Economics of Education Review 109 (2025) 102717

Contents lists available at ScienceDirect

Economics of Education Review

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

Free to improve? The impact of free school attendance in England☆

Marco Bertoni [a], Gabriel Heller-Sahlgren [b], Olmo Silva [c,*]

[a] University of Padova and IZA-Bonn
[b] Research Institute of Industrial Economics (IFN) and London School of Economics (LSE)
[c] LSE and IZA-Bonn

ARTICLE INFO

JEL Classification:
I21
I24
I28

Keywords:
School autonomy
Quasi-markets
Free schools
Achievement

ABSTRACT

We evaluate the impact of attending two secondary free schools in England – new autonomous state-funded start-ups – using admission lotteries and a distance-based regression discontinuity design. We characterise each school's ethos through text analysis of vision statements: one follows a 'no excuses' paradigm common among US charter schools; the other adopts a 'classical liberal', knowledge-rich approach. These features distinguish them from each other and from counterfactual schools attended by rejected applicants. Despite pedagogical differences, both schools significantly improve test scores, reduce absences, and lower student mobility. Our findings support policies promoting horizontal differentiation in publicly funded education.

1. Introduction

Over the past decades, policymakers in several countries – including Canada, Chile, England, Sweden, and the US – have promoted reforms that aim to improve school standards by creating quasi-markets in education. These reforms typically include parental choice, school competition, operational autonomy, and entry of new providers.

In this paper, we focus on the latter element – i.e., the entry of new start-up schools – by analysing the impact of free schools in England. Introduced in 2010, free schools are newly established state-funded schools set up by parents, charities, and other non-governmental groups. Institutionally, free schools share the same characteristics as academies – a type of publicly-funded school that enjoys significant autonomy and operates outside the control of local government, in terms of staff management and pay, length of school terms and days, and pedagogical approaches. However, while academies are conversions from existing state-managed schools, free schools are new schools with no prior institutional history. This start-up (free) vs. conversion

☆ July 2025

We would like to thank the Free Schools Network (FSN) for facilitating discussions with the two free schools that have shared data with us. Our findings represent our views and do not mirror or are influenced by the views of the FSN or the schools involved in our analysis. We are grateful to the Editor (Richard Murphy), to three anonymous reviewers, to seminar participants at the University of New South Wales in Sydney, Ca' Foscari University of Venice, CEP Education Work-in-Progress meeting, CEP 2023 Annual Conference, CESIfo Effee 2022 Conference, EALE 2022 Annual Conference, IFS Education Seminar, SED 2022 Conference, SEHO 2022 Conference, SOLE 2023 Conference, Workshop on 'Human Capital, Labour Markets and Public Policy' in Naples, and Atila Abdulkadiroğlu, Gigi Foster, Björn Öckert, Marco Ovidi, Parag Pathak and Steve Ross for comments and suggestions. We are also truly indebted to Filippo Boeri for helping us structure our text analysis. We are responsible for any error or omission. Office for National Statistics (ONS) Disclaimer: This work was produced using statistical data from ONS. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates. Marco Bertoni acknowledges funding from the European Union – NextGenerationEU, within the PRIN 2022 Project (DM 104 of 02.02.2022) "STEP BY STEP: Educational stages, choices, achievement and long-run outcomes" Mission 4, Component 1, CUP: C53D23002740006 Project code: 2022HLSETF. Olmo Silva acknowledges funding from the ESCR Grant Number ES/T014431/1. Gabriel Heller-Sahlgren acknowledges funding from the ESRC Grant Number ES/J500070/1, the Jan Wallander and Tom Hedelius Foundation, and the former research programme "The Economics of the Service Sector" at the Research Institute of Industrial Economics. The data used in this analysis cannot be shared with the readership of this journal. Pupil and school census data used in this work are provided by the UK Department for Education (DfE) on a case-by-case basis and under a signed agreement that the data will not be shared. The authors remain available for guidance on how to access the data and are happy to share the Stata and R codes written to assemble, organize and analyze the data.

* Corresponding author at: LSE, Houghton Street, WC2A 2AE, London, U.K.
E-mail address: o.silva@lse.ac.uk (O. Silva).

https://doi.org/10.1016/j.econedurev.2025.102717
Received 13 February 2025; Received in revised form 26 August 2025; Accepted 3 September 2025
Available online 30 September 2025

(academy) distinction is central to our analysis.

Like charter schools in the US and Canada, and *friskolor* in Sweden, English free schools embody a policy agenda that seeks to liberalise the supply side of the education quasi-market, an understudied part of the English education reforms enacted over the past two decades.[1] Proponents of such reforms argue that entry of new providers enhances choice, stiffens competition, and promotes horizontal differentiation and pedagogical innovation. But are start-up schools good at educating their own pupils?

This paper seeks to answer this question through a quantitative case study of two high profile secondary free schools that were set up in the early days of the programme. As there is no prior causal research on the effectiveness of English free schools, even evidence from two schools helps fill a key gap. Early research on US charter schools also relied on small samples: Angrist et al. (2010, 2012) analysed KIPP Lynn; Dobbie and Fryer (2011) studied three Promise Academies; Curto and Fryer (2014) examined two schools; and Abdulkadiroğlu et al. (2011) focused on three secondary schools.[2] Yet these studies proved important contributions to the literature. We view our own research in a similar spirit.

Our two free schools differ markedly in both context and educational approach. One is located in a relatively deprived mid-sized town in West Yorkshire; the other in the relatively affluent London borough of Hammersmith and Fulham. Unlike most US charter schools, which typically serve very disadvantaged urban populations, the schools in our sample cater to more diverse and less needy communities. And while both adopt traditional educational models, their high-level philosophies diverge. The West Yorkshire school embraces the 'no excuses' (NE) paradigm common among US charter schools – emphasising high expectations, structured routines, and strict discipline (Angrist et al., 2010, 2013; Dobbie & Fryer, 2013). By contrast, the London school follows a 'classical liberal' (CL), knowledge-rich approach rooted in both the humanities and sciences. The CL model echoes the ethos of schools in the National Heritage Academy network (Dynarski et al., 2018), which emphasise character education and draw on Plato's cardinal virtues, including self-control, perseverance, and respect. It also parallels the Core Knowledge K–8 philosophy studied by Grissmer et al. (2023), which assumes that building a broad base of general knowledge in early education is key to enabling deeper comprehension at later academic stages.

The two free schools also differ in their admission tie-breaking procedures, which underpin our identification strategy. The NE school employs fair banding combined with lotteries: applicants sit a reasoning test, are sorted into ability bands, and places are allocated via random draws within each band to ensure a balanced intake. Conversely, the CL school uses lotteries for children grouped at different distances from the school premises but also reserves some seats for pupils who live nearby, thus creating the possibility to use a distance-based regression discontinuity (RD) design to study the impact of the school on achievement. We exploit the workings of the assignment mechanisms for causal inference, adapting to our context the insights from the recent 'Research Design Meets Market Design' (RDMD) methods introduced by Abdulkadiroğlu et al. (2017, 2021).

To conduct our analysis, we use pupil-level data covering two cohorts of applicants to each school, including admission criteria, applications, and enrolment records. These are linked to administrative datasets containing background characteristics and outcomes. To characterise educational approaches, we complement this with original data: we web-scrape 'vision and ethos' statements for the free schools and for the counterfactual schools attended by rejected applicants.[3] This text-based analysis is guided by two full-day visits at the free schools and by semi-structured interviews carried out with the schools' headteachers and founders.

Our results can be summarised as follows. To begin with, our research design yields treatment and control groups that are balanced across key pre-determined characteristics, including gender, ethnicity, free school meal eligibility, and prior test scores, supporting the internal validity of our setup.

Next, using offers as an instrumental variable (IV) for attendance, we find large, statistically significant effects on exam scores at the end of secondary education (comprising of five years of schooling). One additional year in either free school improves exam scores at age 16 by around 0.08 standard deviations – cumulatively, 0.4 SD over five years. Effects are similar across both cohorts and schools, despite their different pedagogical models. Free school attendance also improves non-test score outcomes (see Jackson, 2018), decreasing pupils' unauthorised absences. Free school pupils are less likely to change school after enrolment than pupils in counterfactual schools – a potential indicator of higher school satisfaction. In contrast, we find little evidence of increased disciplinary sanctions such as temporary exclusions.

We also explore heterogeneity by ethnicity, gender, free school meal (FSM) eligibility, and end-of-primary school scores. We find that the NE school appears more effective for boys, while the CL school benefits White and non-FSM pupils. However, only one of eight subgroup comparisons is statistically significant, so we interpret these patterns cautiously and avoid drawing strong conclusions.

Finally, we explore what distinguishes these free schools from the counterfactual schools attended by rejected applicants. First, we compare 'standard' characteristics – i.e., peer composition, teacher profiles, and school inspectorate ratings. These differ across the two schools, but not consistently. For example, the NE school has smaller class sizes and better inspectorate ratings, but the CL school does not. Teacher qualifications and pay also vary in opposite directions. This inconsistency suggests such factors cannot fully account for the observed impacts.

We then examine pedagogical differences by setting up a quantitative text analysis of schools' web-scraped 'vision and ethos' (VE) statements. We first create a taxonomy that characterises the approaches used by the two schools on the basis of: *i)* key words identified in the VE statements – including schools' 'core values' and 'drivers'; and *ii)* salient elements that emerged during our visits – including from our semi-structured interviews, discussions with students, and experience of class time. We then create a 'dictionary' that allows us to associate such key words to a wide range of synonyms and akin expressions in the VE statements of counterfactual schools. Lastly, we create some synthetic measures that capture whether a key concept is more present at one of the two free schools than what would be expected by looking at the overall set of VE statements of the counterfactuals.

We find that the two free schools share similarities with their counterfactuals along many traits, such as a focus on good manners, autonomy, and purpose in learning, high standards, and high aspirations. However, what clearly stands out relative to the counterfactual schools are the two high-level philosophies that permeate every aspect of the schools' approaches to teaching and learning. The NE school focusses on a 'no excuses, no shortcuts' approach to education, based on strict routines, a culture of success, and hard work. The CL instead focusses on a 'classical liberal' education, centred around a knowledge-rich curriculum, teacher-led learning, and an openly competitive

[1] The impacts of choice, autonomy, and competition have been studied in England (e.g. Gibbons et al., 2008; Eyles & Machin, 2019; Bertoni et al., 2020) and extensively explored in the US and Sweden (e.g. Hoxby, 2000; Cullen et al., 2006; Böhlmark & Lindahl, 2015).

[2] Following these initial studies, researchers then went on to analyse larger school samples, often using observational models (e.g. Angrist et al. 2013; Dobbie and Fryer 2013).

[3] We identify 52 counterfactuals for the NE free school and 96 for the CL one. For comparison, we collect similar data for all 102 secular, secondary free schools opened by the end of our analysis period. Each school has a website with a published 'vision and ethos' statement.

environment. We conclude that the two schools differ in their approaches in this sense with respect to their counterfactuals, but also with respect to each other. Given our evidence on their effectiveness, this is an important result: it shows that, in the context of start-up schools, a CL approach can 'deliver' as much as a NE pedagogy.

Our findings relate to the study of school choice in education. The evaluation of choice-oriented reforms rests on forming an in-depth understanding of their equilibrium effect on all students, both those who opt into new schools and those who remain in traditional settings. In the US context, Chen and Harris (2023) and Ridley and Terrier (2025) report evidence on charter-school effects on achievement for both charter students and those enrolled in traditional public schools (see also Gilraine et al., 2021, and Crema & Singleton, 2025). The role of pedagogical innovation has similarly gained substantial relevance in the debate around school choice and autonomy. The US-based literature has highlighted the importance of the NE approach as a key driver of charter-school effectiveness (Cheng et al., 2017; Cohodes & Parham, 2021).

We offer two key contributions. First, we report credible evidence on how school choice affects those who exercise it, by estimating the direct effect of attending two newly established free schools in England among eligible applicants. Although there is research on US charter schools (reviewed in Cohodes & Parham, 2021), Swedish *friskolor* (e.g., Böhlmark & Lindahl, 2015), and both English grant-maintained schools (Clark, 2009) and academies (see Bertoni et al., 2020; Eyles & Machin, 2019; Eyles et al., 2017), there is no causal evidence on the effectiveness of start-up English free schools.[4] Second, we show that free schools can be equally effective despite different pedagogical models. This supports the idea that autonomous schools can succeed without adhering to a single pedagogy (e.g., a 'no excuses' approach). Additionally, by shedding light on the effectiveness of more traditional and teacher-led educational approaches on different types of pupils, our findings complement the evidence on American charters, which is mostly focused on disadvantaged students (see Angrist at al., 2013, Angrist et al., 2010, and Dobbie & Fryer, 2013).

From a methodological perspective, our study is one of the few that exploit admission lotteries to study school effects on pupil achievement outside the US, in a developed country context. Furthermore, the use of quantitative text analysis to characterise schools' education philosophies is new. In this respect, we share some similarities with Gilraine et al. (2021), who use data from the charter school opening applications submitted to the North Carolina Charter Schools Advisory Board and classify charter schools as either horizontally differentiated or not, compared to public schools in their educational programme, and with Biasi and Ma (2022), who focus on the innovative content of taught courses at university.

While limited to two schools, our findings provide an important piece of evidence for ongoing policy discussions on the organisation of public education systems: they serve as a proof-of-concept that the state needs not directly run schools, as school services can be successfully provided by independent entities. In such a system, the role of the state is mostly to determine the 'rules of the game' in terms of common curricula, assessment, and non-discriminatory admission criteria; to fund schools on the basis of their ability to cater for demand; and to monitor school operations and pupil outcomes. This in turn suggests that our work provides a 'qualified yes' in response to the question of whether bottom-up initiatives that expand horizontal differentiation in school supply can improve pupil performance.

[4] Machin et al. (2020) examine University Technical Colleges – quasi-vocational, often undersubscribed schools that are part of a later phase of the free-school policy. Other studies on free schools focus on social segregation or provide descriptive analyses (e.g. Green et al., 2015; Andrews & Johnes, 2017; Allen & Higham, 2018).

2. Institutional setting

2.1. Education stages and main features of the English education system

Compulsory education in England spans primary (ages 5–11) and secondary (ages 11–16) schooling, organised around Key Stages (KS). Pupils begin at the Foundation Stage (age 4–5), then progress through KS1 (ages 5–7), KS2 (7–11), KS3 (11–14), and KS4 (14–16). Compulsory schooling ends at age 16, typically after KS4.

Pupils are assessed at multiple points. KS1 exams in English and Maths are externally set but marked by teachers. KS2 tests in English, Maths, and Science are externally set and graded. At KS4, pupils sit externally assessed academic (GCSE) and/or vocational (NVQ) exams, with English, Maths, and Science being compulsory. These tests are externally set and assessed.[5] Additional information comes from Ofsted, the school inspectorate, which conducts inspections every 3–5 years and rates schools from 'outstanding' to 'inadequate' across domains like teaching, management, and behaviour. Although Ofsted is a government agency, it operates independently in publishing its reports.

2.2. Quasi-markets for education and autonomous schools in England

Quasi-markets in English education have existed since the 1988 Education Reform Act, which ended automatic state school assignment by residence and granted parents the right to choose schools. It also tied funding more closely to enrolment, incentivising schools to compete for pupils. Under this system, admissions to state schools are based on parental preferences, but popular schools often become oversubscribed. In such cases, priority is typically given to pupils who live nearby, have special educational needs, are in Local Authority (LA) care, or have siblings at the school. Some schools, such as faith or grammar schools, can apply additional selection criteria, including religion or entrance exams. To allocate pupils to their preferred schools, LAs run constrained versions of the Gale-Shapley student-optimal stable mechanism, also known as a Deferred Acceptance (DA) algorithm.

State-funded school types originally included community, voluntary-controlled, foundation, and voluntary-aided schools. Community and voluntary-controlled schools are managed by the LA, which employs the staff, owns the buildings, and handles admissions. Voluntary-aided and foundation schools enjoy more autonomy from the LA, which nevertheless retains powers of oversight. All these schools are funded by the LA using central government grants drawn from general taxation. On top of this, there is also a small fee-charging private school sector, enrolling roughly 5 percent of pupils.

The transition towards an education quasi-market accelerated in the 2000s with the introduction of academies. Initially promoted by the Labour government, these 'sponsored' academies aimed to replace underperforming LA-controlled schools. Sponsors – typically charities or business groups – were given control and allowed to reform staffing, pay, and pedagogical ethos under increased autonomy. The 2010 Academies Act, introduced by the Liberal-Conservative coalition, expanded the model to encourage innovation and competition. It created 'converter' academies: high-performing schools (rated 'good' or 'outstanding' by Ofsted) that could become academies without a sponsor. Unlike other state-funded schools, all academies are financed directly by central government, with no LA intermediary.

The 2010 Academies Act also created the legal basis for 'free schools' – the focus of this study. Unlike academies, which are conversions of existing schools, free schools are entirely new state-funded institutions established by non-profit groups including parents, charities, teachers, and others. Their goal is to expand the supply side of education through

[5] Standardized KS3 national tests (age 14/grade 9) were abolished in 2008, before our cohorts entered secondary school. The current KS3 test is locally set and internally marked by teachers. We exclude it from our analysis.

bottom-up initiatives, increasing choice and competition. Free schools initially open with only the first grade of their educational phase (primary or secondary) and then expand year by year. Once fully established, they operate identically to academies: they enjoy autonomy over teacher pay and conditions, school terms, curriculum, pedagogy, ethos, and admissions – following the non-discriminatory rules set by the School Admissions Code. To open a free school, applicants must demonstrate unmet demand – either for additional school places or for higher-quality provision. As a result, the establishment of free schools is not only related to a need for more school places, pushed by demographic changes. Governance rests with a not-for-profit trust that employs staff and oversees performance. Trusts may run single or multiple schools – either free schools or academies – and in the latter case are known as multi-academy trusts (MATs). Like academies, free schools are funded directly by central government.

Initially a small-scale initiative, the free school programme expanded rapidly: from just 24 schools in 2011 to around 250 by 2015 – a tenfold increase. Growth then slowed, reaching around 600 schools by 2021, with a particularly high concentration in London. Figure AI.1 in the Appendix shows that both the number and enrolment share of secondary free schools rose steeply in the early years.

The profile of founders also shifted. Early free schools were set up by parents and teachers, who were perceived by many as 'disrupting innovators' and 'pioneering educators'. Over time, however, an increasing share have been established and managed by MATs.

Reflecting the policy's aims, the sector remains diverse and includes later variants like University Technical Colleges (UTCs), which are hybrid schools providing a blend of general and vocational education (often specialising in engineering subjects), and Studio Schools, which offer vocational or project-based models. These types are excluded from our study.

Instead, this paper analyses the impact on achievement of two secondary free schools that were examples of the early days of the programme: the schools opened in its first and second year. Below, we describe these schools' key features.

2.3. The 'no excuses' school

The first school we analyse (Dixons Trinity Academy) is located in Bradford, a relatively deprived town in West Yorkshire with a large South Asian British population. Opened in 2012 as part of the second cohort of free schools (fewer than 20 secondary schools nationwide), it remains one of only four in the area. The school admits around 112 pupils per cohort and has consistently been oversubscribed. The school mostly advertised itself through word-of-mouth and on social media, such as Facebook. Prior to opening, the founders also used a consultation period to raise the school profile in the local community. It is supported by the Dixons Trust, a MAT active in Bradford and Leeds focused on improving outcomes for disadvantaged youth. Despite its name, the school is non-denominational (i.e., not religious).

The school adopts an unashamedly 'no excuses' (NE) approach. Quoting from their webpage, its philosophy reads: 'We promise to do whatever it takes to ensure that every student (…) achieves their full potential. We have extremely high expectations, and *just as there are no shortcuts, there are no excuses.* We (…) have a sentence that states the lasting impression we want to leave on the world. Our academy sentence is: "The academy ensured that all students succeeded at university, or a real alternative, thrived in a top job and had a great life."' In 2014, it became the first free school rated 'outstanding' by Ofsted, which praised its disciplined, high-aspiration, results-driven ethos.

We visited the school in late June 2022 and met with both the headteacher and the CEO of the Dixons Trust (in this case, the founder). Our meeting centred around a semi-structured interview and the questionnaire we prepared is presented in Exhibit AI-1, Appendix I. During our discussions, we avoided as much as possible to reveal the details of our results on effectiveness, including the patterns of heterogeneity. We also avoided prompting to minimise answers that would *ex-post* rationalise our findings. The interview lasted for around 1.5h and was followed by an 'experiential' tour of the school, including lunch in the canteen with the pupils and some discussions with the Year-7 and Year-8 head boy / girl.

The school presents itself as a new, purpose-built construction in a relatively deprived and diverse part of Bradford. Banners hung outside the school gate highlight the school philosophy encapsulated in the 'no excuses' and 'in this academy only excellence will do' sentences – see Exhibit AI-2 in Appendix I.

Our interview highlighted similar features. The school approach to study is guided by three core values – hard-work, trust, and fairness – and pupils are told to follow three drivers: mastery, autonomy, and purpose. A key metaphor used by the school to describe learning and organise activities is that of 'climbing the mountain', highlighting how achievements require hard-work, devotion, and a long-term perspective. In line with this, the three schoolhouses are named after three different mountains (Makalu, Aconcagua, and Pelvoux) with the following three mottos: 'no excuses', 'no shortcuts', and '100 % every day'. Lists on the school corridors rank the 'best climbers' of the week, based on behaviour and progression. The three schoolhouses compete with one another and receive rewards and acknowledgments for their success. However, these rewards are deliberately small and low key as the aim of the school is to inspire passion for learning and long-term goals, rather than short-term recompenses. The headteacher and CEO clarified that these pedagogical choices were influenced by research trips taken to visit US charter schools. In particular, the US Knowledge Is Power Program (KIPP) network was mentioned as one of the inspiring models.

Both the headteacher and the CEO emphasised the importance of strict routines that guide school life. Understanding and adhering to such routines is key for both pupils and teachers, and formal induction processes are in place. Prospective pupils are invited to a meeting at the school before they apply and then again upon offer to clarify the school practices and principles. Furthermore, the first week of enrolment is not devoted to formal learning, but to forming an understanding of the culture and the ethos of the school. Throughout their education, pupils who 'forget' values and practices go through a process of 're-induction' via individual or group sessions, and the school has a dedicated re-induction room. While strict routines and behavioural practices guide every aspect of school life, including walking down the corridors, queueing for food in the canteen, and tidying up afterwards, both interviewees emphasised that the militaristic, 'masculine' and harsh approaches taken by some US charters find no place at the school. Instead, they described their approach as 'maternal'.

Importantly, both leaders emphasised the importance of commitment to routines and core values also for staff members. Like pupils, new recruits go through a process of induction – and a process of re-induction when their practices deviate from expectations. To make sure that the pedagogical philosophy is not diluted, the school uses no supply or temporary teachers. Instead, gaps left by staff absences, periods of sick leave, and even maternity/paternity breaks are covered by colleagues, who put in the extra time to preserve the integrity of the school's educational approach. When discussing the teaching approach deployed by the school, we understood this is not focussed on drill or rote learning, nor follows a set of strict guidelines provided by the central management of the trust or senior staff at the school. Teachers are empowered to carry out some 'experimentation' if they wish to do so, although this should not disrupt or undermine the core routines and key values of the school.

Finally, to appeal to a community that might have professional-oriented aspirations for their children, the curriculum is not narrowed down to a limited number of academic topics. However, there is a clear ambition to 'do the key things well' (meaning the core subjects) – and to send as many children as possible to university.

2.4. The 'classical liberal' school

The second school in our study (West London Free School) is located in Hammersmith and Fulham, an inner London borough with both affluent areas and deprived neighbourhoods. The area has a largely White population that includes many European migrants, particularly French, Polish, and Portuguese. Opened in 2011 as part of the first cohort of seven secondary free schools, it remained the only one in the borough throughout our analysis. The school has a capacity of 120 pupils per cohort and has been oversubscribed from the start. To raise visibility, the founders held open days, visited the local primary schools, and 'banged the drum' in the media (one of the founders is a well-known journalist).

Founded by a group of high-profile advocates for traditional, knowledge-rich education, the school follows a 'classical liberal' (CL) philosophy, that is 'a (…) rich education that draws its material and methods from the best and most important work in both the humanities and the sciences. The aim (…) is to prepare children to participate in (…) debates about contemporary issues, as well as the universal questions that have been troubling mankind (…). *We want pupils to leave our school with the confidence that comes from possessing a treasure trove of essential knowledge*, and a deep and lasting respect for reason, evidence, civility, honesty, kindness and the value of hard work and self-discipline.'

The school was inspected by Ofsted in 2013 and 2017 and, in both instances, it was rated as good with outstanding features in the domains of pupils' behaviour and safety. The reports confirm the 'classical liberal philosophy' advertised by the school.

We visited the school in early July 2022 and met with the lead headteacher. Our semi-structured interview was centred around the same questionnaire we used for the NE school and lasted just above 1h. As for the NE school, we avoided prompting and disclosing too many details of our evidence to avoid influencing our interviewee's answers. We also had an opportunity to briefly talk to the joint headteacher and were taken for a tour of the school. As students were away on fieldtrips, we did not have an opportunity to talk to the pupils. Therefore, we visited the school again in September 2022 and observed lessons.

The school is housed in a mid-19th century residence (see Exhibit AI-2 in Appendix I) with a modern annex at the back. The historical nature of the building provides an appropriate background to the classical liberal ethos of the school. Although according to the headteacher the classical liberal label is 'somewhat dated', everything in the school's visual impact betrays this classics-inspired philosophy.

The headteacher described the learning approach of the school as centred around a 'knowledge-rich curriculum' and a 'specify and excel' (i.e., narrow the curriculum, but deepen the level of understanding) philosophy. Its three key values – kindness, hard-work, and high standards – overlap to some extent with those of the NE school. However, in the case of the CL school, the notions of high standards and aspirations are more clearly oriented towards academic excellence and the study of classic and core subjects. Indeed, Latin is compulsory for all students up to the third year of secondary education; the study of arts is focussed on the classics throughout history; music is a well-renowned specialism of the school; and areas that are considered not challenging enough (such as 'drama' and 'home economics') are not offered at the school. The school also places strong emphasis on debating and public speaking, and on providing students with the skills that equip them to join discussions on contemporary issues as well as universal questions about society. According to the headteacher, one reason the school has not been rated 'outstanding' by Ofsted is its limited subject offer, which the inspectorate sees as a constraint on inclusivity. Nonetheless, there were no indications that the school intended to alter its curriculum simply to improve its rating.

The classical liberal foundations are further emphasised in the names of the four schoolhouses – Athenians, Corinthians, Olympians, and Spartans – and all teaching rooms are named after classic and influential scientists or thinkers (e.g., Euclid, Pythagoras, and Turing). Like at the NE school, the schoolhouses compete with one another in terms of achievements and behaviour. But unlike the NE school, the CL school openly values competition and aims to foster a competitive atmosphere.

As for the NE school, the CL one places great emphasis on discipline, routines, and outstanding behaviour. Indeed, the headteacher clarified that 'liberal does not mean lax' – and portrayed the school as 'non-permissive'. Although we were not made aware of specific processes for pupils' inductions, it is clear that students are expected to adhere to a set of well-codified behavioural rules – including how to walk in the corridors, how to address teachers, and how and when to talk to one another.

The latter point is particularly relevant during classroom time, which is always centred around teacher-led lessons. Although the school promotes debating, lessons are primarily characterised by direct, explicit instruction by the teacher. Most discussions are thus between the teacher and individual pupils. Indeed, there is little room for pupil-to-pupil 'chatting', apart from during short periods (about one minute) following a prompt by the teacher. Classes are physically designed in a way that fosters teacher-to-pupil interactions: all desks and chairs are oriented in the same direction, making sure pupils face the teachers and whiteboard.

In terms of its staff, the headteacher reported that the school has had a 'healthy amount of turnover', meaning that staff members who did not endorse and commit to the school ethos left (rather than stay and dilute the core pedagogical values). The headteacher also emphasised the advantage of originally setting up as a brand-new free school in terms of selecting only teachers and personnel that shared the classical-liberal orientations of the original founders. In short, it was clear from our visits that teachers are immersed in a well-defined culture and provided with quite clear guidance about what it means to work at the CL school.

3. Admission rules and identification

3.1. General considerations

Free school applicants are a self-selected group who is likely to differ from the general pupil population in unobservable ways. As a result, comparing them to non-applicants reveals selection patterns rather than causal effects. Even within the set of applicants, admission criteria such as priority by ability or proximity can create a correlation between pupil characteristics and admission chances that bias simple comparisons between those admitted or not. Given that both schools under investigation have been highly oversubscribed since opening, we leverage data on applications and detailed information about the way in which the schools prioritise applicants to design an identification strategy. To solve issues related to the multi-staged nature of the assignment mechanisms and identify pupils who face an identical 'risk' of being admitted to the schools, we borrow some of the key insights of the RDMD literature developed by Abdulkadiroğlu et al. (2017) and Abdulkadiroğlu et al. (2021).

Before detailing the admission rules for each free school, it is worth noting that both schools managed their admissions independently in their first year of operation, while they joined the LA assignment-mechanism from the second year onwards.[6] As a result, we face an additional empirical challenge when dealing with the second cohort (discussed below).

3.2. Admission rules and identification - NE school

When oversubscribed, the school prioritises pupils in local authority care (LAC), those with special educational needs (SEN), and siblings of

[6] Free schools control admissions in their first year but must join the LA process from year two. However, their application and offer timelines have always overlapped with LA procedures.

current students. All other applicants fall into a single group subject to 'fair banding', a system designed to ensure that the admitted cohort reflects the ability distribution of applicants. All applicants take a non-verbal reasoning test and are placed into one of nine ability bands ('stanines'), using pre-set thresholds. If there are more applicants than available places in a given band, offers are made by lottery within that band. The number of seats allocated per band reflects the proportion of applicants in each band.[7] Lottery numbers are assigned at application and drawn unconditionally at random. Within each band, pupils with the highest lottery numbers receive offers first.

In practice, unless they are classified as LAC or SEN, or have siblings at the school, applicants are at the same risk of being admitted within their band, and the lottery ensures that assignment within band is random.[8] As a result, the inclusion of stanine-fixed effects (cohort-specific) in our models implies that we only compare pupils who share the same (random) likelihood of admission when studying the effects of attending the school.

While the first cohort was assigned directly by the school – allowing for clean randomisation within bands – this is not true for the second cohort. Although we still observe all applications, and the school applied the same criteria, it became part of the local authority's centralised DA mechanism. Using this algorithm, applicants are considered for a seat at a school only if they are not offered a seat from a school for which they have a stronger preference. As a consequence, only applicants who did not receive an offer from a school which they prefer to the NE school may receive an offer from the NE school. This means that the set of pupils who receive an offer is endogenously 'truncated' depending on their priority at other schools that they rank above the NE school.

This implies that a necessary condition for the offer-based estimator to recover the causal effect of interest for the second cohort is that receiving an offer from a more preferred school is independent of an applicant's potential outcomes in the counterfactual schools within each risk set. What may render this assumption implausible in our context is that the intensity of parents' preferences for the free school is likely correlated with their expectations about the benefits from attending it.

Importantly, lottery numbers in the NE school's admission process are assigned independently of applicants' preferences, which the school does not observe, and before the LA runs the DA algorithm. These numbers determine priority within stanines but are not influenced by families' rankings of other schools. We therefore use lottery numbers – rather than offers – as an instrument for attendance in the second cohort, conditioning on stanine fixed effects. As it turns out, estimates using this instrument are similar in magnitude to those based on offers, suggesting that any bias introduced by preference-based truncation is limited.

3.3. Admission rules and identification - CL school

In case of oversubscription, the CL school gives priority to pupils in local authority care (LAC), those with special educational needs (SEN), 10 % of places to pupils with musical aptitude, and siblings of current students. Half of the remaining places are allocated to pupils living closest to the school. Of the rest, two-thirds go to pupils within a three-mile radius via lottery, and one-third to those living between three and five miles, also by lottery. If places remain, they are assigned by lottery to applicants living beyond five miles.

In our set-up, we focus on applicants who fall within either the distance or the 'lotteries-within-distance' categories. First, for pupils living within three miles from the school, the risk of being admitted to the school is determined by a distance-based regression discontinuity (RD) design as well as by a lottery among those who are not admitted through the distance mechanism. Thus, pupils living within three miles of the school are in principle at risk of being admitted via both the distance and the lottery categories, and we need to account for the fact that the risk of entering the lottery is conditional on pupils' outcome in the distance category. Second, for pupils in the 3–5-mile distance category, the risk of being admitted to the school is only determined through a lottery. The last admission criterion is instead inactive: due to heavy oversubscription, pupils living beyond five miles from the school were never offered a seat in the period of analysis and are dropped from the sample.[9]

To design an identification strategy for this complex setup, we borrow some of the key insights from the RDMD literature. We start with the distance-based RD: this generates, *ex-post*, a distance cut-off τ that is *ex-ante* unknown to parents. However, pupils who live very close to the school are always admitted and are not 'at risk'. They are therefore dropped from the estimation sample. Similarly, pupils who live too far away from the school are not going to be admitted due to residential proximity – so they are also not at risk of being admitted via the distance criterion. Only applicants who live 'close enough' to the τ distance cut-off are at risk of being admitted through the RD. Following Abdulkadiroğlu et al. (2021), we use an optimal bandwidth estimation strategy to identify the bandwidth δ that determines which pupils are at risk (see Calonico et al., 2014). According to the results in Abdulkadiroğlu et al. (2021), pupils within $[\tau-\delta, \tau+\delta]$ share the same 50/50 risk of being admitted in the RD draw.[10] We group these pupils together by controlling for a $[\tau-\delta, \tau+\delta]$ dummy (cohort-specific) and a piecewise linear spline in distance from cut-off τ. At this stage, it is worth noticing that, *ex-ante*, pupils within $[\tau-\delta, \tau+\delta]$ are also at risk of being admitted via the lottery among applicants who live within three miles of the school. Indeed, the $[\tau-\delta, \tau+\delta]$ dummy controls for the combination of RD and lottery risk of being admitted to the school.[11] Next, we consider applicants who live within three miles of the school, but not close enough to be meaningfully at risk of being admitted via the distance-based RD assignment. That is, pupils within $(\tau+\delta$, 3 miles] of the school. These pupils are exposed to the same lottery risk of admission, and we group them together by controlling for a $(\tau+\delta$, 3 miles] indicator (cohort-specific). Finally, a similar situation applies to pupils within (3 miles, 5 miles] of the school. These pupils are exposed to the same lottery risk of being admitted, and we group them together by controlling for a (3, 5 miles] (cohort-specific) indicator.[12]

The above approach works for the first cohort of applicants, but not for the second cohort, due to the truncation problem that arises from joining the LA admission mechanism (explained above). We again use

[7] Pooling across cohorts, each stanine band received between 4% and 17% of applicants, with all bands oversubscribed. Offers were made proportionally, with no notable differences across cohorts.

[8] No pupils were admitted under the LAC or SEN criteria, and none via the sibling rule in the first cohort. Only a few received sibling priority in the second. As our analysis relies on banding and random allocation, these pupils are excluded.

[9] In our data, no pupils are assigned under the LAC or SEN criteria, and only a few in the second cohort qualify via the sibling rule; these are excluded from the analysis. Fewer than the reserved 10% are admitted for musical aptitude in either cohort and, as they are not subject to the random allocation mechanisms we exploit, they are also excluded. Pupils who live beyond three miles of the school are never admitted through the distance criterion.

[10] The thought experiment here is to fix the admission criteria, school capacity, and pupils' priority groups, and resample the distance tie-breaker within this group.

[11] This approach identifies 30 pupils in the $[\tau-\delta, \tau+\delta]$ group in the first cohort and 42 in the second. About a third receive an offer on average, but this differs by cohort: 55% in the first and 20% in the second, reflecting truncation due to higher-ranked schools in the LA-wide DA assignment, which affects only the second cohort.

[12] Approximately, 65% and 80% of applicants are in the $(\tau+\delta$, 3 miles] area for cohorts 1 and 2, respectively and 25% and 15% in the (3 miles, 5 miles] experiment. The corresponding figures for the likelihood of an offer are 16% and 4%; and 25% and 12%. Once again, numbers for the second cohort are affected by offer truncation.

applicants' lottery numbers as an IV for school offers within the 'lottery' samples. However, as we cannot reconstruct the full outcomes of the LA algorithm, we have to drop pupils in the RD sample in the second cohort. Luckily, there are only 42 pupils for whom this applies – and the similarity of results including/excluding pupils at risk from the RD 'experiment' clearly indicates that they do not drive our findings.

4. Data and descriptive facts

We obtained pupil-level data for the first two cohorts of applicants to each free school, including admission criteria, application records, offer status (and the associated criterion), and enrolment outcomes – covering approximately 2500 students. These records are linked to administrative datasets from the Pupil Level Annual School Census (PLASC) and the National Pupil Database (NPD), which track all pupils in England's state-funded schools. From these sources, we obtain information on applicants' demographic and background characteristics at the time of secondary school application: gender, eligibility for free school meals (FSM), special educational needs (SEN) status, ethnicity (White British), and whether English is their first language (EFL). We also observe school outcomes, including KS1, KS2, and KS4 test scores; the share of days missed due to absences (justified or unjustified); and whether pupils were ever suspended (excluded) during secondary school.[13]

The school-enrolment data allow us to reconstruct the number of years each pupil spent in either of the two free schools. This is the key independent ('treatment') variable in our analysis. We also gathered information on the schools that unsuccessful applicants attend in the first year of secondary education. This allow us to reconstruct the type of institution they attend (single-sex schools, academies, faith schools, or grammar schools); intake composition (the share of female, FSM-eligible, White British, EFL, and SEN pupils); Ofsted inspection outcomes, overall and by domain (management, behaviour, teaching, and achievement),[14] as well as the pupil-teacher ratios, and teachers' qualifications, gender, and gross salary, and teacher/headteacher mobility from the School Workforce Census. These data allow us to shed light on some 'standard' characteristics of counterfactual schools attended by applicants who did not enrol in one of the free schools.

Columns (1)–(3) of Table 1 present key descriptive statistics regarding the application data (Panel A) and the composition (Panel B) of both schools. Panel A shows that, on average, 1-in-3 applicants received an offer from the NE school, and 1-in-4 applicants enrolled in the school. However, these figures are different across the two cohorts: while 1-in-5 applicants received an offer and 1-in-5 applicants attended the school in cohort 1, the corresponding figures for cohort 2 are 1-in-2 and 1-in-5 respectively. For the CL school, on average 1-in-10 pupils received an offer and 1-in-11 pupils enrolled in the school. However, there is again cross-cohort heterogeneity: in cohort 1, 1-in-5 pupils received an offer and 1-in-10 enrolled in the school, while the corresponding figures for cohort 2 are 1-in-8 and 1-in-12.

While some of the variation in the share of offers made reflects changes in the number of applicants, cross-cohort differences in offers also reflect the truncation issue discussed above. Since data on preferences for years before 2014 were not collected by the DfE, we cannot characterise the extent of the truncation problem for the cohorts under investigation. However, using data for 2014 we find the following patterns. About 16 % and 22 % of preferences expressed for the NE and CL schools are first preferences. This evidence suggests that the pool of applicants displays a mixed degree of 'intensity of preference' for the two schools – and that our results are unlikely to only reflect the effect of 'attending your preferred school' (as in Cullen et al., 2006). In terms of preference truncation, among the students who ranked the two schools as their second preference (36 % for the NE school and 23 % for CL school), 42 % and 52 % for the NE school and CL school respectively receive an offer from a school they prefer more. Expectedly, the truncation problem is more important for 3rd and 4th choices (e.g., 82.5 % of third preferences for the NE school are truncated from above, and 74 % for CL school). As discussed, we address this issue by using lottery numbers as instruments for offers.

In terms of composition, Panel B of Table 1 shows that the NE school has a lower share of White British and EFL applicants than the CL one, and lower KS1 and KS2 scores in English, Mathematics, and Science.[15] However, the shares of female, FSM-eligible, and SEN applicants are roughly comparable across the two schools. For the CL school, we also know the home-to-school distance among applicants, one of the admission criteria, which is on average equal to roughly 1.9 km.

Next, Columns (4)–(6) of Table 1 report the national and the 'local comparator' averages of pre-application pupil characteristics. The local comparator areas cover the LA or set of LAs from which each school attracts more than 90 % of their applicants. While this area for the NE school overlaps with the LA in which it is located, where nearly all applicants live, the local comparator for the CL school covers five London LAs. For both schools, the share of female applicants is lower than the national and local-comparator averages. The shares of FSM-eligible applicants are in line with the national average, but lower than the local-comparator average. Both schools have lower shares of White British and EFL applicants compared with the national average. However, the CL school has higher shares of White British and EFL applicants compared with the local-comparator average, while the opposite is true for the NE school. Finally, among applicants to both schools, average KS1 and KS2 scores are higher than both the national and local-comparator averages.[16]

Our data indicate some positive self-selection into application based on prior ability, while socio-economic and ethnic selection appears less pronounced. Although this does not threaten the internal validity of our estimates, it has implications for external validity. This is an 'inescapable' feature of all studies that evaluate school effects in the context of school choice in education quasi-markets and using features of the application process, such as lotteries in the case of US charters. Parents and pupils are likely to self-select into the pool of applicants for different schools based on expected gains from those schools. Applicants are therefore almost by definition different from non-applicants. Nonetheless, these very same issues mean that the pool of applicants is a key population to focus on in the context of market-based approaches to education reforms – including the free-school programme we analyse.

5. Main results

5.1. Balancing tests

5.1.1. Methods

If our identification strategy works, successful applicants should be comparable to unsuccessful applicants along all observable and unobservable dimensions – conditional on cohort-specific 'experiment' fixed effects and the distance-based running variable that determines assignment for pupils exposed to the RD admission risk (see Sections 3.2 and

[13] Applicant data were matched to pupil-census records by the DfE using identifying details (full name, address, date of birth) before being anonymised. Over 95% of pupils were successfully linked. Some additional cases are lost due to missing achievement or background data, yielding a final sample of around 2,200 observations.

[14] We used the data from the most recent inspection before the last cohort in the data finished secondary school.

[15] Pupils can attain KS1 scores between 3 and 21 (representing terms of progress), and KS2 levels between 3, 4, and 5 (with level 4 being the expected attainment at this stage).

[16] Andrews and Johnes (2017) find that free schools tend to open in disadvantaged areas, but that less well-off students from these areas are less likely to attend a free school.

Table 1
Descriptive statistics - pupil-level data.

Variable	(1) Both schools	(2) NE school	(3) CL school	(4) National sample	(5) NE School - Local Comparator Area	(6) CL School - Local Comparator Area
Panel A. Application patterns						
Offer FS	0.203	0.348	0.098	–	–	–
Start FS	0.152	0.237	0.091	–	–	–
Panel B. School composition						
KS4 score (standardized)	0	−0.192	0.142	–	–	–
Female	0.437	0.427	0.445	0.489	0.487	0.487
FSME	0.185	0.180	0.189	0.183	0.226	0.253
White British	0.399	0.133	0.601	0.797	0.499	0.435
English FL	0.501	0.278	0.671	0.854	0.565	0.522
SEN	0.184	0.188	0.181	0.211	0.250	0.222
KS2 - English Level	4.431	4.281	4.552	4.212	4.147	4.262
KS2 - Maths Level	4.436	4.316	4.530	4.231	4.145	4.296
KS2 - Science Level	4.437	4.316	4.534	4.246	4.150	4.276
KS1 - English Level	15.521	14.969	15.927	14.933	14.158	14.718
KS1 - Maths Level	16.223	15.562	16.763	15.763	14.913	15.667
KS1 - Science Level	15.706	15.027	16.262	15.566	14.764	15.372
Distance	–	–	1.921	–	–	–
Observations	2179	914	1265	1543,129	12,583	17,177

Notes: Local Comparator Area refers to the Local Authority (LA) or set of LAs from which each of the two schools attract more than 90 % of their applicants.

3.3).

To provide evidence in support of our identification strategy, we test the balancing of treated (successful) and control (unsuccessful) applicants' observable and pre-determined characteristics by estimating the following model with Ordinary Least Squares (OLS):

$$X_i = \gamma D_i^j + \alpha_{jck} + \phi(d_i) + \varepsilon_i, \ \ if\ 0 < P_i^{jck} < 1 \tag{1}$$

In Eq. (1), P_i^{jck} identifies pupil i's probability of being admitted to free school j (NE, $j = 1$ or CL, $j = 2$) in cohort c (Cohort 1 or 2) via experiment k, representing a stanine in the NE school and the RD group plus the two lottery bands in the CL school. The restriction $0 < P_i^{jck} < 1$ identifies pupils with a non-degenerate admission risk, thereby excluding from our estimation sample pupils with 'reserved' seats at each school and those not at risk (the 'always seated' and the 'never seated' in the RDMD terminology – see Section 3). X_i is a vector of observable and pre-determined pupil characteristics, including female, White British, EFL, FSM, and SEN dummies, as well as KS1 and KS2 scores in English, Mathematics, and Science; D_i^j is a dummy variable equal to 1 if pupil i received an offer from one of the free schools ($j = 1, 2$) and equal to 0 otherwise; α_{jck} are free school j-by-cohort c-by-experiment k dummies, identifying pupils that face the same risk of being admitted to the schools within each cohort. In addition, $\phi(d_i)$ are cohort-specific piecewise linear running-variable controls for distance from the cut-off for the RD draws, d_i, parametrised as follows:

$$\phi(d_i) = \sum_{2ck} L_{i,2ck}\left[\phi_{1,2ck} \times d_i + \phi_{2,2ck} \times d_i \times \ 1(d_i \geq \tau_{2ck})\right] \tag{2}$$

In Eq. (2), $L_i = I(\tau_{2ck} - \delta_{2ck} < d_i < \tau_{2ck} + \delta_{2ck})$. That is, L_i is a dummy variable equal to 1 if pupil i is at risk of being admitted through the RD mechanism at the CL school ($j = 2$) in Cohort $c \in (1,2)$, and belonging to the [τ-δ, τ+δ] category, and equal to 0 otherwise. In practice, τ_{2ck} is observed in the data as the distance of the last pupil admitted, while bandwidth δ is estimated by using the optimal bandwidth estimation method proposed by Calonico et al. (2014). This specification ignores the truncation problem among the second cohort of applicants to both schools. As discussed, we use applicants' lottery number in an IV strategy to test whether this poses a threat to our identification strategy.

5.1.2. Results

Table 2 presents balancing tests pooling both cohorts. Panel A reports combined estimates for both schools, while Panels B and C provide

Table 2
Balancing tests.

Dependent variable:	(1) White	(2) Male	(3) EFL	(4) FSM	(5) SEN	(6) KS2 English	(7) KS2 Maths	(8) KS2 Science	(9) KS1 English	(10) KS1 Maths	(11) KS1 Science
Panel A. Pooling both schools. Observations: 2112											
Free School Offer	−0.014	−0.040	0.007	0.004	−0.018	0.031	−0.002	0.023	0.187	0.059	0.176
	(0.024)	(0.030)	(0.027)	(0.023)	(0.021)	(0.034)	(0.034)	(0.033)	(0.197)	(0.193)	(0.184)
Joint significance	*0.798*										
Panel B. NE school. Observations: 914											
Free School Offer	−0.009	−0.030	−0.014	0.008	−0.012	0.022	0.006	0.028	0.096	−0.063	0.085
	(0.024)	(0.036)	(0.032)	(0.028)	(0.027)	(0.040)	(0.040)	(0.041)	(0.227)	(0.225)	(0.209)
Joint significance	*0.981*										
Panel C. CL School. Observations: 1198											
Free School Offer	−0.024	−0.061	0.049	−0.006	−0.030	0.052	−0.020	0.013	0.388	0.326	0.377
	(0.052)	(0.052)	(0.049)	(0.039)	(0.035)	(0.062)	(0.064)	(0.057)	(0.384)	(0.369)	(0.370)
Joint significance	*0.733*										
Cohort-by-experiment dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort-specific running variable controls	CL only	CL only	CL only	CL only	CL only	CL only	CL only	CL only	CL only	CL only	CL only

Notes: each coefficient comes from a different OLS regression. The dependent variable is reported in each column's heading. The specification adopted is the one in Eq. (1). Joint significance tests are obtained after jointly estimating the models for all outcomes. Robust standard errors in parentheses. ***: $p < 0.01$, **: $p < 0.05$, *: $p < 0.1$.

Table 3
Free school effectiveness – the impact on KS4 test scores.

Sample:	(1) *Both schools*	(2)	(3)	(4) *NE school*	(5)	(6)	(7) *CL school – full sample*	(8)	(9)	(10) *CL school – lottery only*	(11)	(12)
Equation:	Reduced form	1st stage	Structural form	Reduced form	1st stage	Structural form	Reduced form	1st stage	Structural form	Reduced form	1st stage	Structural form
Estimator:	OLS	OLS	TSLS	OLS	OLS	TSLS	OLS	OLS	TSLS	OLS	OLS	TSLS
Dependent variable:	KS4 Score Std.	Years of Exposure	KS4 Score Std.	KS4 Score Std.	Years of Exposure	KS4 Score Std.	KS4 Score Std.	Years of Exposure	KS4 Score Std.	KS4 Score Std.	Years of Exposure	KS4 Score Std.
Free School Offer	0.167***	2.189***		0.171***	1.969***		0.159**	2.675***		0.172**	2.759***	
	(0.044)	(0.135)		(0.055)	(0.163)		(0.075)	(0.234)		(0.081)	(0.250)	
Years of exposure			0.076***			0.087***			0.059**			0.062**
			(0.020)			(0.026)			(0.029)			(0.030)
Kleibergen-Paap F statistic	*264.6*			*146.6*			*131*			*121.9*		
Observations	1914	1914	1914	852	852	852	1062	1062	1062	1005	1005	1005
Cohort-by-experiment dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort-specific running variable controls	Yes	Yes	Yes	No	No	No	No	Yes	Yes	No	No	No
Pupil-level controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: each coefficient comes from a different OLS or TSLS regression. The dependent variable is reported in each column's heading. The specification adopted is the one in Eq. (3). Pupil-level controls include female, White British, EFL, FSM, SEN dummies, as well as the level of end-of-primary scores in English, Maths and Science. The number of observations differs from Table 4 due to the presence of missing values in some pupil-level controls. The Kleibergen-Paap robust first-stage F statistic for the significance of the instrument is reported at the bottom of the table. Robust standard errors in parentheses.
***: $p < 0.01$.
**: $p < 0.05$
*: $p < 0.1$.

separate results for the NE and CL schools. In all cases, treatment and control groups are balanced on observable characteristics, as confirmed by tests for joint significance at the bottom of each panel.

Estimation results obtained for each cohort separately are reported in Tables AII-1 and AII-2 in Appendix II and are in line with those in Table 4. While a few variables are unbalanced, the extensive set of characteristics is jointly balanced. In unreported models, we also found that the balancing estimates were very similar when excluding all pupils in the RD sample, thus only comparing pupils who are admitted through one of the lotteries.

Finally, Table AII-3 in Appendix II reports the results of balancing tests for the lottery-number instrument used to address endogenous truncation in the second cohort. Results exclude RD-assigned pupils at the CL school. The instrument is uncorrelated with most baseline characteristics, and joint significance tests yield p-values of 0.818 (NE) and 0.153 (CL), supporting the validity of our approach.

5.2. Average free-school effects

5.2.1. Methods

We estimate the effects of years of attendance at each free school using the following model:

$$Y_i = \beta YearsOfExposure_i^j + \phi_{jck} + \psi(d_i) + \eta_{jc}X_i + \epsilon_i, \ \ if \ \ 0 < P_i^{jck} < 1 \qquad (3)$$

Eq. (3) is analogous to Eq. (1), but the dependent variable Y_i is pupil i's KS4 average test scores across all subjects taken by students, while the treatment variable of interest is $YearsOfExposure_i^j$, measuring the number of years spent by pupil i in either of the two free schools. To increase the precision of our estimates, we include the controls in vector X_i and allow their coefficients to differ by each school-by-cohort group. Given that actual free school attendance is endogenous, we estimate Eq. (3) with Two-Stage Least Squares (TSLS), using the free-school offer dummy D_i^j as an instrument for $YearsOfExposure_i^j$. The relevant first-stage equation is:

$$YearsOfExposure_i^j = \gamma D_i^j + \psi_{jck} + \mu(d_i) + \kappa_{jc}X_i + \nu_i, \ \ if \ \ 0 < P_i^{jck} < 1 \qquad (4)$$

This specification ignores the truncation problem for cohort 2. We therefore also use the IV strategy that exploits applicants' lottery numbers as instrument (described in Section 3) to assess the importance of this issue and validate our main estimates.

5.2.2. Main results

Table 3 reports results from models pooling both cohorts. Columns (1)–(3) present estimates for both schools combined; Columns (4)–(6) for the NE school; Columns (7)–(9) for the CL school; and Columns (10)–(12) exclude 72 pupils exposed to RD assignment at the CL school. This last sample provides a relevant benchmark for the IV estimates reported in Table 4.

Columns (1), (4), (7), and (10) show reduced-form effects of receiving a free-school offer on standardised KS4 scores. In all samples, offers lead to gains of approximately 0.16 standard deviations, statistically significant and similar across schools.[17]

Columns (2), (5), (8), and (11) report first-stage effects, indicating

[17] KS4 scores range between 0 and 9, with a mean of 5.18 for the NE school and 5.72 for the CL one. The observed standard deviation is 1.67 points for the NE school, and 1.73 points for the CL one. Although we standardise scores pooling schools, effects using school-specific standardised scores are comparable.

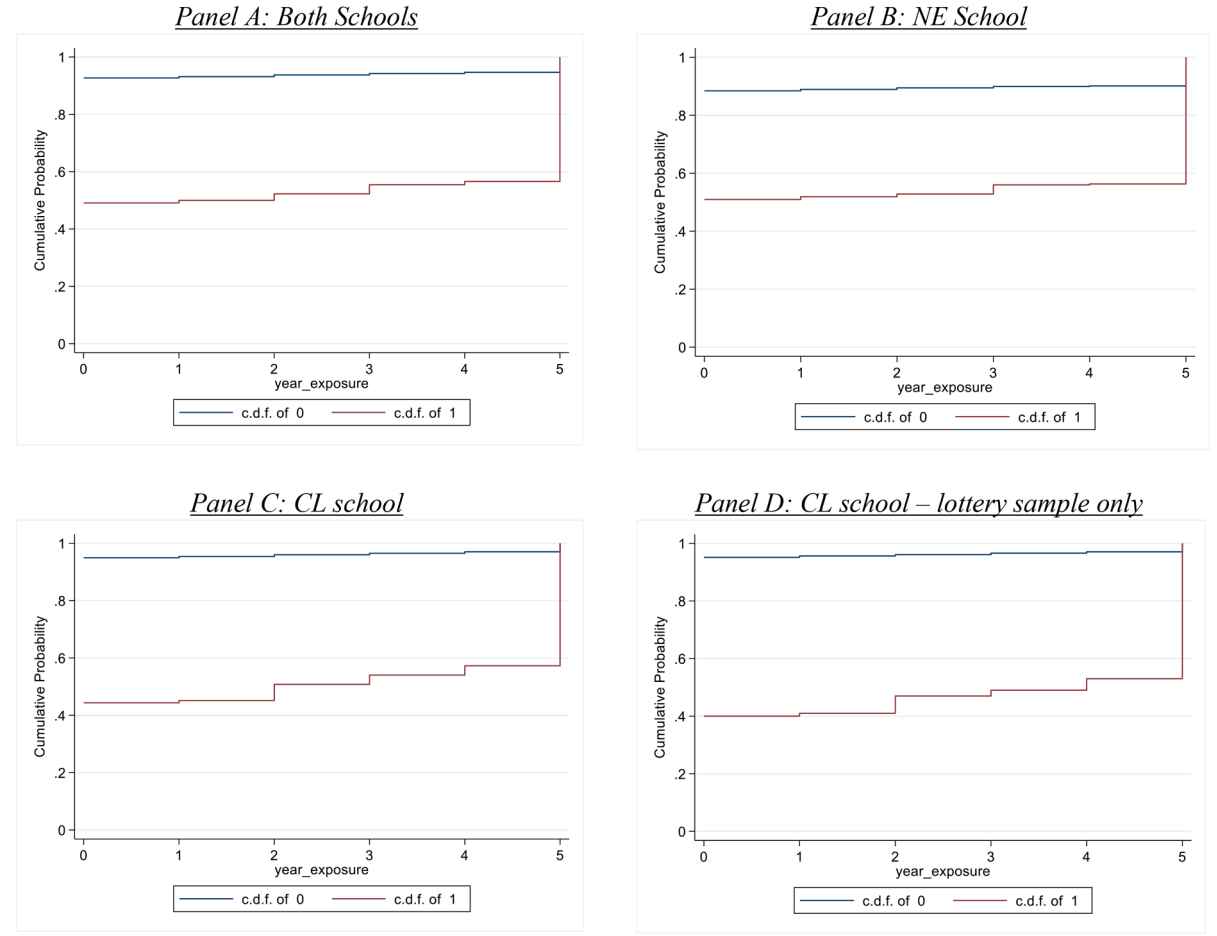

Fig. 1. Cumulative Distribution Functions (CDF) of years of enrolment for offer and non-offer recipients.
Note: the plots report Cumulative Distribution Functions (CDF) of years of enrolment at the free schools depending on the value of the offer instrument.

that offers increase years of enrolment in the assigned free school by about two years on average. These effects are consistent across specifications and strongly significant, with first-stage F-statistics above 100.

In the presence of heterogeneous effects, our IV estimates are informative about the effects on compliers to the instrument – and describing the characteristics of compliers relative to always-takers and never-taker applicants is important to shed light on the external validity of our findings. We follow the approach of Angrist et al. (2023), with results reported in Table AI-1 in Appendix I. For the NE school, differences across groups are minimal across all characteristics. At the CL school, however, always-takers appear more disadvantaged: they are more likely than compliers and never-takers to be male (67 % vs. 54 %), FSM-eligible (30 % vs. 16 %), White British (74 % vs. 59 %), and have SEN status (25 % vs. 13 %). By contrast, prior attainment at the end of primary school is similar across groups.

Finally, as our treatment leverages variation in years of free school enrolment, the estimates obtained using binary offer instruments identify an average of per-year of enrolment effects – with Average Causal Response (ACR) weights that are proportional to differences in the CDF of the years of enrolment for offer recipients and non-offer recipients (Angrist & Imbens, 1995). Fig. 1 displays these CDFs for all applicants, and separately for the NE and CL schools. Fewer than 5 % of non-offer applicants attend the free schools, and most who do complete the full five years. Among offer recipients, nearly 50 % do not enrol, over 40 % stay all five years, and fewer than 10 % attend for a shorter period. Patterns are similar across the two schools, though the CDF for admitted CL applicants starts lower and increases more steeply. This yields a relatively flat ACR weighting function, indicating that estimated effects reflect a broad distribution of treatment intensities.

To conclude, columns (3), (6), (9), and (12) of Table 3 report the TSLS effects of years of free-school exposure on KS4 scores. The results show that spending an additional year in a free school boosts achievement by 0.076 standard deviations. We obtain a marginally higher effect size for the NE school (0.087 standard deviations) than for the CL school (0.059 and 0.062 standard deviations), but these differences are not large. Importantly, we find that excluding pupils at risk of being admitted through the RD mechanism makes no relevant difference for the estimation of the CL-school effect.[18] Cohort-specific results, shown in Appendix Tables AII-4 and AII-5, are consistent with the main

[18] We stress again that excluding students 'at risk' of RD admission is very different from disregarding altogether pupils who might be selected by distance and focussing on the two lottery groups at the CL school. This is because *ex-ante* it is impossible to know who is only at risk of admission from a lottery draw among pupils living within 3 miles of the school. This set can only be identified by using the insights of the RDMD literature and categorising pupils as 'always seated', 'never seated' and 'conditionally seated' for the RD experiment.

Table 4
Addressing truncation - Cohort 2.

Sample:	(1) *Both schools*	(2)	(3)	(4) *NE school*	(5)	(6)	(7) *CL school – lottery only*	(8)	(9)
Equation:	Reduced form	1st stage	Structural form	Reduced form	1st stage	Structural form	Reduced form	1st stage	Structural form
Estimator:	OLS	OLS	TSLS	OLS	OLS	TSLS	OLS	OLS	TSLS
Dependent variable:	KS4 Score Std.	Years of Exposure	KS4 Score Std.	KS4 Score Std.	Years of Exposure	KS4 Score Std.	KS4 Score Std.	Years of Exposure	KS4 Score Std.
Panel A. Offer dummy as IV									
Free School Offer	0.167***	2.334***		0.165**	1.955***		0.175	3.213***	
	(0.063)	(0.178)		(0.078)	(0.205)		(0.120)	(0.352)	
Years of exposure			0.072**			0.084**			0.055
			(0.026)			(0.037)			(0.038)
Kleibergen-Paap F statistic	*172.3*			*83.40*			*83.40*		
Panel B. Lottery number as IV									
Lottery number	−1.766***	−0.150*		−2.736***	−0.278**		−1.149***	−0.069	
	(0.202)	(0.082)		(0.359)	(0.140)		(0.240)	(0.069)	
Years of exposure			0.085*			0.101**			0.060
			(0.046)			(0.048)			(0.102)
Kleibergen-Paap F statistic	*76.67*			*58.12*			*22.91*		
Observations	1138	1138	1138	364	364	364	774	774	774
Cohort-by-experiment dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pupil-level controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: each coefficient comes from a different OLS or TSLS regression. The dependent variable is reported in each column's heading. The specification adopted is the one in Eq. (3) for Panel A, while Panel B uses the lottery number (percentile rank varying between 0 and 1) instead of the offer dummy as instrument for years of exposure. Pupil-level controls include female, White British, EFL, FSM, SEN dummies, as well as the level of end-of-primary scores in English, Maths and Science. For the CL school, the sample excludes students exposed to RD admission risk. The number of observations differ from Table 4 due to the presence of missing values in some pupil-level controls. The Kleibergen-Paap robust first-stage F statistic for the significance of the instrument is reported at the bottom of each panel. Robust standard errors in parentheses.
*** : $p < 0.01$.
** : $p < 0.05$.
* : $p < 0.1$.

findings.[19]

5.2.3. *Dealing with offer truncation*

Next, we investigate the potential problem of offer truncation among applicants in the second cohort. Our findings are presented in Table 4, in which we compare the effects for each school when using the offer dummy (Panel A) or the lottery number (Panel B) as instruments for years of free-school exposure, among pupils in the second cohort only. For the CL school, we drop pupils at risk of being admitted through the RD mechanism as we do not have a credible way to address the issue of truncation. This is not a significant exclusion – approximately 40 pupils – and does not affect our main estimates.

The instrumental variable estimates that we obtain using the offer IV are 0.072 (s.e. 0.026) when pooling the two schools; 0.084 (s.e. 0.037) for the NE school; and 0.055 (s.e. 0.038) for the CL school. The corresponding estimates we obtain using the lottery number IVs are 0.085 (s.e. 0.026), 0.101 (s.e. 0.048), and 0.060 (s.e. 0.102). We therefore conclude that the two approaches lead to economically similar and statistically indistinguishable effects.

To further assess robustness, we extend the lottery number IV to both cohorts for the NE school. Unfortunately, we cannot replicate this for the CL school, as we do not have access to applicants' lottery numbers in the first cohort. The evidence in this respect is reassuring. When pooling the two cohorts for the NE school and instrumenting years of attendance with lottery numbers we find an estimate at 0.088 (s.e. 0.026), virtually identical to the one obtained when instrumenting years of attendance with the offer dummy.

Finally, a remaining concern is that lottery number is continuous while the offer dummy is binary, which could identify different parameters under treatment effect heterogeneity (Heckman & Vytlacil, 2005). To deal with this issue, we use lottery numbers to construct a discrete 'potential offer' variable defined as being above the lottery cutoff of the relevant risk set *ignoring truncation*. This is only feasible for the NE school, where offer risk is independent across stanines, and we can compute potential offers by ordering pupils according to lottery numbers within stanine and assigning offers to the pupils with the highest numbers until capacity is reached (ignoring truncation). On the other hand, for the CL school, receiving an offer in the lottery samples is conditional on the results of the offer draws in the RD sample, so the 'potential offer' approach is difficult to conceptualise. As a result, we do not implement it for the CL school and only focus on the NE one. Table AI-2 in Appendix I reports that the estimated causal effects obtained with the offer IV (0.087, s.e. 0.026), the lottery IV (0.088, s.e. 0.026) and the potential offer IV (0.101, s.e. 0.040). These are all of very similar magnitudes, suggesting offer truncation is not a significant problem in

[19] We also estimated models using free-school attendance at KS4 instead of years of exposure. Results, shown in Table AII-6 (Appendix II), confirm similarly large, positive, and significant effects on KS4 test scores for both schools. To further test robustness, we also: i) excluded pupil-level controls; ii) used KS2 dummies instead of a linear KS2 score; and iii) varied the polynomial order and bandwidth in the RD sample for the CL school. As reported in Table AII-7, the main findings remain robust across all specifications.

Table 5
Heterogeneous effects on KS4 test scores - by pupil characteristics.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample:	White British	Non-White British	Male	Female	FSM	Non-FSM	High KS2	Low KS2
Equation:	Structural form	Structural form	Structural form	Structural form	Structural form	Structural form	Structural form	Structural form
Estimator:	TSLS	TSLS	TSLS	TSLS	TSLS	TSLS	TSLS	TSLS
Panel A: NE school								
Years of exposure	0.102	0.081***	0.117***	0.038	0.058	0.081***	0.066	0.084***
	(0.096)	(0.027)	(0.032)	(0.046)	(0.074)	(0.028)	(0.065)	(0.027)
Observations	112	740	492	360	144	708	347	505
K-P F statistic	*6.010*	*132.2*	*82.31*	*52.65*	*22.38*	*116.1*	*26.55*	*126.5*
P-value of difference	*0.832*		*0.162*		*0.767*		*0.798*	
Panel B: CL school – lottery only								
Years of exposure	0.124***	0.010	0.070	0.055	0.020	0.075**	0.084*	0.051
	(0.045)	(0.046)	(0.043)	(0.045)	(0.100)	(0.032)	(0.049)	(0.036)
Observations	597	408	561	444	193	812	439	566
K-P F statistic	*62.53*	*52.55*	*68.09*	*51.30*	*15.31*	*105.2*	*35.25*	*103.6*
P-value of difference	*0.075*		*0.817*		*0.596*		*0.585*	
Cohort-by-experiment dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pupil-level controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: each coefficient comes from a different TSLS regression. The dependent variable is the standardised KS4 score, and the sample considered is listed in the column heading. The specification adopted is the one in Equation- (3). Pupil-level controls include female, White British, EFL, FSM, SEN dummies, as well as the level of end-of-primary scores in English, Maths and Science. For the CL school, the sample excludes students exposed to RD admission risk. The Kleibergen-Paap (K-P) robust first-stage F statistic for the significance of the instrument is reported at the bottom of each panel. Robust standard errors in parentheses.
*** : $p < 0.01$.
** : $p < 0.05$.
* : $p < 0.1$.

our context.[20]

5.2.5. Results by subject

The effects reported so far concern average pupil achievement across all subjects. In Table AI-3 in Appendix I, we focus on the three core subjects that all students take at the end of secondary education – namely, English, Maths, and Science. We also investigate the impact of attending a free school on the number of exams taken at KS4. We find positive effects of attending the two free schools on scores in each of the three core subjects. Due to data limitations, we only observe science scores for one cohort, and as a result the positive effects for science are imprecisely estimated for the CL school. Importantly, the two free schools not only improve average grades and increase the share of pupils attaining a pass grade, but they also increase the incidence of pupils who achieve top-level grades. Conversely, we find no evidence that the higher scores obtained in the core subjects are achieved at the expense of students entering a significantly lower number of exams. Although the coefficient for the number of exams entered for the CL school pupils has a negative sign (consistent with its educational philosophy described in Section 2.4), it is small and not statistically significant.

To summarise, the two schools increase KS4 achievement by 0.076 standard deviations per year of exposure, with slightly larger effects at the NE school (0.087 SD) than at the CL school (0.062 SD), though this difference is not statistically significant. Overall, our effects are comparable to those of successful US charter schools (Cohodes & Parham, 2021), though smaller than those found for Boston charters (Abdulkadiroğlu et al., 2011).

5.3. Heterogeneous effects by pupil background

Given that our two free schools adopt different pedagogical approaches, we may expect them to benefit different subgroups of pupils. To investigate these issues, in Table 5 we estimate the impact of free-school attendance on KS4 test scores across different subgroups pooling the two cohorts of applicants. Panel A reports the effects for the NE school and Panel B for the CL school.[21] Throughout our discussion, we emphasise the patterns that emerge when comparing point estimates – instead of the statistical significance of any differences – as sample sizes are too small to precisely pin down heterogeneous effects. For the same reason, we do not adjust for multiple testing – e.g., a Bonferroni correction would render all differences insignificant – or apply more flexible methods for heterogeneity detection (e.g. Athey et al., 2019; Chernozhukov et al., 2018), which require larger samples to yield meaningful findings.

For the NE school, we find that effects are larger for males, non-FSM-eligible pupils, non-White British pupils, and low-achieving pupils (based on their KS2 performance) – although none of these differences is statistically significant. We can contrast these findings with the insights from our visit. When asked which groups of pupils benefited more from attending the school, the headteacher suggested 'disadvantaged students'. While this lines up with the breakdown in terms of prior achievement and ethnicity, it does not square with the (slight) difference in favour of non-FSM pupils. Furthermore, when asked whether the school's educational approach particularly favours boys, the trust CEO suggested that this is possible as 'girls manage to stay afloat in the other schools in the area, whereas boys get lost and disengage from education'. Although this *ex-post* rationalisation is in line with our findings, the gender divide was not offered as one of the key patterns of heterogeneity perceived on the ground.

On the other hand, the CL school delivers considerably larger effects among White British than among non-White British pupils. This difference is economically meaningful (0.124 vs. 0.010) and borderline statistically significant, with a p-value (unadjusted for multiple testing) of 0.075. We find no clear differential effects depending on gender, but some (insignificant) evidence that the school is more effective for non-FSME pupils and high-achieving pupils. Does this evidence line up

[20] As reported in Panel C of Table AI-1 in Appendix 1, there are also no marked differences in the characteristics of the compliant subpopulation for the offer vs. potential offer instruments.

[21] We use offer, not lottery numbers, as the instrument for years of attendance to retain both CL cohorts. For consistency, we also exclude pupils admitted based on proximity.

with the information we gathered from our interviews? To a large extent, it does. Indeed, the headteacher suggested that the school has historically been more effective and more attractive for pupils from more affluent families, probably because of its CL pedagogical philosophy.

Irrespective of the perceptions of stakeholders we spoke to – and discounting the power issues that we face – these patterns appear in line with the educational approaches championed by the two schools. Prior research suggests the 'no excuses' model may be particularly effective for ethnic minorities and for boys, the latter of whom may benefit from a more disciplinarian structure, and for low-achieving pupils, who need support in raising expectations and aspirations. In contrast, the 'classical liberal' approach appears more effective among White British pupils, and to some extent among non-FSM and high-achieving students. This suggests that a knowledge-rich, 'specify and excel' curriculum is better suited to pupils from more advantaged backgrounds.

5.4. The free-school effects on behavioural outcomes

Given growing evidence on the importance of school (and teacher) effects on non-cognitive outcomes (e.g., Jackson, 2018), we investigate the impact of the free schools on behavioural outcomes. Specifically, we estimate effects on: *i)* unauthorised absences; *ii)* school exclusions for disciplinary reasons; and *iii)* the probability that pupils sit their KS4 exams at the school where they started Year 7.[22] Our results are presented in Table 6.

We find that pupils enrolling in one of the free schools in Year 7 miss fewer school sessions due to unauthorised absences, irrespective of whether we use the share of sessions missed (impact sizes of around 9 % and 18 % for the NE and CL schools, respectively) or a dummy for whether they have missed any session (impact sizes of 3.6 % and 5.3 % for the NE and CL schools). However, we do not find significant differences in terms of disciplinary exclusions, although there is a marginally significant effect on the probability of any exclusion for the NE school. This null effect may be the result of two forces: better behaviour among pupils on the one hand, and stricter discipline on the other. Still, such exclusions are rare: 14 percent of pupils have been excluded for at least a session, but the average number of sessions missed is less than half a day per year. We may therefore lack power to uncover small effects. Finally, free-school attendees are substantially more likely to stay in the same school until their KS4 exams compared to applicants who enrolled elsewhere (approximately a 25 percentage-point increase over the average among counterfactual students at 69 % – amounting to a 36 % increase in retention). Overall, these results suggest that the positive effects of the two free schools on achievement may be partly explained by their positive effect on behaviour and negative impact on mobility.[23]

6. What sets these two schools apart?

To better understand what might explain the strong achievement gains observed in both free schools, we consider what might set our two schools apart from their counterfactuals and from one another. To do so, we proceed in two steps. First, we examine whether standard school-level characteristics – such as peer composition, staffing profiles, and inspectorate ratings – differ between the free schools and the ones attended by their unsuccessful applicants. Second, we analyse differences in pedagogical philosophy and organisational ethos, using a structured text analysis of schools' VE statements.

Table 6
Mechanisms: effects on behavioural outcomes.

	(1)		(2)	
Sample:	*NE school*		*CL school – lottery sample*	
Equation:	Structural form		Structural form	
Estimator:	TSLS		TSLS	
Outcome variable:	Coefficient	Std. Err.	Coefficient	Std. Err.
Year 7 and KS4 in the same school	0.273***	0.076	0.241***	0.079
Share of classes missed for unauthorised absences	−0.001*	<0.001	−0.002*	0.001
Dummy - positive share of classes missed for unauthorised absences	−0.027*	0.016	−0.033*	0.018
Average yearly # of sessions lost due to disciplinary exclusions	0.014	0.095	0.261	0.173
Dummy - positive # of sessions lost due to disciplinary exclusions	0.024*	0.013	0.019	0.014
Observations	883		1055	
Cohort-by-experiment dummies	Yes		Yes	
Cohort-specific running variable controls	No		No	
Pupil-level controls	Yes		Yes	

Notes: each coefficient comes from a different TSLS regression. The dependent variable is listed in each row, and the sample considered is listed in the column heading. The endogenous variable is "starting Year 7 in the school" for the "same school" regression, and "years of exposure to free schools" for the other variables. The specification adopted is the one in Eq. (3). Pupil-level controls include female, White British, EFL, FSM, SEN dummies, as well as the level of end-of-primary scores in English, Maths and Science. For the CL school, the sample excludes students exposed to RD admission risk. Robust standard errors in parentheses.
***: $p < 0.01$
**: $p < 0.05$.
*: $p < 0.1$.

A limitation of our analysis is that although these investigations reveal significant and informative contrasts, they do not establish which specific features drive our estimated effects. Indeed, institutional features and school 'cultures' are shaped by treatment assignment and cannot be treated as exogenous mediators. For this reason, we refrain from conditioning on these post-treatment variables in our regressions in the style of a 'mediation analysis'. Moreover, other unobserved factors – such as leadership quality or staff commitment – may also play a central role. Our goal in this section is therefore not to pinpoint specific causal pathways, but to characterise how these two schools differ from their counterfactuals and from one another, along dimensions that plausibly matter for student learning.

6.1. School attributes

We start by considering whether the two free schools differ along a wide range of standard school characteristics compared to counterfactual schools attended by rejected applicants. We carry out our analysis at the pupil level and study whether students starting secondary education (in Year 7) at one of the two free schools experience different school characteristics compared to pupils who applied to the free schools but ended up attending other schools. As before, we overcome the endogeneity of the decision to enrol in a free school by using the offer dummies as instruments. We look at four categories of school characteristics: *i)* pupil demographics; *ii)* characteristics of the teaching body; *iii)*

[22] Grade repetition is very rare in England so we cannot look at this outcome. Note that we use a binary treatment variable indicating whether pupils enrolled in the free schools in Year 7 when analysing effects on the probability of remaining at the same school until the KS4 exams.

[23] We also analyse potential heterogeneous effects of the free schools on these behavioural outcomes but fail to detect insightful patterns. Our results are reported in Tables AII-8–AII-10 in Appendix II.

Table 7
Mechanisms: the characteristics of the counterfactual schools.

	(1)		(2)	
Sample:	*NE school*		*CL school – lottery sample*	
Equation:	Structural form		Structural form	
Estimator:	TSLS		TSLS	
Outcome variable:	Coefficient	Std. Err.	Coefficient	Std. Err.
Panel A. Demographics				
Number of pupils	−108.4***	(7.346)	−46.9***	(4.666)
% Female pupils	−8.70***	2.40	−6.40	3.90
% FSM pupils	−4.165***	1.369	−0.235	1.647
% White British pupils	−20.919***	4.097	16.127***	1.857
Average KS2 grades of peers	0.042	(0.039)	−0.135***	(0.038)
Panel B. Features of the teaching body				
Pupil/teacher ratio	−2.795***	0.218	0.524*	0.273
% male teachers	0.217	1.122	12.013***	1.384
% qualified teachers	5.074***	1.027	−18.573***	0.536
Gross salary (£)	−2995***	306	9227***	254
Teacher turnover	0.039***	0.009	−0.063***	0.012
Headteacher turnover	−0.033	0.024	0.843***	0.035
Panel C. School institutional features				
Mixed gender school	0.171***	0.043	0.310***	0.057
Academy/free school	0.413***	0.067	0.581***	0.058
Faith school	−0.309***	0.053	−0.474***	0.055
Selective (grammar) school	0.009	0.026	−0.023	0.021
Panel D. Ofsted reports				
Overall grade	−1.408***	0.166	0.280***	0.085
Management	−1.225***	0.154	0.311***	0.077
Behaviours	−1.206***	0.130	−0.642***	0.079
Teaching	−1.333***	0.157	0.140*	0.084
Achievement	−1.408***	0.166	0.262***	0.085
Observations	837		859	
Cohort-by-experiment dummies	Yes		Yes	
Cohort-specific running variable controls	No		No	
Pupil-level controls	Yes		Yes	

Notes: each coefficient comes from a different TSLS regression. The dependent variable is listed in each row, and the sample considered is listed in the column heading. The regressions exclude pupils for whom the Year 7 school identifier is missing. The specification adopted is the one in Eq. (3). Pupil-level controls include female, White British, EFL, FSM, SEN dummies, as well as the level of end-of-primary scores in English, Maths and Science. For the CL school, the sample excludes students exposed to RD admission risk. Robust standard errors in parentheses.
*** : $p < 0.01$
**: $p < 0.05$.
* : $p < 0.1$.

institutional features; and *iv)* Ofsted reports.[24] Our results are reported in Table 7. Column (1) reports the results for the NE school compared with the counterfactual schools, while column (2) reports the results for the lottery sample at the CL school (to limit problems related to the truncation issue discussed above).[25]

First, free schools are smaller than the counterfactual schools attended by unsuccessful applicants – by 108 pupils for the NE school and 47 for the CL school – reflecting their gradual expansion model. Prior work (e.g., Abdulkadiroğlu, Hu, & Pathak, 2013; Schwartz et al., 2013) suggests that smaller schools can enhance achievement, warranting closer scrutiny of this dimension.

In Fig. 2, we use aggregate (publicly-available) school level data to report the evolution of average school size (Panel A), Key Stage 4 scores (Panel B), and Key Stage 2 scores (Panel C) for pupils enrolled at our two free schools and those in counterfactual schools.[26] School size is measured at the time of Year 7 for pupils carrying out their Key Stage 4 in the year reported on the horizontal axis. As expected, both schools grew steadily after opening. For the NE school, KS4 performance consistently exceeded that of counterfactual schools – even as size increased, with a modest dip post-pandemic. For the CL school, performance started at a similar level but rose markedly over time. KS2 prior attainment remained stable across groups, suggesting intake composition is not driving these trends. These results reduce concerns that our positive effects are short-lived and depend on the fact that the schools are new – with students and teachers being galvanised. If that were the case, school effectiveness should be consistently diluted over time as such novelty effects progressively wear out. We find that this is not the case.[27]

Second, peer demographics differ between free and counterfactual schools, though inconsistently. NE attendees are exposed to fewer FSM-eligible (−4 pp.), White British (−20 pp.), and female pupils (−9 pp.). Their peers' KS2 scores are similar. CL attendees face more White British peers (+16 pp.) but slightly lower average KS2 scores (−0.13 SD).

Third, staff characteristics also vary. NE pupils have lower pupil-teacher ratios (−2.8), more qualified teachers (+5 pp.), but lower average pay (−£3000). The CL school has more male teachers (+12 pp.), fewer qualified staff (−19 pp.), and higher salaries (+£9000). Teacher mobility was initially higher at the NE school and lower at the CL, with reversed patterns for headteachers.[28] Overall, these differences are unlikely to consistently explain the large positive effects on KS4 outcomes, with similar effect sizes across the two schools.

Institutional features also diverge. Compared to free-school attendees, counterfactual pupils are more likely to attend single-sex and faith schools (neither free school is selective or religious), and less likely to attend academies or free schools.[29] However, we see no statistically significant difference in the probability of attending a selective school. Once again, it is unlikely that these differences drive the effects on pupil achievement. Bertoni et al. (2022) study the relative effectiveness of different secondary school types in the England and find that only selective grammar schools deliver higher academic returns on average. Neither of the free schools we analyse is selective, nor are the counterfactual schools.

Finally, applicants not enrolling in the NE school attend schools with significantly worse Ofsted ratings (by 1.2–1.4 points across domains), while CL applicants end up in schools with marginally better ratings – except in behaviour, where the CL school scores 0.64 points higher.

6.2. School pedagogical approaches

We next quantify how the two free schools differ in terms of their distinctive pedagogical approaches. To investigate this issue, we carry out a quantitative text analysis on web-scraped VE statements for the

[24] These regressions are limited to pupils with a valid Year 7 school identifier. Missing identifiers – affecting around 12% of the sample – may reflect short-term mobility or moves to/from independent schools. Tables AII-11 and AII-12 (Appendix II) replicate the balancing and effectiveness estimates within this subsample, yielding results virtually identical to the main findings.

[25] Note that we again instrument attendance using offer dummies rather than lottery numbers, as the latter are unavailable for the second CL cohort. Relying on lottery numbers would substantially reduce precision. However, as previously shown, addressing truncation via lottery instruments does not materially affect our results.

[26] Key Stage scores are standardised within cohort to ensure comparability across years, as test formats changed slightly over time. Also, no data are available for the pandemic-affected years 2020 and 2021.

[27] Notice that we cannot carry out an investigation similar to the one we present in the main analysis for subsequent cohorts as we do not have access to application/admissions data for the two schools.

[28] Due to data limitation, we can only measure teacher/headteacher mobility across two adjacent years.

[29] There are no free schools among the CL counterfactuals, and only two among the NE counterfactuals.

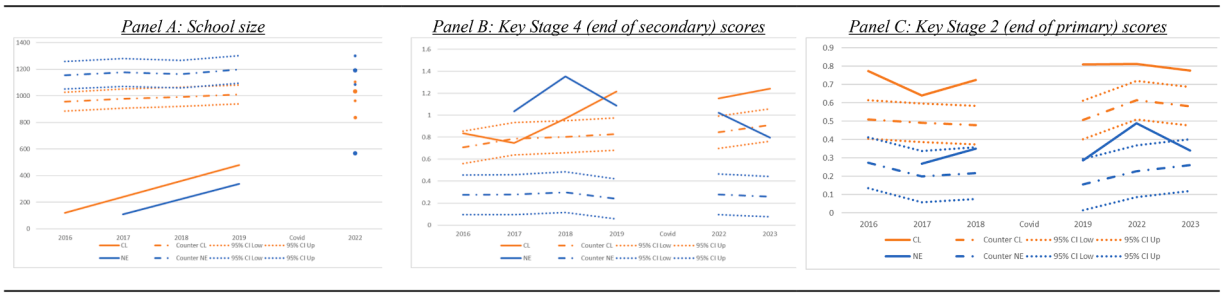

Fig. 2. Size and attainments by school year. Free schools and counterfactual schools.
Notes: the figures report the average school size (Panel A), Key Stage 4 (Panel B), and Key Stage 2 (Panel C) scores of each free school and of the counterfactual schools attended by unadmitted applicants, by school year. For the counterfactual schools, the 95 % confidence intervals for the mean score are also reported. School size is measured with reference to Year 7 for pupils doing Key Stage 4 in the year report in the horizontal axis. All data is obtained for school-level, publicly available data. Information not available for the pandemic years (2020 and 2021) and in 2023 for school size.

two free schools and the counterfactuals in our sample.

To do so, we start by creating a taxonomy of 'concepts' that characterise the pedagogical approaches of the free schools. These are pinned down by key words – including core values and drivers – and key expressions in the VE statements that resonate with the impressions we gathered from our school visits and semi-structured interviews. We then create an extended 'dictionary' that associates such concepts to a wide range of synonyms and akin expressions that can be 'searched' in the statements of counterfactual schools. Finally, we create two synthetic measures that capture whether a key concept – representing an aspect of the free schools' pedagogical approach – is 'over-represented' at a given school compared to what would be expected by looking at the full sample of VE statements (correcting or not for the complexity of the words used). More details are provided in Appendix I.

Descriptive statistics for our indexes are reported in Appendix Table AI-4. The top panel focuses on the NE school, and the bottom on the CL school. Column (1) shows the (standardised) unadjusted index for each key concept at the free schools. Columns (2) and (3) report the mean and median index values among counterfactual schools. Column (4) shows the share of counterfactual schools referencing the concept. Columns (5) and (6) report the free school's rank within its sample (NE plus 52 counterfactuals; CL plus 96), based on the unadjusted and complexity-adjusted indexes, respectively.

Starting with the top panel, we observe that some features associated with the NE school are also common among counterfactuals. For instance, 85 % and 90 % of counterfactual schools reference purpose and autonomy, and the NE school ranks only 24th and 42nd on these dimensions using the complexity-adjusted index. In contrast, features like no excuses and strict routines appear more distinctive: just 7.7 % and 9.6 % of counterfactual schools mention them, while the NE school ranks 1st or 2nd/3rd depending on the index. Although many schools reference hard work and a culture of success – present in 92 % and 98 % of VE statements – the NE school gives these ideas significantly more weight, ranking between 3rd and 11th. This suggests our index effectively captures not only the presence of key concepts but also the intensity with which they are emphasised.

Moving to the bottom panel, we find that some of the practices adopted by the CL school are also quite common. Two of its core values – hard work and high standards – appear in 82 % and 95 % of the VE statements of the counterfactuals, and the school only ranks 45th and 47th (using the complexity adjusted index). On the other hand, the classical liberal and knowledge-rich concepts are unique: they only appear in 2 % and 10 % of the statements – and the school is very highly ranked (1st and 4th looking at Column 6). More strikingly, we find that teacher-led learning does not appear in any of the counterfactuals. Finally, while politeness is a concept that emerges from 86 % of the VE statements, the CL school emphasises it to a larger extent, ranking 4th and 15th in Columns (5) and (6).

As for the 'standard' school attributes considered above, we merge this information to our main dataset at the pupil level and use our regression framework to study whether individuals who do not attend one of the free schools join counterfactuals with similar or different pedagogical approaches. Again, we control for experiment-by-cohort effects and pupil characteristics, and instrument attendance at a free school at the start of secondary education with offer. Our findings are presented in Table 8, where we concentrate on the standardised complexity-adjusted index (results using the simple index are similar). The top panel focusses on the NE school, while the bottom panel reports results for the CL one.

We find that pupils attending the NE school are significantly more likely than those in counterfactual schools to be exposed to a 'no excuses' approach to education, characterised by a strong emphasis on strict routines, culture of success, and hard work, as well as to the values of trust, fairness, and mastery. These are characterised by the school as follows: 'We are loyal (…) and always do what we say we will do' (trust); 'We are open-minded (…) and play by the rules' (fairness); and 'The urge to get better and better (…)' (mastery). Once again, this seems to emphasise a 'no excuses' approach to education.[30]

For pupils attending the CL school, we find that the classical-liberal philosophy – alongside the concepts of knowledge-rich, competitive atmosphere, and debates – are significantly more prevalent than among pupils at counterfactual schools. Expectedly, we find that a teacher-led approach to learning is a key feature of the CL school, while high aspirations seem to be emphasised less than at the counterfactuals. We also find that the concepts of politeness and kindness are over-represented among pupils in the CL school. While the first relates to the disciplined, non-permissive approach to pupils' behaviour at school, the second is more closely related to ideas such as altruism, compassion, and consideration of others (and is one of the school's core values). Finally, hard work and high standards are not substantially more salient elements of the environment faced by pupils at the CL school compared to the counterfactuals, despite both being among its core values.[31]

[30] Other features of the 'no excuse' model seen in US charters – such as extended instructional time and high-dosage tutoring (Angrist et al., 2013; Dobbie & Fryer, 2013) – are rare in England due to institutional constraints, particularly around staffing. Cirin (2014) found that only 8% of academies and 4% of free schools had altered the length of the school day or terms.

[31] We cannot directly observe schools' actual adherence to their aspirational VE statements. Given the growing policy emphasis in England on general principles like aspiration and high standards, other schools may echo these themes in their VEs without fully embracing them. If so, we may underestimate the true differences in educational philosophy between the two free schools and others.

Table 8
Text analysis on schools' 'Vision and Ethos Statements' - pupil-level counterfactual school regressions.

Panel A. NE school.	(1) No excuses	(2) Strict routines	(3) Culture of success	(4) Good manners	(5) Hard work	(6) Trust	(7) Fairness	(8) Autonomy	(9) Purpose	(10) Mastery
Start at FS	0.995***	1.526***	0.482***	−0.092	1.439***	2.131***	1.119***	−0.801***	−0.286	1.311***
	(0.157)	(0.130)	(0.158)	(0.163)	(0.141)	(0.038)	(0.146)	(0.136)	(0.181)	(0.149)
Panel B. CL School – lottery only	Classic liberal	Knowledge-rich	Compet. Atmosphere	Teacher-led learning	Politeness	Debate	High aspirations	Kindness	Hard work	High standards
Start at FS	2.952***	1.293***	2.890***	3.032***	0.856***	2.714***	−0.244*	0.623***	−0.006	−0.280***
	(0.023)	(0.074)	(0.028)	(0.001)	(0.093)	(0.055)	(0.132)	(0.106)	(0.094)	(0.091)
Cohort-by-experiment dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pupil-level control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: each coefficient comes from a different TSLS regression focussing on whether pupils start at one of the two free schools instruments with offer from that school. The dependent variable is an index measuring excess occurrence of a key work (and its synonyms) in the free school statement relative to expected occurrence in body of statements and adjusting for the 'complexity' (rare occurrence) of the words used in the statements. Index values have been standardized. Different key words (and their synonyms) are analysed as reported in each column's heading. The specification adopted is the one in Eq. (3). For the CL school, the sample excludes students exposed to RD admission risk. Pupil-level controls include female, White British, EFL, FSM, SEN dummies, as well as the level of end-of-primary scores in English, Maths and Science. Robust standard errors in parentheses.

***: $p < 0.01$

**: $p < 0.05$.

*: $p < 0.1$.

Table 9
Review of school type ITT effects from small-scale lottery or regression discontinuity studies.

Reference	Country / Region	School Types	School Phase	# Schools	Test scores	ITT Effects (SD)
This study	UK	Free schools vs. others	Secondary	2	KS4 average scores in English, Math, Science	0.17SD (Average English, Math, Science)
Dobbie & Fryer (2011)	USA (NYC)	Exam schools vs. others	Secondary	3	SAT scores in math, ELA writing	No significant effects
Abdulkadiroğlu, et al. (2014)	USA (Boston & NYC)	Exam schools vs. others	Secondary	6	SAT and Regent's math and reading	No significant effects
Dobbie and Fryer (2011)	USA (Harlem)	Harlem Children's Zone Promise Academy vs. others	Primary & Middle	3	Math and ELA scores	0.28 SD (Math); 0.06 SD (ELA)
Dobbie and Fryer (2015)	USA (Harlem)	Harlem Children's Zone Promise Academy vs. others	Primary & Middle	3	Math and ELA scores 6 years after enrollment	0.28 SD (Math); 0.12 SD (ELA)
Angrist et al. (2010)	USA (Massachusetts)	KIPP Lynn vs. local public school	Middle	1	Math and ELA scores	0.43 SD (Math); 0.15 SD (ELA)
Angrist et al. (2011)	USA (Boston)	Boston charters vs. traditional public	Secondary	3	Math and ELA scores	0.25SD (Math); 0.4 SD (ELA)
Angrist et al. (2016)	USA (Boston)	Boston charters vs. traditional public	Secondary	6	Math and ELA scores	0.18-0.22SD (Math); 0.11-0.15SD (ELA)
Curto & Fryer (2014)	USA (Washington DC)	SEED boarding charter vs. others	Secondary	1	Math and ELA scores	0.22SD (Math); 0.20 SD (English)
Grissmer et al. (2023)	USA (Denver, Colorado)	Core Knowledge charters vs. others	Primary & Middle	9	English, math, science	0.14SD(English); 0.08SD (Math); 0.16SD(Science)

Notes: the table reports ITT effects from other studies using credible causal designs – primarily lottery-based or regression discontinuity analyses of specific schools. ELA stands for English Language Achievements.

7. External validity

In this section, we assess the extent to which our findings may generalise to the broader free school sector and discuss how they sit in the context of some related literature.

To begin with, we compare the two schools under study to other secondary free schools operating at the time, using school-level characteristics and VE statements. Table AI-5 (Appendix I) shows that our schools display positive selection in terms of intake ability: average KS2 scores are higher than the sector average. However, differences in terms of other observable dimensions – such as pupil-to-teacher ratios, teacher pay, and inspectorate ratings – are more nuanced. Table AI-6 further highlights the distinctiveness of the two schools' educational philosophies. While a few other free schools share elements of the 'no excuses' approach, only one other school exhibits features resembling the 'classical liberal' model.[32]

Next, we investigate whether value-added models that adjust for a rich set of observables can replicate our quasi-experimental estimates. While the availability of application data and admission tiebreakers allowed us to estimate causal effects among applicants to the two free schools, we do not have access to similar data for other English free schools. In Appendix AI.2, we show that value-added models fail to replicate our quasi-experimental results, underscoring the limits of

[32] This is another high-profile free school: Michaela Community School (in Wembley, North-West London). The school is well-known for its regimented teaching philosophy and combines a 'no excuses' approach with concepts that are similar to those of the CL school in our taxonomy (such as knowledge rich and teacher-led learning).

selection-on-observables approaches in the context of a diversified pedagogical offer. From this perspective, our study serves as a rigorous, scalable proof-of-concept for evaluating horizontally differentiated school models.

Last, it is instructive to situate our findings within the literature on small-scale, school-specific interventions using credible designs. Table 9 summarises findings from ten such studies, mostly conducted in the US charter and exam school sectors.[33] Our estimated Intent To Treat (ITT) effects of being offered a seat at the two free schools (roughly 0.17 SD) are smaller or fall within the range of the effects reported in other lottery- or RD-based studies, such as KIPP Lynn (0.43 SD; Angrist et al., 2010) or Core Knowledge charter schools (0.14–0.16 SD; Grissmer et al., 2023).

Within England, our study is the first lottery-based evaluation of newly created autonomous secondary schools. These early-day 'pioneering' free schools are radically different from the early sponsored academies, studied by Eyles and Machin (2019), as they are not affiliated with a sponsor and do not benefit from the additional resources or expertise brought by such associations. Instead, free schools share many features with converter academies, although they are start-up schools. Available evidence shows that outstanding converter academies can enhance pupil achievement (see Bertoni et al., 2020).

However, methodological considerations caution against a straightforward comparison of these findings. Indeed, in this study, we have exploited lotteries and a distance-based RD design to study the causal effects of free schools among a set of applicants. Conversely, research on academies uses a 'legacy' difference-in-difference approach to by-pass selection issues. This approach pins down the impact of autonomy on pupils who chose these schools prior to their conversion to academy– and is unlikely to encompass match-specific effects that would be estimated among a pool of applicants. Although both sets of estimates are internally valid, they refer to sub-populations with different unobservable propensities to apply to, and attend, an autonomous school – and are therefore also likely to reflect different expected returns from autonomous schools.

Specifically, our estimates are internally valid for a narrow population – i.e., applicants who are lottery-randomized to receive an offer from the two oversubscribed schools (see Weiland et al., 2024, for a discussion on the potential and limits of lottery-based studies of school effects) – and cannot be assumed to generalize to the average student or to the broader free-school sector. Nonetheless, such local impact estimates remain valuable for policy as they highlight the potential for bottom-up school initiatives to deliver meaningful gains.

8. Conclusion

This paper provides novel causal evidence on the effectiveness of newly established English free schools by applying a research design inspired by the US charter school lottery literature. Leveraging admissions lotteries and distance-based regression discontinuity, we show that two early-adopter free schools significantly improved students' academic performance and reduced absenteeism, with similar magnitudes of impact across both schools.[34] These findings fill a gap in the English school-choice literature, where causal evidence on start-up schools remains scarce. In doing so, we also extend the US charter school research tradition – pioneered through small-sample studies of exceptional schools – by analysing a more diverse and less markedly deprived set of pupils, and by examining different educational models.

A key contribution of our study lies in examining the role of educational philosophy. We carry out a novel text-based comparison of school philosophies by analysing 'vision and ethos' statements scraped from the websites of the free schools and the counterfactual ones attended by unsuccessful applicants. This allows us to map key differences in pedagogical orientation between our case-study schools, their counterfactuals, and the broader sector. Although both schools champion a 'traditional education', we detect notable differences in their pedagogical approaches: one adheres to a 'no excuses' model while the other champions a 'classical liberal' philosophy. Both approaches are markedly different from the ones adopted by counterfactual schools; the classic liberal one is also relatively uncommon among other free schools.

Despite these differences, we find that the two free schools deliver comparable positive effects on student outcomes. These findings suggest that autonomous schools do not need to adopt one specific educational philosophy to be effective. Rather, they highlight that excellence in school outcomes can be achieved through different rigorous models when schools are granted sufficient autonomy. In this respect, our results provide nuance to the prevailing emphasis on the 'no excuses' philosophy in the charter-school literature (see Cheng et al., 2017) and align with emerging evidence that alternative models can support high performance.

While our analysis focuses on student-level outcomes from attending free schools, an important avenue for future research concerns the broader systemic effects of free-school entry on existing schools. The existing work on school competition within quasi-markets suggests that school choice may induce improvements in quality among incumbent providers, although effects are often modest and vary by context (Gilraine et al., 2021; Chen & Harris, 2023; Ridley & Terrier, 2025; Crema & Singleton, 2025). For our study context and period, such competitive responses are likely to be limited given that the policy was still in its infancy during the years we study, with few new schools opening per area. As the sector matures and more schools open, examining whether and how free schools influence local school markets would be a valuable extension to better understand the long-term implications of school choice reforms in England.

Declaration of competing interest

The authors declare that they have no relevant or material financial interests that relate to the research described in this paper. Dr Heller-Sahlgren has benefited from financial support for unrelated research projects from the Hans and Barbara Bergström Foundation, the Confederation of Swedish Enterprise, and the Swedish Free Enterprise Foundation, and has undertaken consultancy work for Swedish municipal and independent education providers. None of these organisations had any role in the funding, design, analysis, or reporting of this study. The results were cleared by Office for National Statistics (ONS) for disclosure of sensitive information. Our discussions do not represent the views of the ONS or any other interested party.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.econedurev.2025.102717.

References

Abdulkadiroğlu, A., Angrist, J. D., Dynarski, S. M., Kane, T. J., & Pathak, P. A. (2011). Accountability and Flexibility in Public Schools: Evidence from Boston's Charters and Pilots. *Quarterly Journal of Economics, 126*(2), 669–748.

Abdulkadiroğlu, A., Hu, W., Pathak, P. A. (2013). Small high schools and student achievement: Lottery-based evidence from New York City. NBER Working Paper No. 19576.

[33] Two relevant Dutch studies (Ruijs, 2017; and Oosterbeek et al, 2023), reporting effects of Montessori and academic vs. comprehensive schools, were excluded because they do not report ITT effects on test scores.

[34] Dobbie and Fryer (2020) show that charter schools have the potential to impact university and labour market outcomes. Currently, however, it is not meaningful to analyse longer-term outcomes of free school attendance: the earliest cohort would have completed university in 2020, and the latest in 2022. We thus believe it is better to wait a few years before analysing such outcomes in our context.

Abdulkadiroğlu, A., Angrist, J. D., Narita, Y., & Pathak, P. A. (2017). Research design meets market design: Using centralized assignment for impact evaluation. *Econometrica, 85*(5), 1373–1432.
Abdulkadiroğlu, A., Angrist, J. D., Narita, Y., & Pathak, P. A. (2021). Breaking Ties: Regression Discontinuity Design Meets Market Design. Forthcoming *Econometrica*.
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.
Andrews, J., & Johnes, R. (2017). *Free schools in england*. Education Policy Institute report.
Angrist, J. D., Dynarski, S. M., Kane, T. J., Pathak, P. A., & Walters, C. R. (2010). Inputs and Impacts in Charter Schools: KIPP Lynn. *American Economic Review, 100*(2), 239–243.
Angrist, J. D., Dynarski, S. M., Kane, T. J., Pathak, P. A., & Walters, C. R. (2012). Who Benefits from KIPP?. *The Journal of Policy Analysis and Management*. Fal.
Angrist, J. D., Hull, P., & Walters, C. (2023). *Methods for measuring school effectiveness in handbook of the economics of education, 7* pp. 1–60). Amsterdam: Elsevier.
Angrist, J. D., & Imbens, G. W. (1995). Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American Statistical Association, 90*(430), 431–442.
Angrist, J. D., Pathak, P. A., & Walters, C. R. (2013). Explaining Charter School Effectiveness. *American Economic Journal: Applied Economics, 5*(4), 1–27.
Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized Random Forests. *The Annals of Statistics, 47*(2), 1148–1178.
Bertoni, M., Gibbons, S., & Silva, O. (2020). School choice during a period of radical school reform. Evidence from academy conversion in England. *Economic Policy, 35* (104), 739–795.
Bertoni, M., Gonschorek, F., Klein, T., & Silva, O. (2022). *School types: evidence on the relative effectiveness of different school arrangements in england*. Mimeo: London School of Economics.
Biasi, B., & Ma,. S. (2022). The Education-Innovation Gap. NBER Working Paper No. 9853.
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.
Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica, 82*(6), 2295–2326.
Chen, F., & Harris, D. N. (2023). The market-level effects of charter schools on student outcomes: A national analysis of school districts. *Journal of Public Economics, 228*, 105015.
Cheng, A., Hitt, C., Kisida, B., & Mills, J. N. (2017). "No excuses" charter schools: A meta-analysis of the experimental evidence on student achievement. *Journal of School Choice, 11*(2), 209–238.
Chernozhukov, V., Demirer, M., Duflo, E., & Fernandez-Val, I. (2018). Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in India. NBER Working Paper No. 24678.
Cirin, R. (2014). *Do Academies Make Use of Their Autonomy*. London: Department for Education. Research Report 366.
Clark, D. (2009). The performance and competitive effects of school autonomy. *Journal of political Economy, 117*(4), 745–783.
Crema, A., & Singleton, J. D. (2025). The Spillover Impacts of Charter Schools. *Program on Education Policy and Governance Conference Papers Series. PEPG 25-14*. Program on Education Policy and Governance.
Cohodes, S., & Parham, K. (2021). Charter Schools' Effectiveness, Mechanisms, and Competitive Influence. Forthcoming *Oxford Research Encyclopedia of Economics and Finance*.
Cullen, J. B., Jacob, B. A., & Levitt, S. (2006). The Impact of School Choice on Participants: Evidence From Randomized Lotteries. *Econometrica, 75*(5), 1191–1230.
Curto, V. E., & Fryer, R. G., Jr. (2014). The Potential of Urban Boarding Schools for the Poor: Evidence from SEED. *Journal of Labor Economics, 32*(1).
Dobbie, W., & Fryer, R. G., Jr. (2011). Are High-Quality Schools Enough to Increase Achievement Among the Poor? Evidence from the Harlem Children's Zone. *American Economic Journal: Applied Economics, 3*(3), 158–187.
Dobbie, W., & Fryer, R. G. (2013). Getting Beneath the Veil of Effective Schools: Evidence from New York City. *American Economic Journal: Applied Economics, 5*(4), 28–60.
Dobbie, W., & Fryer, R. G. (2020). Charter Schools and Labor Market Outcomes. *Journal of Labor Economics, 38*(4), 915–957.
Dynarski, S., Hubbard,. D., Jacob,. B., & Robles,. S. (2018). Estimating the Effects of a Large For-Profit Charter School Operator. NBER Working Paper No. 24428.
Eyles, A., Machin, S., & Silva, O. (2017). Academies 2: The new batch. *Fiscal Studies, 39* (1), 121–158.
Eyles, A., & Machin, S. (2019). The Introduction of Academy Schools to England's Education. *Journal of the European Economic Association, 17*(4), 1107–1146.
Gibbons, S., Machin, S., & Silva, O. (2008). Competition, Choice and Pupil Achievement. *Journal of the European Economic Association, 6*(4), 912–947.
Gilraine, M., Petronijevic, U., & Singleton, J. D. (2021). Horizontal differentiation and the policy effect of charter schools. *American Economic Journal: Economic Policy, 13* (3), 239–276.
Green, F., Allen, R., & Jenkins, A. (2015). Are English free schools socially selective? A quantitative analysis. *British Educational Research Journal, 41*(6), 907–924.
Grissmer, D., White, T., Buddin, R., Berends, M., Willingham, D., DeCoster, J., & Evans, T. (2023). *A kindergarten lottery evaluation of core knowledge charter schools: should building general knowledge have a central role in educational and social science research and policy? edworkingpaper no. 23-755*. Annenberg Institute for School Reform at Brown University.
Heckman, J. J., & Vytlacil, E. (2005). Structural equations, treatment effects, and econometric policy evaluation. *Econometrica, 73*(3), 669–738.
Hoxby, C. (2000). Does Competition Among Public Schools Benefit Students and Taxpayers? *American Economic Review, 90*(5), 1209–1238.
Jackson, C. K. (2018). What do test scores miss? The importance of teacher effects on non–test score outcomes. *Journal of Political Economy, 126*(5), 2072–2107.
Ridley, M., & Terrier, C. (2025). Fiscal and education spillovers from charter school expansion. *Journal of Human Resources, 60*(4), 1356–1404.
Schwartz, A. E., Stiefel, L., & Wiswall, M. (2013). Do small schools improve performance in large, urban districts? Causal evidence from New York City. *Journal of Urban Economics, 77*, 27–40.
Weiland, C., Unterman, R., Dynarski, S., Abenavoli, R., Bloom, H., Braga, B., … Weixler, L. (2024). Lottery-Based Evaluations of Early Education Programs: Opportunities and Challenges for Building the Next Generation of Evidence. *AERA Open, 10*.23328584241231933.