	452	0.224	281	-0.050	1,566	-0.049
		(0.884)		(1.012)		(0.971)
Operations	452	0.200	281	-0.135	1,566	-0.082
		(0.900)		(1.052)		(0.977)
People	452	0.227	281	0.344	1,566	0.112
		(0.825)		(0.923)		(1.013)

OUTCOME VARIABLES	TREATMENT		CONTROL		LA MAINTAINED	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	N	**Mean**	**N**	**Mean**	**N**	**Mean**
POLICIES						
Learning Time (hours)						
Mathematics	844	4.165	249	3.868	1,923	3.929
		(1.260)		(0.773)		(0.968)
English	843	4.304	249	4.329	1,923	4.083
		(1.543)		(2.083)		(1.566)
Science	842	4.861	249	4.768	1,919	4.795
		(1.815)		(1.121)		(1.543)
Total	789	26.030	235	26.477	1,811	26.095
		(4.952)		(5.714)		(5.079)
Admissions						
Academic Performance	629	0.218	281	0.214	1,798	0.194
		(0.413)		(0.410)		(0.396)
Feeder Schools	629	0.258	281	0.390	1,779	0.303
		(0.437)		(0.488)		(0.460)
School Philosophy or Religion	605	0.433	281	0.237	1,798	0.287
		(0.496)		(0.425)		(0.452)
School Specialism	629	0.370	281	0.644	1,748	0.463
		(0.483)		(0.479)		(0.499)
Former Pupils and Siblings	629	0.715	281	0.881	1,798	0.772
		(0.451)		(0.324)		(0.419)
Catchment Area	629	0.838	281	0.881	1,798	0.869
		(0.368)		(0.324)		(0.337)
Pupil Assessment						
Mandatory Standardised Tests	602	0.201	281	0.217	1,798	0.317
(More than 2 times a year)		(0.401)		(0.412)		(0.465)
Non-mandatory Standardised	602	0.756	281	0.873	1,773	0.669
Tests (At least once a year)		(0.430)		(0.333)		(0.471)
Teacher-developed Tests	602	0.621	281	0.606	1,798	0.514
(At least monthly)		(0.485)		(0.489)		(0.500)
Teachers Ratings	602	0.489	281	0.801	1,798	0.562
(At least monthly)		(0.500)		(0.400)		(0.496)

OUTCOME VARIABLES	TREATMENT (1) N	(2) Mean	CONTROL (3) N	(4) Mean	LA MAINTAINED (5) N	(6) Mean
LEARNING ENVIRONMENT						
Learning Environment index						
Disciplinary Climate	835	-0.066	242	-0.248	1,910	-0.101
		(1.013)		(1.010)		(1.028)
Inquiry-based Learning	825	0.022	237	-0.233	1,886	-0.002
		(1.079)		(1.099)		(1.029)
Teacher-directed Instruction	823	-0.044	232	-0.191	1,875	-0.025
		(0.965)		(1.028)		(1.046)
Adaptive Instruction	805	-0.023	224	-0.211	1,834	-0.027
		(0.997)		(1.100)		(1.019)
School Climate index						
Pupil behaviour	602	0.315	281	0.786	1,798	0.366
		(0.755)		(0.723)		(0.826)
Teacher behaviour	602	0.422	281	0.443	1,798	0.179
		(0.934)		(0.700)		(0.943)

Note: The treatment group includes pupils enrolled in secondary schools that become academies between 2012/13 and 2014/15, The control group includes pupils enrolled in LA maintained secondary schools that convert to academy status in 2015/16 and 2016/17. The LA maintained group includes all maintained schools in 2014/15. The quantity in parenthesis below the mean (or proportion in the case of categorical variables) is the standard deviation. All statistics, except the number of observations (N) that is unweighted, reflect PISA survey settings. Estimates reflect the number of pupils in schools where variables are observed. The sample is restricted to England (OECD, 2016a).

5.4 METHODOLOGY

This chapter ascertains the distinctive features of secondary schools that become academies, namely their management practices, policies, and learning environment. The differences between schools are captured through several OLS regressions,[159] identifying the likelihood of a cohort member attending a school with a particular characteristic if enrolled in an academy. Having one single cohort at PISA, I am not able to control for temporal differences in school characteristics, focusing instead on the differences between the treatment and the control groups at a given time.

Linear regression (OLS) compares pupils at LA maintained schools that become academies to those at LA maintained schools that do not convert. Causality requires that observed differences be attributable to academy conversion alone, ruling out confounding factors (conditional independence assumption). I control for the pre-treatment school level variables presented in the previous section (see section 5.3.2), attempting an identification strategy based on selection on observables. This implies the observation of all relevant variables. However, the decision to become an academy reflects school dimensions that we are not able to control for and could affect other school characteristics (Andrews et al., 2017). The selection bias, implied in the conversion process, might significantly affect the comparison between academies and LA maintained schools. The estimates are thus interpreted descriptively.

---

[159] Linear probability model when the dependant variable is binary.

The multivariate OLS model is based on the following linear relationship for pupil *i* in school *s*:

$$C_{is} = \alpha_1 + \beta_1 A_s + \sum_{j=1}^{J} \delta_j X_{js} + \varepsilon_{1is} \qquad (5.1)$$

where C_{is} are school variables describing management practices, policies, and learning environment, X_{js} is a set of *j* pre-treatment school characteristics, α_1 is an intercept term and ε_{1is} is an error term. X_{js} includes number of pupils, percentage of boys, of English natives, of SEN and of FSM pupils, KS2 average score, capped average point score, capped value added, percentage obtaining five good GCSEs, pupil/teacher ratio, and the last Ofsted report in 2012 (see detailed list of variables in appendix C Table C4). *A* is a dummy variable equal to 1 for academy attendance (treatment group). β_1 is the coefficient of interest. The error term is assumed to be uncorrelated with treatment, but a degree of autocorrelation between pupils within schools is allowed.

Acknowledging the limitations of the previous model, the main analysis explores the staggered nature of the academy programme. Following Bertoni et al. (2021), Eyles and Machin (2019), Eyles et al. (2017), and Eyles et al. (2018), I use as control group schools that become academies in the following period,[160] assuming that the timing of conversion is as good as random and expecting that prospective academies have common unobservable characteristics to the treatment group (schools that convert between 2012/13 and 2014/15).[161] Despite treatment not being randomly assigned, the

---

[160] After school year 2014/15 when the PISA 2015 survey takes place.

[161] In England, 616 secondary schools (485 converter and 131 sponsor-led) become academies between 2012/13 and 2014/15 (Eyles et al., 2018).

difference between academies that convert before and after a given year can be thought of as being as good as randomly assigned (Angrist & Pischke, 2009). Abdulkadiroğlu et al. (2016) use the same assumption for charter school takeovers. This accounts for unobserved heterogeneity between schools that convert and those that do not, ensuring internal validity.

Using the National Pupil Database (NPD), Eyles and Machin (2019), Eyles et al. (2016b), Eyles et al. (2017), Eyles et al. (2018) and Andrews et al. (2017) show that the timing of conversion is unrelated to school characteristics and that there is no significant pre-treatment difference between current and future academies which have similar pre-conversion trends. Despite the small sample size, I verify this. In fact, policy changes affecting the academy conversion process and eligibility could be an issue. For example, there are no single-academy trusts (SAT) in the control group due to a policy shift in 2016 (Department for Education, 2016a). However, the comparison of pre-treatment characteristics between treatment and control groups is reassuring. Prior to conversion, schools that become academies between 2012/13 and 2014/15 are similar to schools that convert in 2015/16 and 2016/17 (as discussed in next section, see appendix C Table C4).[162] This similarity supports the randomness assumption and adds credibility to the identification strategy.

The main model is based on the following univariate relationship for pupil *i* in school *s*:

$$C_{is} = \alpha_2 + \beta_2 A_s + \varepsilon_{2is} \qquad (5.2)$$

---

[162] The schools that become academies in 2017/18 and 2018/19 are a more heterogeneous group and a less credible counterfactual.

where C_{is} are school variables describing management practices, policies, and learning environment, α_2 is an intercept term, and ε_{2is} is an error term. *A* is a dummy variable equal to 1 for academy attendance (treatment group). β_2 is the coefficient of interest. Separate subgroup analyses for sponsor-led and converter academies and for MATs reveal heterogeneous effects.

5.5 EMPIRICAL RESULTS

Schools that become academies between 2012/13 and 2014/15 are compared to those that become academies in 2015/16 and 2016/17 and with all LA maintained schools, focusing on management practices, school policies, and learning environment.[163]

Table 5.3 shows school level differences between academies and LA maintained schools, controlling for the characteristics of schools in 2012.[164] Considering that predecessor schools may differ beyond observed variables,[165] these differences are interpreted descriptively.

Assuming timing of academy conversion is as if randomly assigned, schools that become academies in the following years provide a more appropriate control group. In fact, when I compare pre-treatment characteristics of schools that convert between 2012/13 and 2014/15 with

[163] The sample is weighted according to PISA 2015 survey design. Estimates are based on pupil level regressions and reflect the percentage of pupils in schools where school level variables are observed.

[164] Estimates are based on complete case analysis. There are 97% complete cases (2,866 pupils). Missing data affects 10 explanatory variables (four schools, 103 pupils).

[165] See appendix C Table C4 for a full list of the explanatory variables used.

those of schools that become academies in 2015/16 and 2016/17, I do not find significant differences.[166] The only exception is the large number of sponsor-led academies having a bad Ofsted report.[167] The observed similarity in 2012 across the two groups reinforces the credibility of the randomness assumption and supports the identification strategy as approximating a causal estimate. Table 5.4 compares outcome variables in current and future academies.

An alternative control group including schools that become academies in 2017/18 and 2018/19 is used in appendix C Table C5. This is a more heterogeneous group of schools and a less credible counterfactual. Most results are non-significant or reflect the main analysis. To clarify main results, additional school features are considered in appendix C Table C6 and Table C7, including variables used in the construction of PISA indices. A separate analysis for converter and sponsor-led academies and those that have joined a MAT is provided.[168]

---

[166] At 5% significance level, see appendix C Table C4.

[167] This is hardly surprising since schools rated "Inadequate" prior to 2012 would have been forced to conversion in the sample period. A similar report post-2012 may have triggered the same intervention post-2015.

[168] In the sample there are no schools that become standalone academies in 2015/16 and 2016/17. In effect, only a few schools become academies as a single-academy trust (SAT) after 2014/15, reflecting a policy shift in 2016 (Department for Education, 2016a). In 2016/17 more than 95% of conversions join a MAT (Neri et al., 2022). Therefore, subgroup analysis focuses on MATs. Multivariate regressions suggest SATs are very similar to LA maintained schools. SATs, however, take less responsibility over resources and the curriculum and use non-mandatory standardised tests more often (see Table 5.3).

Table 5.3 Multivariate comparison of academies with all LA maintained schools

OUTCOME VARIABLES	ACADEMIES		CONVERTER		SPONSOR-LED		SAT		MAT	
	(1) N	(2) Treatment	(3) N	(4) Treatment	(5) N	(6) Treatment	(7) N	(8) Treatment	(9) N	(10) Treatment
MANAGEMENT										
Responsibility index										
Curriculum	2,430	0.015	2,201	-0.308	2,027	0.436	2,043	-0.746**	2,185	0.388
		(0.262)		(0.285)		(0.388)		(0.345)		(0.306)
Resources	2,430	0.048	2,201	-0.056	2,027	0.239	2,043	-0.652*	2,185	0.515
		(0.279)		(0.279)		(0.487)		(0.350)		(0.331)
Leadership index										
Curricular Development	2,354	0.289	2,146	0.213	1,978	0.429	1,988	-0.464	2,136	0.778***
		(0.234)		(0.289)		(0.299)		(0.308)		(0.227)
Instructional	2,304	0.360	2,096	0.342	1,928	0.442	1,938	-0.387	2,086	0.891***
		(0.251)		(0.327)		(0.287)		(0.285)		(0.236)
Professional Development	2,327	0.196	2,119	0.277	1,978	0.205	1,988	0.151	2,109	0.284
		(0.213)		(0.296)		(0.282)		(0.294)		(0.266)
Teacher Participation	2,354	0.277	2,146	0.534	1,978	-0.176	1,988	-0.136	2,136	0.544**
		(0.256)		(0.364)		(0.283)		(0.436)		(0.231)
WMS Management index										
Overall	1,966	0.174	1,808	0.304	1,696	-0.279	1,705	-0.115	1,799	0.293
		(0.201)		(0.245)		(0.367)		(0.375)		(0.217)
Operations	1,966	0.185	1,808	0.324	1,696	-0.270	1,705	0.019	1,799	0.263
		(0.204)		(0.231)		(0.386)		(0.378)		(0.226)
People	1,966	0.048	1,808	0.078	1,696	-0.193	1,705	-0.647	1,799	0.291
		(0.244)		(0.324)		(0.376)		(0.419)		(0.230)

OUTCOME VARIABLES	ACADEMIES		CONVERTER		SPONSOR-LED		SAT		MAT	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	N	Treatment	N	Treatment	N	Treatment	N	Treatment	N	Treatment
POLICIES										
Learning Time (hours)										
Mathematics	2,673	0.179	2,359	0.110	2,196	0.547***	2,158	0.208	2,397	0.316**
		(0.110)		(0.143)		(0.184)		(0.226)		(0.124)
English	2,672	0.195	2,359	0.076	2,195	0.638***	2,158	0.356	2,396	0.214
		(0.129)		(0.155)		(0.200)		(0.220)		(0.159)
Science	2,667	0.120	2,354	0.014	2,191	0.618*	2,153	0.130	2,392	0.212
		(0.176)		(0.201)		(0.332)		(0.251)		(0.221)
Total	2,515	-0.189	2,227	0.201	2,064	-0.717*	2,038	-0.329	2,253	-0.063
		(0.315)		(0.377)		(0.381)		(0.424)		(0.421)
Admissions										
Academic Performance	2,375	0.018	2,146	0.001	1,999	0.023	1,988	-0.111	2,157	0.115
		(0.111)		(0.160)		(0.127)		(0.164)		(0.139)
Feeder Schools	2,356	-0.142	2,127	-0.160	1,980	-0.109	1,969	-0.138	2,138	-0.124
		(0.116)		(0.165)		(0.150)		(0.217)		(0.136)
School Philosophy or Religion	2,375	0.022	2,146	-0.187	1,999	0.289	1,988	-0.230	2,157	0.152
		(0.124)		(0.148)		(0.241)		(0.172)		(0.155)
School Specialism	2,325	-0.187	2,096	-0.132	1,949	-0.264*	1,938	0.022	2,107	-0.282*
		(0.134)		(0.173)		(0.158)		(0.225)		(0.153)
Former Pupils and Siblings	2,375	-0.067	2,146	-0.120	1,999	-0.089	1,988	-0.072	2,157	-0.122
		(0.110)		(0.154)		(0.188)		(0.234)		(0.124)
Catchment Area	2,375	-0.033	2,146	-0.025	1,999	0.027	1,988	0.004	2,157	-0.033
		(0.091)		(0.097)		(0.137)		(0.102)		(0.122)
Pupil Assessment										
Mandatory Standardised Tests	2,348	-0.094	2,146	-0.084	1,972	-0.097	1,988	-0.095	2,130	-0.121
(More than 2 times a year)		(0.105)		(0.144)		(0.210)		(0.191)		(0.140)
Non-mandatory Standardised	2,323	0.100	2,121	0.151	1,947	0.019	1,963	0.337*	2,105	-0.029
Tests (At least once a year)		(0.139)		(0.160)		(0.252)		(0.176)		(0.179)
Teacher-developed Tests	2,348	0.190*	2,146	0.272**	1,972	0.062	1,988	0.259	2,130	0.186
(At least monthly)		(0.102)		(0.128)		(0.169)		(0.166)		(0.134)
Teachers Ratings	2,348	-0.027	2,146	0.135	1,972	-0.213	1,988	-0.013	2,130	-0.005
(At least monthly)		(0.134)		(0.141)		(0.295)		(0.225)		(0.156)

OUTCOME VARIABLES	ACADEMIES (1) N	ACADEMIES (2) Treatment	CONVERTER (3) N	CONVERTER (4) Treatment	SPONSOR-LED (5) N	SPONSOR-LED (6) Treatment	SAT (7) N	SAT (8) Treatment	MAT (9) N	MAT (10) Treatment
LEARNING ENVIRONMENT										
Learning Environment index										
Disciplinary Climate	2,649	0.100	2,339	0.154*	2,176	-0.009	2,138	0.125	2,377	0.100
		(0.066)		(0.089)		(0.115)		(0.097)		(0.094)
Inquiry-based Learning	2,618	0.063	2,312	0.149*	2,151	-0.031	2,111	0.138	2,352	0.048
		(0.065)		(0.082)		(0.106)		(0.114)		(0.073)
Teacher-directed Instruction	2,605	-0.018	2,298	0.035	2,140	-0.086	2,100	0.060	2,338	-0.044
		(0.049)		(0.058)		(0.109)		(0.067)		(0.071)
Adaptive Instruction	2,548	0.009	2,252	-0.003	2,090	0.117	2,057	0.044	2,285	0.019
		(0.051)		(0.066)		(0.084)		(0.059)		(0.084)
School Climate index										
Pupil behaviour	2,348	-0.355*	2,146	-0.425**	1,972	-0.395	1,988	-0.288	2,130	-0.441*
		(0.190)		(0.187)		(0.347)		(0.248)		(0.229)
Teacher behaviour	2,348	0.101	2,146	0.170	1,972	-0.266	1,988	0.545	2,130	-0.250
		(0.327)		(0.296)		(0.722)		(0.351)		(0.402)

Note: The treatment group includes pupils enrolled in secondary schools that become academies between 2012/13 and 2014/15, whereas the control group includes pupils enrolled in LA maintained secondary schools in 2014/15. The treatment group is composed by pupils attending converter and sponsor-led academies that are either a single-academy trust or part of a multi-academy trust. The estimates are based on a pupil level regression model that controls for school level variables in 2012 and reflects the number of pupils in schools where variables are observed. PISA indices are standardised to mean zero and standard deviation one. The quantity in parenthesis below the treatment coefficient is the standard error. All statistics, except the number of observations (N) that is unweighted, reflect PISA survey settings (OECD, 2016a). * significant at 10% level; ** significant at 5% level; *** significant at 1% level or below.

Table 5.4 Comparison of academies with late conversions

OUTCOME VARIABLES	ACADEMIES (1) N	ACADEMIES (2) Treatment	CONVERTER (3) N	CONVERTER (4) Treatment	SPONSOR-LED (5) N	SPONSOR-LED (6) Treatment	difference (6)-(4)	MAT (7) N	MAT (8) Treatment
MANAGEMENT									
Responsibility index									
Curriculum	937	-0.599***	583	-0.770***	354	-0.341	0.430	692	-0.343
		(0.220)		(0.295)		(0.318)	(0.434)		(0.234)
Resources	937	-0.140	583	-0.137	354	-0.158	-0.022	692	0.229
		(0.390)		(0.443)		(0.668)	(0.800)		(0.412)
Leadership index									
Curricular Development	889	0.289	556	-0.203	333	1.067	1.269	671	0.583
		(0.353)		(0.281)		(0.890)	(0.926)		(0.377)
Instructional	889	0.209	556	-0.275	333	0.970	1.245	671	0.652*
		(0.378)		(0.428)		(0.713)	(0.812)		(0.360)
Professional Development	862	-0.116	529	-0.126	333	-0.115	0.011	644	0.109
		(0.374)		(0.364)		(0.704)	(0.740)		(0.427)
Teacher Participation	889	0.409	556	0.480	333	0.285	-0.195	671	0.809**
		(0.326)		(0.456)		(0.418)	(0.691)		(0.337)
WMS Management index									
Overall	733	0.273	450	-0.073	283	0.823	0.895	566	0.546
		(0.343)		(0.377)		(0.734)	(0.844)		(0.362)
Operations	733	0.335	450	-0.019	283	0.897	0.916	566	0.566
		(0.367)		(0.453)		(0.694)	(0.838)		(0.391)
People	733	-0.117	450	-0.275	283	0.130	0.404	566	0.215
		(0.362)		(0.503)		(0.661)	(0.886)		(0.354)

OUTCOME VARIABLES	ACADEMIES (1) N	ACADEMIES (2) Treatment	CONVERTER (3) N	CONVERTER (4) Treatment	SPONSOR-LED (5) N	SPONSOR-LED (6) Treatment	difference (6)-(4)	MAT (7) N	MAT (8) Treatment
POLICIES									
Learning Time (hours)									
Mathematics	1,093	0.297*	633	-0.051	460	0.750***	0.801***	817	0.445**
		(0.173)		(0.224)		(0.207)	(0.297)		(0.178)
English	1,092	-0.026	633	-0.226	459	0.214	0.441	816	0.037
		(0.190)		(0.275)		(0.218)	(0.341)		(0.207)
Science	1,091	0.093	632	-0.145	459	0.425	0.570	816	0.150
		(0.188)		(0.274)		(0.332)	(0.485)		(0.219)
Total	1,024	-0.447	601	0.009	423	-1.140	-1.150	762	-0.385
		(0.547)		(0.603)		(0.894)	(1.022)		(0.539)
Admissions									
Academic Performance	910	0.003	556	0.061	354	-0.085	-0.146	692	0.063
		(0.183)		(0.232)		(0.273)	(0.343)		(0.195)
Feeder Schools	910	-0.132	556	-0.239	354	0.032	0.270	692	-0.129
		(0.209)		(0.272)		(0.282)	(0.363)		(0.222)
School Philosophy or Religion	886	0.196	556	0.034	330	0.456**	0.422	668	0.194
		(0.193)		(0.292)		(0.183)	(0.338)		(0.218)
School Specialism	910	-0.273	556	-0.458**	354	0.033	0.490*	692	-0.383*
		(0.190)		(0.194)		(0.298)	(0.294)		(0.209)
Former Pupils and Siblings	910	-0.165	556	-0.253*	354	-0.026	0.226	692	-0.138
		(0.134)		(0.129)		(0.251)	(0.277)		(0.144)
Catchment Area	910	-0.042	556	-0.096	354	0.051	0.147	692	-0.065
		(0.126)		(0.100)		(0.253)	(0.271)		(0.138)
Pupil Assessment									
Mandatory Standardised Tests	883	-0.016	556	-0.144	327	0.190	0.334	665	-0.022
(More than 2 times a year)		(0.099)		(0.174)		(0.128)	(0.207)		(0.128)
Non-mandatory Standardised	883	-0.118	556	-0.051	327	-0.225	-0.174	665	-0.163
Tests (At least once a year)		(0.152)		(0.219)		(0.148)	(0.263)		(0.180)
Teacher-developed Tests	883	0.015	556	0.128	327	-0.165	-0.293	665	-0.028
(At least monthly)		(0.200)		(0.221)		(0.244)	(0.252)		(0.203)
Teachers Ratings	883	-0.311*	556	-0.227	327	-0.444	-0.217	665	-0.361**
(At least monthly)		(0.167)		(0.178)		(0.320)	(0.348)		(0.180)

OUTCOME VARIABLES	ACADEMIES (1) N	(2) Treatment	CONVERTER (3) N	(4) Treatment	SPONSOR-LED (5) N	(6) Treatment	difference (6)-(4)	MAT (7) N	(8) Treatment
LEARNING ENVIRONMENT									
Learning Environment index									
Disciplinary Climate	1,077	0.182	619	0.205*	458	0.157	-0.048	805	0.190
		(0.119)		(0.109)		(0.207)	(0.209)		(0.131)
Inquiry-based Learning	1,062	0.255**	609	0.190	453	0.339*	0.150	796	0.261**
		(0.112)		(0.143)		(0.191)	(0.239)		(0.111)
Teacher-directed Instruction	1,055	0.147	602	0.227**	453	0.050	-0.177	788	0.159
		(0.115)		(0.107)		(0.221)	(0.262)		(0.118)
Adaptive Instruction	1,029	0.188	589	0.133*	440	0.257	0.124	766	0.218*
		(0.117)		(0.079)		(0.218)	(0.217)		(0.121)
School Climate index									
Pupil behaviour	883	-0.471	556	-0.293	327	-0.760**	-0.467	665	-0.552*
		(0.289)		(0.441)		(0.306)	(0.480)		(0.287)
Teacher behaviour	883	-0.021	556	0.339	327	-0.594	-0.934	665	-0.269
		(0.299)		(0.322)		(0.628)	(0.723)		(0.385)

Note: The treatment group includes pupils enrolled in secondary schools that become academies between 2012/13 and 2014/15, whereas the control group includes pupils enrolled in LA maintained secondary schools that convert to academy status in 2015/16 and 2016/17. The treatment group is composed by pupils attending converter and sponsor-led academies that are either a single-academy trust or part of a multi-academy trust. The estimates are based on a pupil level regression model that reflects the number of pupils in schools where variables are observed. PISA indices are standardised to mean zero and standard deviation one. The quantity in parenthesis below the treatment coefficient is the standard error. All statistics, except the number of observations (N) that is unweighted, reflect PISA survey settings (OECD, 2016a). * significant at 10% level; ** significant at 5% level; *** significant at 1% level or below.

5.5.1 School Management Practices

When we control for the characteristics of schools in 2012, management-related outcomes are very similar in academies and LA maintained schools (see Table 5.3). Headteachers in MAT academies, though, show higher level of leadership in curricular development, instruction, and teacher participation, as initially suggested.

Focusing on late conversions, Table 5.4 shows that academies take significantly less responsibility over curriculum.[169] Some caution is required in interpreting this result. Basically, there is no variation over the control group (see Table 5.2). So, I will focus analysis on underlying within-school responsibilities.

In England, school governance is generally undertaken by the school governing board and the headteacher. The headteacher, appointed by the governing board, is responsible for daily management (Long, 2019b). The school governing board provides strategic direction and holds the headteacher to account (Long, 2019b).[170]

The responsibilities of the headteacher, the school governing board, and teachers at academies are distinct from LA maintained schools (see appendix C Table C7). School governing boards at academies have a lesser

---

[169] The estimate for curriculum responsibility index is -0.77 SD lower for converter academies.

[170] The school governing board (board of trustees in academies) includes representatives of teachers and parents as well as of the LA in the case of maintained schools (Long, 2019b) At sponsor-led academies, sponsors nominate most of the board (Duchini et al., 2023; Gibson, 2015). In MATs, the board of trustees decides if each school has a governing board and its decision-making powers (Long, 2019b; A. West & Wolfe, 2019) and the chief executive officer (CEO) provides executive leadership (Culpin & Male, 2022).

role in personnel management, whereas headteachers are less influential over budgeting, discipline, and assessment. Pupils attending academies are 38% less likely to be in a school where the governing board takes a role in firing teachers[171] and 25% less likely that it determines their salary.[172] They are also 12% and 15% less likely to be in a school where the headteacher sets disciplinary[173] and pupil assessment policies[174] and 16% less likely that he is responsible for the budget.[175]

Since the likelihood of other school agents participating in those decisions is not significantly altered at academies, it is reasonable to assume a different relative role of the headteacher and the school governing board in the governance of schools. Responsibilities such as firing teachers and setting their salary are shared in most schools between the headteacher and the governing board. However, the number of academies where this happens is much lower. The school governing board is not involved in these tasks in over a third of academies. The headteacher, on the other hand, is responsible for it in more than one in five academies. This suggests headteachers are taking full responsibility for this in a significant number of schools. Other tasks such as budgeting or setting assessment and disciplinary policies are typically

[171] Headteachers and governing boards are responsible for firing teachers in most schools. LAs also take a role in 14% of maintained schools and 10% of academies.

[172] In most schools the governing board and the headteacher establish teachers' salary increase together. LAs take a role in about 10% of schools.

[173] In almost every school headteachers take a role over pupil disciplinary policy. Most schools also involve teachers and the school governing board in this. A few schools say LAs have responsibility over disciplinary matters.

[174] In above 85% of schools, headteachers set assessment policies. In most schools, teachers are also involved in this. School governing boards have a role over assessment policies in over a third of schools.

[175] In more than four in five schools, headteachers set the budget. Most schools also involve the school governing board in this process. 27% of maintained schools acknowledge the involvement of LAs.

undertaken by the headteacher. In some academies, though, the governing board or teachers are responsible for these tasks.

Overall results are largely confirmed when focusing on academies that have joined a MAT. School governing boards at MATs are 34% less likely to fire teachers and 30% less likely to determine their salary (see appendix C Table C7). This might reflect new organisational structures where MAT central bodies take further responsibility (Chapman, 2013). These structures, however, are not specified in the PISA school questionnaire.

Organizational differences at sponsor-led academes are more pronounced (see appendix C Table C7).[176] The governing board at these schools is 45% and 56% less likely to fire teachers and set the school budget. The headteacher is 27% less likely to be involved in pupil admissions.[177] Interestingly, both the headteacher and the school governing board are less likely to set disciplinary policies (19% and 47%) and assessment policies (24% and 50%), suggesting a greater relative role of teachers. However, teachers at sponsor-led academies are 64% less likely to be involved in recruiting new teachers than their colleagues at LA maintained schools.[178]

Turning now to the leadership indices, academies are very similar to late conversions. Nevertheless, Table 5.4 shows that leadership in instruction and in teacher participation are significant features of schools that have joined

---

[176] In line with overall results, the school governing board at converter academies is 34% less likely to be involved in firing teachers and 22% less likely to determine teachers' salary increase, and the headteacher is 18% less likely to set the school budget.

[177] Most schools include headteachers in the admission process. The school governing board and teachers are also involved in many schools. LAs are responsible for pupil admissions for 60% of maintained schools and 23% of academies.

[178] Headteachers are responsible for hiring teachers in almost every school. Most school also involve school governing boards in this decision.

a MAT, as suggested in multivariate regressions. The estimates for the instructional leadership index and the teacher participation index are 0.65 of a standard deviation (SD) and 0.81 SD higher for academies belonging to a MAT. In effect, headteachers at these schools are 33% more likely to discuss school goals with teachers and 35% more likely to promote evidence-based practices (see appendix C Table C6).

Regarding the WMS management index, academies seem very similar to LA maintained schools that convert later. Table 5.4 suggests, nonetheless, that academies belonging to a MAT have somewhat more structured management practices, especially in operational domains. These schools have 0.55 SD higher management scores than LA maintained schools, but this estimate is marginally not significant.[179]

5.5.2 School Policies

When we control for pre-treatment characteristics, school policies are similar in academies and LA maintained schools (see Table 5.3). The only significant difference is the more frequent use of teacher developed tests at academies, particularly at converter academies. At sponsor-led academies, though, pupils devote more time to mathematics, English, and science and have slightly less total learning time. These academies are less likely to select pupils based on aptitude for the school specialism.

Looking into school time per subject, pupils attending academies spend 18 more minutes per week learning mathematics than their peers at late

---

[179] Coefficient level of significance 13%.

conversions (see Table 5.4).[180] Pupils at sponsor-led academies study mathematics 45 more minutes per week than their peers at maintained schools (significantly more than at converter academies). Despite this difference in mathematics learning time, total learning time is not significantly different at academies. Since the time devoted to English and science is not affected, it is safe to assume other subjects have lesser importance at these schools. Basically, either these pupils are taking fewer options or academies allocate less teaching time to the other subjects.

Regarding assessment policies, the main analysis does not confirm the more frequent use of teacher developed tests at academies (see Table 5.4). Teacher ratings, though, are less frequently used. Pupils at academies that joined a MAT are 36% less likely to have monthly teacher ratings than their peers at maintained schools. Still, assessment practices look broadly similar.

On pupil admission policies, Table 5.4 shows there are no significant differences between academies and LA maintained schools that convert after the sample period. Nevertheless, converter academies select pupils 46% less often based on interest in a special programme (significantly less that sponsor-led academies)[181] and are 25% less likely to prioritize relatives of current or former pupils. On the other hand, sponsor-led academies are 46% more likely to select pupils based on the school philosophy or faith.

---

[180] Pupils at MATs devote 27 more minutes per week to mathematics.
[181] MAT academies are also less likely to use the school specialism for admissions.

5.5.3 Learning Environment

When we control for baseline school characteristics, the outcomes describing the learning environment are very similar in academies and LA maintained schools (see Table 5.3). Nevertheless, academies report fewer issues with pupil behaviour. As suggested earlier, pupil behaviour and disciplinary climate are better at converter academies. Besides, teachers at these schools use problem-based learning more frequently.

The comparison between academies and LA maintained schools that convert in the following period is more revealing (see Table 5.4). Pupils at academies are more likely to be exposed to problem-based learning. Those at academies that joined a MAT report having classes using inquiry-based learning and adaption of instruction 0.26 SD and 0.22 SD more often respectively. These schools also have fewer issues with pupil behaviour. The estimate for pupil-related behaviour hindering learning is -0.55 SD lower for pupils attending MAT academies.

As suggested in multivariate regressions, there is a better classroom disciplinary climate at converter academies (see Table 5.4). The estimate for the disciplinary climate index is 0.21 SD higher for pupils at these academies than for their peers at LA maintained schools. Teachers at converter academies are more likely to use teacher-directed instruction and adaptive instruction. The estimate for the teacher-directed instruction index is 0.23 SD higher and for the adaptive instruction index is 0.13 SD higher for converter academies. The previous suggestion of more frequent inquiry-based learning and better pupil behaviour at these schools is not established here. Besides, headteachers at converter academies are significantly more concerned with teacher absenteeism. Indeed, they are 51% more likely to flag it as hindering learning than headteachers at LA maintained schools (see appendix C Table C6).

Teachers at sponsor-led academies, on the other hand, are more likely to use problem-based learning (see Table 5.4). The estimate for the inquiry-based learning index is 0.34 SD higher for pupils at sponsor-led academies than for their peers at LA maintained schools. Moreover, sponsor-led academies face fewer issues with pupil behaviour. The estimate for pupil-related behaviour hindering learning is -0.76 SD lower for these academies. Furthermore, pupils attending sponsor-led academies are 52% less likely to be in a school where headteachers report that teachers do not meet individual pupils' needs. Nevertheless, headteachers at these schools are 25% more likely to report staff resisting change.

5.6 Discussion

Autonomy has been granted to state-funded schools to raise educational standards and develop a self-improving school-led system (Chapman & Salokangas, 2012; Greany & Waterhouse, 2016; Woods & Simkins, 2014). There is some evidence suggesting academisation has an impact on exam performance (Andrews et al., 2017; Eyles et al., 2016b, 2017, 2018; Eyles & Machin, 2019) and on pupils' decision-making skills, self-esteem, and subject choice (see chapters 3 and 4). Previous studies reveal, however, that the effect of autonomy varies widely across schools (Andrews et al., 2017). More important than a school having autonomy is how a school uses that autonomy. This chapter, using school level information from PISA 2015 and staggered treatment, highlights distinctive features of academies that can help explain those effects.

The school governing board in academies plays a lesser role in personnel management and the headteacher is less influential over pupil discipline and assessment. Sponsor-led academies have a particularly distinctive organisation, possibly due to restructuring under an external sponsor. The school governing board in these academies is less likely to have a role over the management of teachers or the school budget, teachers are less involved in recruitment, and the headteacher is less likely to be responsible for discipline and assessment policies. Cirin (2014) also shows that sponsor-led academies make important functioning changes after conversion. Accordingly, staff resisting change is flagged by headteachers at sponsor-led academies as a cause of concern. This is consistent with the observation of Keddie (2019).

A distinctive feature of academies that have joined MATs is a strong leadership in instruction and in teacher participation. Headteachers within MATs are more likely to discuss school goals with teachers and more likely to promote evidence-based practices. The WMS management index also suggests, as Bloom et al. (2015), that academies belonging to a MAT have higher management quality. Bloom et al. (2015) point to possible economies of scale but attribute performance of autonomous schools to leadership and accountability. Some MATs have indeed developed a professional school governance focused on results (Healey, 2022). MATs are typically led by former headteachers of outstanding schools (Simon et al., 2021). Besides, many MAT academies have changed leadership (Cirin, 2017). In effect, schools that have joined a MAT often experience staff renewal (Keddie, 2019). Qualitative data also indicate that MAT academies operate in a more standardised environment (Keddie, 2019; Wiborg et al., 2018). A change in leadership coupled with standardised managerial processes could partly explain improvement in management and leadership within MATs.

Schools have been encouraged to join MATs expecting better management and leadership (Greany & Waterhouse, 2016; Neri & Pasini, 2018). Good management is associated with higher attainment (Bloom et al., 2015; Leaver et al., 2022). Leadership should stimulate school performance (OECD, 2016c). This is consistent with some evidence of higher performance at MAT academies (Neri et al., 2022; Neri & Pasini, 2018). The higher management quality and strong leadership of these academies supports to some extent the policy encouraging schools to join MATs. There is no evidence, however, of better human resource management in MATs, confirming Bryson et al. (2023).

Regarding teaching, pupils at academies are more frequently engaged in problem-based learning and spend more time on mathematics than their peers at LA maintained schools. They are also less likely to receive monthly teacher ratings. To make sense of these differences, let us look to converter and sponsor-led academies separately.

Teachers at converter academies are more likely to use teacher-directed instruction and adaptive instruction. Advantaged schools tend to use teacher-directed instruction more often and autonomous schools to use more frequently adaptive instruction (OECD, 2016c). These methods might explain some cognitive effects of conversion (Andrews et al., 2017) or be related to the prominence of science subjects and facilitating subjects at these schools (see chapter 4). It is noteworthy, though, that problem-based learning is not used more at converter academies, given reported improvements in decision-making skills (see chapter 3). This pedagogical approach is designed to promote problem-solving and decision making skills (Arends, 2006; Davis & Harden, 1999; Hmelo-Silver, 2004; OECD, 2013, 2014, 2016c). The speculation that those improvements reflect the use of these educational methods is not confirmed here.

On the other hand, sponsor-led academies make frequent use of problem-based learning. This is more time-consuming and complicated to implement, especially if covering a long curriculum (OECD, 2016c). So its frequent use is contrary to the speculation that ranking maximization strategies underlie positive effects of academy conversion on exam scores (Andrews et al., 2017; Eyles et al., 2016a, 2016b; Eyles & Machin, 2019). Interestingly, chapter 3 has not identified significant effects of sponsor-led academies on decision-making skills which should be fostered by these educational methods (Arends, 2006). The high percentage of teachers without Qualified Teacher Status at these schools could have contributed to this (Martindale, 2022). The emphasis on problem-based learning might reflect a focus on specific subjects. Conversely, chapter 4 shows that GCSE subject choices are similar at sponsor-led academies and at LA maintained schools. Still, pupils at these academies spend considerably more time studying mathematics. This investment in mathematics may have contributed to improvements in exam performance and could be associated with a wider pedagogical strategy.

Besides, studying mathematics is correlated with attending a high-ranking university and offers a wage premium (Dilnot, 2018; Wolf, 2011).

Learning requires an orderly and positive environment (OECD, 2016c). The learning environment at academies and at LA maintained schools seems very similar. Academies that have joined a MAT, though, have fewer issues with pupil behaviour, possibly owing to good management (Leaver et al., 2022). Effective school leaders foster a positive school climate and control disruptive behaviour (OECD, 2016c). Sponsor-led academies also report fewer issues with pupil behaviour. Perhaps this is because teachers are more likely to meet the individual needs of pupils. Or it could reflect the positive effects of academy conversion on social behaviour (Frostick et al., 2018) and on self-esteem (see chapter 3). Some sponsor-led academies, for example, use pupil reward schemes (J. P. Morris, 2020) or adopt a rigorous disciplinary policy (Machin & Sandi, 2020). A renovated workforce could also contribute to this (Duchini et al., 2023; Martindale, 2022). On the other hand, there is a better classroom disciplinary climate at converter academies, typical of advantaged schools (OECD, 2016c), that certainly contributes to their overall effectiveness (see chapter 3).

School admission policies are of interest because of increased social stratification between sponsor-led and converter academies (Braz, 2018; Eyles et al., 2018). Converter academies select pupils less often based on the school specialism or on siblings' attendance. Sponsor-led academies are more likely to use instead faith or school philosophy as admission criteria. These differences do not seem to explain intake profiles. Faith selection, for example, is usually associated with a more advantaged intake (Allen & West, 2009), but sponsor-led academies have less advantaged intakes (Braz, 2018; Eyles et al., 2018). Perhaps parental preferences explain school segregation, as suggested by Bertoni et al. (2021) or academies have kept predecessor school's admissions criteria (A. West, 2014). In any case, converter and sponsor-led academies have distinctive admission policies that shape pupil intakes and influence the learning environment we observe at academies.

6. CONCLUSION

This thesis studies the impact of academy conversion on several pupil outcomes and school features in England, revealing effects of school autonomy and education quasi-markets. Each chapter explores variations associated with the timing of academy conversion to estimate effects on outcomes not previously studied in the literature. By focusing on staggered treatment and legacy enrolment and using rich longitudinal data, I attempt to approximate causality and avoid selection issues which are common in the literature.

A recurrent conclusion across chapters 3 and 4 is that pupil outcomes at academies and at LA maintained schools are similar on average. The effect of academy conversion appears when sponsor-led academies are separated from converter academies. Pupils at sponsor-led academies have higher self-esteem. Pupils at converter academies have better decision-making skills and are more likely to take science subjects and facilitating subjects at secondary school. So, academy conversion affects pupil outcomes differently depending on the academy route taken. These results are discussed at the end of each chapter.

Academies have created a distinct set of management practices, policies and school climates that can help explain these results, as discussed at the end of chapter 5. Converter and sponsor-led academies have different admission policies. Teachers at converter academies manage classrooms more effectively and are more likely to use teacher-directed instruction and adaptive instruction. On the other hand, pupils at sponsor-led academies spend considerably more time studying mathematics and engage in problem-based learning more frequently. Headteachers at these academies are less likely to report that pupil behaviour hinders learning.

Understanding the options schools take is critical if one wants to influence the outcomes of an autonomous school system. This is particularly relevant, given the limited data we have on what is going on in schools. The signs of a positive disciplinary climate and good pupil behaviour at academies are encouraging. There is, however, wide variation across academies. Further research on school practices would help understanding these results and disseminating best practices. Besides, considering the known effects of academy conversion on pupil outcomes and on the social composition of schools, a diverging school system in terms of management, school policy, and learning environment raises concerns over equal access and social mobility that must be addressed. Indeed, some pupils may thrive on certain subjects and pedagogical approaches that are specially worked by some academies. Addressing this concern with the many possibilities of school choice raises other issues. In effect, given the expansion of academy chains, local school choice may be limited.

Current government policy encourages schools to join a MAT (Department for Education, 2022). The evidence suggesting MATs have strong leadership and more structured management practices is reassuring. However, there are concerns over the possible emergence of local education monopolies and monopolistic practices (Wilkins, 2017). If large networks with highly prescriptive practices dominate the education market, school competition and the performance of academies will be affected. Further research on the effects of MATs on school competition and performance is required. Moreover, MAT leadership is far more challenging than leading an individual school, requiring system leadership (Culpin & Male, 2022). Academy chains are taking responsibilities in the education system that rely on the professional and personal qualities of school leaders, not as exposed to local democratic accountability or parents' participation (Healey, 2022). In this context, understanding the incentives driving education quasi-markets and where they are leading schools becomes increasingly important.

Increased social stratification between sponsor-led academies and converter academies, with the latter having more advantaged intakes (Braz, 2018; Eyles et al., 2018), raises important equity concerns. The improvement in problem-solving skills of pupils attending converter academies, coupled with a more science-based and academically oriented curriculum, gives these pupils an additional advantage in pursuing high-status careers (Dilnot, 2016). Since these effects are not observed at sponsor-led academies, that typically attract more disadvantaged intakes, the academy programme faces obvious issues regarding social mobility. Moreover, this raises questions on how reported gains at national exams from sponsor-led academies are attained (Andrews et al., 2017). Further research is required to reveal teaching and learning practices within schools and investigate school maximising ranking performance strategies. A reappraisal of quasi-market incentives and mechanisms currently driving the school system should follow.

Looking into non-cognitive outcomes leads us to a similar conclusion. Pupils at converter academies do not significantly improve their non-cognitive skills in relation to their peers. The same can be said of sponsor-led academies, except for self-esteem. Apparently, quasi-market incentives do not encourage academies to target non-cognitive outcomes, at least as they are captured in this study. Since non-cognitive skills are critical for later life outcomes (Kautz et al., 2014), schools promoting them should be encouraged. It would be important, however, to check if these results hold in another cohort and using different outcomes. Moreover, this work focuses on short-term effects after academy conversion. Future research should study long-term effects, allowing schools more time to implement changes and take advantage of autonomy. Further research is also required to evaluate the impact of academy attendance on later life outcomes, including high education participation, labour market, criminality, and health. School outcomes have social and economic consequences. Policy makers must take responsibility over the outcomes delivered by the school system, setting the purpose of state education and designing school incentives accordingly.

REFERENCES

Abadie, A., Athey, S., Imbens, G., & Wooldridge, J. (2017). *When Should You Adjust Standard Errors for Clustering?* (w24003). National Bureau of Economic Research. https://doi.org/10.3386/w24003

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. https://doi.org/10.1093/qje/qjr017

Abdulkadiroğlu, A., Angrist, J. D., Hull, P. D., & Pathak, P. A. (2016). Charters without Lotteries: Testing Takeovers in New Orleans and Boston. *American Economic Review*, *106*(7), 1878–1920. https://doi.org/10.1257/aer.20150479

Adamecz-Volgyi, A., Henderson, M., & Shure, N. (2023). Intergenerational educational mobility – The role of non-cognitive skills. *Education Economics*, 1–20. https://doi.org/10.1080/09645292.2023.2176826

Agalioti-Sgompou, V., Mark Atkinson, David Church, Jon Johnson, Tarek Mostafa, Tom Murphy, Andrew Peters, & Rachel Rosenberg. (2017). *MCS6 Derived Variables. User guide*. Centre for Longitudinal Studies.

Alexiadou, N., Dovemark, M., Erixon-Arreman, I., Holm, A.-S., Lundahl, L., & Lundström, U. (2016). Managing inclusion in competitive school systems: The cases of Sweden and England. *Research in Comparative and International Education*, *11*(1), 13–33. https://doi.org/10.1177/1745499916631065

Alldritt, C., & Taylor, R. (2015). *Lack of options: How a pupil's academic choices are affected by where they live* [OPSN Report]. OPSN.

Allen, R., & Burgess, S. (2010). *The future of competition and accountability in education* (p. 30). 2020 Public Services Trust.

Allen, R., & Burgess, S. (2012). *How should we treat under-performing schools? A regression discontinuity analysis of school inspections in England* (Working Paper 287; CMPO Working Paper). Centre for Market and Public Organisation. http://search.proquest.com/docview/1697784980/?pq-origsite=primo

Allen, R., & Burgess, S. (2013). Evaluating the provision of school performance information for school choice. *Economics of Education Review*, *34*, 175–190. https://doi.org/10.1016/j.econedurev.2013.02.001

Allen, R., Burgess, S., & McKenna, L. (2013). The short-run impact of using lotteries for school admissions: Early results from Brighton and Hove's reforms: *Lotteries for school admissions*. *Transactions of the Institute of British Geographers*, *38*(1), 149–166. https://doi.org/10.1111/j.1475-5661.2012.00511.x

Allen, R., & Higham, R. (2018). Quasi-markets, school diversity and social selection: Analysing the case of free schools in England, five years on. *London Review of Education*, *16*(2), 191–213. https://doi.org/10.18546/LRE.16.2.02

Allen, R., & West, A. (2009). Religious schools in London: School admissions, religious composition and selectivity. *Oxford Review of Education*, *35*(4), 471–494. https://doi.org/10.1080/03054980903072611

Anders, J., Henderson, M., Moulton, V., & Sullivan, A. (2018a). Incentivising Specific Combinations of Subjects – Does It Make Any Difference to University Access? *National Institute Economic Review*, *243*(1), R37–R52. https://doi.org/10.1177/002795011824300113

Anders, J., Henderson, M., Moulton, V., & Sullivan, A. (2018b). The role of schools in explaining individuals’ subject choices at age 14. *Oxford Review of Education*, *44*(1), 75–93. https://doi.org/10.1080/03054985.2018.1409973

Andrews, J., Perera, N., Eyles, A., Sahlgren, G., Machin, S., Sandi, M., & Silva, O. (2017). *The impact of academies on educational outcomes* (p. 98). Education Policy Institute.

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. *Journal of Labor Economics*, *34*(2), 275–318.

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist’s companion*. Princeton University Press.

Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, B., & Wong, B. (2012). Science Aspirations, Capital, and Family Habitus: How Families Shape Children’s Engagement and Identification With Science. *American Educational Research Journal*, *49*(5), 881–908. https://doi.org/10.3102/0002831211433290

Arends, R. (2006). *Learning to teach* (7th ed). McGraw Hill.

Armstrong, R. A. (2014). When to use the Bonferroni correction. *Ophthalmic and Physiological Optics*, *34*(5), 502–508. https://doi.org/10.1111/opo.12131

Bannink, R., Pearce, A., & Hope, S. (2016). Family income and young adolescents’ perceived social position: Associations with self-esteem and life satisfaction in the UK Millennium Cohort Study. *Archives of Disease in Childhood*, *101*(10), 917–921. https://doi.org/10.1136/archdischild-2015-309651

Barker, B., & Hoskins, K. (2017). Can high-performing academies overcome family background and improve social mobility? *British Journal of Sociology of Education*, *38*(2), 221–240. https://doi.org/10.1080/01425692.2015.1073104

Barrance, R., & Elwood, J. (2018). Young people's views on choice and fairness through their experiences of curriculum as examination specifications at GCSE. *Oxford Review of Education*, *44*(1), 19–36. https://doi.org/10.1080/03054985.2018.1409964

Barrows, H. (2002). Is it Truly Possible to Have Such a Thing as dPBL? *Distance Education*, *23*(1), 119–122. https://doi.org/10.1080/01587910220124026

Bastelaer, A. van, Lemaître, G., & Marianna, P. (1997). *The Definition of Part-Time Work for the Purpose of International Comparisons* (OECD Labour Market and Social Policy Occasional Papers 22; OECD Labour Market and Social Policy Occasional Papers, Vol. 22). OECD. https://doi.org/10.1787/132721856632

Belfield, C., Boneva, T., Rauh, C., & Shaw, J. (2020). What Drives Enrolment Gaps in Further Education? The Role of Beliefs in Sequential Schooling Decisions. *Economica*, *87*(346), 490–529. https://doi.org/10.1111/ecca.12307

Bertoni, M., Gibbons, S., & Silva, O. (2021). School choice during a period of radical school reform. Evidence from academy conversion in England. *Economic Policy*, *35*(104), 739–795. https://doi.org/10.1093/epolic/eiaa023

Bertoni, M., Heller-Sahlgren, G., & Silva, O. (2023). *Free to improve? The impact of free school attendance in England* (Discussion Paper 1946; CEP, p. 68). CEP.

Bibby, P., & Shepherd, J. (2004). *Developing a New Classification of Urban and Rural Areas for Policy Purposes – the Methodology* (p. 30). Department for Environment, Food and Rural Affairs.

Bleemer, Z., & Zafar, B. (2018). Intended college attendance: Evidence from an experiment on college returns and costs. *Journal of Public Economics*, *157*, 184–211. https://doi.org/10.1016/j.jpubeco.2017.11.002

Bloom, N., Lemos, R., Sadun, R., & Van Reenen, J. (2015). Does Management Matter in schools? *The Economic Journal*, *125*(584), 647–674. https://doi.org/10.1111/ecoj.12267

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. https://doi.org/10.1111/ecca.12130

Bolton, P. (2012). *Sponsored academies: Statistics : February 2012*. UK Parliament.

Bolton, P. (2015). *Converter academies: Statistics* (Briefing Paper 6233; p. 9). House of Commons Library.

Boyask, R. (2018). Primary school autonomy in the context of the expanding academies programme. *Educational Management Administration & Leadership*, *46*(1), 107–123. https://doi.org/10.1177/1741143216670649

Braconier, H. (2012). *Reforming Education in England* (OECD Economics Department Working Papers 939). https://doi.org/10.1787/5k9gsh772h9q-en

Brandt, V., Kerner auch Koerner, J., & Palmer-Cooper, E. (2019). The Association of Non-obscene Socially Inappropriate Behavior With Attention-Deficit/Hyperactivity Disorder Symptoms, Conduct Problems, and Risky Decision Making in a Large Sample of Adolescents. *Frontiers in Psychiatry*, *10*, 660. https://doi.org/10.3389/fpsyt.2019.00660

Braz, N. (2018). *The Equity Impact of Academies: Pupil Intake* [MSc]. UCL.

Breen, R., & Goldthorpe, J. H. (1997). Explaining educational differentials towards a formal rational action theory. *Rationality and Society*, *9*(3), 275–305.

Breen, R., Werfhorst, H. G. van de, & Jaeger, M. M. (2014). Deciding under Doubt: A Theory of Risk Aversion, Time Discounting Preferences, and Educational

Decision-making. *European Sociological Review*, *30*(2), 258–270. https://doi.org/10.1093/esr/jcu039

Bryson, A., Stokes, L., & Wilkinson, D. (2023). Is pupil attainment higher in well-managed schools? *Education Economics*, *31*(1), 129–144. https://doi.org/10.1080/09645292.2022.2042202

Burgess, S., Greaves, E., Vignoles, A., & Wilson, D. (2015). What Parents Want: School Preferences and School Choice. *The Economic Journal*, *125*(587), 1262–1289. https://doi.org/10.1111/ecoj.12153

Cambridge Cognition. (2019). *CANTAB® [Cognitive assessment software]* [Computer software]. www.cambridgecognition.com/products/cognitive-research/

Carneiro, R. (2003). *Fundamentos da educação e da aprendizagem: 21 ensaios para o século 21* (2a. ed). Fundacão Manuel Leão.

Carroll, M., & Gill, T. (2018). *Uptake of GCSE subjects 2017* (120; Statistics Report Series). Cambridge Assessment.

Centre for Longitudinal Studies. (2020). *Millennium Cohort Study User Guide (Surveys 1-5)* (Ninth). Centre for Longitudinal Studies.

Chapman, C. (2013). Academy Federations, Chains, and Teaching Schools in England: Reflections on Leadership, Policy, and Practice. *Journal of School Choice*, *7*(3), 334–352. https://doi.org/10.1080/15582159.2013.808936

Chapman, C., & Salokangas, M. (2012). Independent state-funded schools: Some reflections on recent developments. *School Leadership & Management*, *32*(5), 473–486. https://doi.org/10.1080/13632434.2012.731329

Children’s Society. (2018). *The Good Childhood Report 2018*. The Children’s Society.

Chng, E., Yew, E. H. J., & Schmidt, H. G. (2011). Effects of tutor-related behaviours on the process of problem-based learning. *Advances in Health Sciences Education*, *16*(4), 491–503. https://doi.org/10.1007/s10459-011-9282-7

Chulkov, D., & Nizovtsev, D. (2015). *Problem-Based Learning in Managerial Economics with an Integrated Case*. *16*(1), 11.

Church, D. (2017). *Guide to geographic identifiers*. 14.

Cirin, R. (2014). *Do academies make use of their autonomy?* (p. 53) [Research Report]. Department for Education.

Cirin, R. (2017). *Academy trust survey 2017* (p. 52) [Research Report]. Department for Education.

Closs, S. J. (1976). *APU vocabulary test: Specimen set*. Hodder and Stoughton.

Codiroli, N. (2017). Who studies STEM subjects at A level and degree in England? An investigation into the intersections between students' family background, gender and ethnicity in determining choice. *British Educational Research Journal*, *43*(3), 528–553. https://doi.org/10.1002/berj.3270

Cohort and Longitudinal Studies Enhancement Resources. (2018). *Consistent School Database. [Data collection].*

Constantinides, M. (2022). High-stakes accountability policies and local adaptation: Exploring how school principals respond to multiple policy demands. *School Leadership & Management*, *42*(2), 170–187. https://doi.org/10.1080/13632434.2021.2016687

Culpin, J., & Male, T. (2022). Examining the competencies required for leadership of multi-academy trusts. Implications for a case study trust. *School Leadership & Management*, *42*(3), 293–308. https://doi.org/10.1080/13632434.2022.2088491

Dahlgren, M. A., & Dahlgren, L. O. (2002). Portraits of PBL: students' experiences of the characteristics of problem-based learning in physiotherapy, computer engineering and psychology. *Instructional Science*, *30*, 111–127.

Davies, P., Adnett, N., & Turnbull, A. (2003). Market forces and diversity: Some evidence from the 14-19 curriculum. *Journal of Curriculum Studies*, *35*(4), 479–498. https://doi.org/10.1080/00220270305519

Davies, P., Diamond, C., & Perry, T. (2021). Implications of autonomy and networks for costs and inclusion: Comparing patterns of school spending under different governance systems. *Educational Management Administration & Leadership*, *49*(1), 128–144. https://doi.org/10.1177/1741143219888738

Davies, P., Telhaj, S., Hutton, D., Adnett, N., & Coe, R. (2008). Socioeconomic background, gender and subject choice in secondary schooling. *Educational Research*, *50*(3), 235–248. https://doi.org/10.1080/00131880802309358

Davis, M. H., & Harden, R. M. (1999). AMEE Medical Education Guide No. 15: Problem-based learning: a practical guide. *Medical Teacher*, *21*(2), 130–140. https://doi.org/10.1080/01421599979743

De Philippis, M. (2021). STEM Graduates and Secondary School Curriculum: Does Early Exposure to Science Matter? *Journal of Human Resources*, 1219-10624R1. https://doi.org/10.3368/jhr.1219-10624R1

Deakin, J., Aitken, M., Robbins, T., & Sahakian, B. J. (2004). Risk taking during decision-making in normal volunteers changes with age. *Journal of the International Neuropsychological Society*, *10*(4), 590–598. https://doi.org/10.1017/S1355617704104104

Dearden, L., Emmerson, C., Frayne, C., & Meghir, C. (2009). Conditional Cash Transfers and School Dropout Rates. *The Journal of Human Resources*, *44*(4), 827–857.

Dearden, L., Sibieta, L., & Sylva, K. (2011). The socio-economic gradient in early child outcomes: Evidence from the Millennium Cohort Study. *Longitudinal and Life Course Studies*, *2*(1), 19–40.

Delavande, A., & Fumagalli, L. (2019). *Information, Expectations and Transition to Higher Education* (p. 53). Nuffield Foundation.

Delors, J. (Ed.). (1996). *Learning, the treasure within: Report to UNESCO of the International Commission on Education for the Twenty-first Century*. Unesco Pub.

Deming, D. J. (2017). The Growing Importance of Social Skills in the Labor Market*. *The Quarterly Journal of Economics*, *132*(4), 1593–1640. https://doi.org/10.1093/qje/qjx022

Department for Education. (2012). *Schools, pupils and their characteristics: January 2012* (Statistical First Release SFR10/2012).

Department for Education. (2013a). *Academies annual report 2011/12*. Stationery Office.

Department for Education. (2013b). *Schools, pupils and their characteristics: January 2013* (Statistical First Release SFR21).

Department for Education. (2014). *The national curriculum in England: Key stages 3 and 4 framework document*.

Department for Education. (2016a). *Educational excellence everywhere*. The Stationery Office.

Department for Education. (2016b). *How Progress 8 and Attainment 8 measures are calculated* (p. 5).

Department for Education. (2017). *Achievement of 15-Year- Olds in England: PISA 2015 Collaborative Problem Solving National Report* (Research Brief DFE-RB765).

Department for Education. (2018). *Revised GCSE and equivalent results in England, 2016 to 2017* (SFR 01/2018).

Department for Education. (2022). *Opportunity for all: Strong schools with great teachers for your child*. Dandy Booksellers Ltd.

Department for Work and Pensions. (2012). *An analysis of the income distribution 1994/95 – 2010/11* (23; Households Below Average Income).

Dilnot, C. (2016). How does the choice of A-level subjects vary with students' socio-economic status in English state schools? *British Educational Research Journal*, *42*(6), 1081–1106. https://doi.org/10.1002/berj.3250

Dilnot, C. (2018). The relationship between A-level subject choice and league table score of university attended: The 'facilitating', the 'less suitable', and the counter-intuitive. *Oxford Review of Education*, *44*(1), 118–137. https://doi.org/10.1080/03054985.2018.1409976

Dobbie, W., & Fryer, R. G. (2015). The Medium-Term Impacts of High-Achieving Charter Schools. *Journal of Political Economy*, *123*(5), 985–1037. https://doi.org/10.1086/682718

Duchini, E., Lavy, V., Machin, S., & Telhaj, S. (2023). *School Management Takeover, Leadership Change, and Personnel Policy* (w31994; p. w31994). National Bureau of Economic Research. https://doi.org/10.3386/w31994

DuGoff, E. H., Schuler, M., & Stuart, E. A. (2014). Generalizing Observational Study Results: Applying Propensity Score Methods to Complex Surveys. *Health Services Research*, *49*(1), 284–303. https://doi.org/10.1111/1475-6773.12090

Elliott, C. D. (1996). *BAS II : British ability scales second edition* (2nd ed.). Windsor : NFER-Nelson.

Elwick, A. (2018). New forms of government school provision – an international comparison. *Journal of Education Policy*, *33*(2), 206–225. https://doi.org/10.1080/02680939.2017.1329551

Esper, T. (2023). Quasi-markets, accountability and innovation: Analyzing the case of Free Schools in Englandv through teachers' perspectives. *Revista Española de Educación Comparada*, *42*, 193–220. https://doi.org/10.5944/reec.42.2023.34408

Eyles, A., Hupkau, C., & Machin, S. (2016a). Academies, charter and free schools: Do new school types deliver better outcomes? *Economic Policy*, *31*(87), 453–501. https://doi.org/10.1093/epolic/eiw006

Eyles, A., Hupkau, C., & Machin, S. (2016b). School reforms and pupil performance. *Labour Economics*, *41*, 9–19. https://doi.org/10.1016/j.labeco.2016.05.004

Eyles, A., & Machin, S. (2019). The Introduction of Academy Schools to England's Education. *Journal of the European Economic Association*, *17*(4), 1107–1146. https://doi.org/10.1093/jeea/jvy021

Eyles, A., Machin, S., & McNally, S. (2017). Unexpected school reform: Academisation of primary schools in England. *Journal of Public Economics*, *155*, 108–121. https://doi.org/10.1016/j.jpubeco.2017.09.004

Eyles, A., Machin, S., & Silva, O. (2018). Academies 2 - The New Batch: The Changing Nature of Academy Schools in England: The changing nature of academy schools in England. *Fiscal Studies*, *39*(1), 121–158. https://doi.org/10.1111/j.1475-5890.2017.12146

Fenwick-Sehl, L. (2013). Lessons from Elsewhere? The Evolution of the Labour Academy School Concept, 1997–2010. *Research in Comparative and International Education*, *8*(2), 176–192. https://doi.org/10.2304/rcie.2013.8.2.176

Fitzsimons, E., Agalioti-Sgompou, V., Calderwood, L., Gilbert, E., Haselden, L., Johnson, J., Smith, K., & Millennium Cohort Team. (2020). *Millennium Cohort Study Sixth Survey 2015-2016 User Guide* (Second). Centre for Longitudinal Studies.

Fitzsimons, E., Haselden, L., Smith, K., Gilbert, E., Calderwood, L., Agalioti-Sgompou, V., Veeravalli, S., Silverwood, R., & Ploubidis, G. (2020). *Millennium Cohort Study Age 17 Sweep (MCS7): User Guide* (First). Centre for Longitudinal Studies.

Flannery, D., & O'Donoghue, C. (2013). The demand for higher education: A static structural approach accounting for individual heterogeneity and nesting patterns. *Economics of Education Review*, *34*, 243–257. https://doi.org/10.1016/j.econedurev.2012.12.001

Flouri, E., Moulton, V., & Ploubidis, G. B. (2019). The role of intelligence in decision-making in early adolescence. *British Journal of Developmental Psychology*, *37*(1), 101–111. https://doi.org/10.1111/bjdp.12261

Flouri, E., & Papachristou, E. (2019). Peer problems, bullying involvement, and affective decision-making in adolescence. *British Journal of Developmental Psychology*, *37*(4), 466–485. https://doi.org/10.1111/bjdp.12287

Flouri, Eirini. (2006). Parental interest in children’s education, children’s self-esteem and locus of control, and later educational attainment: Twenty-six year follow-up of the 1970 British Birth Cohort. *British Journal of Educational Psychology*, *76*(1), 41–55. https://doi.org/10.1348/000709905X52508

Ford, T., Degli Esposti, M., Crane, C., Taylor, L., Montero-Marín, J., Blakemore, S.-J., Bowes, L., Byford, S., Dalgleish, T., Greenberg, M. T., Nuthall, E., Phillips, A., Raja, A., Ukoumunne, O. C., Viner, R. M., Williams, J. M. G., Allwood, M., Aukland, L., Casey, T., … Wainman, B. (2021). The Role of Schools in Early Adolescents’ Mental Health: Findings From the MYRIAD Study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *60*(12), 1467–1478. https://doi.org/10.1016/j.jaac.2021.02.016

Francis, B. (2000). The Gendered Subject: Students’ subject preferences and discussions of gender and subject ability. *Oxford Review of Education*, *26*(1), 35–48. https://doi.org/10.1080/030549800103845

Frostick, C., Tong, J., Moore, D., Renton, A., & Netuveli, G. (2018). The impact of academies on school connectedness, future aspirations and mental health in adolescents from areas of deprivation in London. *Pastoral Care in Education*, *36*(4), 325–342. https://doi.org/10.1080/02643944.2018.1528626

Gabay-Egozi, L., Shavit, Y., & Yaish, M. (2010). Curricular Choice: A Test of a Rational Choice Model of Education. *European Sociological Review*, *26*(4), 447–463. https://doi.org/10.1093/esr/jcp031

Gelman, A., Hill, J., & Yajima, M. (2012). Why We (Usually) Don’t Have to Worry About Multiple Comparisons. *Journal of Research on Educational Effectiveness*, *5*(2), 189–211. https://doi.org/10.1080/19345747.2011.618213

Gibbons, S., Machin, S., & Silva, O. (2008). Choice, Competition, and Pupil Achievement. *Journal of the European Economic Association*, *6*(4), 912–947. https://doi.org/10.1162/JEEA.2008.6.4.912

Gibbons, S., Machin, S., & Silva, O. (2013). Valuing school quality using boundary discontinuities. *Journal of Urban Economics*, *75*, 15–28. https://doi.org/10.1016/j.jue.2012.11.001

Gibbons, S., & Silva, O. (2008). Urban density and pupil attainment. *Journal of Urban Economics*, *63*(2), 631–650. https://doi.org/10.1016/j.jue.2007.04.006

Gibson, M. T. (2015). Ethos and vision realization in sponsored academy schools. *Management in Education*, *29*(1), 9–13. https://doi.org/10.1177/0892020614559237

Gillies, R. M., & Boyle, M. (2010). Teachers' reflections on cooperative learning: Issues of implementation. *Teaching and Teacher Education*, *26*(4), 933–940. https://doi.org/10.1016/j.tate.2009.10.034

Goodman, A., Lamping, D. L., & Ploubidis, G. B. (2010). When to Use Broader Internalising and Externalising Subscales Instead of the Hypothesised Five Subscales on the Strengths and Difficulties Questionnaire (SDQ): Data from British Parents, Teachers and Children. *Journal of Abnormal Child Psychology*, *38*(8), 1179–1191. https://doi.org/10.1007/s10802-010-9434-x

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry*, *38*(5), 581–586.

Goodman, R. (2001). Psychometric Properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(11), 1337–1345. https://doi.org/10.1097/00004583-200111000-00015

Graaff, E., & Kolmos, A. (2003). Characteristics of problem-based learning. *International Journal of Engineering Education*, *19*(5), 657–662.

Granoulhac, F. (2017). Excellence at What Cost? Austerity and the Reform of the School System in England (2010-2016). *The Political Quarterly*, *88*(3), 434–442. https://doi.org/10.1111/1467-923X.12375

Greany, T., Cowhitt, T., & Downey, C. (2023). Do informal networks become formalised over time? Analysing school networks and multi-academy trust membership in England using ego-centric analysis. *Journal of Educational Change*. https://doi.org/10.1007/s10833-023-09485-7

Greany, T., & Waterhouse, J. (2016). Rebels against the system: Leadership agency and curriculum innovation in the context of school autonomy and accountability in England. *International Journal of Educational Management*, *30*(7), 1188–1206. https://doi.org/10.1108/IJEM-11-2015-0148

Green, F., Allen, R., & Jenkins, A. (2015). Are English free schools socially selective? A quantitative analysis. *British Educational Research Journal*, *41*(6), 907–924. https://doi.org/10.1002/berj.3190

Green, F., Machin, S., Murphy, R., & Zhu, Y. (2011). The Changing Economic Advantage from Private Schools. *Economica*, *79*, 658–679. https://doi.org/10.1111/j.1468-0335.2011.00908.x

Hagenaars, A., Vos, K. de, & Zaidi, M. A. (with Europäische Kommission). (1994). *Poverty statistics in the late 1980s: Research based on micro-data*. Off. of Official Publ. of the Europ. Communities.

Hanushek, E. A., Link, S., & Woessmann, L. (2013). Does school autonomy make sense everywhere? Panel estimates from PISA. *Journal of Development Economics*, *104*, 212–232. https://doi.org/10.1016/j.jdeveco.2012.08.002

Harris, N. (2012). Local Authorities and the Accountability Gap in a Fragmenting Schools System: Local Authorities, Accountability and Schools. *The Modern Law Review*, *75*(4), 511–546. https://doi.org/10.1111/j.1468-2230.2012.00913.x

Harrison, A., Francesconi, M., & Flouri, E. (2022). Does decision-making at age 11 predict prodromal eating pathology at ages 14 and 17? A prospective, observational, UK population-based cohort study. *BMJ Open*, *12*(9), e051351. https://doi.org/10.1136/bmjopen-2021-051351

Hatton, A., Hampson, R., & Drake, R. (2019). *An analysis of the performance of sponsored academies* [Analytical research report]. Department for Education.

He, Q., Stockford, I., & Meadows, M. (2018). Inter-subject comparability of examination standards in GCSE and GCE in England. *Oxford Review of Education*, *44*(4), 494–513. https://doi.org/10.1080/03054985.2018.1430562

Healey, K. B. (2022). Engaging parents in multi academy trusts (MATs) as local civic knowledge experts. *Management in Education*, 089202062211447. https://doi.org/10.1177/08920206221144704

Hedges, S., & Speckesser, S. (2017). *Peer Effects and Social Influence in Post-16 Educational Choice* (Discussion Paper 008; p. 20). Centre for Vocational Education Research.

Henderson, M., Anders, J., Green, F., & Henseke, G. (2020). Private schooling, subject choice, upper secondary attainment and progression to university. *Oxford Review of Education*, *46*(3), 295–312. https://doi.org/10.1080/03054985.2019.1669551

Henderson, M., Sullivan, A., Anders, J., & Moulton, V. (2018). Social Class, Gender and Ethnic Differences in Subjects Taken at Age 14. *The Curriculum Journal*, *29*(3), 298–318. https://doi.org/10.1080/09585176.2017.1406810

Hetherington, J. E., & Forrester, G. (2023). Brand advantage, risk mitigation, and the illusion of democracy: Approaches to school governance. *Educational Management Administration & Leadership*, 17411432231194852. https://doi.org/10.1177/17411432231194852

Hicks, T. (2014). Partisan Governance and Policy Implementation: The Politics of Academy Conversion Amongst English Schools. *Public Administration*, *92*(4), 995–1016. https://doi.org/10.1111/padm.12100

Hicks, T. (2015). Inequality, marketisation and the left: Schools policy in England and Sweden: Inequality, marketisation and the left. *European Journal of Political Research*, *54*(2), 326–342. https://doi.org/10.1111/1475-6765.12086

Hicks, T. (2017). *Consequences of School Autonomy: Evidence from Mass "Academy" Conversion in the English Schooling System* (p. 72).

Higham, R. (2014). 'Who owns our schools?' An analysis of the governance of free schools in England. *Educational Management Administration & Leadership*, *42*(3), 404–422. https://doi.org/10.1177/1741143214522203

Higham, R., & Earley, P. (2013). School Autonomy and Government Control: School Leaders' Views on a Changing Policy Landscape in England. *Educational Management Administration & Leadership*, *41*(6), 701–717. https://doi.org/10.1177/1741143213494191

Hmelo-Silver, C. E. (2004). Problem-Based Learning: What and How Do Students Learn? *Educational Psychology Review*, *16*(3), 235–266. https://doi.org/10.1023/B:EDPR.0000034022.16470.f3

Holmlund, H., & Silva, O. (2014). Targeting Noncognitive Skills to Improve Cognitive Outcomes: Evidence from a Remedial Education Intervention. *Journal of Human Capital*, *8*(2), 126–160. https://doi.org/10.1086/676460

Hutchings, M., & Francis, B. (2018). *Chain effects 2018: The impact of academy chains on low-income pupils*. The Sutton Trust.

Iannelli, C. (2013). The role of the school curriculum in social mobility. *British Journal of Sociology of Education*, *34*(5–6), 907–928. https://doi.org/10.1080/01425692.2013.816031

Jæger, M. M. (2007). Economic and Social Returns To Educational Choices: Extending the Utility Function. *Rationality and Society*, *19*(4), 451–483. https://doi.org/10.1177/1043463107083739

Jerrim, J., Lopez-Agudo, L. A., Marcenaro-Gutierrez, O. D., & Shure, N. (2017). What happens when econometrics and psychometrics collide? An example using the PISA data. *Economics of Education Review*, *61*, 51–58. https://doi.org/10.1016/j.econedurev.2017.09.007

Jerrim, J., & Shure, N. (2016). *Achievement of 15-Year- Olds in England: PISA 2015 National Report* (p. 246). Department for Education.

Jin, W., Muriel, A., & Sibieta, L. (2011). *Subject and course choices at ages 14 and 16 amongst young people in England: Insights from behavioural economics* (DFE-RR160). Department for Education.

Johnson, D. W., & Johnson, R. T. (2014). Cooperative Learning in 21st Century. [Aprendizaje cooperativo en el siglo XXI]. *Anales de Psicología*, *30*(3), 841–851. https://doi.org/10.6018/analesps.30.3.201241

Kautz, T., Heckman, J., Diris, R., Weel, B., & Borghans, L. (2014). *Fostering and Measuring Skills: Improving Cognitive and Non-cognitive Skills to Promote*

Lifetime Success (OECD Education Working Papers 110; OECD Education Working Papers, Vol. 110). https://doi.org/10.1787/5jxsr7vr78f7-en

Keddie, A. (2019). Context *matters*: Primary schools and academies reform in England. *Journal of Education Policy*, *34*(1), 6–21. https://doi.org/10.1080/02680939.2017.1402959

Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., Howes, M. J., Normand, S.-L. T., Manderscheid, R. W., Walters, E. E., & Zaslavsky, A. M. (2003). Screening for Serious Mental Illness in the General Population. *ARCH GEN PSYCHIATRY*, *60*, 6.

Kolenikov, S. (2010). Resampling Variance Estimation for Complex Survey Data. *The Stata Journal: Promoting Communications on Statistics and Stata*, *10*(2), 165–199. https://doi.org/10.1177/1536867X1001000201

Kraft, M. A. (2020). Interpreting Effect Sizes of Education Interventions. *Educational Researcher*, *49*(4), 241–253. https://doi.org/10.3102/0013189X20912798

Krowka, S., Hadd, A., & Marx, R. (2017). *“No excuses” charter schools for increasing math and literacy achievement in primary and secondary education*. The Campbell Collaboration.

Leaver, C., Lemos, R., & Scur, D. (2022). *Understanding school management with public data: A new measurement approach and applications* (p. 82).

Leroux, G. (2015). *Choosing to succeed. Do parents pick the right schools?* Social Market Foundation.

Leuven, E., & Sianesi, B. (2003). *PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing* (Version 3.0.0.) [Computer software]. http://ideas.repec.org/c/boc/bocode/s432001.html

Liu, Y., Bessudnov, A., Black, A., & Norwich, B. (2020). School autonomy and educational inclusion of children with special needs: Evidence from England. *British Educational Research Journal*, *46*(3), 532–552. https://doi.org/10.1002/berj.3593

Long, R. (2017, March). *GCSE, AS and A level reform (England)*. UK Parliament.

Long, R. (2019a, September). *English Baccalaureate*. UK Parliament.

Long, R. (2019b, December). *School Governance*. UK Parliament.

Long, R., Danechi, S., & Maisuria, A. (2023). *Grammar schools in England* (Briefing Paper 07070). House of Commons Library.

Lucas, S. R. (2001). Effectively Maintained Inequality: Education Transitions, Track Mobility, and Social Background Effects. *American Journal of Sociology*, *106*(6), 1642–1690. https://doi.org/10.1086/321300

Machin, S., & McNally, S. (2012). The Evaluation of English Education Policies. *National Institute Economic Review*, *219*, 815–825.

Machin, S., & Sandi, M. (2020). Autonomous Schools and Strategic Pupil Exclusion. *The Economic Journal*, *130*(625), 125–159. https://doi.org/10.1093/ej/uez041

Machin, S., & Silva, O. (2013). *School Structure, School Autonomy and the Tail* (Special Paper 29; CEP Special Paper, p. 18). CEP.

Male, T. (2022). The rise and rise of academy trusts: Continuing changes to the state-funded school system in England. *School Leadership & Management*, 1–21. https://doi.org/10.1080/13632434.2022.2095996

Marcus, J., Reif, S., Wuppermann, A., & Rouche, A. (2020). Increased instruction time and stress-related health problems among school children. *Journal of Health Economics*, *70*, 102256. https://doi.org/10.1016/j.jhealeco.2019.102256

Martindale, N. (2019). Does outsourcing school systems degrade education workforces? Evidence from 18,000 English state schools. *British Journal of Sociology of Education*, *40*(8), 1015–1036. https://doi.org/10.1080/01425692.2019.1647092

Martindale, N. (2022). Austerity, outsourcing and the state school workforce: Trends from 20,000 English state schools. *British Journal of Sociology of Education*, *43*(3), 451–474. https://doi.org/10.1080/01425692.2021.2018650

Ministry of Housing, Communities and Local Government. (2020). *English Housing Survey. Headline Report, 2018-19* (English Housing Survey). Ministry of Housing, Communities and Local Government.

Morris, J. P. (2020). Is this the culture of academies? Utilising the cultural web to investigate the organisational culture of an academy case study. *Educational Management Administration & Leadership*, *48*(1), 164–185. https://doi.org/10.1177/1741143218788580

Morris, R. (2014). The admissions criteria of secondary Free Schools. *Oxford Review of Education*, *40*(3), 389–409. https://doi.org/10.1080/03054985.2014.921614

Morris, R., & Perry, T. (2019). Private schools for free? Parents' reasons for choosing a new Free School for their child. *Cambridge Journal of Education*, *49*(5), 535–552. https://doi.org/10.1080/0305764X.2019.1571561

Moulton, V., Sullivan, A., Henderson, M., & Anders, J. (2018). Does what you study at age 14–16 matter for educational transitions post-16? *Oxford Review of Education*, *44*(1), 94–117. https://doi.org/10.1080/03054985.2018.1409975

National Audit Office. (2018). *Converting maintained schools to academies* (p. 59). The National Audit Office.

Neri, L., & Pasini, E. (2018). *Heterogeneous Effects of School Autonomy: The Case of English State Schools* (p. 40).

Neri, L., Pasini, E., & Silva, O. (2022). *The Organizational Economics of School Chains* (Discussion Paper 15442). IZA.

Norwich, B., & Black, A. (2015). The placement of secondary school students with Statements of special educational needs in the more diversified system of English secondary schooling: *British Journal of Special Education*, *42*(2), 128–151. https://doi.org/10.1111/1467-8578.12097

OECD. (2013). *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*. OECD. https://doi.org/10.1787/9789264190511-en

OECD. (2014). *PISA 2012 Results: Creative Problem Solving (Volume V): Students' Skills in Tackling Real-Life Problems*. OECD. https://doi.org/10.1787/9789264208070-en

OECD. (2016a). *PISA 2015. [Data collection].*

OECD. (2016b). *PISA 2015 Results (Volume I): Excellence and Equity in Education*. OECD. https://doi.org/10.1787/9789264266490-en

OECD. (2016c). *PISA 2015 Results (Volume II): Policies and Practices for Successful Schools*. OECD. https://doi.org/10.1787/9789264267510-en

OECD. (2017a). *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic, Financial Literacy and Collaborative Problem Solving*. OECD. https://doi.org/10.1787/9789264281820-en

OECD. (2017b). *PISA 2015 Results (Volume III): Students' Well-Being*. OECD. https://doi.org/10.1787/9789264273856-en

OECD. (2017c). *PISA 2015 Results (Volume V): Collaborative Problem Solving*. OECD. https://doi.org/10.1787/9789264285521-en

OECD. (2017d). *PISA 2015 Technical Report*.

OECD. (2019). *Balancing School Choice and Equity: An International Perspective Based on Pisa*. OECD. https://doi.org/10.1787/2592c974-en

Office for National Statistics. (2014). *2011 Census Variable and Classification Information*.

Office of the Deputy Prime Minister & Neighbourhood Renewal Unit. (2004). *The English indices of deprivation 2004.* Office of the Deputy Prime Minister.

Opposs, D. (2016). *A policy position for Ofqual on inter-subject comparability* (58/16; Ofqual Board). Ofqual.

Patalay, P., & Fitzsimons, E. (2016). Correlates of Mental Illness and Wellbeing in Children: Are They the Same? Results From the UK Millennium Cohort Study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *55*(9), 771–783. https://doi.org/10.1016/j.jaac.2016.05.019

Pawson, E., Fournier, E., Haigh, M., Muniz, O., Trafford, J., & Vajoczki, S. (2006). Problem-based Learning in Geography: Towards a Critical Assessment of its Purposes, Benefits and Risks. *Journal of Geography in Higher Education*, *30*(1), 103–116. https://doi.org/10.1080/03098260500499709

Pinto, J., & Santos, L. (2006). *Modelos de avaliação das aprendizagens*. Universidade Aberta.

Platt, L., & Centre for Longitudinal Studies. (2014). *Millennium Cohort Study: Initial findings from the age 11 survey*. Centre for Longitudinal Studies, Institution of Education.

http://www.cls.ioe.ac.uk/page.aspx?&sitesectionid=1330&sitesectiontitle=MCS+age+11+initial+findings

Pring, R. (2018). Philosophical debates on curriculum, inequalities and social justice. *Oxford Review of Education*, *44*(1), 6–18. https://doi.org/10.1080/03054985.2018.1409963

Rayner, S. M., Courtney, S. J., & Gunter, H. M. (2018). Theorising systemic change: Learning from the academisation project in England. *Journal of Education Policy*, *33*(1), 143–162. https://doi.org/10.1080/02680939.2017.1327084

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. https://doi.org/10.1016/j.econedurev.2018.02.004

Ridgeway, G., Kovalchik, S. A., Griffin, B. A., & Kabeto, M. U. (2015). Propensity Score Analysis with Survey Weighted Data. *Journal of Causal Inference*, *3*(2), 237–249. https://doi.org/10.1515/jci-2014-0039

Rogers, R. D., Owen, A. M., Middleton, H. C., Williams, E. J., Pickard, J. D., Sahakian, B. J., & Robbins, T. W. (1999). Choosing between Small, Likely Rewards and Large, Unlikely Rewards Activates Inferior and Orbital Prefrontal Cortex. *The Journal of Neuroscience*, *20*(19), 9029–9038.

Rose, D., & Pevalin, D. J. (Eds.). (2003). *A researcher's guide to the national statistics socio-economic classification*. Sage Publications.

Rosenberg, M. (1965). *Society and the Adolescent Self-Image*. Princeton University Press. http://www.jstor.org/stable/j.ctt183pjjh

Salokangas, M., & Chapman, C. (2014). Exploring governance in two chains of academy schools: A comparative case study. *Educational Management*

Administration & Leadership, *42*(3), 372–386. https://doi.org/10.1177/1741143214521588

Sass, T. R., Zimmer, R. W., Gill, B. P., & Booker, T. K. (2016). Charter High Schools' Effects on Long-Term Attainment and Earnings. *Journal of Policy Analysis and Management*, *35*(3), 683–706. https://doi.org/10.1002/pam.21913

Schul, J. E. (2011). Revisiting an Old Friend: The Practice and Promise of Cooperative Learning for the Twenty-First Century. *The Social Studies*, *102*(2), 88–93. https://doi.org/10.1080/00377996.2010.509370

Shah, Q. (2018). Academisation: Who now holds the key to the secret garden? *Power and Education*, *10*(3), 215–232. https://doi.org/10.1177/1757743818764746

Sharan, Y. (2014). Learning to cooperate for cooperative learning. [Aprendiendo a cooperar en el aprendizaje cooperativo]. *Anales de Psicología*, *30*(3), 802–807. https://doi.org/10.6018/analesps.30.3.201211

Sharpe, D. (2015). Chi-Square Test is Statistically Significant: Now What? *Practical Assessment, Research, and Evaluation*, *20*(8). https://doi.org/10.7275/TBFA-X148

Siddiqui, N., & Ventista, O. M. (2018). A review of school-based interventions for the improvement of social emotional skills and wider outcomes of education. *International Journal of Educational Research*, *90*, 117–132. https://doi.org/10.1016/j.ijer.2018.06.003

Simkins, T. (2015). School restructuring in England: New school configurations and new challenges. *Management in Education*, *29*(1), 4–8. https://doi.org/10.1177/0892020614559235

Simon, C. A., James, C., & Simon, A. (2021). The Growth of Multi-Academy Trusts in England: Emergent Structures and the Sponsorship of Underperforming

Schools. *Educational Management Administration & Leadership*, *49*(1), 112–127. https://doi.org/10.1177/1741143219893099

Smyth, E. (2018). Working at a different level? Curriculum differentiation in Irish lower secondary education. *Oxford Review of Education*, *44*(1), 37–55. https://doi.org/10.1080/03054985.2018.1409967

Sullivan, A. (2009). Academic self-concept, gender and single-sex schooling. *British Educational Research Journal*, *35*(2), 259–288. https://doi.org/10.1080/01411920802042960

Sullivan, A., Joshi, H., & Leonard, D. (2010). Single-Sex Schooling and Academic Attainment at School and Through the Lifecourse. *American Educational Research Journal*, *47*(1), 6–36. https://doi.org/10.3102/0002831209350106

Sullivan, A., Ketende, S., & Joshi, H. (2013). Social Class and Inequalities in Early Cognitive Scores. *Sociology*, *47*(6), 1187–1206. https://doi.org/10.1177/0038038512461861

Sullivan, A., Parsons, S., Wiggins, R., Heath, A., & Green, F. (2014). Social origins, school type and higher education destinations. *Oxford Review of Education*, *40*(6), 739–763. https://doi.org/10.1080/03054985.2014.979015

Sullivan, A., Zimdars, A., & Heath, A. (2010). The social structure of the 14–16 curriculum in England. *International Studies in Sociology of Education*, *20*(1), 5–21. https://doi.org/10.1080/09620211003655572

Telhaj, S., Machin, S. J., & Eyles, A. (2022). “Super Heads”, School Autonomy and Leadership Pay. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4035237

Todd, P. E., & Wolpin, K. I. (2003). On the Specification and Estimation of the Production Function for Cognitive Achievement. *The Economic Journal*, *113*(485), F3–F33. https://doi.org/10.1111/1468-0297.00097

Tripney, J., Newman, M., Bangpan, M., Niza, C., MacKintosh, M., & Sinclair, J. (2010). *Subject Choice in STEM: Factors influencing young people (aged 14-19) in education : a systematic review of the UK literature.* Wellcome Trust. http://wellcomelibrary.org/item/b21248758

University of London, Institute of Education, Centre for Longitudinal Studies. (2020a). Millennium Cohort Study: Fourth Survey, 2008. [Data collection]. 8th Edition. *UK Data Service*, *SN: 6411*. https://doi.org/10.5255/UKDA-SN-6411-8

University of London, Institute of Education, Centre for Longitudinal Studies. (2020b). Millennium Cohort Study: Seventh Survey, 2018. [Data collection]. *UK Data Service*, *SN: 8682*. https://doi.org/10.5255/UKDA-SN-8682-1

University of London, Institute of Education, Centre for Longitudinal Studies. (2020c). Millennium Cohort Study: Sixth Survey, 2015. [Data collection]. 7th Edition. *UK Data Service*, *SN: 8156*. https://doi.org/10.5255/UKDA-SN-8156-7

University of London, Institute of Education, Centre for Longitudinal Studies. (2020d). Millennium Cohort Study: Third Survey, 2006. [Data collection]. 8th Edition. *UK Data Service*, *SN: 5795*. https://doi.org/10.5255/UKDA-SN-5795-5

University of London, Institute of Education, Centre for Longitudinal Studies. (2022). Millennium Cohort Study: Fifth Survey, 2012. [Data collection]. 6th Edition. *UK Data Service*, *SN: 7464*. https://doi.org/10.5255/UKDA-SN-7464-5

VanderWeele, T. J., & Mathur, M. B. (2019). Some desirable properties of the bonferroni correction: Is the bonferroni correction really so bad? *American Journal of Epidemiology*, *188*(3), 617–618. https://doi.org/10.1093/aje/kwy250

Vidal Rodeiro, C., & Vitello, S. (2021). Progression to post-16 education in England: The role of vocational qualifications. *Research Papers in Education*, 1–23. https://doi.org/10.1080/02671522.2021.1961295

Vidal Rodeiro, C., & Williamson, J. (2019). Meaningful destinations: Using national data to investigate how different education pathways support young people's progression in England. *Research Papers in Education*, *34*(6), 725–748. https://doi.org/10.1080/02671522.2018.1536889

Walford, G. (2014). From city technology colleges to free schools: Sponsoring new schools in England. *Research Papers in Education*, *29*(3), 315–329. https://doi.org/10.1080/02671522.2014.885731

Walters, C. R. (2018). The Demand for Effective Charter Schools. *Journal of Political Economy*, *126*(6), 2179–2223.

Werfhorst, H. G. V. D., Sullivan, A., & Cheung, S. Y. (2003). Social Class, Ability and Choice of Subject in Secondary and Tertiary Education in Britain. *British Educational Research Journal*, *29*(1), 41–62.

West, A. (2014). Academies in England and independent schools (*fristående skolor*) in Sweden: Policy, privatisation, access and segregation. *Research Papers in Education*, *29*(3), 330–350. https://doi.org/10.1080/02671522.2014.885732

West, A. (2015). Education policy and governance in England under the Coalition Government (2010–15): Academies, the pupil premium, and free early education. *London Review of Education*, *13*(2), 21–36. https://doi.org/10.18546/LRE.13.2.03

West, A., & Bailey, E. (2013). The Development of the Academies Programme: ‘Privatising’ School-Based Education in England 1986–2013. *British Journal of Educational Studies*, *61*(2), 137–159. https://doi.org/10.1080/00071005.2013.789480

West, A., & Nikolai, R. (2017). The Expansion of “Private” Schools in England, Sweden and Eastern Germany: A Comparative Perspective on Policy Development, Regulation, Policy Goals and Ideas. *Journal of Comparative Policy Analysis: Research and Practice*, *19*(5), 452–469. https://doi.org/10.1080/13876988.2016.1262549

West, A., & Wolfe, D. B. (2019). Academies, autonomy, equality and democratic accountability: Reforming the fragmented publicly funded school system in England. *London Review of Education*. https://doi.org/10.18546/LRE.17.1.06

West, M. R., Kraft, M. A., Finn, A. S., Martin, R. E., Duckworth, A. L., Gabrieli, C. F. O., & Gabrieli, J. D. E. (2016). Promise and Paradox: Measuring Students’ Non-Cognitive Skills and the Impact of Schooling. *Educational Evaluation and Policy Analysis*, *38*(1), 148–170. https://doi.org/10.3102/0162373715597298

Wiborg, S., Green, F., Taylor-Gooby, P., & Wilde, R. J. (2018). Free Schools in England: ‘Not Unlike other Schools’? *Journal of Social Policy*, *47*(1), 119–137. https://doi.org/10.1017/S004727941700023X

Wilkerson, L. (1995). Identification of skills for the problem-based tutor: Student and faculty perspectives. *Instructional Science*, *22*(4), 303–315. https://doi.org/10.1007/BF00891783

Wilkins, A. (2017). Rescaling the local: Multi-academy trusts, private monopoly and statecraft in England. *Journal of Educational Administration and History*, *49*(2), 171–185. https://doi.org/10.1080/00220620.2017.1284769

Williams, B. (2023). The 2010–15 coalition government and the legacy of free schools in England. *Policy Futures in Education*, 14782103231176360. https://doi.org/10.1177/14782103231176360

Wilson, K., Wolfe, B., & Haveman, R. (2005). The Role of Expectations in Adolescent Schooling Choices: Do Youths Respond to Economic Incentives? *Economic Inquiry*, *43*(3), 467–492. https://doi.org/10.1093/ei/cbi032

Wolf, A. (2011). *Review of Vocational Education – The Wolf Report* (DFE-00031). Department for Education.

Woods, P. (1976). The Myth of Subject Choice. *The British Journal of Sociology*, *27*(2), 130. https://doi.org/10.2307/590022

Woods, P., & Simkins, T. (2014). Understanding the local: Themes and Issues in the Experience of Structural Reform in England. *Educational Management Administration & Leadership*, *42*(3), 324–340. https://doi.org/10.1177/1741143214521587

Wooldridge, J. M. (2014). *Introduction to econometrics* (Europe, Middle East and Africa ed). Cengage Learning.

Wright, S. P. (1992). Adjusted P-Values for Simultaneous Inference. *Biometrics*, *48*(4), 1005. https://doi.org/10.2307/2532694

Zeldow, B., & Hatfield, L. A. (2021). Confounding and regression adjustment in DIFFERENCE-IN-DIFFERENCES studies. *Health Services Research*, *56*(5), 932–941. https://doi.org/10.1111/1475-6773.13666

APPENDICES

APPENDIX A SUPPLEMENTAL MATERIAL TO CHAPTER 3

Table A1 Sample selection checks

PROBIT INDEPENDENT VARIABLES	SAMPLE SELECTION z-scores
CM ethnic group (ref. White)	
Mixed	-0.110
	(0.0957)
Indian	-0.273
	(0.207)
Pakistani and Bangladeshi	0.212*
	(0.123)
Black	-0.253*
	(0.145)
Other ethnic group	-0.0528
	(0.128)
SEN statement (age 11)	0.137
	(0.1000)
Free School Meal Eligible (age 11)	0.0636
	(0.0601)
Occupation (Main) (ref. Managerial and professional)	
Intermediate	0.0644
	(0.0554)
Small employers and self-employed	0.0539
	(0.0815)
Lower supervisory and technical	-0.0570
	(0.142)
Semi-routine and routine	0.224***
	(0.0670)
Not in work nor on leave	-0.0841
	(0.0651)
Occupation (Partner) (ref. Managerial and professional)	
Intermediate	0.0941
	(0.100)
Small employers and self-employed	0.0480
	(0.0658)
Lower supervisory and technical	0.188**
	(0.0882)
Semi-routine and routine	0.0514
	(0.0709)
Not in work nor on leave	0.0985
	(0.0827)
Education (Main) (ref. NVQ equiv. level 1)	
NVQ equiv. level 2 (incl. GSCE A* – C)	-0.133**
	(0.0667)
NVQ equiv. level 3 (incl. A levels)	-0.0930
	(0.0763)
NVQ equiv. level 4 (incl.BA degree)	-0.181**
	(0.0713)
NVQ equiv. level 5 (incl. postgraduate degree)	-0.135
	(0.102)
Education (Partner) (ref. NVQ equiv. level 1)	
NVQ equiv. level 2 (incl. GSCE A* – C)	0.109
	(0.108)
NVQ equiv. level 3 (incl. A levels)	0.0152
	(0.118)

PROBIT INDEPENDENT VARIABLES	SAMPLE SELECTION z-scores
NVQ equiv. level 4 (incl.BA degree)	-0.0772
	(0.102)
NVQ equiv. level 5 (incl. postgraduate degree)	-0.210*
	(0.116)
IMD Income Domain (ref. Most Deprived Decile)	
Least Deprived Decile	-0.394***
	(0.133)
ONS Urban Code	0.241**
	(0.112)
Word Reading Standardised (age 7)	-0.00281**
	(0.00124)
SDQ Total Difficulties (age 7)	-0.000445
	(0.0226)
Observations	**7,814**

Note: The quantity in parenthesis below the coefficient is the (robust) standard error reflecting MCS survey settings. A dummy variable identifying missing cases for each variable is included. (University of London, Institute of Education, Centre for Longitudinal Studies, 2020a, 2022). * significant at 10% level; ** significant at 5% level; *** significant at 1% level or below.

Table A2 Descriptive statistics pre-treatment (all explanatory variables)

	CONTROL		TREATMENT			SPONSOR-LED			CONVERTER		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
VARIABLES	**N**	**Mean**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**
AGE 11											
CM CHARACTERISTICS											
CM Sex (Male)	2,563	0.531	562	0.525	-0.00525	199	0.480	-0.0509	363	0.555	0.0247
		[0.502]		[0.498]	(0.0368)		[0.471]	(0.0582)		[0.513]	(0.0386)
Age at MCS5 interview	2,562	11.14	562	11.17	0.0296	199	11.17	0.0298	363	11.17	0.0294
(10th of year)		[0.342]		[0.317]	(0.0215)		[0.301]	(0.0275)		[0.326]	(0.0284)
CM ethnic group	2,476		543			195			348		
White		0.775		0.819	0.0432		0.822	0.0468		0.816	0.0408
		[0.420]		[0.384]	(0.0443)		[0.360]	(0.0598)		[0.399]	(0.0554)
Mixed		0.059		0.041	-0.0185		0.038	-0.0210		0.042	-0.0169
		[0.238]		[0.197]	(0.0133)		[0.181]	(0.0178)		[0.208]	(0.0163)
Indian		0.023		0.011	-0.0125		0.006	-0.0171**		0.014	-0.00939
		[0.151]		[0.102]	(0.00819)		[0.072]	(0.00851)		[0.119]	(0.00960)
Pakistani and Bangladeshi		0.071		0.087	0.0158		0.089	0.0175		0.086	0.0147
		[0.259]		[0.281]	(0.0379)		[0.268]	(0.0538)		[0.289]	(0.0490)
Black		0.042		0.023	-0.0190		0.023	-0.0186		0.023	-0.0192
		[0.202]		[0.149]	(0.0116)		[0.142]	(0.0141)		[0.154]	(0.0148)
Other ethnic group		0.029		0.020	-0.00910		0.021	-0.00761		0.019	-0.0101
		[0.169]		[0.139]	(0.00698)		[0.136]	(0.00792)		[0.140]	(0.00871)
Language spoken household	2,563	0.875	562	0.886	0.0114	199	0.898	0.0231	363	0.879	0.00369
(English only)		[0.333]		[0.317]	(0.0397)		[0.285]	(0.0565)		[0.337]	(0.0483)
SEN statement	2,558	0.056	560	0.035	-0.0209	199	0.014	-0.042***	361	0.049	-0.00704
		[0.232]		[0.184]	(0.0131)		[0.111]	(0.00881)		[0.223]	(0.0195)
CM's general level of health	2,563		562			199			363		
Excellent		0.526		0.543	0.0167		0.479	-0.0479		0.586	0.0591*
		[0.503]		[0.497]	(0.0303)		[0.471]	(0.0480)		[0.509]	(0.0349)
Very Good		0.320		0.287	-0.0333		0.320	6.33e-05		0.265	-0.0551
		[0.470]		[0.452]	(0.0274)		[0.440]	(0.0394)		[0.456]	(0.0335)
Good		0.117		0.114	-0.00272		0.120	0.00350		0.110	-0.00681
		[0.323]		[0.317]	(0.0177)		[0.307]	(0.0284)		[0.323]	(0.0228)
Fair		0.029		0.048	0.0186*		0.066	0.0371**		0.036	0.00655
		[0.170]		[0.213]	(0.0111)		[0.235]	(0.0172)		[0.192]	(0.0124)

VARIABLES	CONTROL (1) N	(2) Mean	TREATMENT (3) N	(4) Mean	(5) diff	SPONSOR-LED (6) N	(7) Mean	(8) diff	CONVERTER (9) N	(10) Mean	(11) diff
Poor		0.007		0.008	0.000617		0.014	0.00721		0.003	-0.00371
		[0.085]		[0.088]	(0.00499)		[0.112]	(0.0109)		[0.061]	(0.00348)
COGNITIVE SKILLS											
SWM Strategy	2,423	0.006	542	0.028	0.0214	195	0.116	0.110	347	-0.033	-0.0389
		[0.975]		[0.911]	(0.0522)		[0.800]	(0.0795)		[0.981]	(0.0621)
SWM Total errors	2,423	0.004	542	0.022	0.0186	195	0.074	0.0705	347	-0.013	-0.0167
		[0.996]		[0.933]	(0.0570)		[0.890]	(0.0958)		[0.959]	(0.0705)
FAMILY CONTEXT											
One parent/carer in Household	2,563	0.306	562	0.269	-0.0368	199	0.325	0.0197	363	0.232	-0.074***
		[0.464]		[0.443]	(0.0280)		[0.442]	(0.0467)		[0.436]	(0.0283)
No grandparents in Household	2,563	0.957	562	0.961	0.00455	199	0.975	0.0179	363	0.952	-0.00423
		[0.205]		[0.193]	(0.0107)		[0.148]	(0.0133)		[0.220]	(0.0145)
Number of siblings in Household	2,563	1.604	562	1.737	0.133	199	1.875	0.270**	363	1.647	0.0428
		[1.174]		[1.203]	(0.0832)		[1.138]	(0.125)		[1.236]	(0.103)
Age (Main)	2,563	39.007	562	38.543	-0.464	199	37.232	-1.775***	363	39.403	0.396
		[6.448]		[6.046]	(0.436)		[5.285]	(0.494)		[6.390]	(0.586)
Age (Partner)	1,711	41.957	387	41.540	-0.417	128	40.106	-1.851**	259	42.368	0.410
		[7.449]		[6.732]	(0.544)		[6.023]	(0.807)		[6.999]	(0.646)
Main has longstanding Illness	2,561	0.206	562	0.209	0.00350	199	0.242	0.0365	363	0.188	-0.0182
		[0.407]		[0.406]	(0.0223)		[0.404]	(0.0350)		[0.403]	(0.0224)
Main's mental health (Kessler 6 Scale)	2,372	0.099	520	0.073	-0.0265	183	0.258	0.159	337	-0.048	-0.148*
		[1.082]		[1.051]	(0.0715)		[1.121]	(0.133)		[0.963]	(0.0770)
SOCIOECONOMIC STATUS											
OECD equiv weekly family income (log)	2,563	5.829	562	5.813	-0.0159	199	5.698	-0.131*	363	5.889	0.0598
		[0.478]		[0.474]	(0.0487)		[0.432]	(0.0683)		[0.485]	(0.0659)
Free School Meal Eligible	2,555	0.248	560	0.220	-0.0278	198	0.284	0.0362	362	0.178	-0.0698**
		[0.435]		[0.413]	(0.0265)		[0.425]	(0.0547)		[0.395]	(0.0291)
Housing Tenure	2,537		557			196			361		
Rent privately		0.186		0.145	-0.0404*		0.158	-0.0273		0.137	-0.0489**
		[0.393]		[0.353]	(0.0224)		[0.348]	(0.0417)		[0.355]	(0.0212)
Owner occupier		0.511		0.560	0.0496		0.437	-0.0739*		0.639	0.129***
		[0.505]		[0.498]	(0.0351)		[0.472]	(0.0442)		[0.496]	(0.0426)
LA or HA rent		0.304		0.295	-0.00918		0.405	0.101**		0.224	-0.0797*
		[0.464]		[0.457]	(0.0354)		[0.467]	(0.0445)		[0.431]	(0.0440)

	CONTROL		TREATMENT			SPONSOR-LED			CONVERTER		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
VARIABLES	**N**	**Mean**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**
Combined labour market	2,563		562			199			363		
Both in work		0.433		0.445	0.0123		0.354	-0.0792		0.505	0.0724
		[0.499]		[0.496]	(0.0392)		[0.451]	(0.0588)		[0.516]	(0.0473)
Main in work, partner not		0.044		0.043	-0.000719		0.052	0.00855		0.037	-0.00681
		[0.205]		[0.202]	(0.0114)		[0.210]	(0.0166)		[0.194]	(0.0143)
Partner in work, main not		0.153		0.184	0.0310		0.195	0.0418		0.177	0.0240
		[0.362]		[0.387]	(0.0265)		[0.374]	(0.0481)		[0.394]	(0.0274)
Both not in work		0.065		0.059	-0.00589		0.074	0.00925		0.049	-0.0158
		[0.248]		[0.235]	(0.0155)		[0.247]	(0.0245)		[0.223]	(0.0206)
Main in work or on		0.171		0.155	-0.0163		0.161	-0.00989		0.151	-0.0204
leave, no partner		[0.379]		[0.361]	(0.0201)		[0.347]	(0.0284)		[0.370]	(0.0240)
Main not in work nor		0.134		0.114	-0.0205		0.164	0.0295		0.081	-0.0533**
on leave, no partner		[0.343]		[0.317]	(0.0246)		[0.349]	(0.0391)		[0.282]	(0.0243)
Full-time work (Main)	2,552		557			198			359		
Part-time work		0.389		0.388	-0.000444		0.365	-0.0233		0.403	0.0147
		[0.492]		[0.487]	(0.0333)		[0.455]	(0.0556)		[0.508]	(0.0388)
Full-time work		0.256		0.250	-0.00580		0.198	-0.0575*		0.284	0.0284
		[0.440]		[0.433]	(0.0248)		[0.377]	(0.0301)		[0.467]	(0.0311)
Not in work nor on		0.355		0.362	0.00624		0.436	0.0807		0.312	-0.0431
leave		[0.483]		[0.481]	(0.0396)		[0.468]	(0.0646)		[0.480]	(0.0480)
Full-time work (Partner)	1,698		383			128			255		
Part-time work		0.088		0.067	-0.0202		0.077	-0.0107		0.062	-0.0257
		[0.295]		[0.260]	(0.0159)		[0.263]	(0.0369)		[0.257]	(0.0173)
Full-time work		0.751		0.788	0.0365		0.720	-0.0315		0.827	0.0763*
		[0.452]		[0.425]	(0.0349)		[0.444]	(0.0505)		[0.404]	(0.0424)
Not in work nor on		0.161		0.145	-0.0163		0.204	0.0422		0.111	-0.0506
leave		[0.384]		[0.366]	(0.0286)		[0.398]	(0.0458)		[0.335]	(0.0381)
Occupation (Main)	2,519		553			197			356		
Managerial and		0.194		0.158	-0.0357*		0.096	-0.098***		0.199	0.00545
Professional		[0.399]		[0.364]	(0.0201)		[0.278]	(0.0202)		[0.413]	(0.0229)
Intermediate		0.150		0.154	0.00390		0.158	0.00828		0.151	0.000993
		[0.360]		[0.360]	(0.0206)		[0.344]	(0.0352)		[0.370]	(0.0205)
Small employers and		0.060		0.066	0.00658		0.021	-0.038***		0.096	0.0364
self-employed		[0.239]		[0.248]	(0.0188)		[0.136]	(0.00959)		[0.304]	(0.0274)

VARIABLES	CONTROL (1) N	(2) Mean	TREATMENT (3) N	(4) Mean	(5) diff	SPONSOR-LED (6) N	(7) Mean	(8) diff	CONVERTER (9) N	(10) Mean	(11) diff
Lower supervisory		0.017		0.028	0.0107		0.042	0.0245		0.019	0.00150
and technical		[0.131]		[0.164]	(0.00961)		[0.188]	(0.0195)		[0.140]	(0.00786)
Semi-routine and		0.219		0.232	0.0124		0.247	0.0277		0.222	0.00222
routine		[0.417]		[0.421]	(0.0256)		[0.406]	(0.0414)		[0.429]	(0.0303)
Not in work nor on		0.360		0.362	0.00214		0.435	0.0752		0.313	-0.0465
leave		[0.484]		[0.479]	(0.0396)		[0.467]	(0.0641)		[0.479]	(0.0484)
Occupation (Partner)	1,658		375			125			250		
Managerial and		0.325		0.308	-0.0163		0.160	-0.165***		0.395	0.0708
Professional		[0.489]		[0.478]	(0.0444)		[0.361]	(0.0311)		[0.521]	(0.0516)
Intermediate		0.056		0.052	-0.00325		0.085	0.0293		0.033	-0.0223
		[0.239]		[0.231]	(0.0160)		[0.274]	(0.0371)		[0.191]	(0.0145)
Small employers and		0.171		0.183	0.0125		0.188	0.0171		0.181	0.00986
self-employed		[0.393]		[0.401]	(0.0300)		[0.385]	(0.0470)		[0.410]	(0.0341)
Lower supervisory		0.091		0.130	0.0390*		0.164	0.0737**		0.109	0.0187
and technical		[0.300]		[0.348]	(0.0231)		[0.365]	(0.0362)		[0.333]	(0.0248)
Semi-routine and		0.193		0.179	-0.0143		0.196	0.00315		0.169	-0.0245
routine		[0.412]		[0.397]	(0.0284)		[0.391]	(0.0343)		[0.399]	(0.0371)
Not in work nor on		0.165		0.147	-0.0177		0.207	0.0419		0.112	-0.0525
leave		[0.387]		[0.367]	(0.0295)		[0.399]	(0.0461)		[0.336]	(0.0394)
Education (Main)	2,144		461			147			314		
NVQ equiv. level 1		0.124		0.116	-0.00826		0.136	0.0120		0.105	-0.0193
(incl. GSCE D – G)		[0.334]		[0.323]	(0.0183)		[0.330]	(0.0311)		[0.318]	(0.0215)
NVQ equiv. level 2		0.334		0.388	0.0536*		0.452	0.118**		0.353	0.0188
(incl. GSCE A* – C)		[0.479]		[0.492]	(0.0316)		[0.478]	(0.0510)		[0.495]	(0.0368)
NVQ equiv. level 3		0.175		0.193	0.0185		0.203	0.0286		0.188	0.0131
(incl. A levels)		[0.385]		[0.399]	(0.0241)		[0.387]	(0.0475)		[0.405]	(0.0237)
NVQ equiv. level 4		0.296		0.242	-0.0535*		0.181	-0.115***		0.275	-0.0202
(incl. BA degree)		[0.463]		[0.433]	(0.0290)		[0.370]	(0.0332)		[0.463]	(0.0347)
NVQ equiv. level 5		0.071		0.061	-0.0103		0.028	-0.043***		0.079	0.00764
(incl. postgraduate)		[0.261]		[0.242]	(0.0133)		[0.158]	(0.0136)		[0.279]	(0.0171)
Education (Partner)	1,383		326			98			228		
NVQ equiv. level 1		0.093		0.114	0.0210		0.146	0.0527		0.097	0.00412
(incl. GSCE D – G)		[0.302]		[0.330]	(0.0265)		[0.340]	(0.0495)		[0.318]	(0.0277)
NVQ equiv. level 2		0.364		0.272	-0.0920**		0.389	0.0250		0.209	-0.154***
(incl. GSCE A* – C)		[0.499]		[0.461]	(0.0397)		[0.470]	(0.0505)		[0.437]	(0.0428)

	CONTROL		TREATMENT			SPONSOR-LED			CONVERTER		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
VARIABLES	**N**	**Mean**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**
NVQ equiv. level 3		0.154		0.247	0.0928***		0.289	0.134***		0.225	0.0707*
(incl. A levels)		[0.375]		[0.447]	(0.0308)		[0.437]	(0.0450)		[0.449]	(0.0374)
NVQ equiv. level 4		0.298		0.255	-0.0434		0.120	-0.179***		0.327	0.0289
(incl. BA degree)		[0.475]		[0.452]	(0.0365)		[0.313]	(0.0281)		[0.504]	(0.0420)
NVQ equiv. level 5		0.090		0.112	0.0215		0.057	-0.0331		0.141	0.0507*
(incl. postgraduate)		[0.297]		[0.327]	(0.0228)		[0.224]	(0.0322)		[0.374]	(0.0264)
HOME ENVIRONMENT											
How many books at home	2,536		557			196			361		
0-10		0.158		0.203	0.0457		0.218	0.0604		0.194	0.0364
		[0.368]		[0.404]	(0.0310)		[0.393]	(0.0552)		[0.409]	(0.0298)
11-25		0.171		0.180	0.00827		0.231	0.0598		0.147	-0.0247
		[0.380]		[0.385]	(0.0263)		[0.401]	(0.0460)		[0.365]	(0.0265)
26-100		0.339		0.325	-0.0132		0.333	-0.00569		0.321	-0.0180
		[0.478]		[0.470]	(0.0244)		[0.449]	(0.0330)		[0.482]	(0.0300)
101-200		0.176		0.149	-0.0266		0.103	-0.074***		0.179	0.00331
		[0.385]		[0.357]	(0.0186)		[0.289]	(0.0233)		[0.396]	(0.0234)
201-500		0.115		0.103	-0.0119		0.090	-0.0255		0.112	-0.00317
		[0.322]		[0.305]	(0.0171)		[0.272]	(0.0243)		[0.326]	(0.0205)
More than 500		0.041		0.039	-0.00230		0.026	-0.0156		0.047	0.00617
		[0.201]		[0.194]	(0.00946)		[0.150]	(0.0158)		[0.219]	(0.0113)
Visited library at least monthly	2,563	0.279	562	0.222	-0.0575*	199	0.253	-0.0257	363	0.201	-0.0784**
		[0.452]		[0.414]	(0.0298)		[0.410]	(0.0433)		[0.413]	(0.0367)
Attended religious service monthly	2,561	0.275	562	0.198	-0.0769**	199	0.170	-0.104**	363	0.216	-0.0589
		[0.450]		[0.398]	(0.0360)		[0.355]	(0.0488)		[0.425]	(0.0459)
Watching TV or videos	2,563		562			199			363		
Up to one hour		0.126		0.133	0.00716		0.127	0.000925		0.137	0.0113
		[0.334]		[0.339]	(0.0214)		[0.314]	(0.0282)		[0.355]	(0.0281)
1 hour to less than 3 hours		0.686		0.707	0.0211		0.642	-0.0437		0.750	0.0636*
		[0.467]		[0.454]	(0.0315)		[0.452]	(0.0389)		[0.447]	(0.0339)
More than 3 hours		0.188		0.160	-0.0283		0.231	0.0427		0.113	-0.075***
		[0.393]		[0.366]	(0.0267)		[0.397]	(0.0425)		[0.327]	(0.0264)
Anyone helps with homework	2,559		561			199			362		
Always		0.213		0.220	0.00664		0.254	0.0409		0.198	-0.0159
		[0.413]		[0.413]	(0.0220)		[0.411]	(0.0362)		[0.411]	(0.0250)

VARIABLES	CONTROL (1) N	(2) Mean	TREATMENT (3) N	(4) Mean	(5) diff	SPONSOR-LED (6) N	(7) Mean	(8) diff	CONVERTER (9) N	(10) Mean	(11) diff
Usually		0.257		0.243	-0.0144		0.259	0.00161		0.232	-0.0249
		[0.440]		[0.428]	(0.0233)		[0.413]	(0.0385)		[0.436]	(0.0264)
Sometimes		0.413		0.420	0.00741		0.394	-0.0188		0.437	0.0247
		[0.496]		[0.492]	(0.0283)		[0.461]	(0.0396)		[0.512]	(0.0327)
Never or almost never		0.117		0.117	0.000349		0.093	-0.0237		0.133	0.0161
		[0.324]		[0.321]	(0.0222)		[0.274]	(0.0352)		[0.350]	(0.0228)
Regular bedtime on weekdays (usually, always)	2,563	0.869	562	0.904	0.0354*	199	0.887	0.0182	363	0.915	0.0467**
		[0.340]		[0.294]	(0.0186)		[0.299]	(0.0275)		[0.288]	(0.0212)
SCHOOL-LEVEL VARIABLES											
School fees applicable	2,483	0.006	549	0.007	0.00132	194	0.001	-0.0049**	355	0.011	0.00530
		[0.076]		[0.083]	(0.00390)		[0.027]	(0.00188)		[0.107]	(0.00604)
Single sex school	2,563	0.008	562	0.006	-0.00124	199	0.007	-0.00118	363	0.006	-0.00128
		[0.088]		[0.080]	(0.00426)		[0.076]	(0.00717)		[0.082]	(0.00473)
School with a particular faith or religion	2,477	0.266	549	0.250	-0.0159	194	0.164	-0.102***	355	0.306	0.0395
		[0.447]		[0.432]	(0.0416)		[0.351]	(0.0368)		[0.474]	(0.0566)
Attended more than 2 primary schools	2,245	0.058	506	0.046	-0.0116	181	0.039	-0.0189	325	0.051	-0.00665
		[0.238]		[0.212]	(0.0115)		[0.184]	(0.0163)		[0.230]	(0.0170)
GEOGRAPHICAL VARIABLES											
IMD Income Domain	2,555		562			199			363		
Most Deprived Decile		0.159		0.152	-0.00736		0.209	0.0491		0.115	-0.0444
		[0.369]		[0.358]	(0.0334)		[0.383]	(0.0702)		[0.329]	(0.0321)
Least Deprived Decile		0.056		0.063	0.00682		0.014	-0.042***		0.095	0.0388
		[0.231]		[0.242]	(0.0193)		[0.110]	(0.0121)		[0.302]	(0.0245)
IMD Health Deprivation	2,555		562			199			363		
Most Deprived Decile		0.158		0.160	0.00279		0.140	-0.0173		0.174	0.0159
		[0.367]		[0.366]	(0.0348)		[0.328]	(0.0522)		[0.391]	(0.0449)
Least Deprived Decile		0.074		0.062	-0.0122		0.005	-0.069***		0.099	0.0253
		[0.264]		[0.241]	(0.0286)		[0.065]	(0.0173)		[0.309]	(0.0398)
IMD Education Skills	2,555		562			199			363		
Most Deprived Decile		0.090		0.037	-0.053***		0.048	-0.0418		0.029	-0.06***
		[0.288]		[0.188]	(0.0186)		[0.202]	(0.0282)		[0.174]	(0.0176)
Least Deprived Decile		0.112		0.112	0.000884		0.124	0.0121		0.105	-0.00649
		[0.317]		[0.315]	(0.0294)		[0.311]	(0.0503)		[0.317]	(0.0336)

	CONTROL		TREATMENT			SPONSOR-LED			CONVERTER		
VARIABLES	(1) N	(2) Mean	(3) N	(4) Mean	(5) diff	(6) N	(7) Mean	(8) diff	(9) N	(10) Mean	(11) diff
IMD Barriers to Housing	2,555		562			199			363		
Most Deprived Decile		0.181		0.228	0.0469		0.303	0.122		0.179	-0.00217
		[0.388]		[0.419]	(0.0444)		[0.433]	(0.0842)		[0.396]	(0.0466)
Least Deprived Decile		0.054		0.067	0.0130		0.004	-0.050***		0.109	0.0545
		[0.228]		[0.250]	(0.0281)		[0.060]	(0.00978)		[0.321]	(0.0438)
IMD Crime Domain	2,555		562			199			363		
Most Deprived Decile		0.092		0.110	0.0182		0.109	0.0172		0.111	0.0188
		[0.291]		[0.312]	(0.0376)		[0.294]	(0.0444)		[0.324]	(0.0557)
Least Deprived Decile		0.069		0.067	-0.00110		0.043	-0.0259		0.084	0.0152
		[0.255]		[0.250]	(0.0235)		[0.191]	(0.0307)		[0.286]	(0.0320)
IMD Living Environment	2,555		562			199			363		
Most Deprived Decile		0.112		0.084	-0.0279		0.131	0.0187		0.054	-0.0585**
		[0.318]		[0.277]	(0.0350)		[0.318]	(0.0648)		[0.232]	(0.0279)
Least Deprived Decile		0.074		0.113	0.0388		0.017	-0.057***		0.176	0.102*
		[0.263]		[0.315]	(0.0365)		[0.121]	(0.0129)		[0.393]	(0.0526)
Urban (ONS Code)	1,741	0.807	324	0.827	0.0198	129	0.912	0.104***	195	0.753	-0.0545
		[0.390]		[0.362]	(0.0506)		[0.251]	(0.0365)		[0.439]	(0.0852)
AGE 7											
COGNITIVE SKILLS											
Word Reading Standardised	2,312	111.316	515	110.365	-0.951	187	108.681	-2.635**	328	111.529	0.213
		[17.611]		[17.698]	(1.020)		[16.708]	(1.282)		[18.256]	(1.550)
Pattern Construction T-Scores	2,306	52.096	513	52.535	0.439	187	51.412	-0.683	326	53.313	1.218
		[11.153]		[11.265]	(0.869)		[10.381]	(1.128)		[11.808]	(1.161)
Maths Standardised	2,312	96.721	514	96.644	-0.0768	187	93.497	-3.224	327	98.825	2.104
		[16.054]		[15.932]	(1.612)		[14.367]	(2.486)		[16.643]	(1.626)
NON-COGNITIVE SKILLS											
SDQ Total Difficulties	2,240	0.136	501	0.049	-0.0871	175	0.115	-0.0207	326	0.004	-0.132
		[1.088]		[0.953]	(0.0587)		[0.795]	(0.0688)		[1.056]	(0.0874)
SDQ Pro-social	2,267	-0.043	503	-0.037	0.00609	176	-0.067	-0.0238	327	-0.017	0.0261
		[1.067]		[1.045]	(0.0592)		[0.910]	(0.0805)		[1.134]	(0.0678)
AGE 5											
COGNITIVE SKILLS											
Picture Similarities T-scores	2,408	54.254	528	53.027	-1.226	188	51.177	-3.077**	340	54.273	0.0190
		[11.532]		[10.859]	(0.962)		[9.349]	(1.520)		[11.696]	(0.860)

	CONTROL		TREATMENT			SPONSOR-LED			CONVERTER		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**
VARIABLES	**N**	**Mean**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**
Pattern Construction	2,409	49.804	529	49.359	-0.444	189	48.783	-1.020	340	49.748	-0.0552
T-scores		[10.660]		[10.554]	(0.763)		[9.824]	(0.974)		[11.020]	(0.941)
NON-COGNITIVE SKILLS											
SDQ Total Difficulties	2,262	0.146	496	0.103	-0.0424	174	0.441	0.295***	322	-0.130	-0.276***
		[1.087]		[0.985]	(0.0935)		[0.909]	(0.0957)		[0.960]	(0.100)
SDQ Pro-social	2,292	-0.093	501	-0.029	0.0644	176	-0.053	0.0397	325	-0.011	0.0816
		[1.109]		[0.988]	(0.0669)		[0.841]	(0.0831)		[1.089]	(0.0830)
CHILDCARE											
CM attended	2,399	0.926	527	0.938	0.0124	189	0.931	0.00537	338	0.943	0.0172
childcare providers		[0.265]		[0.243]	(0.0167)		[0.240]	(0.0278)		[0.243]	(0.0152)
Grandparents look	2,400	0.262	527	0.289	0.0273	189	0.281	0.0190	338	0.295	0.0331
after CM weekdays		[0.444]		[0.458]	(0.0237)		[0.426]	(0.0320)		[0.479]	(0.0316)
How often read to CM	2,400		526			188			338		
Less often		0.080		0.072	-0.00859		0.098	0.0179		0.053	-0.0268
		[0.275]		[0.260]	(0.0163)		[0.282]	(0.0340)		[0.237]	(0.0180)
At least weekly		0.461		0.448	-0.0135		0.489	0.0283		0.419	-0.0422
		[0.504]		[0.502]	(0.0273)		[0.473]	(0.0410)		[0.519]	(0.0308)
Daily		0.459		0.481	0.0221		0.412	-0.0461		0.528	0.0690**
		[0.504]		[0.505]	(0.0290)		[0.466]	(0.0444)		[0.525]	(0.0318)

Note: The treatment group includes pupils enrolled in 2012/13 at LA maintained secondary schools that become academies in 2013/14 and 2014/15, whereas the control group includes pupils enrolled at LA maintained secondary schools that do not convert to academy status. The treatment group is composed by pupils attending prospective sponsor-led and converter academies. T-scores in ability tests are adjusted for age and mean scores of norming group (mean of 50 and standard deviation of 10). The quantity in parenthesis below the mean (or proportion in the case of categorical variables) is the standard deviation. All statistics, except the number of observations (N) that is unweighted, reflect MCS survey settings. The difference (diff) refers to the difference in means between the treatment group and the control group and the quantity in parenthesis below is the (robust) standard error clustered at the school level and reflecting MCS survey settings. (University of London, Institute of Education, Centre for Longitudinal Studies, 2020a, 2020d, 2022). * significant at 10% level; ** significant at 5% level; *** significant at 1% level or below.

Table A3 Chi-square test (categorical variables)

	TREATMENT		SPONSOR-LED		CONVERTER	
VARIABLES	F-stat	p-value	F-stat	p-value	F-stat	p-value
AGE 11						
CM CHARACTERISTICS						
CM sex	0.020	0.887	0.765	0.382	0.407	0.524
CM ethnic group	1.011	0.378	0.668	0.537	0.511	0.641
Language spoken in household	0.077	0.782	0.139	0.709	0.006	0.940
SEN statement	1.922	0.166	10.108	0.0016	0.118	0.731
CM's general level of health	1.133	0.339	1.799	0.132	1.125	0.342
FAMILY CONTEXT						
Parents/carers in household	1.639	0.201	0.183	0.669	5.962	0.0151
Grandparents in household	0.168	0.683	0.139	0.287	0.090	0.764
Main has longstanding illness	0.025	0.875	1.195	0.275	0.634	0.427
SOCIOECONOMIC STATUS						
Free school meals eligible	1.047	0.307	0.471	0.493	4.784	0.0293
Housing tenure	1.468	0.232	2.322	0.103	4.722	0.0156
Combined labour market status	0.531	0.708	0.720	0.569	1.354	0.253
Full-time work (Main)	0.025	0.960	1.454	0.235	0.617	0.509
Full-time work (Partner)	0.779	0.444	0.420	0.647	1.659	0.197
Occupation (Main)	0.682	0.608	3.208	0.0203	0.971	0.416
Occupation (Partner)	0.664	0.619	2.941	0.0190	1.061	0.372
Education (Main)	1.511	0.206	3.376	0.0181	0.376	0.784
Education (Partner)	3.439	0.0113	4.923	0.0022	3.805	0.0071
HOME ENVIRONMENT						
How many books at home	1.090	0.360	1.861	0.129	0.640	0.642
How often visited library	3.299	0.0701	0.333	0.564	3.769	0.0529
How often attended church	3.792	0.0522	3.175	0.0755	1.453	0.229
Hours per weekday watching TV	0.592	0.551	0.741	0.464	2.943	0.0539
How often helps homework	0.110	0.947	0.481	0.657	0.583	0.624
Regular bedtime on weekdays	3.098	0.0792	0.398	0.528	3.690	0.0555
SCHOOL-LEVEL VARIABLES						
School fees applicable	0.129	0.720	4.310	0.0386	1.213	0.271
Single sex or mixed school	0.075	0.784	0.024	0.878	0.063	0.801
School with a particular faith	0.142	0.707	5.845	0.0161	0.519	0.472
No. of primary schools attended	0.893	0.345	0.954	0.329	0.140	0.709
GEOGRAPHICAL VARIABLES						
IMD Income domain	0.232	0.957	1.664	0.158	0.782	0.566
IMD Health deprivation/disability	1.049	0.387	2.776	0.0190	0.924	0.458
IMD Education skills and training	1.220	0.292	1.071	0.374	1.221	0.293
IMD Barriers to housing/services	0.341	0.905	2.080	0.0739	1.122	0.346
IMD Crime domain	0.386	0.875	1.687	0.149	0.345	0.863
IMD Living environment domain	1.064	0.382	1.563	0.169	2.193	0.0507
ONS Rural urban code	0.146	0.702	6.367	0.0121	0.474	0.492
AGE 5						
CHILDCARE						
CM attended childcare providers	0.502	0.479	0.035	0.851	1.136	0.287
Who looks after CM weekdays	1.363	0.244	0.365	0.546	1.149	0.284
How often read to CM	0.339	0.708	0.512	0.579	2.332	0.101

Note: The treatment group includes pupils enrolled in 2012/13 at LA maintained secondary schools that become academies in 2013/14 and 2014/15, whereas the control group includes pupils enrolled at LA maintained secondary schools that do not convert to academy status. The treatment group is composed by pupils attending prospective sponsor-led and converter academies. The p-value of the Chi-square test, reflecting MCS survey settings, is presented for each categorical variable (University of London, Institute of Education, Centre for Longitudinal Studies, 2020d, 2022).

Table A4 List of independent variables per method

VARIABLES	(1) OLS	(2) PSM	(3) DID	(4) DID_PSM	(5) DID_IV
AGE 11					
CM CHARACTERISTICS					
CM Sex	✓	✓	✓	✓	✓
Age at MCS5 interview (10th of year)	✓	✓	✓	✓	✓
CM ethnic group	✓	✓	✓	✓	✓
Language spoken in household	✓	✓	✓	✓	✓
SEN statement	✓	✓	✓	✓	✓
CM's general level of health	✓	✓	✓	✓	✓
COGNITIVE SKILLS					
SWM Strategy	✓	✓	✓	✓	✓
SWM Total errors	✓	✓	✓	✓	✓
FAMILY CONTEXT					
Parents/carers in household	✓	✓	✓	✓	✓
Grandparents in household	✓	✓	✓	✓	✓
Number of siblings in household	✓	✓	✓	✓	✓
Age (Main)	✓	✓	✓	✓	✓
Age (Partner)	✓	✓	✓	✓	✓
Whether Main has longstanding illness	✓	✓	✓	✓	✓
Main’s mental health (Kessler 6 Scale)	✓	✓	✓	✓	✓
SOCIOECONOMIC STATUS					
OECD equiv weekly family income (log)	✓	✓	✓	✓	✓
Free School Meal Eligible	✓	✓	✓	✓	✓
Housing Tenure	✓	✓	✓	✓	✓
Combined labour market status	✓	✓	✓	✓	✓
Full-time work (Main)	✓	✓	✓	✓	✓
Full-time work (Partner)	✓	✓	✓	✓	✓
Occupation (Main)	✓	✓	✓	✓	✓
Occupation (Partner)	✓	✓	✓	✓	✓
Education (Main)	✓	✓	✓	✓	✓
Education (Partner)	✓	✓	✓	✓	✓
HOME ENVIRONMENT					
How many books at home	✓	✓	✓	✓	✓
How often visited library	✓	✓	✓	✓	✓
How often attended religious service	✓	✓	✓	✓	✓
Hours spent watching TV or videos	✓	✓	✓	✓	✓
How often anyone helps with homework	✓	✓	✓	✓	✓
Regular bedtime on weekdays	✓	✓	✓	✓	✓
SCHOOL-LEVEL VARIABLES					
School fees applicable	✓	✓	✓	✓	✓
Single sex or mixed school	✓	✓	✓	✓	✓
School with a particular faith or religion	✓	✓	✓	✓	✓
Number of Primary Schools attended	✓	✓	✓	✓	✓
GEOGRAPHICAL VARIABLES					
IMD Income Domain	✓	✓	✓	✓	✓
IMD Health Deprivation and Disability	✓	✓	✓	✓	✓
IMD Education Skills and Training Domain	✓	✓	✓	✓	✓
IMD Barriers to Housing and Services	✓	✓	✓	✓	✓
IMD Crime Domain	✓	✓	✓	✓	✓
IMD Living Environment Domain	✓	✓	✓	✓	✓
ONS Rural Urban Code	✓	✓	✓	✓	✓

VARIABLES	(1) OLS	(2) PSM	(3) DID	(4) DID_PSM	(5) DID_IV
AGE 7					
COGNITIVE SKILLS					
Word Reading Standardised	✓	✓	✓	✓	✓
Pattern Construction T-Scores	✓	✓	✓	✓	✓
Maths Standardised	✓	✓	✓	✓	✓
NON-COGNITIVE SKILLS					
SDQ Total Difficulties	✓	✓	✓	✓	✓
SDQ Pro-social	✓	✓	✓	✓	✓
AGE 5					
COGNITIVE SKILLS					
Picture Similarities T-scores	✓	✓	✓	✓	✓
Pattern Construction T-scores	✓	✓	✓	✓	✓
NON-COGNITIVE SKILLS					
SDQ Total Difficulties	✓	✓	✓	✓	✓
SDQ Pro-social	✓	✓	✓	✓	✓
CHILDCARE					
CM attended childcare providers	✓	✓	✓	✓	✓
Who looks after CM weekdays	✓	✓	✓	✓	✓
How often read to CM	✓	✓	✓	✓	✓

Note: OLS regression and propensity score matching (PSM) only use pre-treatment independent variables. Difference in differences (DID) models, based on panel data, require variables to be measured before and after treatment. The PSM column identifies matching variables (except MCS survey weights and strata). Weights from PSM are used in the matched difference in differences (DID_PSM) method. The DID model using instrumental variables (DID_IV) corrects for non-compliance with treatment assignment. (University of London, Institute of Education, Centre for Longitudinal Studies, 2020a, 2020d, 2022).

Table A5 Alternative matching methods

VARIABLES	(1) EPANECHINIKOV KERNEL BANDWIDTH OF 0.06 ATT	(2) EPANECHINIKOV KERNEL BANDWIDTH OF 0.01 ATT	(3) NEAREST NEIGHBOUR ATT	(4) MANALANOBIS ATT
COGNITIVE SKILLS				
Word Activity	-0.040	-0.029	-0.014	-0.027
CGT Quality of Decision Making	0.092	0.095	0.138	-0.001
CGT Risk Adjustment	0.118	0.127	0.178	0.004
NON-COGNITIVE SKILLS				
CGT Risk Taking	-0.054	-0.073	-0.075	-0.056
CGT Deliberation Time	-0.162	-0.184	-0.129	0.021
SDQ Total Difficulties	-0.015	-0.023	-0.049	0.228
SDQ Pro-social	0.098	0.086	0.157	0.049
Rosenberg Self-esteem Scale	0.016	-0.008	0.018	-0.134
Observations	**2,255**	**2,243**	**2,258**	**2,258**
R^2	**0.02**	**0.01**	**0.11**	**0.32**

Note: The treatment group includes pupils enrolled in 2012/13 at LA maintained secondary schools that become academies in 2013/14 and 2014/15, whereas the control group includes pupils enrolled at LA maintained secondary schools that do not convert to academy status. The treatment group is composed by pupils attending prospective sponsor-led and converter academies. All outcome variables are standardised to mean zero and standard deviation one. ATT is the average treatment effect on the treated. Sample size is restricted to complete cases for all outcome variables. R^2 refers to the probit estimation of the propensity score on the matched sample (Leuven & Sianesi, 2003). MCS survey settings are not applied to propensity score matching, but MCS survey weights and strata are included among matching variables. (University of London, Institute of Education, Centre for Longitudinal Studies, 2020a, 2020c, 2020d, 2022).

Table A6 Sensitivity analysis

VARIABLES	N	(1) DID Treatment	N	(2) TEACHER Treatment	(3) SCHOOL Treatment	(4) GRAMMAR Treatment	(5) ACADEMIC Treatment	(6) LEVEL Treatment	(7) QUINTILES Treatment	(8) MEDIAN Treatment	(9) RAW DID Treatment
COGNITIVE SKILLS											
Word Activity	2,842	0.0574			0.0624	0.0556	0.0567	0.0572	0.0564	0.0540	0.0670
		(0.0898)			(0.0907)	(0.0898)	(0.0897)	(0.0898)	(0.0897)	(0.0894)	(0.0922)
CGT Quality of Decision Making	2,750	0.0703			0.0648	0.0647	0.0752	0.0703	0.0692	0.0669	0.0750
		(0.0661)			(0.0646)	(0.0660)	(0.0667)	(0.0663)	(0.0660)	(0.0664)	(0.0678)
CGT Risk Adjustment	2,749	0.0389			0.0450	0.0342	0.0330	0.0390	0.0382	0.0390	0.0431
		(0.0733)			(0.0752)	(0.0737)	(0.0740)	(0.0732)	(0.0734)	(0.0727)	(0.0727)
NON-COGNITIVE SKILLS											
CGT Risk Taking	2,749	-0.0859			-0.0846	-0.0830	-0.0810	-0.0844	-0.0854	-0.0856	-0.0844
		(0.0609)			(0.0607)	(0.0610)	(0.0611)	(0.0609)	(0.0615)	(0.0608)	(0.0624)
CGT Deliberation Time	2,750	-0.0797			-0.0798	-0.0751	-0.0830	-0.0805	-0.0761	-0.0784	-0.0863
		(0.0614)			(0.0622)	(0.0622)	(0.0621)	(0.0615)	(0.0618)	(0.0618)	(0.0639)
SDQ Total Difficulties	2,961	-0.0279	2,200	-0.00866	-0.0313	-0.0302	-0.0278	-0.0279	-0.0278	-0.0261	-0.0334
		(0.0482)		(0.0801)	(0.0478)	(0.0485)	(0.0482)	(0.0482)	(0.0480)	(0.0482)	(0.0472)
SDQ Pro-social	2,979	0.0511	2,203	0.0855	0.0512	0.0527	0.0465	0.0515	0.0533	0.0522	0.0479
		(0.0636)		(0.0829)	(0.0638)	(0.0635)	(0.0635)	(0.0635)	(0.0634)	(0.0635)	(0.0665)
Rosenberg Self-esteem Scale	2,696	-0.00832			-0.00903	-0.000474	-0.00432	-0.00712	-0.00802	-0.00648	0.00512
		(0.0645)			(0.0653)	(0.0656)	(0.0657)	(0.0647)	(0.0646)	(0.0648)	(0.0758)
Model includes:											
Teacher reported SDQ (outcome)				✓							
Pupil level variables		✓		✓	✓	✓	✓	✓	✓	✓	
School level variables		✓		✓		✓	✓	✓	✓	✓	
Grammar school						✓					
Academic qualifications							✓				
Income level								✓			
Income quintiles									✓		
Income median										✓	

Note: The treatment group includes pupils enrolled in 2012/13 at LA maintained secondary schools that become academies in 2013/14 and 2014/15, whereas the control group includes pupils enrolled at LA maintained secondary schools that do not convert to academy status. The treatment group is composed by pupils attending prospective sponsor-led and converter academies. Different specifications are presented in comparison to the original difference in differences model (DID, repeated here for convenience) and follow it unless stated otherwise. Sample size is the same as in the DID model, except when using teacher reported outcome variables. All outcome variables are standardised to mean zero and standard deviation one. The quantity in parenthesis below the treatment coefficient is the (robust) standard error clustered at the school level and reflecting MCS survey settings. (University of London, Institute of Education, Centre for Longitudinal Studies, 2020a, 2020c, 2020d, 2022). * significant at 10% level; ** significant at 5% level; *** significant at 1% level or below.

Table A7 Sensitivity analysis: Sponsor-led Academies

VARIABLES	N	(1) DID Treatment	N	(2) TEACHER Treatment	(3) SCHOOL Treatment	(4) GRAMMAR Treatment	(5) ACADEMIC Treatment	(6) LEVEL Treatment	(7) QUINTILES Treatment	(8) MEDIAN Treatment	(9) RAW DID Treatment
COGNITIVE SKILLS											
Word Activity	2,519	-0.0478			-0.0526	-0.0464	-0.0461	-0.0494	-0.0472	-0.0497	-0.0604
		(0.142)			(0.141)	(0.142)	(0.142)	(0.142)	(0.142)	(0.141)	(0.137)
CGT Quality of	2,428	-0.0987			-0.0928	-0.0951	-0.0945	-0.0998	-0.103	-0.101	-0.0746
Decision Making		(0.120)			(0.121)	(0.120)	(0.121)	(0.120)	(0.120)	(0.121)	(0.122)
CGT Risk	2,427	-0.174*			-0.176*	-0.171*	-0.176*	-0.174*	-0.177*	-0.172*	-0.183*
Adjustment		(0.0965)			(0.0981)	(0.0964)	(0.0989)	(0.0965)	(0.0967)	(0.0960)	(0.0954)
NON-COGNITIVE SKILLS											
CGT Risk Taking	2,427	-0.0891			-0.0802	-0.0916	-0.0840	-0.0890	-0.0909	-0.0898	-0.0536
		(0.0981)			(0.0967)	(0.0983)	(0.100)	(0.0981)	(0.0991)	(0.0981)	(0.101)
CGT Deliberation	2,428	-0.0459			-0.0383	-0.0501	-0.0519	-0.0461	-0.0433	-0.0465	-0.0475
Time		(0.0923)			(0.0911)	(0.0924)	(0.0939)	(0.0921)	(0.0960)	(0.0926)	(0.0872)
SDQ Total	2,617	0.0125	1,935	0.0335	0.00629	0.0141	0.0129	0.0137	0.0123	0.0143	-0.0351
Difficulties		(0.0792)		(0.134)	(0.0791)	(0.0791)	(0.0788)	(0.0794)	(0.0788)	(0.0793)	(0.0670)
SDQ Pro-social	2,634	0.0513	1,938	0.0213	0.0567	0.0499	0.0437	0.0522	0.0537	0.0528	0.0578
		(0.129)		(0.101)	(0.128)	(0.129)	(0.128)	(0.129)	(0.128)	(0.129)	(0.135)
Rosenberg Self-	2,377	0.196**			0.196**	0.193**	0.203**	0.198**	0.197**	0.198**	0.226**
esteem Scale		(0.0797)			(0.0795)	(0.0798)	(0.0793)	(0.0803)	(0.0807)	(0.0802)	(0.0875)
Model includes:											
Teacher reported SDQ (outcome)				✓							
Pupil level variables		✓		✓	✓	✓	✓	✓	✓	✓	
School level variables		✓		✓		✓	✓	✓	✓	✓	
Grammar school						✓					
Academic qualifications							✓				
Income level								✓			
Income quintiles									✓		
Income median										✓	

Note: The treatment group includes pupils enrolled in 2012/13 at LA maintained secondary schools that become academies in 2013/14 and 2014/15, whereas the control group includes pupils enrolled at LA maintained secondary schools that do not convert to academy status. The treatment group is composed by pupils attending prospective sponsor-led academies. Different specifications are presented in comparison to the original difference in differences model (DID, repeated here for convenience) and follow it unless stated otherwise. Sample size is the same as in the DID model, except when using teacher reported outcome variables. All outcome variables are standardised to mean zero and standard deviation one. The quantity in parenthesis below the treatment coefficient is the (robust) standard error clustered at the school level and reflecting MCS survey settings. (University of London, Institute of Education, Centre for Longitudinal Studies, 2020a, 2020c, 2020d, 2022). * significant at 10% level; ** significant at 5% level; *** significant at 1% level or below.

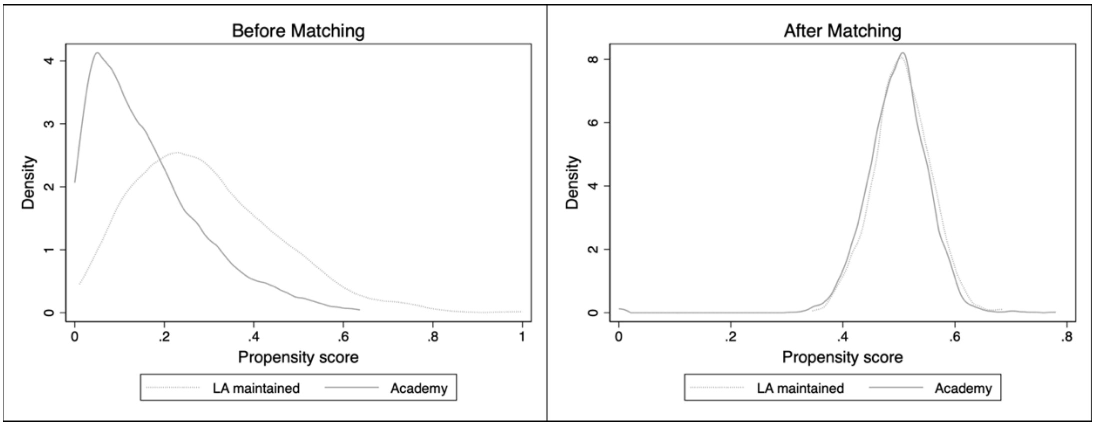
Table A8 Sensitivity analysis: Converter Academies

VARIABLES	N	(1) DID Treatment	N	(2) TEACHER Treatment	(3) SCHOOL Treatment	(4) GRAMMAR Treatment	(5) ACADEMIC Treatment	(6) LEVEL Treatment	(7) QUINTILES Treatment	(8) MEDIAN Treatment	(9) RAW DID Treatment
COGNITIVE SKILLS											
Word Activity	2,652	0.133			0.147	0.130	0.133	0.134	0.132	0.129	0.157*
		(0.0900)			(0.0917)	(0.0897)	(0.0904)	(0.0901)	(0.0898)	(0.0899)	(0.0935)
CGT Quality of Decision Making	2,566	0.181***			0.174***	0.170***	0.185***	0.183***	0.184***	0.179***	0.179***
		(0.0560)			(0.0545)	(0.0565)	(0.0556)	(0.0558)	(0.0557)	(0.0558)	(0.0585)
CGT Risk Adjustment	2,565	0.183**			0.199**	0.174**	0.174**	0.183**	0.184**	0.182**	0.200**
		(0.0807)			(0.0839)	(0.0819)	(0.0809)	(0.0806)	(0.0807)	(0.0795)	(0.0810)
NON-COGNITIVE SKILLS											
CGT Risk Taking	2,565	-0.0893			-0.0896	-0.0829	-0.0853	-0.0862	-0.0870	-0.0874	-0.106
		(0.0685)			(0.0701)	(0.0682)	(0.0683)	(0.0688)	(0.0690)	(0.0684)	(0.0735)
CGT Deliberation Time	2,566	-0.107			-0.112	-0.0972	-0.106	-0.109	-0.103	-0.104	-0.113
		(0.0828)			(0.0826)	(0.0840)	(0.0830)	(0.0829)	(0.0831)	(0.0830)	(0.0844)
SDQ Total Difficulties	2,772	-0.0492	2,075	-0.0300	-0.0524	-0.0530	-0.0481	-0.0502	-0.0494	-0.0486	-0.0323
		(0.0601)		(0.0925)	(0.0596)	(0.0612)	(0.0596)	(0.0600)	(0.0601)	(0.0599)	(0.0606)
SDQ Pro-social	2,787	0.0553	2,078	0.133	0.0523	0.0587	0.0517	0.0557	0.0580	0.0568	0.0413
		(0.0528)		(0.103)	(0.0520)	(0.0529)	(0.0527)	(0.0528)	(0.0528)	(0.0526)	(0.0522)
Rosenberg Self-esteem Scale	2,537	-0.130			-0.132	-0.115	-0.130	-0.129	-0.131	-0.129	-0.131
		(0.0803)			(0.0815)	(0.0829)	(0.0810)	(0.0805)	(0.0802)	(0.0805)	(0.0920)
Model includes:											
Teacher reported SDQ (outcome)				✓							
Pupil level variables		✓		✓	✓	✓	✓	✓	✓	✓	
School level variables		✓		✓		✓	✓	✓	✓	✓	
Grammar school						✓					
Academic qualifications							✓				
Income level								✓			
Income quintiles									✓		
Income median										✓	

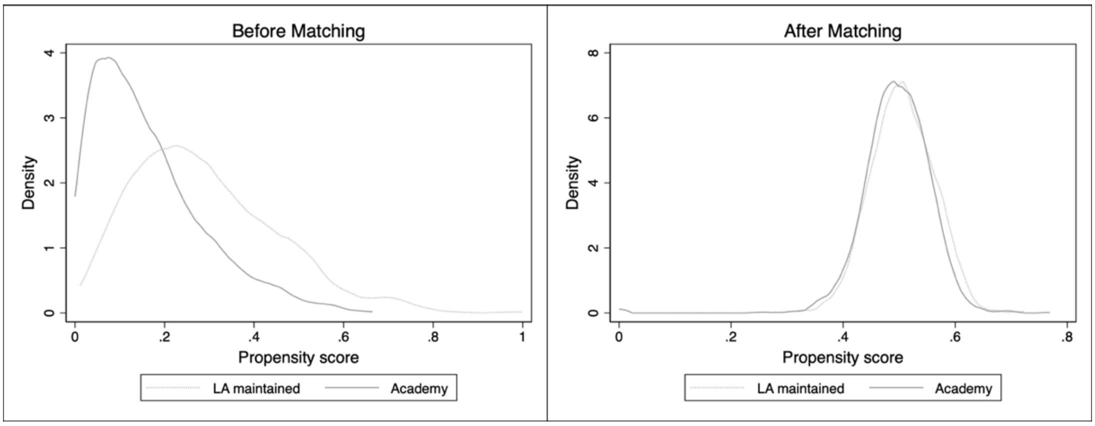
Note: The treatment group includes pupils enrolled in 2012/13 at LA maintained secondary schools that become academies in 2013/14 and 2014/15, whereas the control group includes pupils enrolled at LA maintained secondary schools that do not convert to academy status. The treatment group is composed by pupils attending prospective converter academies. Different specifications are presented in comparison to the original difference in differences model (DID, repeated here for convenience) and follow it unless stated otherwise. Sample size is the same as in the DID model, except when using teacher reported outcome variables. All outcome variables are standardised to mean zero and standard deviation one. The quantity in parenthesis below the treatment coefficient is the (robust) standard error clustered at the school level and reflecting MCS survey settings. (University of London, Institute of Education, Centre for Longitudinal Studies, 2020a, 2020c, 2020d, 2022). * significant at 10% level; ** significant at 5% level; *** significant at 1% level or below.

Figure A1 Density of propensity scores before and after matching

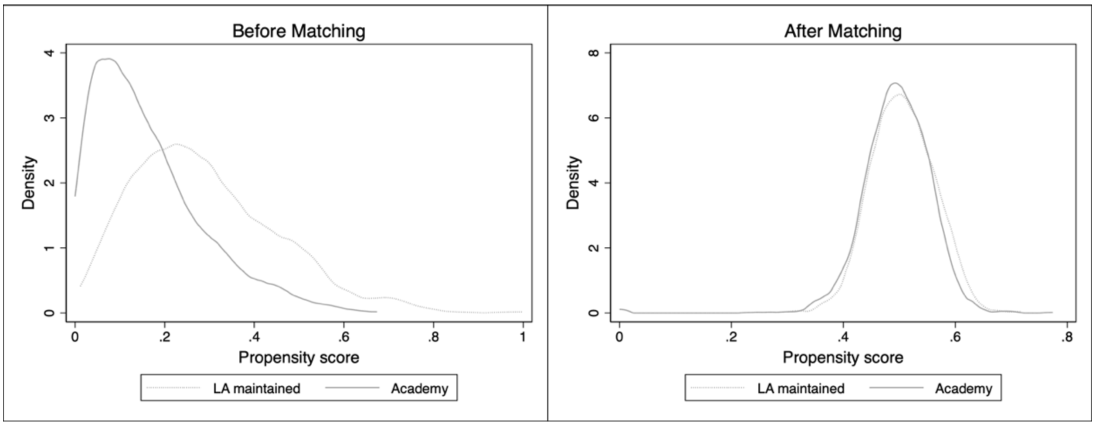
Word Activity



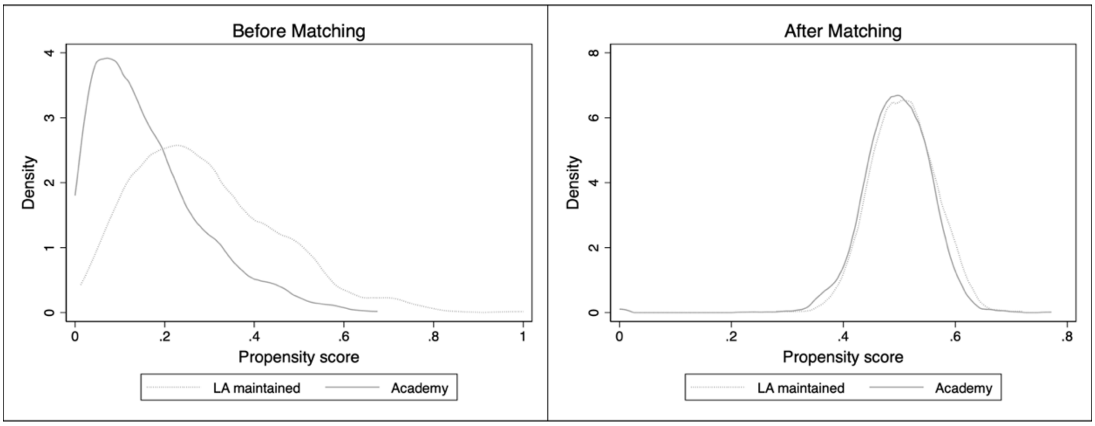
CGT Quality of Decision Making



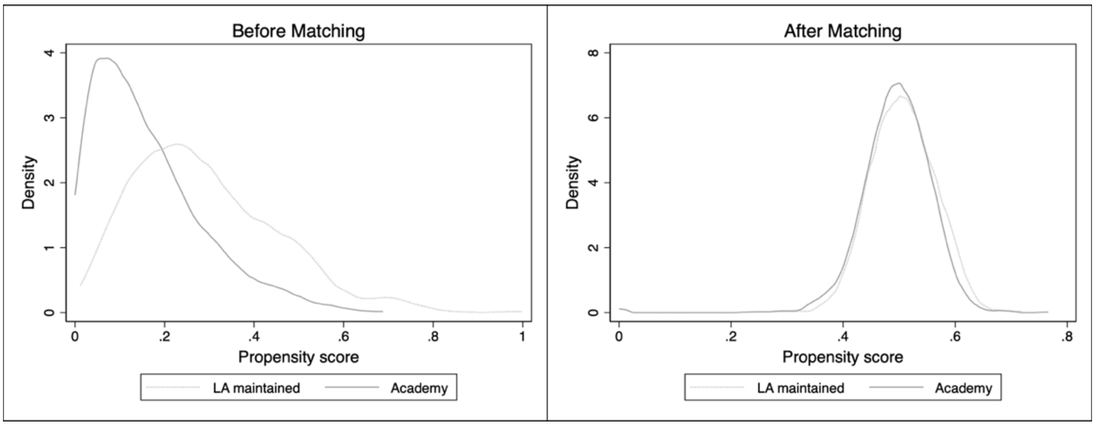
CGT Risk Adjustment



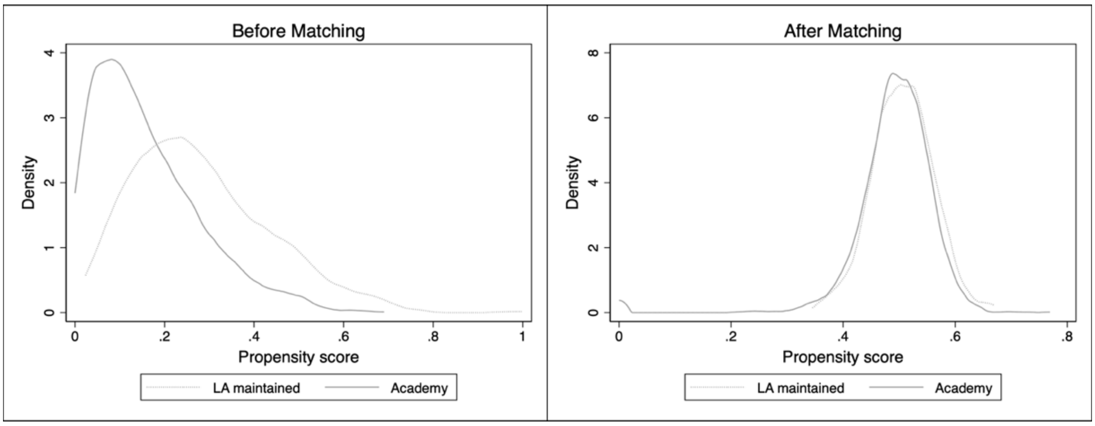
CGT Risk Taking



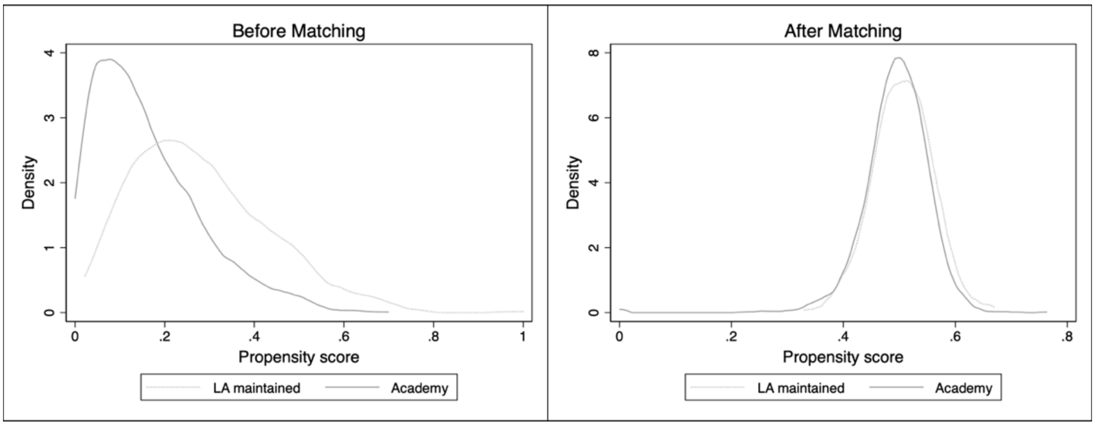
CGT Deliberation Time



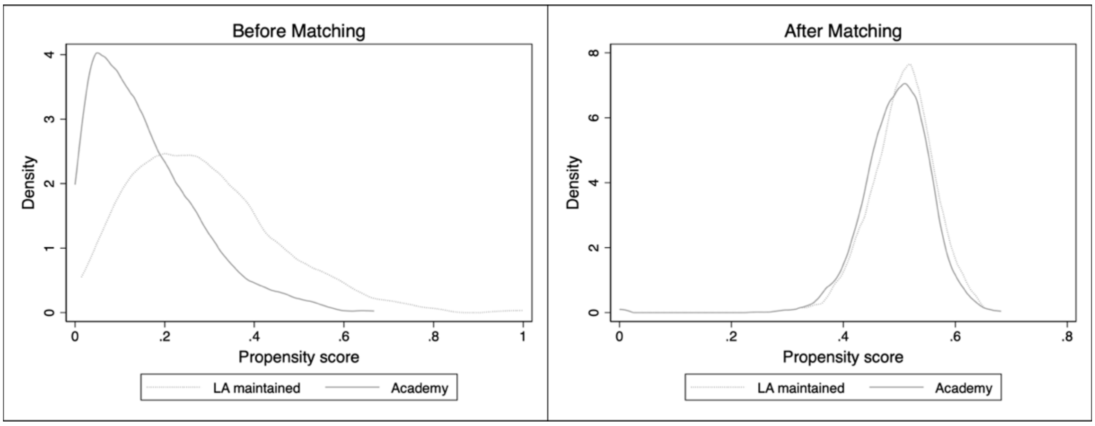
SDQ Total Difficulties



SDQ Prosocial



Rosenberg Self-esteem Scale



Note: The treatment group includes pupils enrolled in 2012/13 at LA maintained secondary schools that become academies in 2013/14 and 2014/15, whereas the control group includes pupils enrolled at LA maintained secondary schools that do not convert to academy status. The treatment group is composed by pupils attending prospective sponsor-led and converter academies. Because of missing values, the propensity score is calculated separately for each outcome variable (Leuven & Sianesi, 2003). Matching variables include MCS survey weights and strata. The propensity score density is displayed before and after matching. (University of London, Institute of Education, Centre for Longitudinal Studies, 2020a, 2020c, 2020d, 2022).

APPENDIX B SUPPLEMENTAL MATERIAL TO CHAPTER 4

Table B1 List of subjects per outcome variable (age 14)

SUBJECT NAME	(1) EBacc	(2) Facilitating	(3) STEM	(4) Vocational
GCSE Subjects				
Additional Applied Science				✓
Additional Mathematics	✓			
Additional Science	✓	✓	✓	
Additional Science (Modular)	✓	✓	✓	
Applied Art and Design				✓
Applied Business				✓
Applied ICT				✓
Applied Physical Education				✓
Applied Science				✓
Biology	✓	✓	✓	
Chemistry	✓	✓	✓	
Combined Science	✓	✓	✓	
Computer Science	✓	✓	✓	
Engineering				✓
Further Additional Science	✓	✓	✓	
Further Mathematics	✓			
Geography	✓	✓		
Health and Social Care				✓
History	✓	✓		
Information and Communication Tech			✓	
Language: Welsh	✓	✓		
Language: English	✓			
Language: English Language	✓			
Language: English Literature	✓			
Language: French	✓	✓		
Language: German	✓	✓		
Language: Irish	✓	✓		
Language: Latin	✓	✓		
Language: Spanish	✓	✓		
Language: Urdu	✓	✓		
Language: Welsh First Language	✓	✓		
Mathematics	✓			
Mathematics - Linear	✓			
Physics	✓	✓	✓	
Science	✓	✓	✓	
Science (Modular)	✓	✓	✓	
Statistics			✓	
IGSCE Subjects				
Biology	✓	✓	✓	
Chemistry	✓	✓	✓	
Geography	✓	✓		
History	✓	✓		
Language: English - First Language	✓			
Language: English Literature	✓			
Language: French	✓	✓		
Language: German	✓	✓		
Language: Spanish	✓	✓		
Mathematics	✓			

SUBJECT NAME	(1) EBacc	(2) Facilitating	(3) STEM	(4) Vocational
Physics	✓	✓	✓	
Science	✓	✓	✓	
Vocational qualifications				
BTEC Level 1 and 2				✓
Cambridge National Award				✓
City and Guilds Certificate				✓
Essential Skills Level 1 and 2				✓
NVQ/SVQ				✓
QCF Diploma				✓
QCF Award				✓
QCF Certificate				✓
Skills for Life Level 1 and 2				✓
Other Qualification Level 1 and 2				✓

Table B2 Sample selection

LOGIT INDEPENDENT VARIABLES	DEPENDENT VARIABLE: SAMPLE SELECTION Odds ratios
Verbal Similarities Ability Score (age 11)	0.944 (0.0452)
SDQ Total Difficulties (age 7)	0.982 (0.0509)
CM Sex (Male)	1.014 (0.0738)
CM ethnic group (ref. White)	
Mixed	0.823 (0.129)
Indian, Pakistani and Bangladeshi	1.139 (0.193)
Black	0.631** (0.138)
Other ethnic group	0.894 (0.193)
Free School Meal Eligible (age 11)	1.148 (0.175)
OECD equiv weekly family income (log, age 11)	1.070 (0.205)
Number of siblings in household (age 11)	1.021 (0.0391)
Parents Occupation (ref. Managerial and professional)	
Intermediate	1.168 (0.136)
Small employers and self-employed	1.162 (0.156)
Lower supervisory and technical	1.075 (0.189)
Semi-routine and routine	1.289* (0.172)
Not in work nor on leave	0.809 (0.144)
Parents Education (ref. NVQ equiv. level 1)	
NVQ equiv. level 2 (incl. GSCE A* – C)	0.945 (0.197)
NVQ equiv. level 3 (incl. A levels)	1.021 (0.225)
NVQ equiv. level 4 (incl.BA degree)	0.978 (0.204)
NVQ equiv. level 5 (incl. postgraduate degree)	0.801 (0.188)
IMD Income Domain (ref. Most Deprived Decile)	
Least Deprived Decile	0.441*** (0.110)
Observations	**6,592**

Note: The quantity in parenthesis below the estimator is the (robust) standard error reflecting MCS survey setting. Estimates are based on 70 imputations using chained equations. (University of London, Institute of Education, Centre for Longitudinal Studies, 2020a, 2022). * significant at 10% level; ** significant at 5% level; *** significant at 1% level or below.

Table B3 Descriptive statistics pre-treatment (all explanatory variables)

	CONTROL		TREATMENT			SPONSOR-LED			CONVERTER		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
VARIABLES	**N**	**Mean**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**
AGE 11											
CM CHARACTERISTICS											
CM Sex (Male)	2,070	0.508	454	0.494	-0.0135	161	0.411	-0.0966**	293	0.538	0.0305
		(0.539)		(0.527)	(0.0361)		(0.525)	(0.0452)		(0.522)	(0.0430)
Age at MCS5 interview (10th of year)	2,070	11.14	454	11.17	0.0240	161	11.18	0.0341	293	11.16	0.0186
		(0.362)		(0.325)	(0.0222)		(0.340)	(0.0296)		(0.317)	(0.0283)
CM ethnic group	2,068		454			161			293		
White		0.828		0.859	0.0313		0.871	0.0431		0.853	0.0251
		(0.407)		(0.367)	(0.0365)		(0.358)	(0.0442)		(0.371)	(0.0479)
Mixed		0.0512		0.0395	-0.0117		0.0307	-0.0205		0.0441	-0.00710
		(0.238)		(0.205)	(0.0162)		(0.184)	(0.0161)		(0.215)	(0.0212)
Indian, Pakistani and Bangladeshi		0.0776		0.0772	-0.00040		0.0665	-0.0112		0.0829	0.00529
		(0.289)		(0.281)	(0.0314)		(0.266)	(0.0414)		(0.289)	(0.0405)
Black		0.0231		0.0115	-0.0116*		0.0150	-0.00814		0.00962	-0.0135*
		(0.162)		(0.112)	(0.00670)		(0.130)	(0.00940)		(0.102)	(0.00793)
Other ethnic group		0.0203		0.0127	-0.00757		0.0170	-0.00329		0.0104	-0.00983
		(0.152)		(0.118)	(0.00552)		(0.138)	(0.00835)		(0.106)	(0.00601)
Language spoken in household (English only)	2,070	0.908	454	0.922	0.0137	161	0.933	0.0252	293	0.916	0.00760
		(0.311)		(0.283)	(0.0308)		(0.266)	(0.0401)		(0.291)	(0.0393)
CM's general level of health	2,070		454			161			293		
Excellent		0.562		0.570	0.00821		0.440	-0.121***		0.638	0.0767**
		(0.535)		(0.522)	(0.0327)		(0.530)	(0.0458)		(0.503)	(0.0341)
Very Good		0.292		0.281	-0.0116		0.323	0.0301		0.259	-0.0337
		(0.491)		(0.474)	(0.0288)		(0.499)	(0.0403)		(0.459)	(0.0338)
Good		0.116		0.0817	-0.0346**		0.117	0.000358		0.0632	-0.053***
		(0.346)		(0.289)	(0.0175)		(0.343)	(0.0358)		(0.255)	(0.0164)
Fair or Poor		0.0297		0.0677	0.0380*		0.120	0.0907*		0.0398	0.0101
		(0.183)		(0.265)	(0.0220)		(0.347)	(0.0538)		(0.205)	(0.0155)
Verbal Similarities Ability Score	2,041	-0.0149	447	-0.178	-0.164*	159	-0.259	-0.244	288	-0.136	-0.121
		(1.049)		(1.020)	(0.0893)		(1.033)	(0.166)		(1.011)	(0.0994)
CGT Quality of Decision Making	1,968	0.118	442	0.0822	-0.0356	158	0.0604	-0.0573	284	0.0940	-0.0237
		(1.039)		(1.011)	(0.0677)		(0.993)	(0.0922)		(1.020)	(0.0888)

	CONTROL		TREATMENT			SPONSOR-LED			CONVERTER		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
VARIABLES	**N**	**Mean**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**
CGT Risk Taking	1,968	0.0348	442	0.0413	0.00646	158	0.0206	-0.0142	284	0.0525	0.0177
		(1.042)		(0.974)	(0.0562)		(0.970)	(0.0923)		(0.976)	(0.0692)
FAMILY CONTEXT											
One parent/carer in household	2,070	0.248	454	0.235	-0.0126	161	0.293	0.0445	293	0.205	-0.0428
		(0.466)		(0.447)	(0.0345)		(0.485)	(0.0697)		(0.423)	(0.0312)
At least one grandparent in Household	2,070	0.0348	454	0.0448	0.0100	161	0.0234	-0.0114	293	0.0562	0.0213
		(0.198)		(0.218)	(0.0160)		(0.161)	(0.0110)		(0.241)	(0.0241)
Number of siblings in Household	2,070	1.516	454	1.617	0.101	161	1.693	0.177	293	1.577	0.0612
		(1.164)		(1.194)	(0.0775)		(1.301)	(0.121)		(1.133)	(0.0960)
Age (Main)	2,070	40.31	454	40.01	-0.302	161	39.11	-1.206*	293	40.49	0.177
		(6.512)		(6.240)	(0.573)		(5.684)	(0.657)		(6.448)	(0.774)
Whether Main has longstanding Illness	2,068	0.199	454	0.219	0.0194	161	0.206	0.00687	293	0.225	0.0260
		(0.431)		(0.436)	(0.0308)		(0.432)	(0.0420)		(0.437)	(0.0384)
Main's mental health (Kessler 6 Scale)	1,849	0.00280	410	0.0818	0.0790	141	0.398	0.395**	269	-0.0853	-0.0881
		(1.059)		(1.103)	(0.0868)		(1.347)	(0.197)		(0.899)	(0.0687)
SOCIOECONOMIC STATUS											
OECD equiv weekly family income (log)	2,070	5.972	454	5.934	-0.0378	161	5.813	-0.159*	293	5.998	0.0264
		(0.492)		(0.487)	(0.0528)		(0.510)	(0.0857)		(0.461)	(0.0660)
Free School Meal Eligible	2,065	0.175	452	0.191	0.0160	160	0.283	0.108*	292	0.143	-0.0323
		(0.410)		(0.414)	(0.0307)		(0.480)	(0.0650)		(0.366)	(0.0302)
Housing Tenure	2,051		450			159			291		
Rent privately		0.150		0.156	0.00543		0.203	0.0526		0.131	-0.0195
		(0.386)		(0.382)	(0.0280)		(0.428)	(0.0581)		(0.353)	(0.0276)
Owner occupier		0.649		0.659	0.00981		0.552	-0.0969		0.715	0.0663
		(0.515)		(0.499)	(0.0402)		(0.529)	(0.0603)		(0.472)	(0.0495)
LA or HA rent		0.200		0.185	-0.0152		0.245	0.0443		0.154	-0.0467
		(0.432)		(0.409)	(0.0277)		(0.457)	(0.0346)		(0.377)	(0.0357)
Average hours worked (Main and Partner)	2,065	24.98	453	24.54	-0.435	160	20.74	-4.238**	293	26.54	1.566
		(15.69)		(15.93)	(1.387)		(15.65)	(2.140)		(15.67)	(1.632)
Parents Occupation	2,048		453			160			293		
Managerial and professional		0.420		0.366	-0.0543		0.207	-0.213***		0.450	0.0295
		(0.534)		(0.508)	(0.0405)		(0.432)	(0.0329)		(0.521)	(0.0464)
Intermediate		0.137		0.141	0.00364		0.169	0.0319		0.126	-0.0113
		(0.372)		(0.367)	(0.0255)		(0.400)	(0.0418)		(0.348)	(0.0266)

	CONTROL		TREATMENT			SPONSOR-LED			CONVERTER		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**
VARIABLES	**N**	**Mean**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**
Small employers and		0.103		0.117	0.0140		0.0979	-0.00509		0.127	0.0240
self-employed		(0.329)		(0.339)	(0.0212)		(0.317)	(0.0249)		(0.349)	(0.0281)
Lower supervisory and		0.0370		0.0500	0.0130		0.0694	0.0324		0.0398	0.00275
technical		(0.204)		(0.230)	(0.0147)		(0.271)	(0.0303)		(0.205)	(0.0131)
Semi-routine and routine		0.145		0.154	0.00882		0.206	0.0604		0.127	-0.0183
		(0.381)		(0.381)	(0.0265)		(0.431)	(0.0367)		(0.349)	(0.0313)
Not in work nor on leave		0.157		0.172	0.0149		0.251	0.0938		0.130	-0.0266
		(0.394)		(0.398)	(0.0362)		(0.462)	(0.0689)		(0.353)	(0.0389)
Parents Education	1,854		403			131			272		
NVQ equiv. level 1		0.0620		0.0579	-0.00401		0.0562	-0.00574		0.0588	-0.00321
(incl. GSCE D – G)		(0.256)		(0.244)	(0.0167)		(0.244)	(0.0269)		(0.244)	(0.0199)
NVQ equiv. level 2		0.232		0.233	0.000919		0.370	0.138***		0.169	-0.0625**
(incl. GSCE A* – C)		(0.449)		(0.442)	(0.0334)		(0.511)	(0.0506)		(0.390)	(0.0269)
NVQ equiv. level 3		0.160		0.191	0.0317		0.238	0.0786		0.170	0.00993
(incl. A levels)		(0.390)		(0.411)	(0.0315)		(0.451)	(0.0492)		(0.390)	(0.0348)
NVQ equiv. level 4		0.404		0.361	-0.0429		0.253	-0.151***		0.412	0.00752
(incl.BA degree)		(0.522)		(0.502)	(0.0360)		(0.460)	(0.0363)		(0.511)	(0.0415)
NVQ equiv. level 5		0.142		0.157	0.0143		0.0833	-0.0590		0.191	0.0483
(incl. postgraduate degree)		(0.372)		(0.380)	(0.0266)		(0.292)	(0.0439)		(0.408)	(0.0296)
HOME ENVIRONMENT											
How many books at home	2,050		450			159			291		
0-10		0.124		0.158	0.0336		0.216	0.0924		0.127	0.00246
		(0.356)		(0.384)	(0.0272)		(0.438)	(0.0634)		(0.348)	(0.0252)
11-25		0.136		0.147	0.0111		0.173	0.0367		0.134	-0.00239
		(0.370)		(0.373)	(0.0214)		(0.402)	(0.0295)		(0.356)	(0.0284)
26-100		0.341		0.344	0.00261		0.369	0.0275		0.331	-0.0105
		(0.512)		(0.500)	(0.0280)		(0.513)	(0.0464)		(0.492)	(0.0318)
101-200		0.205		0.166	-0.0395		0.114	-0.091***		0.193	-0.0121
		(0.436)		(0.391)	(0.0249)		(0.338)	(0.0256)		(0.413)	(0.0313)
201-500		0.141		0.144	0.00353		0.114	-0.0263		0.160	0.0193
		(0.376)		(0.370)	(0.0270)		(0.339)	(0.0446)		(0.384)	(0.0315)
More than 500		0.0530		0.0417	-0.0113		0.0142	-0.039***		0.0562	0.00324
		(0.242)		(0.210)	(0.0112)		(0.126)	(0.0114)		(0.241)	(0.0140)
Visited library at least monthly	2,070	0.262	454	0.219	-0.0431	161	0.232	-0.0303	293	0.212	-0.0499
		(0.474)		(0.436)	(0.0294)		(0.450)	(0.0416)		(0.428)	(0.0394)

	CONTROL		TREATMENT			SPONSOR-LED			CONVERTER		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**
VARIABLES	**N**	**Mean**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**
Attended religious service	2,070	0.269	454	0.177	-0.092***	161	0.121	-0.148***	293	0.206	-0.0630
at least once a month		(0.479)		(0.402)	(0.0346)		(0.349)	(0.0450)		(0.424)	(0.0426)
Hours watching TV or videos	2,070		454			161			293		
Up to one hour		0.138		0.122	-0.0162		0.0981	-0.0398		0.134	-0.00375
		(0.372)		(0.345)	(0.0238)		(0.317)	(0.0451)		(0.357)	(0.0251)
1 hour to less than 3 hours		0.686		0.721	0.0352		0.650	-0.0357		0.758	0.0727**
		(0.501)		(0.473)	(0.0312)		(0.509)	(0.0465)		(0.448)	(0.0308)
More than 3 hours		0.177		0.158	-0.0190		0.252	0.0755		0.108	-0.069***
		(0.411)		(0.384)	(0.0306)		(0.463)	(0.0568)		(0.324)	(0.0239)
Help with homework	2,067		453			161			292		
Always		0.213		0.196	-0.0165		0.226	0.0129		0.181	-0.0321
		(0.442)		(0.419)	(0.0230)		(0.446)	(0.0330)		(0.403)	(0.0292)
Usually		0.277		0.261	-0.0161		0.273	-0.00398		0.254	-0.0226
		(0.483)		(0.463)	(0.0245)		(0.475)	(0.0352)		(0.456)	(0.0290)
Sometimes		0.417		0.439	0.0222		0.391	-0.0256		0.464	0.0475
		(0.532)		(0.523)	(0.0309)		(0.521)	(0.0443)		(0.522)	(0.0390)
Never or almost never		0.0938		0.104	0.0105		0.111	0.0167		0.101	0.00717
		(0.315)		(0.322)	(0.0234)		(0.335)	(0.0529)		(0.315)	(0.0199)
Regular bedtime on weekdays	2,070	0.888	454	0.912	0.0249	161	0.873	-0.0142	293	0.933	0.0456***
(Usually, always)		(0.341)		(0.298)	(0.0225)		(0.355)	(0.0555)		(0.262)	(0.0162)
PRIMARY SCHOOL											
Single sex school	2,070	0.00777	454	0.00753	-0.00024	161	0.00867	0.000902	293	0.00693	-0.00084
		(0.0947)		(0.0911)	(0.00516)		(0.0989)	(0.00937)		(0.0869)	(0.00560)
Faith school	2,012	0.300	442	0.286	-0.0141	157	0.202	-0.0972**	285	0.329	0.0289
		(0.495)		(0.475)	(0.0468)		(0.431)	(0.0392)		(0.489)	(0.0617)
Attended more than two	1,846	0.0414	413	0.0425	0.00111	148	0.0345	-0.00693	265	0.0469	0.00544
primary schools		(0.220)		(0.220)	(0.0113)		(0.201)	(0.0149)		(0.229)	(0.0161)
GEOGRAPHICAL											
IMD Income Domain	2,063		454			161			293		
Most Deprived Decile		0.105		0.113	0.00740		0.168	0.0623		0.0838	-0.0217
		(0.332)		(0.334)	(0.0268)		(0.399)	(0.0665)		(0.290)	(0.0261)
Least Deprived Decile		0.0748		0.0855	0.0106		0.0129	-0.062***		0.124	0.0490
		(0.284)		(0.295)	(0.0282)		(0.120)	(0.0153)		(0.345)	(0.0355)

	CONTROL		TREATMENT			SPONSOR-LED			CONVERTER		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
VARIABLES	**N**	**Mean**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**	**N**	**Mean**	**diff**
IMD Health Deprivation	2,063		454			161			293		
Most Deprived Decile		0.119		0.125	0.00526		0.114	-0.00499		0.130	0.0107
		(0.350)		(0.348)	(0.0313)		(0.340)	(0.0547)		(0.352)	(0.0385)
Least Deprived Decile		0.0790		0.0690	-0.0100		0.00835	-0.071***		0.101	0.0221
		(0.291)		(0.267)	(0.0315)		(0.0971)	(0.0191)		(0.316)	(0.0417)
IMD Education Skills	2,063		454			161			293		
Most Deprived Decile		0.0828		0.0362	-0.047***		0.0581	-0.0247		0.0246	-0.058***
		(0.298)		(0.197)	(0.0179)		(0.250)	(0.0348)		(0.162)	(0.0152)
Least Deprived Decile		0.128		0.121	-0.00649		0.126	-0.00137		0.119	-0.00920
		(0.360)		(0.344)	(0.0342)		(0.355)	(0.0536)		(0.339)	(0.0414)
IMD Barriers to Housing	2,063		454			161			293		
Most Deprived Decile		0.118		0.185	0.0663		0.258	0.140*		0.146	0.0275
		(0.349)		(0.409)	(0.0419)		(0.467)	(0.0724)		(0.370)	(0.0531)
Least Deprived Decile		0.0750		0.0945	0.0195		0.00817	-0.067***		0.140	0.0652
		(0.284)		(0.308)	(0.0394)		(0.0960)	(0.0154)		(0.364)	(0.0562)
IMD Crime Domain	2,063		454			161			293		
Most Deprived Decile		0.0651		0.0811	0.0159		0.0795	0.0143		0.0819	0.0167
		(0.266)		(0.288)	(0.0345)		(0.289)	(0.0307)		(0.287)	(0.0466)
Least Deprived Decile		0.0830		0.0854	0.00244		0.0444	-0.0386		0.107	0.0242
		(0.298)		(0.295)	(0.0271)		(0.220)	(0.0284)		(0.324)	(0.0370)
IMD Living Environment	2,063		454			161			293		
Most Deprived Decile		0.0866		0.0821	-0.00442		0.162	0.0755		0.0398	-0.0467*
		(0.304)		(0.289)	(0.0330)		(0.393)	(0.0741)		(0.205)	(0.0247)
Least Deprived Decile		0.0808		0.134	0.0534		0.0187	-0.062***		0.195	0.114**
		(0.294)		(0.359)	(0.0414)		(0.145)	(0.0154)		(0.415)	(0.0555)
AGE 7											
CM CHARACTERISTICS											
Word Reading Standardised	1,899	113.4	423	112.6	-0.780	154	110.1	-3.309***	269	114.0	0.602
		(18.43)		(18.67)	(1.169)		(18.62)	(1.141)		(18.54)	(1.541)
Pattern Construction T-Scores	1,899	53.31	421	54.12	0.816	154	52.71	-0.593	267	54.90	1.592
		(11.78)		(12.06)	(0.764)		(12.85)	(0.966)		(11.54)	(1.000)
Maths Standardised	1,901	98.61	422	98.07	-0.542	154	93.63	-4.983**	268	100.5	1.901
		(16.88)		(16.36)	(1.518)		(15.66)	(2.236)		(16.13)	(1.432)
SDQ Total Difficulties	1,831	-0.0227	407	-0.0108	0.0119	142	0.0999	0.123	265	-0.0705	-0.0478
		(1.079)		(1.020)	(0.0572)		(0.936)	(0.0822)		(1.057)	(0.0709)

	CONTROL		TREATMENT			SPONSOR-LED			CONVERTER		
VARIABLES	**(1) N**	**(2) Mean**	**(3) N**	**(4) Mean**	**(5) diff**	**(6) N**	**(7) Mean**	**(8) diff**	**(9) N**	**(10) Mean**	**(11) diff**
AGE 5											
CM CHARACTERISTICS											
Picture Similarities T-scores	1,960	55.14	430	54.32	-0.825	152	51.44	-3.700**	278	55.90	0.754
		(11.70)		(11.81)	(1.000)		(10.84)	(1.602)		(11.97)	(0.824)
Pattern Construction T-scores	1,960	50.85	431	50.39	-0.458	153	49.56	-1.291	278	50.85	0.00176
		(11.26)		(10.87)	(0.790)		(10.52)	(1.015)		(11.03)	(0.947)
SDQ Total Difficulties	1,835	-0.0125	402	-0.0229	-0.0104	139	0.280	0.292***	263	-0.191	-0.178**
		(1.081)		(1.001)	(0.0841)		(1.037)	(0.107)		(0.931)	(0.0801)
CHILDCARE											
CM attended childcare	1,954	0.930	429	0.942	0.0112	153	0.950	0.0196	276	0.937	0.00651
Providers		(0.280)		(0.256)	(0.0143)		(0.238)	(0.0185)		(0.265)	(0.0163)
Grandparents looks after CM	1,955	0.280	429	0.317	0.0369	153	0.292	0.0126	276	0.330	0.0505
(weekdays)		(0.493)		(0.507)	(0.0295)		(0.496)	(0.0346)		(0.513)	(0.0388)
How often read to CM	1,955		428			152			276		
Less often		0.0621		0.0469	-0.0152		0.0672	0.00509		0.0357	-0.0264**
		(0.265)		(0.230)	(0.0117)		(0.272)	(0.0256)		(0.202)	(0.0112)
At least weekly		0.446		0.462	0.0162		0.532	0.0862*		0.423	-0.0226
		(0.546)		(0.543)	(0.0318)		(0.542)	(0.0461)		(0.539)	(0.0326)
Daily		0.492		0.491	-0.00102		0.401	-0.0912**		0.541	0.0490
		(0.549)		(0.544)	(0.0310)		(0.533)	(0.0448)		(0.543)	(0.0328)

Note: The treatment group includes pupils enrolled in 2012/13 at LA maintained secondary schools that become academies in 2013/14 and 2014/15, whereas the control group includes pupils enrolled in LA maintained secondary schools that do not convert to academy status. The treatment group is composed by pupils attending prospective sponsor-led and converter academies. T-scores in ability tests are adjusted for age and mean scores of norming group (mean of 50 and standard deviation of 10). The quantity in parenthesis below the mean (or proportion in the case of categorical variables) is the standard deviation. All statistics, except the number of observations (N) that is unweighted, reflect MCS survey settings. The difference (diff) refers to the difference in means between treatment group and control group and the quantity in parenthesis below is the (robust) standard error clustered at the school level and reflecting MCS survey setting (University of London, Institute of Education, Centre for Longitudinal Studies, 2020a, 2020d, 2022). * significant at 10% level; ** significant at 5% level; *** significant at 1% level or below.

Table B4 Chi-square test (categorical variables)

VARIABLES	TREATMENT F-stat	TREATMENT p-value	SPONSOR-LED F-stat	SPONSOR-LED p-value	CONVERTER F-stat	CONVERTER p-value
AGE 17						
OUTCOME VARIABLES						
English Baccalaureate	0.0313	0.860	3.601	0.0585	0.718	0.397
STEM subjects	1.214	0.271	3.078	0.0802	5.939	0.0153
Facilitating subjects	0.00348	0.953	1.721	0.190	2.073	0.151
Vocational subjects	0.376	0.540	0.519	0.472	2.125	0.146
AGE 11						
CM CHARACTERISTICS						
CM Sex	0.140	0.708	4.435	0.0359	0.500	0.480
CM ethnic group	0.532	0.627	0.346	0.730	0.394	0.715
Language spoken in household	0.177	0.674	0.292	0.589	0.0352	0.851
CM's general level of health	3.406	0.0210	6.051	0.00169	2.735	0.0476
FAMILY CONTEXT						
Parents/carers in household	0.130	0.719	0.448	0.503	1.695	0.194
Grandparents in household	0.475	0.491	0.774	0.380	1.159	0.282
Main has longstanding illness	0.418	0.518	0.0274	0.869	0.493	0.483
SOCIOECONOMIC STATUS						
Free School Meal Eligible	0.286	0.593	3.707	0.0550	1.005	0.317
Housing Tenure	0.132	0.864	1.489	0.228	1.192	0.299
Parents Occupation	0.516	0.740	3.745	0.00800	0.388	0.811
Parents Education	0.539	0.681	3.443	0.0148	1.265	0.285
HOME ENVIRONMENT						
How many books at home	0.929	0.451	2.657	0.0439	0.136	0.971
How often visited library	1.956	0.163	0.493	0.483	1.411	0.236
How often attend religious service	5.793	0.0166	5.967	0.0150	1.925	0.166
Hours watching TV or videos	0.491	0.600	1.223	0.289	3.300	0.0377
How often anyone help homework	0.367	0.756	0.135	0.852	0.830	0.470
Regular bedtime on weekdays	1.032	0.310	0.0724	0.788	5.888	0.0157
SCHOOL-LEVEL VARIABLES						
Single sex or mixed school	0.00206	0.964	0.0102	0.920	0.0206	0.886
School with a particular faith	0.0893	0.765	5.158	0.0237	0.228	0.634
Number of Primary Schools attend	0.00988	0.921	0.187	0.666	0.125	0.724
GEOGRAPHICAL VARIABLES						
IMD Income Domain	0.133	0.988	1.610	0.166	0.597	0.717
IMD Health Deprivation	0.979	0.431	3.426	0.00423	1.024	0.399
IMD Education Skills and Training	1.184	0.311	0.822	0.522	1.276	0.265
IMD Barriers to Housing	0.728	0.620	2.682	0.0245	0.901	0.480
IMD Crime Domain	0.371	0.905	1.541	0.188	0.603	0.701
IMD Living Environment Domain	1.082	0.372	1.944	0.0965	2.386	0.0259
AGE 5						
CHILDCARE						
CM attended childcare providers	0.570	0.451	0.922	0.338	0.152	0.697
Who looks after CM weekdays	1.607	0.206	0.134	0.715	1.795	0.181
How often read to CM	0.655	0.515	1.809	0.165	2.416	0.0919

Note: The treatment group includes pupils enrolled in 2012/13 at LA maintained secondary schools that become academies in 2013/14 and 2014/15, whereas the control group includes pupils enrolled in LA maintained secondary schools that do not convert to academy status. The treatment group is composed by pupils attending prospective sponsor-led and converter academies. The F-statistic and p-value of the Chi-square test, reflecting MCS survey setting, is presented for each categorical variable (University of London, Institute of Education, Centre for Longitudinal Studies, 2020d, 2020b, 2022).

Table B5 Complete case analysis

OUTCOME	(1) OLS Treatment	(2) OLS/IV Treatment	(3) LOGIT Treatment Odds	(4) PSM Treatment Odds
AGE 17 **ACADEMIES**				
English Baccalaureate	0.00860 (0.0410)	0.00905 (0.0431)	1.020 (0.240)	1.051 (0.266)
STEM subjects	0.0645** (0.0276)	0.0678** (0.0290)	1.365** (0.183)	1.316 (0.233)
Facilitating Subjects	0.0197 (0.0239)	0.0207 (0.0252)	1.090 (0.200)	1.176 (0.227)
Vocational Subjects	0.00315 (0.0327)	0.00332 (0.0343)	0.999 (0.163)	0.978 (0.178)
First stage		0.951*** (0.0117)		
Observations	**1,728**	**1,728**	**1,728**	**1,728**
SPONSOR-LED				
English Baccalaureate	-0.00730 (0.0572)	-0.00773 (0.0606)	0.821 (0.329)	0.709 (0.432)
STEM subjects	-0.0503 (0.0485)	-0.0533 (0.0513)	0.778 (0.192)	0.691 (0.301)
Facilitating Subjects	-0.0176 (0.0462)	-0.0187 (0.0489)	0.766 (0.226)	0.706 (0.344)
Vocational Subjects	-0.000979 (0.0636)	-0.00104 (0.0674)	0.958 (0.281)	1.189 (0.594)
First stage		0.944*** (0.0165)		
Observations	**1,508**	**1,508**	**1,508**	**1,505**
CONVERTER				
English Baccalaureate	0.0201 (0.0540)	0.0211 (0.0566)	1.129 (0.329)	1.109 (0.378)
STEM subjects	0.111*** (0.0319)	0.116*** (0.0334)	1.736*** (0.275)	1.574* (0.367)
Facilitating Subjects	0.0378 (0.0300)	0.0395 (0.0314)	1.310 (0.326)	1.280 (0.334)
Vocational Subjects	0.00337 (0.0350)	0.00352 (0.0366)	1.003 (0.179)	0.893 (0.212)
First stage		0.956*** (0.0140)		
Observations	**1,625**	**1,625**	**1,625**	**1,621**

Note: The treatment group includes pupils enrolled in 2012/13 at LA maintained secondary schools that become academies in 2013/14 and 2014/15, whereas the control group includes pupils enrolled in LA maintained secondary schools that do not convert to academy status. The treatment group is composed by pupils attending prospective sponsor-led and converter academies. The quantity in parenthesis below the treatment coefficient is the (robust) standard error clustered at the school level and reflecting MCS survey setting. The instrumental variables estimates is based on a two stages linear regression. Propensity score matching (PSM) uses an Epanechinikov kernel with a bandwidth of 0.06 (Leuven & Sianesi, 2003). MCS survey weights are included in matching variables and multiplied by propensity score for matched logit regressions (Ridgeway et al., 2015). For PSM, standard errors are bootstrapped based on 1,000 replications (Kolenikov, 2010). Bootstrap samples respect school clustering. (University of London, Institute of Education, Centre for Longitudinal Studies, 2020a, 2020b, 2020c, 2022). * significant at 10% level; ** significant at 5% level; *** significant at 1% level or below.

Table B6 Sensitivity analysis

OUTCOME	(1) OLS Treatment	(2) OLS/IV Treatment	(3) LOGIT Treatment Odds	(4) LOGIT Treatment Odds	(5) PSM Treatment Odds
ACADEMIES					
English Baccalaureate	0.0162 (0.0329)	0.0171 (0.0347)	1.075 (0.233)	0.968 (0.201)	1.050 (0.173)
STEM subjects	0.0457* (0.0257)	0.0483* (0.0271)	1.271* (0.166)	1.205 (0.158)	1.169 (0.180)
Facilitating subjects	0.0231 (0.0234)	0.0244 (0.0247)	1.148 (0.171)	1.127 (0.171)	1.117 (0.173)
Vocational subjects	-0.0252 (0.0314)	-0.0266 (0.0333)	0.881 (0.133)	0.881 (0.138)	0.867 (0.154)
First stage		0.946*** (0.0105)			
Observations	**2,524**	**2,524**	**2,524**	**2,524**	**2,524**
SPONSOR-LED					
English Baccalaureate	0.0127 (0.0422)	0.0133 (0.0445)	1.033 (0.340)	1.082 (0.361)	0.954 (0.269)
STEM subjects	-0.0558 (0.0376)	-0.0587 (0.0396)	0.766 (0.157)	0.788 (0.162)	0.704 (0.178)
Facilitating subjects	-0.0418 (0.0458)	-0.0440 (0.0482)	0.738 (0.194)	0.752 (0.202)	0.805 (0.260)
Vocational subjects	-0.00872 (0.0613)	-0.00917 (0.0645)	0.953 (0.263)	0.940 (0.263)	1.149 (0.376)
First stage		0.950*** (0.0138)			
Observations	**2,231**	**2,231**	**2,231**	**2,231**	**2,231**
CONVERTER					
English Baccalaureate	0.0205 (0.0430)	0.0217 (0.0456)	1.097 (0.299)	0.927 (0.238)	1.119 (0.221)
STEM subjects	0.0890*** (0.0325)	0.0942*** (0.0343)	1.582*** (0.259)	1.467** (0.247)	1.531** (0.299)
Facilitating subjects	0.0571** (0.0288)	0.0604** (0.0306)	1.482** (0.294)	1.449* (0.297)	1.404* (0.243)
Vocational subjects	-0.0352 (0.0346)	-0.0373 (0.0368)	0.832 (0.143)	0.830 (0.148)	0.816 (0.168)
First stage		0.945*** (0.0128)			
Observations	**2,363**	**2,363**	**2,363**	**2,363**	**2,363**
Model includes secondary school level variables					
Single sex school				✓	
Faith school				✓	
Grammar school				✓	

Note: The treatment group includes pupils enrolled in 2012/13 at LA maintained secondary schools that become academies in 2013/14 and 2014/15, whereas the control group includes pupils enrolled in LA maintained secondary schools that do not convert to academy status. The treatment group is composed by pupils attending prospective sponsor-led and converter academies. The quantity in parenthesis below the treatment coefficient is the (robust) standard error clustered at the school level and reflecting MCS survey setting. The instrumental variables estimates is based on a two stages linear regression. Propensity score matching (PSM) uses an Epanechinikov kernel with a bandwidth of 0.06 (Leuven & Sianesi, 2003). MCS survey weights are included in matching variables and multiplied by propensity score for matched logit regressions (Ridgeway et al., 2015). Models' estimates are based on 70 imputations. The multiple imputation model uses chained equations. (University of London, Institute of Education, Centre for Longitudinal Studies, 2020a, 2020b, 2020c, 2022). * significant at 10% level; ** significant at 5% level; *** significant at 1% level or below.

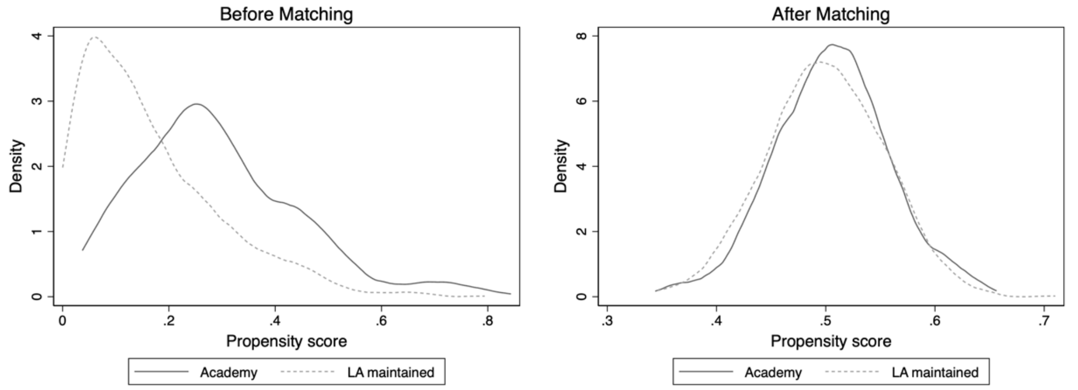
Table B7 Alternative matching methods (complete case analysis)

VARIABLES	(1) EPANECHINIKOV KERNEL BANDWIDTH OF 0.01 ATT	(2) EPANECHINIKOV KERNEL BANDWIDTH OF 0.06 ATT	(3) NEAREST NEIGHBOUR ATT	(4) MANALANOBIS ATT
ACADEMIES				
English Baccalaureate	0.0233	0.0250	0.0588	-0.0991
STEM subjects	0.0807	0.0792	0.0836	0.00929
Facilitating subjects	0.0371	0.0294	0.0526	-0.0650
Vocational subjects	-0.00553	-0.0135	0.0155	0.0805
Observations	**1,728**	**1,728**	**1,728**	**1,728**
R^2	**0.008**	**0.010**	**0.087**	**0.252**
SPONSOR-LED ACADEMIES				
English Baccalaureate	-0.0231	0.0211	-0.0194	-0.0583
STEM subjects	-0.0769	-0.0791	-0.0680	-0.0777
Facilitating subjects	-0.0677	-0.0655	-0.0583	-0.0777
Vocational subjects	0.0751	0.0647	0.0388	0.146
Observations	**1,508**	**1,508**	**1,508**	**1,508**
R^2	**0.072**	**0.096**	**1.000**	**1.000**
CONVERTER ACADEMIES				
English Baccalaureate	0.0256	0.0412	0.0364	-0.100
STEM subjects	0.123	0.121	0.123	0.0318
Facilitating subjects	0.0485	0.0647	0.0864	-0.0409
Vocational subjects	-0.0239	-0.0395	-0.0545	0.0455
Observations	**1,625**	**1,625**	**1,625**	**1,625**
R^2	**0.017**	**0.027**	**0.119**	**0.243**

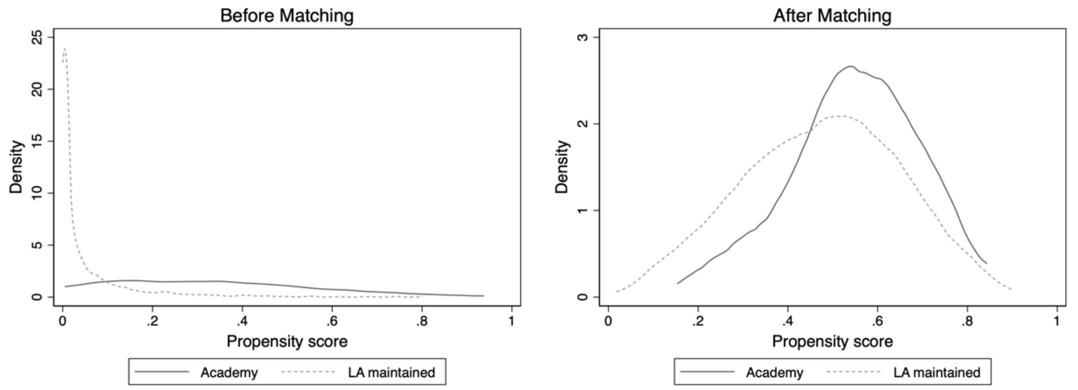
Note: The treatment group includes pupils enrolled in 2012/13 at LA maintained secondary schools that become academies in 2013/14 and 2014/15, whereas the control group includes pupils enrolled in LA maintained secondary schools that do not convert to academy status. The treatment group is composed by pupils attending prospective sponsor-led and converter academies. ATT is the average treatment effect on the treated. R^2 refers to probit estimation of the propensity score on the matched samples (Leuven & Sianesi, 2003). MCS survey settings are not applied to propensity score matching, but MCS survey weights are included among matching variables. (University of London, Institute of Education, Centre for Longitudinal Studies, 2020a, 2020c, 2022).

Figure B1 Density of propensity scores before and after matching (complete cases)

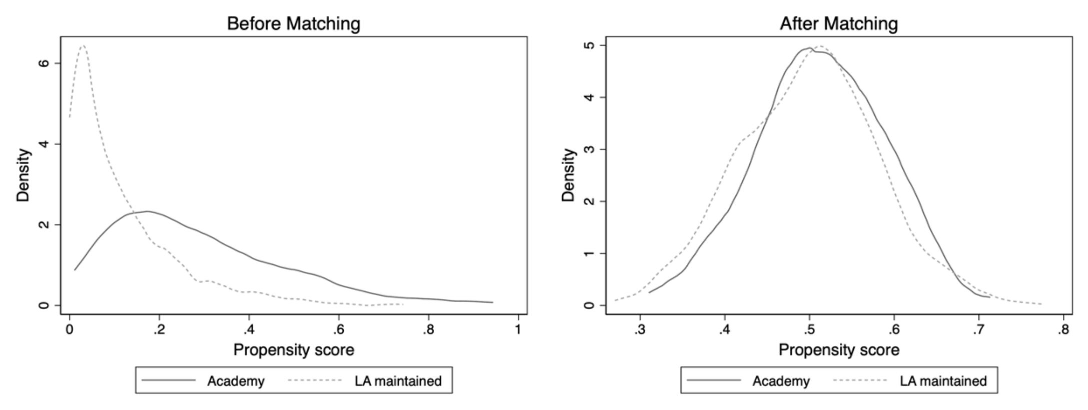
Academies



Sponsor-led Academies



Converter Academies



Note: The treatment group includes pupils enrolled in 2012/13 at LA maintained secondary schools that become academies in 2013/14 and 2014/15, whereas the control group includes pupils enrolled in LA maintained secondary schools that do not convert to academy status. The treatment group is composed by pupils attending prospective sponsor-led and converter academies. Matching variables include MCS survey weights. The propensity score density is displayed before and after matching. (University of London, Institute of Education, Centre for Longitudinal Studies, 2020a, 2020c, 2020d, 2022).

APPENDIX C SUPPLEMENTAL MATERIAL TO CHAPTER 5

Table C1 Detailed list of outcome variables

OUTCOME VARIABLES	
MANAGEMENT	
Responsibility for Curriculum index (School questionnaire)	Ratio of "yes" responses for the school governing board, the headteacher and teachers to "yes" responses for LA and the government
Establishing student assessment policies (set of binary variables: "no" / "yes")	Headteacher, Teachers, School governing board, LA, DfE
Choosing which textbooks are used (set of binary variables: "no" / "yes")	Headteacher, Teachers, School governing board, LA, DfE
Determining course content (set of binary variables: "no" / "yes")	Headteacher, Teachers, School governing board, LA, DfE
Deciding which courses are offered (set of binary variables: "no" / "yes")	Headteacher, Teachers, School governing board, LA, DfE
Responsibility for Resources index (School questionnaire)	Ratio of "yes" responses for the school governing board, the headteacher and teachers to "yes" responses for LA and the government
Selecting teachers for hire (set of binary variables: "no" / "yes")	Headteacher, Teachers, School governing board, LA, DfE
Firing teachers (set of binary variables: "no" / "yes")	Headteacher, Teachers, School governing board, LA, DfE
Establishing teachers' starting salaries (set of binary variables: "no" / "yes")	Headteacher, Teachers, School governing board, LA, DfE
Determining teachers' salary increases (set of binary variables: "no" / "yes")	Headteacher, Teachers, School governing board, LA, DfE
Formulating the school budget (set of binary variables: "no" / "yes")	Headteacher, Teachers, School governing board, LA, DfE
Deciding on budget allocations within the school (set of binary variables: "no" / "yes")	Headteacher, Teachers, School governing board, LA, DfE
Curricular Development index (School questionnaire)	Frequency of listed activities and behaviour in school during last academic year
I use student performance results to develop the school's educational goals.	did not occur, 1-2 times during the year, 3-4 times during the year, once a month, once a week, more than once a week
I make sure that the professional development activities of teachers are in accordance with the teaching goals of the school.	did not occur, 1-2 times during the year, 3-4 times during the year, once a month, once a week, more than once a week
I ensure that teachers work according to the school's educational goals.	did not occur, 1-2 times during the year, 3-4 times during the year, once a month, once a week, more than once a week
I discuss the school's academic goals with teachers at faculty meetings.	did not occur, 1-2 times during the year, 3-4 times during the year, once a month, once a week, more than once a week
Instructional Leadership index (School questionnaire)	Frequency of listed activities and behaviour in school during last academic year
I promote teaching practices based on recent educational research.	did not occur, 1-2 times during the year, 3-4 times during the year, once a month, once a week, more than once a week
I praise teachers whose students are actively participating in learning.	did not occur, 1-2 times during the year, 3-4 times during the year, once a month, once a week, more than once a week
I draw teachers' attention to the importance of pupils' development of critical and social capacities.	did not occur, 1-2 times during the year, 3-4 times during the year, once a month, once a week, more than once a week
Professional Development index (School questionnaire)	Frequency of listed activities and behaviour in school during last academic year
When a teacher has problems in his\her classroom, I take the initiative to discuss matters.	did not occur, 1-2 times during the year, 3-4 times during the year, once a month, once a week, more than once a week
I pay attention to disruptive behaviour in classrooms.	did not occur, 1-2 times during the year, 3-4 times during the year, once a month, once a week, more than once a week
When a teacher brings up a classroom problem, we solve the problem together.	did not occur, 1-2 times during the year, 3-4 times during the year, once a month, once a week, more than once a week

OUTCOME VARIABLES	
Teachers Participation index (School questionnaire)	Frequency of listed activities and behaviour in school during last academic year
I provide staff with opportunities to participate in school decision-making.	did not occur, 1-2 times during the year, 3-4 times during the year, once a month, once a week, more than once a week
I engage teachers to help build a school culture of continuous improvement.	did not occur, 1-2 times during the year, 3-4 times during the year, once a month, once a week, more than once a week
I ask teachers to participate in reviewing management practices.	did not occur, 1-2 times during the year, 3-4 times during the year, once a month, once a week, more than once a week
WMS Management index (School questionnaire)	see Table C2
POLICIES	
Learning Time (Student questionnaire)	
Mathematics	hours per week
English	hours per week
Science	hours per week
Total	hours per week
Student Admission to School (School questionnaire)	binary variables: “never” / “always, sometimes”
Student's record of academic performance	never, always, sometimes
Recommendation of feeder schools	never, always, sometimes
Endorsement of instructional or religious philosophy	never, always, sometimes
Student requires or is interested in a special programme	never, always, sometimes
Preference for family members of current/former pupils	never, always, sometimes
Residence in a particular area	never, always, sometimes
Student Assessment (School questionnaire)	
Mandatory Standardised Tests (binary variable: “less often” / “more than 2 times a year”)	never, 1-2 times a year, 3-5 times a year, monthly, more than once a month
Non-mandatory Standardised Tests (binary variable: “never” / “at least once a year”)	never, 1-2 times a year, 3-5 times a year, monthly, more than once a month
Teacher-developed Tests (binary variable: “less often” / “at least monthly”)	never, 1-2 times a year, 3-5 times a year, monthly, more than once a month
Teachers Ratings (binary variable: “less often” / “at least monthly”)	never, 1-2 times a year, 3-5 times a year, monthly, more than once a month
LEARNING ENVIRONMENT	
Disciplinary climate index (Student questionnaire)	Frequency of listed activities and behaviour in science classes
Students don't listen to what the teacher says.	every lesson, most lessons, some lessons, never or hardly ever
There is noise and disorder.	every lesson, most lessons, some lessons, never or hardly ever
The teacher waits long for students to quiet down.	every lesson, most lessons, some lessons, never or hardly ever
Students cannot work well.	every lesson, most lessons, some lessons, never or hardly ever
Students don't start working for a long time after the lesson begins.	every lesson, most lessons, some lessons, never or hardly ever
Inquiry-based teaching and learning practices index (Student questionnaire)	Frequency of listed activities and behaviour in science classes
Students are given opportunities to explain their ideas.	all lessons, most lessons, some lessons, never or hardly ever
Students spend time in the laboratory doing practical experiments.	all lessons, most lessons, some lessons, never or hardly ever
Students are required to argue about science questions.	all lessons, most lessons, some lessons, never or hardly ever
Students are asked to draw conclusions from an experiment they have conducted.	all lessons, most lessons, some lessons, never or hardly ever
The teacher explains how an idea can be applied to a number of different phenomena.	all lessons, most lessons, some lessons, never or hardly ever
Students are allowed to design their own experiments.	all lessons, most lessons, some lessons, never or hardly ever
There is a class debate about investigations.	all lessons, most lessons, some lessons, never or hardly ever
The teacher clearly explains relevance of concepts to our lives.	all lessons, most lessons, some lessons, never or hardly ever

OUTCOME VARIABLES	
Teacher-directed instruction index (Student questionnaire)	Frequency of listed activities and behaviour in science classes
The teacher explains scientific ideas.	never or almost never, some lessons, many lesson, every lesson or almost every lesson
A whole class discussion takes place with the teacher.	never or almost never, some lessons, many lesson, every lesson or almost every lesson
The teacher discusses our questions.	never or almost never, some lessons, many lesson, every lesson or almost every lesson
The teacher demonstrates an idea.	never or almost never, some lessons, many lesson, every lesson or almost every lesson
Adaption of instruction index (Student questionnaire)	Frequency of listed activities and behaviour in science classes
The teacher adapts the lesson to my class needs and knowledge.	never or almost never, some lessons, many lesson, every lesson or almost every lesson
The teacher provides individual help when a student has difficulties understanding a topic or task.	never or almost never, some lessons, many lesson, every lesson or almost every lesson
The teacher changes the structure of the lesson on a topic that most students find difficult to understand.	never or almost never, some lessons, many lesson, every lesson or almost every lesson
Student related factors affecting school climate index (School questionnaire)	Extent that learning of students is hindered by listed phenomena
Student truancy (binary variable: “very little, not at all” / “a lot, to some extent”)	not at all, very little, to some extent, a lot
Students skipping classes (binary variable: “very little, not at all” / “a lot, to some extent”)	not at all, very little, to some extent, a lot
Students lacking respect for teachers (binary variable: “very little, not at all” / “a lot, to some extent”)	not at all, very little, to some extent, a lot
Student use of alcohol or illegal drugs (binary variable: “very little, not at all” / “a lot, to some extent”)	not at all, very little, to some extent, a lot
Students intimidating or bullying other students (binary variable: “very little, not at all” / “a lot, to some extent”)	not at all, very little, to some extent, a lot
Teacher related factors affecting school climate index (School questionnaire)	Extent that learning of students is hindered by listed phenomena
Teachers not meeting individual students' needs (binary variable: “very little, not at all” / “a lot, to some extent”)	not at all, very little, to some extent, a lot
Teacher absenteeism (binary variable: “very little, not at all” / “a lot, to some extent”)	not at all, very little, to some extent, a lot
Staff resisting change (binary variable: “very little, not at all” / “a lot, to some extent”)	not at all, very little, to some extent, a lot
Teachers being too strict with students (binary variable: “very little, not at all” / “a lot, to some extent”)	not at all, very little, to some extent, a lot
Teachers not being well prepared for classes (binary variable: “very little, not at all” / “a lot, to some extent”)	not at all, very little, to some extent, a lot

Table C2 List of PISA variables used in WMS Management Index

PISA 2015 QUESTION	VALUE LABEL	INDEX SCORE
OPERATIONS		
1. Standardisation of Instructional Processes		
Does improvement exist at school? Implementation of a standardised policy for science subjects	no yes, based on school initiative yes, this is mandatory	0 1 1
2. Personalisation of Instruction and Learning		
Are standardized tests used in school? To adapt teaching to the students' needs	no yes	0 1
Are teacher-developed tests used in school? To adapt teaching to the students' needs	no yes	0 1
3. Data-driven Planning and Student Transitions		
Are standardized tests used in school? To inform parents about child's progress	no yes	0 1
Are teacher-developed tests used in school? To inform parents about child's progress	no yes	0 1
Are standardized tests used in school? To make decisions about students' retention or promotion	no yes	0 1
Are teacher-developed tests used in school? To make decisions about students' retention or promotion	no yes	0 1
Does improvement exist at school? Systematic recording of data such as attendance and professional development	no yes, based on school initiative yes, this is mandatory	0 1 1
Does improvement exist at school? Systematic recording of student test results and graduation rates	no yes, based on school initiative yes, this is mandatory	0 1 1
4. Adopting Educational Best Practices		
Frequency of the last academic year. I promote teaching practices based on recent educational research.	did not occur 1-2 times during the year 3-4 times during the year once a month once a week more than once a week	0 0.2 0.4 0.6 0.8 1
Our school invites specialists to conduct in-service training for teachers.	no yes	0 1
Are standardized tests used in school? To identify aspects of instruction or curriculum that should be improved	no yes	0 1
Are teacher-developed tests used in school? To identify aspects of instruction or curriculum that should be improved	no yes	0 1
Does improvement exist at school? Teacher mentoring	no yes, based on school initiative yes, this is mandatory	0 1 1
5. Continuous Improvement		
Frequency of the last academic year. When a teacher has problems in his\her classroom, I take the initiative to discuss matters.	did not occur 1-2 times during the year 3-4 times during the year once a month once a week more than once a week	0 0.2 0.4 0.6 0.8 1
Frequency of the last academic year. I engage teachers to help build a school culture of continuous improvement.	did not occur 1-2 times during the year 3-4 times during the year once a month once a week more than once a week	0 0.2 0.4 0.6 0.8 1
Internal evaluation \ Self-evaluation	no yes, based on school initiative yes, this is mandatory	0 1 1
Does improvement exist at school? External evaluation	no yes, based on school initiative yes, this is mandatory	0 1 1
Does improvement exist at school? Seeking written feedback from students (e.g. regarding lessons, teachers, resources)	no yes, based on school initiative yes, this is mandatory	0 1 1
7. Performance Review		
During the last academic year, used to monitor the practice of teachers? Tests or assessments of student achievement.	no yes	0 1

PISA 2015 QUESTION	VALUE LABEL	INDEX SCORE
During the last academic year, used to monitor the practice of teachers? Teacher peer review	no yes	0 1
During the last academic year, used to monitor the practice of teachers? Principal or senior staff observation of lessons	no yes	0 1
During the last academic year, used to monitor the practice of teachers? Observation of classes by inspectors.	no yes	0 1
8. Performance Dialogue		
Frequency of the last academic year. I ask teachers to participate in reviewing management practices.	did not occur 1-2 times during the year 3-4 times during the year once a month once a week more than once a week	0 0.2 0.4 0.6 0.8 1
Frequency of the last academic year. When a teacher brings up a classroom problem, we solve the problem together	did not occur 1-2 times during the year 3-4 times during the year once a month once a week more than once a week	0 0.2 0.4 0.6 0.8 1
9. Consequence Management		
We put measures derived from the results of external evaluations into practice promptly.	no yes	0 1
The impetus triggered by the external evaluation "disappeared" very quickly at our school.	no yes	1 0
10. Target Balance		
Frequency of the last academic year. I make sure that the professional development activities of teachers are in accordance with the teaching goals of the school.	did not occur 1-2 times during the year 3-4 times during the year once a month once a week more than once a week	0 0.2 0.4 0.6 0.8 1
Are standardized tests used in school? To compare the school to district or national performance	no yes	0 1
Are teacher-developed tests used in school? To compare the school to district or national performance	no yes	0 1
Are standardized tests used in school? To monitor the school's progress from year to year	no yes	0 1
Are teacher-developed tests used in school? To monitor the school's progress from year to year	no yes	0 1
Are standardized tests used in school? To compare the school with other schools	no yes	0 1
Are teacher-developed tests used in school? To compare the school with other schools	no yes	0 1
11. Target Inter-connection		
Frequency of the last academic year. I discuss the school's academic goals with teachers at faculty meetings.	did not occur 1-2 times during the year 3-4 times during the year once a month once a week more than once a week	0 0.2 0.4 0.6 0.8 1
13. Target Stretch		
Frequency of the last academic year. I use student performance results to develop the school's educational goals	did not occur 1-2 times during the year 3-4 times during the year once a month once a week more than once a week	0 0.2 0.4 0.6 0.8 1
14. Clarity and Comparability of Targets		
Frequency of the last academic year. I ensure that teachers work according to the school's educational goals.	did not occur 1-2 times during the year 3-4 times during the year once a month once a week more than once a week	0 0.2 0.4 0.6 0.8 1
Achievement data used in any of the following accountability procedures? Achievement data are posted publicly	no yes	0 1
Achievement data used in any of the following accountability procedures? Achievement data tracked over time by administrative authority.	no yes	0 1

PISA 2015 QUESTION	VALUE LABEL	INDEX SCORE
Does improvement exist at school? Written specification of the school's curricular profile and educational goals	no yes, based on school initiative yes, this is mandatory	0 1 1
Does improvement exist at school? Written specification of student performance standards	no yes, based on school initiative yes, this is mandatory	0 1 1
PEOPLE		
15. Rewarding High Performers		
Frequency of the last academic year. I praise teachers whose students are actively participating in learning.	did not occur 1-2 times during the year 3-4 times during the year once a month once a week more than once a week	0 0.2 0.4 0.6 0.8 1
Are standardized tests used in school? To make judgements about teachers' effectiveness	no yes	0 1
Are teacher-developed tests used in school? To make judgements about teachers' effectiveness	no yes	0 1
20. Attracting Talent, Creating a Distinctive Employee Value Proposition		
Teaching staff in your school has attended a programme of professional development? All teaching staff (%)	0 1-25 26-50 51-75 76-100	0 0.25 0.50 0.75 1
Teaching staff in your school has attended a programme of profess development? Science teaching staff (%)	0 1-25 26-50 51-75 76-100	0 0.25 0.50 0.75 1

Note: This version of the WMS management index for PISA 2015 is modelled on the work of Leaver et al. (2022) for PISA 2012. Several questions used by them are not available in PISA 2015, particularly those now included in the optional teacher questionnaire that is not employed in England. Some variables have slightly different labels in PISA 2015 and are coded accordingly. A few additional variables from the school questionnaire, not available in PISA 2012 but fitting WMS topics as outlined in Bloom et al. (2015) are included. The average of included questions is the topic's score. Each topic score is standardised. The management index is the average of the topics' scores.

Table C3 Descriptive statistics by school type in 2015

OUTCOME VARIABLES	LA MAINTAINED		ACADEMIES		CONVERTER		SPONSOR-LED		SAT		MAT	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean
MANAGEMENT												
Responsibility index												
Curriculum	1,826	0.003	1,945	0.058	1,279	-0.106	666	0.360*	715	-0.230	1,230	0.223
		(0.994)		(0.965)		(1.035)		(0.730)		(1.098)		(0.837)
Resources	1,826	-0.222	1,945	0.169**	1,279	0.242**	666	0.035	715	0.020	1,230	0.254***
		(1.071)		(0.916)		(0.883)		(0.960)		(1.014)		(0.844)
Leadership index												
Curricular Development	1,798	-0.124	1,871	0.281**	1,226	0.168	645	0.485***	662	-0.150	1,209	0.511***
		(1.037)		(0.975)		(0.976)		(0.939)		(0.835)		(0.966)
Instructional	1,748	-0.053	1,848	0.185	1,203	0.099	645	0.336*	662	-0.332	1,186	0.470***
		(0.962)		(1.040)		(1.078)		(0.951)		(1.021)		(0.936)
Professional Development	1,798	0.034	1,818	0.207	1,173	0.056	645	0.470**	662	-0.061	1,156	0.357*
		(0.880)		(0.980)		(0.769)		(1.222)		(0.768)		(1.052)
Teacher Participation	1,798	-0.095	1,822	0.157	1,202	0.108	620	0.248	662	-0.189	1,160	0.350**
		(0.901)		(1.068)		(1.072)		(1.056)		(1.060)		(1.023)
WMS Management index												
Overall	1,566	-0.049	1,616	0.313*	1,072	0.185	544	0.560***	586	-0.068	1,030	0.524***
		(0.971)		(0.970)		(0.993)		(0.873)		(1.045)		(0.856)
Operations	1,566	-0.082	1,616	0.292*	1,072	0.138	544	0.589***	586	-0.102	1,030	0.511***
		(0.977)		(1.015)		(1.044)		(0.883)		(1.115)		(0.881)
People	1,566	0.112	1,616	0.264	1,072	0.307	544	0.182	586	0.113	1,030	0.348
		(1.013)		(0.780)		(0.746)		(0.836)		(0.736)		(0.791)

OUTCOME VARIABLES	LA MAINTAINED (1) N	(2) Mean	ACADEMIES (3) N	(4) Mean	CONVERTER (5) N	(6) Mean	SPONSOR-LED (7) N	(8) Mean	SAT (9) N	(10) Mean	MAT (11) N	(12) Mean
POLICIES												
Learning Time (hours)												
Mathematics	1,923	3.929	2,431	3.999	1,584	3.801	847	4.355***	839	3.812	1,592	4.092*
		(0.968)		(1.127)		(1.053)		(1.166)		(1.256)		(1.044)
English	1,923	4.083	2,430	4.130	1,585	3.970	845	4.416**	838	4.022	1,592	4.183
		(1.566)		(1.525)		(1.274)		(1.861)		(1.585)		(1.492)
Science	1,919	4.795	2,427	4.824	1,582	4.876	845	4.732	837	4.780	1,590	4.846
		(1.543)		(1.714)		(1.709)		(1.718)		(1.809)		(1.664)
Total	1,811	26.095	2,292	26.235	1,510	26.076	782	26.527	797	26.046	1,495	26.329
		(5.079)		(5.582)		(5.392)		(5.905)		(5.811)		(5.461)
Admissions												
Academic Performance	1,798	0.194	1,893	0.306	1,227	0.307	666	0.304	663	0.345	1,230	0.285
		(0.396)		(0.461)		(0.461)		(0.460)		(0.475)		(0.452)
Feeder Schools	1,779	0.303	1,893	0.309	1,227	0.248	666	0.416	663	0.334	1,230	0.296
		(0.460)		(0.462)		(0.432)		(0.493)		(0.472)		(0.456)
School Philosophy or Religion	1,798	0.287	1,869	0.249	1,227	0.179	642	0.375	663	0.263	1,206	0.241
		(0.452)		(0.432)		(0.384)		(0.484)		(0.440)		(0.428)
School Specialism	1,748	0.463	1,893	0.362	1,227	0.285*	666	0.499	663	0.425	1,230	0.330
		(0.499)		(0.481)		(0.451)		(0.500)		(0.494)		(0.470)
Former Pupils and Siblings	1,798	0.772	1,893	0.789	1,227	0.792	666	0.785	663	0.816	1,230	0.776
		(0.419)		(0.408)		(0.406)		(0.411)		(0.387)		(0.417)
Catchment Area	1,798	0.869	1,893	0.878	1,227	0.925	666	0.795	663	0.925	1,230	0.853
		(0.337)		(0.328)		(0.263)		(0.404)		(0.263)		(0.354)
Pupil Assessment												
Mandatory Standardised Tests	1,798	0.317	1,891	0.170**	1,252	0.130**	639	0.246	688	0.165*	1,203	0.173*
(More than 2 times a year)		(0.465)		(0.376)		(0.336)		(0.430)		(0.371)		(0.379)
Non-mandatory Standardised	1,773	0.669	1,865	0.664	1,226	0.592	639	0.795	662	0.694	1,203	0.647
Tests (At least once a year)		(0.471)		(0.472)		(0.491)		(0.404)		(0.461)		(0.478)
Teacher-developed Tests	1,798	0.514	1,865	0.554	1,226	0.581	639	0.505	662	0.612	1,203	0.523
(At least monthly)		(0.500)		(0.497)		(0.493)		(0.500)		(0.487)		(0.499)
Teachers Ratings	1,798	0.562	1,865	0.436	1,226	0.422	639	0.462	662	0.490	1,203	0.407
(At least monthly)		(0.496)		(0.496)		(0.494)		(0.499)		(0.500)		(0.491)

OUTCOME VARIABLES	LA MAINTAINED		ACADEMIES		CONVERTER		SPONSOR-LED		SAT		MAT	
	(1) N	(2) Mean	(3) N	(4) Mean	(5) N	(6) Mean	(7) N	(8) Mean	(9) N	(10) Mean	(11) N	(12) Mean
LEARNING ENVIRONMENT												
Learning Environment index												
Disciplinary Climate	1,910	-0.101	2,429	-0.031	1,581	0.001*	848	-0.089	834	-0.055	1,595	-0.019
		(1.028)		(1.010)		(0.978)		(1.063)		(0.967)		(1.030)
Inquiry-based Learning	1,886	-0.002	2,408	0.003	1,570	0.006	838	-0.003	827	-0.081	1,581	0.044
		(1.029)		(1.017)		(0.999)		(1.049)		(1.068)		(0.990)
Teacher-directed Instruction	1,875	-0.025	2,403	-0.005	1,569	0.008	834	-0.029	825	-0.027	1,578	0.005
		(1.046)		(1.007)		(0.970)		(1.070)		(0.942)		(1.037)
Adaptive Instruction	1,834	-0.027	2,353	0.020	1,539	0.003	814	0.050	815	-0.052	1,538	0.055
		(1.019)		(1.004)		(0.978)		(1.048)		(0.975)		(1.016)
School Climate index												
Pupil behaviour	1,798	0.366	1,865	0.199	1,226	0.053*	639	0.467	662	0.107	1,203	0.249
		(0.826)		(0.933)		(0.971)		(0.791)		(1.034)		(0.870)
Teacher behaviour	1,798	0.179	1,840	0.164	1,201	0.176	639	0.141	637	0.353	1,203	0.066
		(0.943)		(0.943)		(0.911)		(0.998)		(0.791)		(0.999)

Note: The quantity in parenthesis below the mean (or proportion in the case of categorical variables) is the standard deviation. All statistics, except the number of observations (N) that is unweighted, reflect PISA survey settings. Estimates reflect the number of pupils in schools where variables are observed. The sample is restricted to England (OECD, 2016a). * significant at 10% level; ** significant at 5% level; *** significant at 1% level or below (in relation to LA maintained schools).

Table C4 Differences in school characteristics in 2012

	ACADEMIES		CONVERTER		SPONSOR-LED		MAT	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
	N	**diff.**	**N**	**diff.**	**N**	**diff.**	**N**	**diff.**
Number of pupils on roll	1,156	-1.743	675	121.180	481	-175.007	869	-70.604
		(116.873)		(124.298)		(257.697)		(129.511)
Boys (%)	1,156	-0.005	675	-0.024	481	0.020	869	-0.020
		(0.035)		(0.048)		(0.049)		(0.046)
FSM eligible (%)	1,132	-0.027	675	-0.010	457	-0.065	845	0.004
		(0.052)		(0.046)		(0.105)		(0.051)
English natives (%)	1,132	0.019	675	-0.083*	457	0.179	845	0.034
		(0.080)		(0.049)		(0.174)		(0.076)
SEN statement (%)	1,132	-0.003	675	-0.006	457	0.000	845	0.001
		(0.007)		(0.009)		(0.011)		(0.008)
KS2 average score	1,132	-0.166	675	-0.181	457	-0.001	845	-0.650
		(0.467)		(0.668)		(0.408)		(0.464)
Obtaining 5+ A*-C or equiv (including English and Mathematics) (%)								
2012	1,156	-0.001	675	0.039	481	-0.033	869	-0.053
		(0.044)		(0.053)		(0.046)		(0.049)
2011	1,127	-0.019	675	0.004	452	-0.034	840	-0.066
		(0.039)		(0.050)		(0.034)		(0.043)
2010	1,127	-0.018	675	-0.004	452	-0.020	840	-0.074
		(0.045)		(0.060)		(0.039)		(0.047)
2009	1,127	0.022	675	0.034	452	0.020	840	-0.028
		(0.042)		(0.058)		(0.047)		(0.045)
Value-added based on best 8 GCSE	1,132	-2.332	675	5.178	457	-11.882	845	-5.366
		(6.071)		(8.493)		(9.152)		(6.561)
Total average point score (capped)	1,156	-1.756	675	1.167	481	-1.731	869	-7.789
		(7.289)		(8.390)		(9.971)		(7.666)
Pupil/Teacher ratio	1,104	0.129	675	0.127	429	-0.025	817	0.381
		(0.381)		(0.597)		(0.923)		(0.440)

	ACADEMIES		CONVERTER		SPONSOR-LED		MAT	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
	N	**diff.**	**N**	**diff.**	**N**	**diff.**	**N**	**diff.**
Last Ofsted report (2012)								
Outstanding	1,132	-0.033	675	-0.091	457	0.083	845	-0.146*
		(0.088)		(0.159)		(0.070)		(0.084)
Good	1,132	-0.058	675	0.128	457	-0.325	845	-0.118
		(0.187)		(0.208)		(0.298)		(0.205)
Requires improvement	1,132	-0.080	675	-0.037	457	-0.157	845	0.005
		(0.176)		(0.179)		(0.304)		(0.192)
Inadequate	1,132	0.170***	675	N/A	457	0.400***	845	0.259***
		(0.053)		N/A		(0.107)		(0.076)

Note: The treatment group includes pupils enrolled in secondary schools that become academies between 2012/13 and 2014/15, whereas the control group includes pupils enrolled in LA maintained secondary schools that convert to academy status in 2015/16 and 2016/17. The treatment group is composed by pupils attending converter and sponsor-led academies that are either a single-academy trust or part of a multi-academy trust. The estimates are based on a pupil level regression model that reflects the number of pupils in schools where variables are observed. The quantity in parenthesis below the difference coefficient is the standard error. All statistics, except the number of observations (N) that is unweighted, reflect PISA survey settings (OECD, 2016a). * significant at 10% level; ** significant at 5% level; *** significant at 1% level or below.

Table C5 Sensitivity analysis using extended control group

OUTCOME VARIABLES	ACADEMIES (1) N	ACADEMIES (2) Treatment	CONVERTER (3) N	CONVERTER (4) Treatment	SPONSOR-LED (5) N	SPONSOR-LED (6) Treatment	MAT (7) N	MAT (8) Treatment
MANAGEMENT								
Responsibility index								
Curriculum	1,191	-0.388	760	-0.594*	431	-0.062	946	-0.132
		(0.259)		(0.342)		(0.404)		(0.267)
Resources	1,191	-0.080	760	-0.154	431	0.006	946	0.289
		(0.321)		(0.408)		(0.571)		(0.336)
Leadership index								
Curricular Development	1,143	0.331	733	0.045	410	0.819	925	0.626**
		(0.267)		(0.257)		(0.664)		(0.284)
Instructional	1,143	0.316	733	0.155	410	0.554	925	0.759**
		(0.332)		(0.389)		(0.578)		(0.308)
Professional Development	1,116	-0.068	706	-0.112	410	-0.019	898	0.157
		(0.306)		(0.346)		(0.563)		(0.360)
Teacher Participation	1,143	0.459	733	0.482	410	0.355	925	0.859***
		(0.293)		(0.407)		(0.418)		(0.278)
WMS Management index								
Overall	967	0.267	607	0.102	360	0.533	800	0.539*
		(0.296)		(0.344)		(0.565)		(0.315)
Operations	967	0.304	607	0.121	360	0.605	800	0.534
		(0.307)		(0.380)		(0.556)		(0.335)
People	967	-0.012	607	-0.027	360	-0.018	800	0.320
		(0.309)		(0.410)		(0.449)		(0.290)

OUTCOME VARIABLES	ACADEMIES (1) N	(2) Treatment	CONVERTER (3) N	(4) Treatment	SPONSOR-LED (5) N	(6) Treatment	MAT (7) N	(8) Treatment
POLICIES								
Learning Time (hours)								
Mathematics	1,408	0.242	860	-0.098	548	0.725***	1,132	0.390**
		(0.156)		(0.189)		(0.229)		(0.161)
English	1,406	0.053	858	-0.183	548	0.342*	1,130	0.115
		(0.167)		(0.227)		(0.200)		(0.179)
Science	1,405	0.054	858	-0.210	547	0.512	1,130	0.111
		(0.189)		(0.234)		(0.316)		(0.216)
Total	1,317	-0.408	814	-0.186	503	-0.847	1,055	-0.347
		(0.404)		(0.488)		(0.660)		(0.421)
Admissions								
Academic Performance	1,164	-0.128	733	-0.099	431	-0.173	946	-0.069
		(0.148)		(0.199)		(0.219)		(0.160)
Feeder Schools	1,164	-0.185	733	-0.261	431	-0.057	946	-0.182
		(0.148)		(0.207)		(0.231)		(0.159)
School Philosophy or Religion	1,140	0.137	733	0.059	407	0.280	922	0.135
		(0.159)		(0.218)		(0.231)		(0.187)
School Specialism	1,164	-0.169	733	-0.297	431	0.101	946	-0.279*
		(0.150)		(0.195)		(0.237)		(0.166)
Former Pupils and Siblings	1,164	0.021	733	-0.079	431	0.213	946	0.048
		(0.141)		(0.172)		(0.226)		(0.147)
Catchment Area	1,164	-0.033	733	-0.110	431	0.133	946	-0.056
		(0.106)		(0.095)		(0.215)		(0.118)
Pupil Assessment								
Mandatory Standardised Tests	1,137	0.060	733	-0.006	404	0.190	919	0.053
(More than 2 times a year)		(0.090)		(0.131)		(0.128)		(0.118)
Non-mandatory Standardised	1,137	-0.039	733	-0.035	404	-0.049	919	-0.084
Tests (At least once a year)		(0.140)		(0.188)		(0.210)		(0.168)
Teacher-developed Tests	1,137	0.141	733	0.180	404	0.074	919	0.098
(At least monthly)		(0.151)		(0.181)		(0.208)		(0.157)
Teachers Ratings	1,137	-0.153	733	-0.030	404	-0.348	919	-0.203
(At least monthly)		(0.169)		(0.193)		(0.272)		(0.179)

OUTCOME VARIABLES	ACADEMIES (1) N	ACADEMIES (2) Treatment	CONVERTER (3) N	CONVERTER (4) Treatment	SPONSOR-LED (5) N	SPONSOR-LED (6) Treatment	MAT (7) N	MAT (8) Treatment
LEARNING ENVIRONMENT								
Learning Environment index								
Disciplinary Climate	1,387	0.110	839	0.121	548	0.129	1,115	0.118
		(0.086)		(0.088)		(0.171)		(0.100)
Inquiry-based Learning	1,371	0.121	830	0.097	541	0.172	1,105	0.128
		(0.088)		(0.104)		(0.175)		(0.087)
Teacher-directed Instruction	1,362	0.046	823	0.083	539	-0.029	1,095	0.057
		(0.076)		(0.076)		(0.156)		(0.079)
Adaptive Instruction	1,329	0.107	805	0.044	524	0.189	1,066	0.137*
		(0.072)		(0.073)		(0.151)		(0.076)
School Climate index								
Pupil behaviour	1,137	-0.345	733	-0.271	404	-0.514	919	-0.425*
		(0.247)		(0.341)		(0.350)		(0.255)
Teacher behaviour	1,137	0.052	733	0.353	404	-0.491	919	-0.196
		(0.302)		(0.256)		(0.671)		(0.393)

Note: The treatment group includes pupils enrolled in secondary schools that become academies between 2012/13 and 2014/15, whereas the control group includes pupils enrolled in LA maintained secondary schools that convert to academy status between 2015/16 and 2018/19. The treatment group is composed by pupils attending converter and sponsor-led academies that are either a single-academy trust or part of a multi-academy trust. The estimates are based on a pupil level regression model that reflects the number of pupils in schools where variables are observed. PISA indices are standardised to mean zero and standard deviation one. The quantity in parenthesis below the treatment coefficient is the standard error. All statistics, except the number of observations (N) that is unweighted, reflect PISA survey settings (OECD, 2016a). * significant at 10% level; ** significant at 5% level; *** significant at 1% level or below.

Table C6 Analysis using variables of PISA indices

	ACADEMIES		CONVERTER		SPONSOR-LED		MAT	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	N	Treatment	N	Treatment	N	Treatment	N	Treatment
LEADERSHIP INDEX								
Curricular Development (frequency)								
I use pupil performance results to develop the school's educational goals	889	-0.142 (0.171)	556	-0.113 (0.227)	333	-0.190 (0.291)	671	-0.036 (0.186)
I make sure that the professional development activities of teachers are in accordance with the teaching goals of the school	889	0.029 (0.158)	556	-0.044 (0.207)	333	0.146 (0.303)	671	0.116 (0.162)
I ensure that teachers work according to the school's educational goals	863	0.134 (0.150)	530	-0.060 (0.063)	333	0.386 (0.285)	645	0.146 (0.150)
I discuss the school's academic goals with teachers at faculty meetings	889	0.209 (0.134)	556	0.097 (0.223)	333	0.386 (0.285)	671	0.325*** (0.116)
Instructional (frequency)								
I promote teaching practices based on recent educational research	889	0.182 (0.183)	556	0.049 (0.219)	333	0.386 (0.285)	671	0.348** (0.164)
I praise teachers whose pupils are actively participating in learning	889	0.116 (0.091)	556	0.034 (0.035)	333	0.248 (0.254)	671	0.116 (0.091)
Professional Development (frequency)								
When a teacher has problems in his classroom, I take the initiative to discuss matters	889	-0.080 (0.140)	556	-0.148 (0.105)	333	0.029 (0.259)	671	-0.035 (0.103)
When a teacher brings up a classroom problem, we solve the problem together	862	0.076 (0.141)	529	0.064 (0.205)	333	0.094 (0.252)	644	0.126 (0.140)
Teacher Participation (frequency)								
I engage teachers to help build a school culture of continuous improvement	889	-0.106 (0.151)	556	-0.107 (0.226)	333	-0.110 (0.075)	671	0.080 (0.121)
I ask teachers to participate in reviewing management practices	889	0.154 (0.180)	556	0.124 (0.217)	333	0.197 (0.275)	671	0.237 (0.199)

	ACADEMIES		CONVERTER		SPONSOR-LED		MAT	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	N	Treatment	N	Treatment	N	Treatment	N	Treatment
SCHOOL CLIMATE INDEX								
Pupil behaviour								
Pupil truancy	883	-0.128	556	-0.053	327	-0.248	665	-0.179
		(0.133)		(0.153)		(0.254)		(0.121)
Pupils skipping classes	883	-0.035	556	0.083	327	-0.224	665	-0.086
		(0.095)		(0.082)		(0.238)		(0.083)
Pupils lacking respect for teachers	883	-0.162	556	-0.123	327	-0.226	665	-0.158
		(0.149)		(0.195)		(0.300)		(0.158)
Pupils use of alcohol or illegal drugs	883	-0.076	556	-0.124	327	N/A	665	-0.076
		(0.080)		(0.117)		N/A		(0.080)
Pupils intimidating or bullying other pupils	883	-0.052	556	0.071	327	-0.248	665	-0.095
		(0.101)		(0.071)		(0.254)		(0.091)
Teacher behaviour								
Teachers not meeting individual pupils' needs	883	-0.118	556	0.133	327	-0.520***	665	-0.217
		(0.189)		(0.224)		(0.167)		(0.197)
Teacher absenteeism	883	0.246*	556	0.508***	327	-0.172	665	0.171
		(0.145)		(0.122)		(0.306)		(0.166)
Staff resisting change	883	0.170	556	0.124	327	0.246**	665	0.028
		(0.168)		(0.225)		(0.114)		(0.169)
Teachers being too strict with pupils	883	-0.008	556	N/A	327	-0.023	665	0.042
		(0.106)		N/A		(0.297)		(0.126)
Teachers not being well prepared for classes	883	-0.094	556	N/A	327	-0.247	665	-0.044
		(0.122)		N/A		(0.317)		(0.143)

Note: The treatment group includes pupils enrolled in secondary schools that become academies between 2012/13 and 2014/15, whereas the control group includes pupils enrolled in LA maintained secondary schools that convert to academy status in 2015/16 and 2016/17. The treatment group is composed by pupils attending converter and sponsor-led academies that are either a single-academy trust or part of a multi-academy trust. The estimates are based on a pupil level regression model that reflects the number of pupils in schools where variables are observed. The variables of the climate index represent the percentage of pupils in schools where headteachers say activity occurs a lot or to some extent. The variables of the leadership index represent the percentage of pupils in schools where headteachers say activity occurs at least once a month. The quantity in parenthesis below the treatment coefficient is the standard error. All statistics, except the number of observations (N) that is unweighted, reflect PISA survey settings (OECD, 2016a). * significant at 10% level; ** significant at 5% level; *** significant at 1% level or below.

Table C7 Within school responsibilities analysis

	ACADEMIES		CONVERTER		SPONSOR-LED		MAT	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
RESPONSIBILITY FOR	**N**	**Treatment**	**N**	**Treatment**	**N**	**Treatment**	**N**	**Treatment**
Selecting teachers for hire:								
Headteacher	937	-0.049	583	-0.082	354	0.000	692	N/A
		(0.048)		(0.077)		(0.000)		N/A
Teachers	937	-0.196	583	0.084	354	-0.635**	692	-0.235
		(0.160)		(0.218)		(0.267)		(0.170)
School Governing Board	937	-0.136	583	-0.170	354	-0.074	692	-0.065
		(0.207)		(0.244)		(0.319)		(0.206)
Firing teachers:								
Headteacher	937	-0.106	583	-0.176	354	0.000	692	N/A
		(0.070)		(0.110)		(0.000)		N/A
School Governing Board	937	-0.382***	583	-0.338**	354	-0.450***	692	-0.340***
		(0.106)		(0.149)		(0.150)		(0.128)
Determining teachers' salary increase:								
Headteacher	937	0.033	583	0.030	354	0.038	692	0.093
		(0.143)		(0.210)		(0.254)		(0.140)
Teachers	937	0.038	583	N/A	354	0.096	692	0.061
		(0.056)		N/A		(0.141)		(0.090)
School Governing Board	937	-0.249**	583	-0.219*	354	-0.288	692	-0.298**
		(0.111)		(0.121)		(0.269)		(0.131)
Formulating the school budget:								
Headteacher	937	-0.156**	583	-0.181*	354	-0.120	692	-0.077
		(0.075)		(0.100)		(0.112)		(0.073)
School Governing Board	937	-0.221	583	-0.001	354	-0.559***	692	-0.252
		(0.148)		(0.218)		(0.152)		(0.163)
Establishing pupil disciplinary policies:								
Headteacher	937	-0.123**	583	-0.082	354	-0.186**	692	-0.119*
		(0.061)		(0.077)		(0.094)		(0.065)
Teachers	937	0.033	583	0.038	354	0.028	692	0.044
		(0.211)		(0.277)		(0.315)		(0.216)
School Governing Board	937	-0.184	583	0.004	354	-0.469*	692	-0.257
		(0.202)		(0.265)		(0.253)		(0.204)

RESPONSIBILITY FOR	ACADEMIES (1) N	ACADEMIES (2) Treatment	CONVERTER (3) N	CONVERTER (4) Treatment	SPONSOR-LED (5) N	SPONSOR-LED (6) Treatment	MAT (7) N	MAT (8) Treatment
Establishing pupil assessment policies:								
Headteacher	937	-0.145**	583	-0.085	354	-0.237**	692	-0.152*
		(0.067)		(0.077)		(0.117)		(0.082)
Teachers	937	-0.034	583	-0.133	354	0.127	692	-0.019
		(0.201)		(0.240)		(0.310)		(0.202)
School Governing Board	937	-0.070	583	0.207	354	-0.495*	692	-0.193
		(0.214)		(0.279)		(0.280)		(0.208)
Approving pupils for admission to the school:								
Headteacher	937	-0.059	583	0.073	354	-0.272*	692	-0.016
		(0.179)		(0.262)		(0.150)		(0.186)
Teachers	937	0.010	583	0.053	354	-0.058	692	0.051
		(0.189)		(0.228)		(0.301)		(0.203)
School Governing Board	937	0.184	583	0.193	354	0.173	692	0.253
		(0.157)		(0.240)		(0.304)		(0.172)
Determining course content:								
Headteacher	937	-0.106	583	-0.213	354	0.068	692	-0.037
		(0.133)		(0.186)		(0.068)		(0.137)
Teachers	937	-0.076	583	-0.082	354	-0.068	692	-0.044
		(0.054)		(0.077)		(0.068)		(0.044)
Deciding which courses are offered:								
Headteacher	937	0.045	583	-0.014	354	0.137	692	0.106
		(0.086)		(0.163)		(0.242)		(0.085)
Teachers	937	-0.149	583	-0.162	354	-0.121	692	-0.135
		(0.138)		(0.105)		(0.267)		(0.114)
School Governing Board	937	0.042	583	0.073	354	N/A	692	0.033
		(0.113)		(0.188)		N/A		(0.121)

Note: The treatment group includes pupils enrolled in secondary schools that become academies between 2012/13 and 2014/15, whereas the control group includes pupils enrolled in LA maintained secondary schools that convert to academy status in 2015/16 and 2016/17. The treatment group is composed by pupils attending converter and sponsor-led academies that are either a single-academy trust or part of a multi-academy trust. The estimates are based on a pupil level regression model that reflects the number of pupils in schools where variables are observed. The quantity in parenthesis below the treatment coefficient is the standard error. All statistics, except the number of observations (N) that is unweighted, reflect PISA survey settings (OECD, 2016a). * significant at 10% level; ** significant at 5% level; *** significant at 1% level or below.